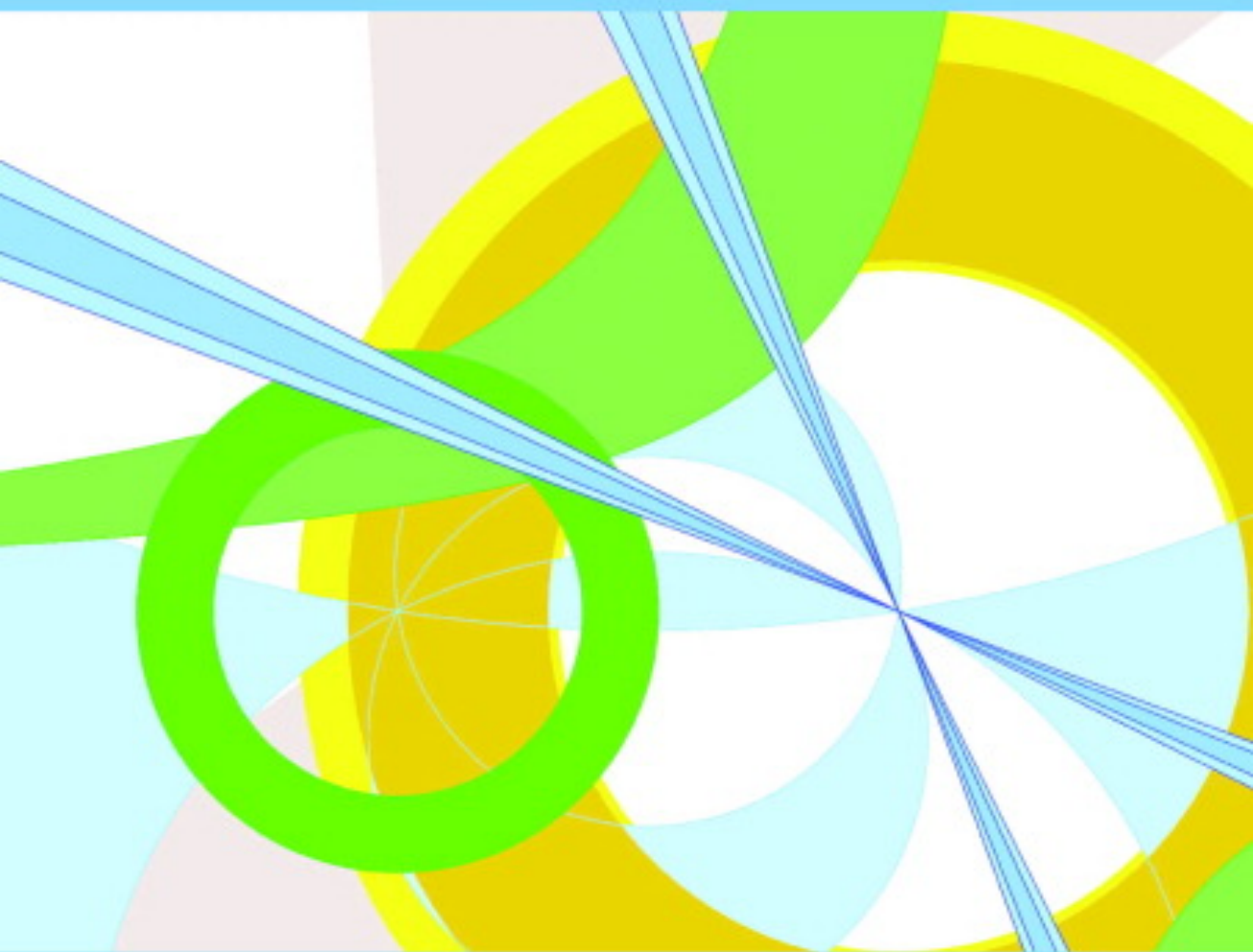


The **Standard Model**

A Primer



Cliff Burgess and **Guy Moore**

CAMBRIDGE

The Standard Model: A Primer

The standard model provides the modern understanding of all of the interactions of subatomic particles, except those due to gravity. The theory has emerged as the best distillation of decades of research.

This book uses the standard model as a vehicle for introducing quantum field theory. In doing this the book also introduces much of the phenomenology on which this model is based. The book uses a modern approach, emphasizing effective field theory techniques, and contains brief discussions of some of the main proposals for going beyond the standard model, such as seesaw neutrino masses, supersymmetry, and grand unification.

Requiring only a minimum of background material, this book is ideal for graduate students in theoretical and experimental particle physics. The book concentrates on getting students to the level of being able to use this theory by doing real calculations with the minimum of formal development. It does so without taking any shortcuts which would leave an incomplete understanding. The book contains several problems, with password-protected solutions available to lecturers at www.cambridge.org/9780521860369

GUY MOORE is Assistant Professor in the Department of Physics at McGill University, Canada. He received his Ph.D. from Princeton University and has held postdoctoral positions at McGill University and the University of Washington, Seattle.

CLIFF BURGESS is a professor in the Department of Physics and Astronomy at McMaster University in Canada, and an Associate Member at the Perimeter Institute for Theoretical Physics. He was made James McGill Professor of Physics at McGill University in 2003, and received a Killam Research Fellowship in 2005.

The Standard Model: A Primer

C. P. Burgess and Guy D. Moore

McGill University
Montreal



CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo
Cambridge University Press
The Edinburgh Building, Cambridge CB2 2RU, UK
Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521860369

© C. P. Burgess and G. D. Moore 2007

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2007

Printed in the United Kingdom at the University Press, Cambridge

A catalog record for this publication is available from the British Library

ISBN-13 978-0-521-86036-9 hardback
ISBN-10 0-521-86036-9 hardback

Cambridge University Press has no responsibility for the persistence or accuracy of
URLs for external or third-party Internet websites referred to in this publication,
and does not guarantee that any content on such websites is, or will remain,
accurate or appropriate.

Contents

<i>List of illustrations</i>	<i>page</i> viii
<i>List of tables</i>	ix
<i>Preface</i>	x
<i>Acknowledgments</i>	xv
Part I Theoretical framework	1
1 Field theory review	3
1.1 Hilbert space, creation and annihilation operators	3
1.2 General properties of interactions	7
1.3 Free field theory	15
1.4 Implications of symmetries	29
1.5 Renormalizable interactions	34
1.6 Some illustrative examples	38
1.7 Problems	46
2 The standard model: general features	53
2.1 Particle content	53
2.2 The Lagrangian	58
2.3 The perturbative spectrum	61
2.4 Interactions	71
2.5 Symmetry properties	84
2.6 Problems	104
3 Cross sections and lifetimes	111
3.1 Scattering states and the S -matrix	111
3.2 Time-dependent perturbation theory	115
3.3 Decay rates and cross sections	118

	Part II Applications: leptons	125
4	Elementary boson decays	127
4.1	Z^0 decay	127
4.2	W^\pm decays	142
4.3	Higgs decays	146
4.4	Problems	151
5	Leptonic weak interactions: decays	153
5.1	Qualitative features	153
5.2	The calculation	160
5.3	The large-mass expansion	163
5.4	Feynman rules	169
5.5	Problems	183
6	Leptonic weak interactions: collisions	188
6.1	The Mandelstam variables	189
6.2	e^+e^- annihilation: calculation	192
6.3	e^+e^- annihilation: applications	197
6.4	The Z boson resonance	202
6.5	t -channel processes: crossing symmetry	209
6.6	Interference: Møller scattering	212
6.7	Processes involving photons	214
6.8	Problems	223
7	Effective Lagrangians	231
7.1	Physics below M_W : the spectrum	232
7.2	The Fermi theory	233
7.3	Physics below M_W : qualitative features	239
7.4	Running couplings	242
7.5	Higgsless effective theory	257
7.6	Problems	269
	Part III Applications: hadrons	273
8	Hadrons and QCD	275
8.1	Qualitative features of the strong interactions	276
8.2	Heavy quarks	284
8.3	Light quarks	290
8.4	Problems	318
9	Hadronic interactions	320
9.1	Quasi-elastic scattering	321
9.2	Hard inelastic scattering: partons	329

9.3	Soft inelastic scattering: low-energy mesons	350
9.4	Neutral-meson mixing	367
9.5	Problems	386
Part IV Beyond the standard model		393
10	Neutrino masses	395
10.1	The kinematics of massive neutrinos	396
10.2	Neutrino oscillations	398
10.3	Neutrinoless double-beta decay	418
10.4	Gauge-invariant formulations	421
10.5	Problems	433
11	Open questions, proposed solutions	434
11.1	Effective theories (again)	435
11.2	Dimension zero: cosmological-constant problem	437
11.3	Dimension two: hierarchy problem	439
11.4	Dimension four: triviality, θ_{QCD} , flavor problems	455
11.5	Dimension six: baryon-number violation	471
11.6	Problems	481
Appendix A Experimental values for the parameters		484
Appendix B Symmetries and group theory review		487
Appendix C Lorentz group and the Dirac algebra		494
Appendix D ξ-gauge Feynman rules		507
Appendix E Metric convention conversion table		518
<i>Select bibliography</i>		525
<i>Index</i>		535

Illustrations

5.1	Differential τ , μ decay rate, as function of the charged lepton energy	page 167
5.2	The Feynman graph for $Z^0 \rightarrow f\bar{f}$.	182
5.3	The Feynman graph for the decay $\mu \rightarrow e\bar{\nu}\nu$.	183
6.1	The Feynman graphs for the process $e^+e^- \rightarrow f\bar{f}$.	193
6.2	Additional graphs for $e^+e^- \rightarrow e^+e^-$ and $e^+e^- \rightarrow \nu_e\bar{\nu}_e$.	193
6.3	The measured pair-production ratio, R_H .	199
6.4	Important corrections near $s = M_Z^2$.	203
6.5	The Feynman graph for $e^-f \rightarrow e^-f$.	209
6.6	The “uncrossed” and “crossed” graphs for e^-e^- scattering.	212
6.7	Feynman graphs for Compton scattering.	214
6.8	Diagram for scattering with a photon emission.	218
6.9	A radiative correction to e^+e^- annihilation.	221
6.10	e^+e^- hadronic cross section near the Z^0 resonance.	223
7.1	The tree graph that generates the Fermi Lagrangian.	235
7.2	Higher-order tree graphs that contribute to the effective Lagrangian.	238
7.3	One-loop scattering processes.	243
9.1	Kinematic variables for deep inelastic scattering.	331
9.2	The Drell–Yan process.	340
9.3	Heavy-quark production.	341
9.4	Proton–parton distribution functions at two scales.	349
9.5	Feynman graphs which dominate $\pi\pi$ scattering.	357
9.6	Feynman graphs dominating nucleon Noether currents.	363
9.7	Box diagrams for $\Delta S = \pm 2$ interactions.	375
9.8	Potentially large higher-order diagrams for $\Delta S = \pm 2$ interactions.	378
9.9	Diagrams introducing CP-violation into the $\Delta S = \pm 1$ interactions.	379
9.10	The “unitarity triangle.”	384
10.1	Helicity suppression in neutrino physics.	413
10.2	Two neutrino and neutrinoless double-beta decay.	419
10.3	Integrating out the right-handed neutrino.	429
10.4	Integrating out the Φ field.	431
11.1	$\Delta B \neq 0$ operator from a dimension-5 SUSY coupling.	475
11.2	Coupling unification in the MSSM.	480

Tables

2.1	Neutral-current charges of the fermions	<i>page</i> 84
4.1	Fermion neutral-current coupling constants	137
4.2	Computed and measured Z^0 branching fractions	138
4.3	Computed and measured W^+ branching fractions	146
7.1	Computed and measured W^\pm mass	252
8.1	Zero-point momenta for the heavier $q\bar{q}$ systems	285
8.2	C and P eigenvalues for quark–antiquark bound states	289
9.1	Theory vs. experiment for low-energy pion scattering	359
A.1	Particle masses	485
A.2	Numerical values of coupling constants	485
E.1	Metric-convention conversion table	521
E.2	Metric-convention conversion table for Feynman rules	522

Preface

The standard model of particle physics, developed in the 1960s and 1970s, has stood for 30 years as “the” theory of particle physics, passing numerous stringent tests. In fact, while many people believe that the standard model is not a complete description of particle physics, it is expected to be, at worst, incomplete rather than wrong; that is, the standard model is at worst a subset of the true theory of particle physics.

For this reason, a good working knowledge of the standard model and its phenomenology is essential for the modern particle physicist. The goal of this book is to provide all the tools for a working, quantitative knowledge of the standard model, with the minimum of formal developments. It presents everything needed to understand the particle spectrum of the standard model, and how to compute decay rates and cross sections at leading order in the weak coupling expansion (tree level). We assume a solid quantum-mechanics background, up to and including canonical quantization and the Dirac equation, but we do not assume familiarity with formal quantum field theory (renormalization, path integrals, generating functionals).

As we see it, this book fills two gaps in the existing literature. The first of these concerns the balance between theoretical sophistication and phenomenological utility. Most treatments of the standard model appear at the end of quantum field theory books. This is rational in the sense that the reader then has the complete set of tools to compute standard-model at the loop level. This approach has its merits; both authors learned the standard model in this way. Unfortunately, for many, especially experimental practitioners, the quantum field theory preliminaries may be too burdensome. Also, such books frequently do not present the standard model in complete detail, and they generally develop little of its phenomenology. The opposite style of approach is a more “cookbook” book, which introduces quantum field theory at the tree level, typically using electrodynamics as

an example, and again presents the standard model at the end. Generally these treatments are incomplete and abbreviated. The intention of this book is to be similar to the latter type of book, except that the presentation of the standard model is complete and contains a discussion of the model's phenomenology and a complete presentation of its Feynman rules.

Our philosophy is that it is important for a particle physicist to have a complete and quantitative knowledge of the standard model; indeed, for many, this is much more important than having a good background in formal quantum field theory. One cannot present the standard model in detail without *some* quantum field theory; but one can get surprisingly far without understanding the details of renormalization and loop effects. Of course, especially for theorists, a good knowledge of quantum field theory is also necessary; indeed, it should be obvious to the reader, at many points in the text, that more formal development is needed to compute to high accuracy. Knowing the material in this book may help the student of more formal quantum field theory by motivating and providing context for that study. Conversely, a student already proficient in quantum field theory can use this book as a succinct presentation of the standard model, and will have the tools to fill in the gaps left in the presentation, where loop corrections are required.

The second gap which we believe this book fills concerns the modern theoretical framework within which the standard model rests: the framework of *effective field theories*. Today we understand the theories we construct to describe nature – including the standard model – to be effective theories which capture the low-energy limit of some more fundamental, microscopic physics. Effective field theories capture a basic experimental fact: although nature comes to us with many scales, it can be understood one scale at a time. For instance, atomic physics can be understood with only limited knowledge of nuclei, and it can because short-distance physics tends to *decouple* from long-distance physics. In the modern understanding it is this observation which ultimately explains the otherwise puzzling requirement of renormalizability which our fundamental theories generally have. This book starts by using the standard model to build up the tools of effective field theory, by showing how and why scattering amplitudes simplify in the low-energy limit. Later chapters then exploit these tools to categorize the kinds of new physics which might ultimately replace the standard model, starting with a discussion of neutrino oscillations and ending with a broad survey of such new physics topics as supersymmetry and grand unified theories.

The first chapter of this book is devoted to introducing the field theory concepts we will need to present the standard model. We present the allowed

fields that can make up a quantum field theory (scalars, fermions, and gauge bosons), with particular emphasis on Majorana fermions and on the gauge principle, which appear to play especially important roles in the standard model. We introduce the required rules for formulating the theory's Lagrangian – the “basic principles,” such as Lorentz invariance, locality, unitarity, and renormalizability. We see what kinds of interactions are allowed, given the available fields and these basic principles. Then we give a few illustrative examples, including QED and QCD. Supplementary material on group theory, the Lorentz group, and spinors is provided in two appendices.

The second chapter introduces the standard model itself. We present the gauge group and the field content. The Lagrangian then follows as the most general Lagrangian consistent with these fields and with basic principles. This section then explores the consequences, determining the mass eigenstates and their interactions. We present in complete detail what the interaction Hamiltonian of the model is in the mass basis. We also briefly discuss the symmetries of the model, especially the accidental global symmetries of baryon and lepton number, and very briefly discuss anomalies and gauge anomaly cancellation.

The third chapter discusses the S matrix formalism in just enough detail to define and motivate decay rates and cross sections, and to show how they are to be computed in the interaction picture. Together, the first three chapters represent an introduction to the framework of the standard model.

Next, we start using the standard-model interactions to compute processes, introducing the needed technology as we go with the philosophy of “learning by doing” and using specific examples to figure out the patterns. We begin with the simplest processes in the standard model, the decays of heavy bosons, in Chapter 4. The rates of Z^0 , W^\pm , and Higgs-boson decays can be computed using interaction picture perturbation theory and an expansion of the fields in creation and annihilation operators, without much difficulty. In Chapter 5, where we consider the decays of leptons lighter than the W boson mass, we first encounter virtual intermediate particles, requiring the introduction of the propagator. After these examples it is possible to generalize the procedure for computing a decay process. This allows us to introduce the Feynman rules. Chapter 5 ends with a complete presentation of the unitary gauge Feynman rules of the standard model, sufficient for tree level analysis. (The R_ξ gauge Feynman rules appear in Appendix D.)

In Chapter 6 we address scattering processes, concentrating on fermion–fermion scattering. We discuss s -channel scattering in some length, especially near the Z^0 pole, where we first discover the necessity of including loop

corrections. We also introduce crossing symmetry and interference between diagrams, external photon states, and initial state radiation.

In Chapter 7 we introduce the notion of effective field theories, using the Fermi theory as the main example. This is especially important as the standard model itself is probably just an effective theory for some more inclusive theory, which is manifested at higher energies. We also present some of the most important results of loop corrections, particularly the running of gauge couplings with scale.

Chapter 8 begins the discussion of hadrons. We motivate why the running of couplings causes the confinement of quarks and gluons within hadrons, and we describe and motivate the spectrum of heavy-light and light-light mesons and of baryons, emphasizing the use of approximate symmetries.

Chapter 9 discusses hadronic interactions. It explains why both the low- and high-energy regimes are somewhat tractable, but the intermediate energy regime is not. We discuss deep inelastic scattering and the partonic structure of hadrons, up to and including the Altarelli–Parisi (DGLAP) equations. Then we discuss chiral perturbation theory, leptonic meson decays, and oscillation phenomena in the K and B meson systems.

The last part of the book gives a brief survey of what may lie beyond the standard model. We begin in Chapter 10 with a discussion of neutrino masses. Technically, these cannot lie beyond the standard model, because they have been observed, and the meaning of the standard model must be enlarged to accommodate them. However, as we discuss, there are two viable ways to do so, Majorana neutrino masses and Dirac neutrino masses, and we do not (yet) know which is correct. We discuss the Majorana possibility at some length in the context of non-renormalizable field theories. We discuss oscillation phenomena in some length, including the MSW effect, and briefly cover neutrinoless double beta decay. We also give examples of high-energy physics that could lead to the non-renormalizable operator responsible for Majorana neutrino masses.

Finally, Chapter 11 discusses what *may* lie beyond the standard model. We organize this material in terms of problems with the standard model, which can in turn be organized in terms of the dimensionality of the operator presenting the problem. The hierarchy problem appears because of the dimension-2 Higgs mass term, and may be solved by supersymmetry. The strong CP problem appears because of the dimension-4 Θ term in QCD, and may be solved by the axion mechanism. The baryon-number conservation “problem” (opportunity) arises because of the possibility of dimension-6 operators in the standard model; these might arise at an interesting level within grand unified theories.

Throughout this book we have used the metric convention, $Z_{\mu\nu} = \text{Diag}[-1 + 1 + 1 + 1]$, which is the less common convention within the phenomenology community. However, to ease the text's use, we present in Appendix E a clear discussion of how to convert between conventions, culminating in a metric convention conversion table.

In our experience it is possible to cover most of this book in a high-paced, one-semester first-year graduate level course. To do so, it is necessary to shave some corners. Most of Chapter 1, and Chapter 2 through Section 2.4, are essential, but Section 2.5 can be skipped without too much loss to the continuity. Similarly Section 4.2 and Section 4.3 can be given as problems instead of covered as sections. Chapter 5 and Chapter 6 should be covered in full, but then material from the remaining chapters can be picked and chosen as time and interest allow. The material in Chapter 10 does not rely on Chapter 8 or Chapter 9. A full year course should quite easily be able to cover all of the material in this book.

Acknowledgments

In writing this book we have drawn heavily on the insight, goodwill, and friendship of many people. In particular, we wish to thank our teachers of field theory – Bryce De Witt, Willy Fischler, Joe Polchinski, Curt Callan, David Gross, Larry Yaffe, and, especially, Steven Weinberg – for shaping the way we think about this subject.

Collaborators and students too numerous to name have continued to help deepen our understanding in the course of a lifetime of conversations about physics. Special thanks go to Joaquim Matias, Fernando Quevedo, and Kai Zuber for their comments on parts of early drafts.

Most importantly, we thank our families (Caroline, Andrew, Ian, Matthew, Michael, Clara, Bettina) for their continuing support and their tolerance for the time taken away from them for writing.

Part I

Theoretical framework

1

Field theory review

Quantum field theory is the language in terms of which the laws of physics are cast, and so we start with a whirlwind summary of some of its main features. Interspersed amongst the introductory topics in this chapter we also discuss some of the more general features that are usually demanded of any reasonable field-theoretic description of nature.

1.1 Hilbert space, creation and annihilation operators

Quantum field theories are special kinds of quantum mechanical theories which describe the behavior of particles. As quantum mechanical theories, their most basic objects are the Hilbert space of possible states \mathcal{H} , and the Hamiltonian H which describes time evolution in that Hilbert space.

The possible kinds of states are zero-particle states, one-particle states, two-particle states, and so on. Therefore, the Hilbert space in which all operators live is the tensor product of the zero-particle space with the one-particle space with the two-particle space, and so on:

$$\mathcal{H} = \mathcal{H}_0 \otimes \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \quad (1.1)$$

Here

$$\mathcal{H}_0 = \{|0\rangle\} \quad (1.2)$$

denotes the one-dimensional space spanned by the zero-particle state or vacuum: $|0\rangle$.

$$\mathcal{H}_1 = \{|\mathbf{p}, k\rangle\} \quad (1.3)$$

is similarly the span of all one-particle states with the basis states chosen to be eigenstates of linear momentum. Here \mathbf{p} represents the momentum of a state, and k denotes all of the other particle labels.

The space of N -particle states is constructed as the tensor product of N copies of the one-particle space. For instance, \mathcal{H}_2 is the set of all two-particle states,

$$\mathcal{H}_2 = \{|\mathbf{p}_1, k_1; \mathbf{p}_2, k_2\rangle = \pm|\mathbf{p}_2, k_2; \mathbf{p}_1, k_1\rangle\} \quad (1.4)$$

etc. The sign, \pm , is $+$ for bosons and $-$ for fermions. A Hilbert space constructed in this way is conventionally referred to as a Fock space.

It is convenient to express the operators that act within this space in terms of a basic set of *creation* and *annihilation* operators in the following way. The *annihilation operator*, $a_{\mathbf{p}k}$, is the operator that removes the particle with quantum numbers \mathbf{p} and k from a given state. If the state on which $a_{\mathbf{p}k}$ acts does not contain the particle in question then the operator is defined to give zero. That is,

$$\begin{aligned} a_{\mathbf{p}k}|0\rangle &= 0 \\ a_{\mathbf{p}k}|\mathbf{q}, l\rangle &= 2E_{\mathbf{p}}(2\pi)^3\delta^3(\mathbf{p} - \mathbf{q})\delta_{kl}|0\rangle \\ a_{\mathbf{p}k}|\mathbf{q}, l; \mathbf{k}, m\rangle &= 2E_{\mathbf{p}}(2\pi)^3\delta^3(\mathbf{p} - \mathbf{q})\delta_{kl}|\mathbf{k}, m\rangle \\ &\quad \pm 2E_{\mathbf{p}}(2\pi)^3\delta^3(\mathbf{p} - \mathbf{k})\delta_{km}|\mathbf{q}, l\rangle \end{aligned} \quad (1.5)$$

and so on. Here, $E_{\mathbf{p}}$ is the energy of a particle of spatial momentum \mathbf{p} , namely, $\sqrt{\mathbf{p}^2 + m^2}$, with m the mass of a particle with labels k . The sign in this last result is \pm according to the statistics of particles $|\mathbf{p}, k\rangle$ and $|\mathbf{q}, l\rangle$. Here and throughout, we use units for which $\hbar = C = 1$. The normalization is chosen to make Lorentz invariance more manifest, as discussed below.

This definition implies that the Hermitian conjugate, $a_{\mathbf{p}k}^*$, of $a_{\mathbf{p}k}$ is a *creation operator* for the same particle type; i.e.

$$a_{\mathbf{p}i}^*|0\rangle = |\mathbf{p}, i\rangle \quad (1.6)$$

$$a_{\mathbf{p}i}^*|\mathbf{q}, j\rangle = |\mathbf{p}, i; \mathbf{q}, j\rangle \quad (1.7)$$

etc. (Our notation is to use an asterisk for complex conjugation of c -numbers and Hermitian conjugation of operators, and to reserve a dagger, † , for Hermitian conjugation of matrices.)

These definitions, together with the normalization convention

$$\langle\mathbf{p}, i|\mathbf{q}, j\rangle = 2E_{\mathbf{p}}(2\pi)^3\delta^3(\mathbf{p} - \mathbf{q})\delta_{ij} \quad (1.8)$$

imply the following properties. For bosons,

$$|\mathbf{p}, i; \mathbf{q}, j\rangle = |\mathbf{q}, j; \mathbf{p}, i\rangle \quad (1.9)$$

$$[a_{\mathbf{p}i}, a_{\mathbf{q}j}] = [a_{\mathbf{p}i}^*, a_{\mathbf{q}j}^*] = 0 \quad (1.10)$$

$$[a_{\mathbf{p}i}, a_{\mathbf{q}j}^*] = 2E_{\mathbf{p}}(2\pi)^3\delta^3(\mathbf{p} - \mathbf{q})\delta_{ij} \quad (1.11)$$

and for fermions,

$$|\mathbf{p}, i; \mathbf{q}, j\rangle = -|\mathbf{q}, j; \mathbf{p}, i\rangle \quad (1.12)$$

$$\{a_{\mathbf{p}i}, a_{\mathbf{q}j}\} = \{a_{\mathbf{p}i}^*, a_{\mathbf{q}j}^*\} = 0 \quad (1.13)$$

$$\{a_{\mathbf{p}i}, a_{\mathbf{q}j}^*\} = 2E_{\mathbf{p}}(2\pi)^3 \delta^3(\mathbf{p} - \mathbf{q}) \delta_{ij} \quad (1.14)$$

in which $[A, B] = AB - BA$ and $\{A, B\} = AB + BA$.

A few comments are in order about the field normalizations above. First, note that momentum integrations $dp/2\pi$ always have factors of 2π in the denominator, and momentum delta functions $2\pi\delta(p - q)$ always have factors of 2π multiplying them. Following these rules,

- momentum space and energy integrations always involve $\int d^3\mathbf{p}/(2\pi)^3$, $\int dE/2\pi$;
- delta functions are always of form $(2\pi)^3\delta^3(\mathbf{p}-\mathbf{q})$ or $(2\pi)\delta(E_1 - E_2)$,

accounts for all 2π factors we will ever encounter.

Second, the momentum delta functions we have written are accompanied by factors of $2E_{\mathbf{p}}$, and the same $2E_{\mathbf{p}}$ appears in the denominator in momentum integrations. This normalization, called relativistic normalization, is convenient in a Lorentz invariant theory, because it makes it easier to make Lorentz invariance manifest. Note in particular, that

$$\int \frac{d^3\mathbf{p}}{2E_{\mathbf{p}}(2\pi)^3} = \int \frac{d^4p}{(2\pi)^4} 2\pi\delta(p^2 + m^2)\theta(p^0) \quad (1.15)$$

which is manifestly Lorentz invariant. [Note that our metric convention is that $\eta_{\mu\nu} = \text{Diag}[-1, 1, 1, 1]$, so $p^2 \equiv Z_{\mu\nu}p^\mu p^\nu = -(p^0)^2 + \mathbf{p}^2$.] This expression can be verified by performing the p^0 integration, using the δ function. Its Lorentz invariance is not quite manifest, since the step function $\theta(p^0)$ does not look invariant, as it refers to the time component; but the $2\pi\delta(p^2 + m^2)$ forces p^μ to be timelike for $m^2 > 0$ and lightlike for $m^2 = 0$, which ensures that the sign of p^0 does not change under (orthochronous) Lorentz transformations. Throughout this book, whenever there is an integral $\int d^3p/(2\pi)^3 2E_{\mathbf{p}}$, we will always implicitly define $p^0 = E_{\mathbf{p}}$ inside the integral.

The fundamental claim now to be made is that *any* operator acting on our Hilbert space, \mathcal{H} , can be written as a linear combination of monomials of the a s and a^* s; i.e.,

$$\mathcal{O} = A_{0,0} + \sum_i \int \frac{d^3p}{2E_{\mathbf{p}}(2\pi)^3} [A_{0,1}(\mathbf{p}, i) a_{\mathbf{p}i} + A_{1,0}(\mathbf{p}, i) a_{\mathbf{p}i}^*] \quad (1.16)$$

$$+ \sum_{ij} \int \frac{d^3p d^3q}{4E_{\mathbf{p}}E_{\mathbf{q}}(2\pi)^6} [A_{0,2}(\mathbf{p}, i; \mathbf{q}, j) a_{\mathbf{p}i} a_{\mathbf{q}j} + A_{1,1}(\mathbf{p}, i; \mathbf{q}, j) a_{\mathbf{p}i}^* a_{\mathbf{q}j} \\ + A_{2,0}(\mathbf{p}, i; \mathbf{q}, j) a_{\mathbf{p}i}^* a_{\mathbf{q}j}^*] + \dots \quad (1.17)$$

The operators, \mathcal{O} , are in one-to-one correspondence with the coefficient functions $\{A_{0,0}, A_{1,0}(\mathbf{p}, i), A_{0,1}(\mathbf{p}, i), \dots\}$. This can be shown inductively by explicitly solving for these coefficients in terms of the matrix elements of \mathcal{O} : $\langle\psi|\mathcal{O}|\phi\rangle$. For example $\langle 0|\mathcal{O}|0\rangle = A_{0,0}$, $\langle 0|\mathcal{O}|\mathbf{p}, i\rangle = A_{0,1}(\mathbf{p}, i)$, and so on.

In particular, the Hamiltonian for a system of free particles has a simple expression in terms of the a s and a^* s:

$$H_0 = E_0 + \sum_i \int \frac{d^3p}{2E_{\mathbf{p}}(2\pi)^3} \varepsilon(\mathbf{p}, i) a_{\mathbf{p}i}^* a_{\mathbf{p}i}. \quad (1.18)$$

To learn the interpretation of the coefficients E_0 and $\varepsilon(\mathbf{p}, i)$, calculate the action of H_0 on various states. On the vacuum H_0 gives

$$H_0|0\rangle = E_0|0\rangle \quad (1.19)$$

since $a_{\mathbf{p}i}|0\rangle = 0$. E_0 is clearly the energy of the no-particle state $|0\rangle$, i.e. the vacuum energy. Similarly,

$$H_0|\mathbf{q}, j\rangle = [E_0 + \varepsilon(\mathbf{q}, j)]|\mathbf{q}, j\rangle \quad (1.20)$$

and

$$H_0|\mathbf{q}_1, j_1; \dots; \mathbf{q}_N, j_N\rangle = \left[E_0 + \sum_{k=1}^N \varepsilon(\mathbf{q}_k, j_k) \right] |\mathbf{q}_1, j_1; \dots; \mathbf{q}_N, j_N\rangle \quad (1.21)$$

etc. The many-particle momentum eigenstates, $|\mathbf{q}_1, j_1; \dots; \mathbf{q}_N, j_N\rangle$ are also eigenstates of the energy, H_0 , with eigenvalue

$$E = E_0 + \sum_{k=1}^N \varepsilon(\mathbf{q}_k, j_k). \quad (1.22)$$

This implies that the energy of a single-particle state $|\mathbf{p}, i\rangle$ relative to the vacuum is $\varepsilon(\mathbf{p}, i)$. Relativistic kinematics then determines the momentum-dependence of ε on \mathbf{p} as

$$\varepsilon(\mathbf{p}, i) = \sqrt{\mathbf{p}^2 + m_i^2} = E_{\mathbf{p}} \quad (1.23)$$

where m_i is the mass of particle type i . Notice that the energy of a many-particle state relative to the vacuum is just the sum of the single-particle energies, showing that the particles described by H_0 do not interact.

We emphasize that this is a special property of free field theories; in general, even if single-particle states are eigenstates of the Hamiltonian, many-particle states are in general not eigenstates of the Hamiltonian. This means that they can undergo non-trivial time evolution. Indeed, almost all interesting phenomena in particle physics arise from the fact that many-particle states are not eigenstates of the Hamiltonian.

1.2 General properties of interactions

We are interested in writing down a Hamiltonian

$$H = H_0 + H_{\text{int}} \quad (1.24)$$

that describes the interactions of the particles we know. The present section is devoted to summarizing the minimal requirements for a physically reasonable theory. These properties translate into a set of restrictions on what form will be allowed for H . The purpose of this process is to arrive at the general class of theories from which the standard model is to be chosen. Being aware of the alternatives available gives some feeling for which features may be changed and which are inviolable.

We now return to a statement of these requirements. A sketch of their justification is given in the next subsection, but for a complete discussion the reader should consult a field theory text.

1.2.1 Physical constraints on H

The basic principles we demand of any candidate physical theory are:

- (i) Unitarity: (i.e. conservation of probability)

The requirement here is to ensure that time evolution preserve the property that the sum of probabilities over all mutually exclusive events gives one. This requires that the time-evolution operator

$$U = e^{-iHt} \quad (1.25)$$

be unitary. Equivalently the Hamiltonian must be Hermitia:

$$H = H^*. \quad (1.26)$$

- (ii) Cluster decomposition: (i.e. locality)

This requirement is that physics be independent at different points in space at a given time. Specifically we require that amplitudes (and hence probabilities) for events that are well separated from one another factorize into a product of independent amplitudes. Such a factorization is what would be expected for statistically independent events.

The condition that physics at spatially separated positions be independent comes in two parts. The first is that physical observables must commute at spatially separated points and the second is that time evolution must preserve this property. We consider each of these in turn:

(a) Microcausality

The first condition is to require that physical observables may be separately measurable at different positions and equal times. In a quantum theory we must therefore demand that all physical observables commute at space-like separations. That is:

$$[A(x), B(y)] = 0 \quad \text{for} \quad (x - y)^2 > 0. \quad (1.27)$$

Condition (1.27) is sometimes referred to as the requirement of microcausality.

(b) Locality

We next require that this property, that spatially separated physical amplitudes must factorize, be preserved by time evolution, provided, of course, that no physical signals propagate from one point to the other. Since the time-evolution operator, Eq. (1.25), is the exponential of the Hamiltonian, the property that it factorizes turns out to require that the Hamiltonian should be the sum of those for each of the spatially separated regions. The Hamiltonian must therefore have the form

$$H = \int d^3x \mathcal{H}(\mathbf{x}, t) \quad (1.28)$$

which boils down to requiring that the total energy be a sum of the energy of the degrees of freedom at each point. This is consistent with the intuition that the degrees of freedom at each point of space at a given time are independent, since the total energy for a set of independent systems is the sum of the energies of the independent constituents.

(iii) Invariance under Lorentz transformations and translations (Poincaré invariance)

Here we build in the requirements of special relativity and translation invariance in space and time. In quantum mechanics this implies the existence of corresponding conserved charges, P^μ and $J^{\mu\nu} = -J^{\nu\mu}$ (with $\mu, \nu = 0, 1, 2, 3$), representing four-momentum and angular momentum respectively. In particular, the total energy is given by

$$H = P^0$$

The particle states transform under unitary representations of the Poincaré group given by the operators:

$$U(a, \omega) = \exp \left[-ia_\mu P^\mu + \frac{i}{2} \omega_{\mu\nu} J^{\mu\nu} \right] \quad (1.29)$$

generated by these conserved charges. The states, $|\mathbf{p}, \sigma, j\rangle$, may then be labelled by their three-momenta, \mathbf{p} , mass, m , total spin, s , and spin-projection, σ , together with any other internal labels, j . The labels m and s are generally not explicitly indicated.

The Minkowski-space conventions used in what follows are:

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \quad (1.30)$$

$$P^\mu = (E, \mathbf{p}), \quad \text{and} \quad P_\mu = (-E, \mathbf{p}) \quad (1.31)$$

$$x^\mu = (t, \mathbf{x}), \quad \text{and} \quad x_\mu = (-t, \mathbf{x}) \quad (1.32)$$

$$\epsilon^{0123} = +1 \quad (1.33)$$

implying that the invariant product $x^2 = -(x^0)^2 + \mathbf{x}^2$ is negative for timelike vectors and positive for space-like vectors. We provide a review of Lorentz symmetry in Appendix C.

(iv) Stability:

The final condition to be imposed is that the spectrum of H be bounded from below. This is necessary if the vacuum state, defined as the state of lowest energy, is to exist.

1.2.2 Renormalizability

A further condition to be imposed on the standard model that is not as fundamental as those just described is the requirement of renormalizability. In fact, perfectly good theories, such as general relativity, are not renormalizable and yet are still very successful at accounting for experiments. Some explanation is therefore required to justify this demand.

The physical motivation comes from the idea that physical theories generically come with an implicit minimum distance, d , (or maximum energy, Λ) beyond which they are not meant to apply. For example, the quantum electrodynamics of electrons and photons is only physically correct up to an energy of twice the mass of the lightest particle that is heavier than the electron: $\Lambda = 2m_\mu$, i.e. twice the muon mass. At energies higher than this, muons can no longer be neglected, since they can be pair-produced in the final state even if they are not present initially. The correct theory for physics at energies above Λ becomes the quantum electrodynamics of photons, electrons and muons. This theory is in turn only valid up to the next threshold, the pion mass, and so on.

Classically, it is not important to specify this “cut-off” carefully. In a quantum theory, however, since all states can contribute to any given process as intermediate (or “virtual”) particles, any quantum calculation will depend explicitly on the cut-off scale, Λ . This may be seen, for example, by considering the expression, in time-independent perturbation theory, for the quadratic energy shift due to a perturbing Hamiltonian,

$$\delta E_\psi = \sum_n' \frac{|\langle \psi | H | n \rangle|^2}{E_\psi - E_n} \quad |n\rangle \neq |\psi\rangle \quad (1.34)$$

Clearly any state, $|n\rangle$, contributes to Eq. (1.34) regardless of its energy. Given our ignorance of the spectrum above the energy Λ , it only makes sense to include those states with energy less than Λ in this sum. The result therefore depends explicitly on Λ in a potentially complicated way.

If detailed knowledge of physics at the Λ scale is necessary in order to calculate probability amplitudes for processes at energies lower than Λ , then the theory is called *non-renormalizable*. These theories have less predictive power, since predictions depend on physics at the scale Λ , about which we are by assumption quite ignorant.

In *renormalizable* theories, on the other hand, Λ only appears in physical predictions (for large Λ) through a small number of parameters, such as the masses and charges of some or all of the particles involved. All other processes may then be computed in terms of these parameters. Once the few Incalculable parameters are taken from experiment, definite predictions may be made.

Whether or not a renormalizable theory should be expected to describe a given system depends therefore on the properties of the system. Physically, successful description in terms of a renormalizable theory is equivalent to the statement that the physics of interest, at energies $E \ll \Lambda$, is largely insensitive to the higher-energy physics appropriate to the scale Λ . In general, a renormalizable description of the physics at an energy E is justified to the extent that contributions of order E/Λ are not important. Otherwise non-renormalizable interactions must be included.

As an example, consider the theory describing the energy levels of the hydrogen atom. Neglecting the structure of the nucleus, this theory is given by the quantum electrodynamics of pointlike electrons, protons, and photons. Ignoring nuclear structure (such as the proton magnetic moment) means neglecting powers of $E_{\text{atom}}/M_{\text{proton}}$, and the resulting theory is renormalizable. Within this theory atomic physics depends only on the electron and proton mass and charge. If we demand accuracy higher than $E_{\text{atom}}/M_{\text{proton}}$,

the proton structure cannot be ignored, leading to a non-renormalizable description.

An example of a situation for which no renormalizable theory should be expected is provided by the theory describing the nuclear scattering of the deuteron. Suppose that in this theory we wish to ignore the fact that the deuteron consists of a proton and neutron bound by these same nuclear interactions, instead taking the deuteron as a point particle. The corresponding theory that describes the scattering data cannot be renormalizable. This reflects the fact that in this case the scale, Λ , of the physics being neglected (the nuclear binding) and the scale, E , of the physics being studied (the nuclear scattering) are essentially the same. Non-renormalizability is the theory's way of telling us that effects of order E/Λ cannot be neglected.

Turning this argument around, we can use the renormalizability of a theory to tell us what the next scale, Λ , of new physics is. If we succeed in describing all data at presently accessible energies, E , in terms of a renormalizable theory then we learn that the scale of any new physics can be large: $\Lambda \gg E$. If a non-renormalizable theory is required, we learn that we are still missing some fundamental ingredients.

This physical picture implies that renormalizability is the minimal criterion for a theory which purports to describe *all* of the physics appropriate to any given scale. Demanding renormalizability for the standard model then amounts to the assumption that no hitherto unknown particles or interactions are required to understand present experiments. As judged by the splendid success of the standard model, this turns out to be a fairly good assumption. The sole exception (at the time of this writing) is the physics of neutrino oscillations, which appears to demand new physics; this can be understood within the standard model as the existence of non-renormalizable interactions. We return to this point at some length in Chapter 10 (and more generally to the issue of renormalizability and high-dimension operators in Chapter 7). Note, for the current purposes, that the scale required to explain neutrino masses is $\Lambda \sim 10^{14}$ GeV. This is so much higher than the intrinsic scales in the electroweak theory that, if the standard model is correct up to this scale, there are virtually no other consequences of the high-energy physics expected, and therefore we are (otherwise) very well justified in treating the standard model as a renormalizable theory.

1.2.3 Canonical quantization

We now turn to the problem of how to ensure that a given set of interactions incorporates the properties listed above. The most efficient way to do so is

to set up the formalism in terms of the action

$$S = \int L(t) dt \quad (1.35)$$

rather than the Hamiltonian. The conditions listed above for H then become relatively simple conditions for S .

H is related to S by the usual canonical methods. That is, given a set of physical variables q^i and a Lagrangian, $L(q, \dot{q})$, define the canonical momenta by

$$p_i = \frac{\partial L}{\partial \dot{q}^i} \quad (1.36)$$

The Hamiltonian, H , is then given by

$$H = \sum_i p_i \dot{q}^i - L \quad (1.37)$$

In this last expression, Eq. (1.36) is supposed to be inverted to allow the elimination of \dot{q}^i in favor of p_i . The formalism may be generalized in the case when this cannot be done, or when L depends on higher time-derivatives of q such as \ddot{q}^i etc.

We consider the implications for s of each of the properties of the previous sections in turn.

(i) Unitarity:

H is real provided that the action, S , is real.

(ii) Locality:

In order for H to be a local function,

$$H = \int d^3x \mathcal{H}(\mathbf{x}, t) \quad (1.38)$$

we require that L must also be expressed as an integral over a *Lagrangian density*:

$$L = \int d^3x \mathcal{L}(\mathbf{x}, t) \quad (1.39)$$

$$\text{so} \quad S = \int d^4x \mathcal{L}(\mathbf{x}, t) \quad (1.40)$$

It is a customary abuse of language in quantum field theory to refer to the Lagrangian density as the Lagrangian.

Recall that \mathcal{H} and \mathcal{L} , like any operators, are to be expressed in terms of the creation and annihilation operators, $a_{\mathbf{p}i}$ and $a_{\mathbf{p}i}^*$. But \mathcal{H} and \mathcal{L} are built of operators at a single spacetime point, which means

that they must be built from the Fourier transforms of $a_{\mathbf{p}i}$ and $a_{\mathbf{p}i}^*$:

$$A_\alpha(\mathbf{x}, t) = \sum_k \int \frac{d^3\mathbf{p}}{2E_{\mathbf{p}}(2\pi)^3} u_\alpha(\mathbf{p}, k) a_{\mathbf{p}k} e^{ipx} \quad (1.41)$$

In this equation $px = p^\mu x_\mu = -p^0 x^0 + \mathbf{p} \cdot \mathbf{x}$, with $p^0 = E_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2}$. α denotes any labels that distinguish the fields due to one particle type from another. The coefficients $u_\alpha(\mathbf{p}, k)$ ensure that both sides of the equation transform the same way under Lorentz transformations. The Lagrange density then becomes

$$\mathcal{L} = \mathcal{L}(x, A_\alpha(x), \partial_\mu A_\alpha(x), \dots) \quad (1.42)$$

We return to the related consequences of causality after first considering Poincaré invariance.

(iii) Translation invariance:

Translation invariance implies that \mathcal{L} depends on the spacetime coordinates \mathbf{x} and t only implicitly through its dependence on $A_\alpha(x)$ and its derivatives:

$$\mathcal{L}(x, A_\alpha(x), \partial_\mu A_\alpha(x), \dots) = \mathcal{L}(A_\alpha(x), \partial_\mu A_\alpha(x), \dots) \quad (1.43)$$

(iv) Lorentz invariance:

Noether's theorem (see Subsection 1.4.2) allows the construction of the conserved charges P^μ and $J^{\mu\nu}$ provided that the action, S , is invariant under Poincaré transformations (unless there is an anomaly, see Subsection 2.5.3). From Eq. (1.40) this implies that \mathcal{L} must be constructed out of the $A_\alpha(x)$ in such a way as to be a Lorentz scalar. In order to do so it is convenient to choose the fields, $A_\alpha(x)$, to transform in (finite-dimensional) representations of the Lorentz group

$$U(\omega) A_\alpha(x) U(\omega)^* = D_{\alpha\beta} A_\beta(\exp[\omega] \cdot x) \quad (1.44)$$

This, together with the transformation law for the single-particle states, determines the coefficients, $u_\alpha(\mathbf{p}, k)$ appearing in Eq. (1.41). This is the main topic of Section 1.3. \mathcal{L} must then be constructed from various combinations of the fields, their derivatives and the invariant tensors $\eta_{\mu\nu}$ and $\epsilon_{\mu\nu\lambda\rho}$.

(v) Causality:

Causality implies that bilinears of fields, such as the Hamiltonian density, must commute at spacelike separations. This is a strong condition, since the fields defined by Eq. (1.41) satisfy

$$[A_\alpha(\mathbf{x}, t), A_\alpha^*(\mathbf{y}, t)] \neq 0 \quad (1.45)$$

Causality is ensured provided that, for each particle, there exists another particle (its antiparticle) of equal mass and spin, described by the field

$$B_\alpha(x) = \sum_k \int \frac{d^3p}{2E_{\mathbf{p}}(2\pi)^3} v_\alpha(\mathbf{p}, k) b_{\mathbf{p}k} e^{ipx} \quad (1.46)$$

\mathcal{L} must depend on the fields $A(x)$ and $B(x)$ only through the combination

$$\phi_\alpha(x) = A_\alpha(x) + \xi B_\alpha^*(x) \quad (1.47)$$

in which ξ is a phase, since in this case

$$[\phi_\alpha(\mathbf{x}, t), \phi_\alpha^*(\mathbf{y}, t)] = 0 \quad (1.48)$$

In general the antiparticle need not be distinct from the particle. If the particle and antiparticle are identical, $a_{\mathbf{p}k} = b_{\mathbf{p}k}$, then ξ can be chosen such that $\phi = \phi^*$.

This observation has three physical consequences.

- (a) Antiparticles exist and couple with a strength identical to particles. This is called *crossing symmetry*. Since H_{int} involves $a_{\mathbf{p}k}$ and $b_{\mathbf{p}k}$ only in the schematic combination $a_{\mathbf{p}k} + b_{\mathbf{p}k}^*$ there are *no* interactions that can conserve the total number of particles.
- (b) For fermions the fields must anticommute at spacelike separations. For general spins the condition that bilinears, such as H_0 , commute for space-like separations implies that integer-spin particles must be bosons and half-integer-spin particles must be fermions – the spin-statistics theorem.
- (c) The behavior of particles and antiparticles under symmetries such as parity or gauge transformations are related. In particular the electric charge of a particle is the opposite of that of the antiparticle.
- (vi) Stability:
The generalization of the canonical method to theories with higher time derivatives shows that the Hamiltonian is in this case generically linear in one of its variables. Such a Hamiltonian cannot be bounded from below. Stability then implies that the Lagrangian must be a function of at most one time derivative of the fields. In practice, this forbids the appearance of more than quadratic powers of derivatives of fields.

(vii) Renormalizability:

Renormalizability may be summarized as the requirement that all parameters that appear in the Lagrangian must have positive dimension in powers of mass. That is to say, if the operator \mathcal{O} appears in \mathcal{L} with a coefficient c :

$$\mathcal{L} = c\mathcal{O} \quad (1.49)$$

then c must have dimension M^d for $d \geq 0$. Since all of the constituents, $A_\alpha(x)$ and ∂_μ , of \mathcal{O} each have dimension M^p for $p > 0$ and \mathcal{L} has dimension M^4 , this severely limits the allowed interactions to only include operators for which $d \leq 4$. Generally, all such interactions which are consistent with the assumed symmetries must be included.

1.3 Free field theory

In this book we will generally be interested in theories which, at least on some energy scale, can be described in terms of *weakly coupled* particles; that is, by a Hamiltonian which is dominated by a “free theory” piece H_0 , with interactions H_I which can be treated by perturbation theory. The standard model turns out to be such a theory, and most of the tools we have available to study quantum field theories are based on this assumption. In most of this book we will only treat corrections to the free-theory approximation at the leading order, that is, at the lowest power in the interaction Hamiltonian H_I at which the phenomena of interest happen.

To proceed with this project we first need to see what the most general free field theories can look like. We focus on particles with spins zero through one since all known non-gravitational experiments appear to be describable in terms of these, and since renormalizability seems to require an interacting field theory to be composed of such particles.

Recall that the Hamiltonian for a system of free particles is given by

$$H_0 = E_0 + \sum_i \int \frac{d^3\mathbf{p}}{2E_{\mathbf{p}}(2\pi)^3} E_{\mathbf{p}} a_{\mathbf{p}i}^* a_{\mathbf{p}i} \quad (1.50)$$

which is quadratic in the operators $a_{\mathbf{p}i}$. We wish to construct the corresponding Lagrangian in terms of the fields, $A_\alpha(x)$. Since the fields are linear in the creation and annihilation operators the desired Lagrangian density, \mathcal{L}_0 , must also be at most quadratic in the A_α s and their derivatives.

The discussion will use properties of the Lorentz group, which are reviewed in Appendix C.

1.3.1 Spin-zero particles

Spin-zero particles are described by fields that transform as scalars under Lorentz transformations. That is,

$$U(\omega)\phi(x)U(\omega)^* = \phi(\Lambda \cdot x) \quad (1.51)$$

where $\Lambda^\mu{}_\nu = (\exp \omega)^\mu{}_\nu$ is a Lorentz-transformation matrix. In terms of creation and annihilation operators

$$\phi(x) = \int \frac{d^3\mathbf{p}}{2E_{\mathbf{p}}(2\pi)^3} \left[a_{\mathbf{p}} e^{ipx} + \bar{a}_{\mathbf{p}}^* e^{-ipx} \right] \quad (1.52)$$

in which

$$E_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2} \quad (1.53)$$

and $px = -E_{\mathbf{p}}x^0 + \mathbf{p} \cdot \mathbf{x}$. The field $\phi(x)$ has been chosen real, as may be done without loss of generality because any complex field can always be decomposed into its real and imaginary parts. The energy relation, Eq. (1.53), implies that the four-momentum p^μ satisfies

$$p^\mu p_\mu = -E_{\mathbf{p}}^2 + \mathbf{p}^2 = -m^2 \quad (1.54)$$

which becomes the Klein–Gordon equation

$$(-\partial_\mu \partial^\mu \phi + m^2 \phi) = 0 \quad (1.55)$$

in position space. In the canonical approach these conditions are derived as equations of motion from the action rather than the representation theory of the Poincaré group.

We now consider the most general possible theory of several scalars, and show that it always reduces to a set of independent scalars, with potentially different masses. Consider then, a system of N types of spinless particles. Such a system may be described in terms of N real fields, $\phi^i(x)$, with $i = 1, \dots, N$. The most general Lagrangian that is Poincaré invariant, involves only two time derivatives (stability), and is quadratic in these N fields, is

$$\mathcal{L}_0 = -\frac{1}{2}A_{ij}\partial_\mu\phi^i\partial^\mu\phi^j - \frac{1}{2}B_{ij}\phi^i\phi^j - C \quad (1.56)$$

A sum from 1 to N is implied over repeated indices.

A term such as

$$D_{ij}\phi^i\partial^\mu\partial_\mu\phi^j$$

is not included since it is equivalent to

$$-D_{ij}\partial^\mu\phi^i\partial_\mu\phi^j$$

after an integration by parts. This Lagrangian is real (unitarity) provided that the (symmetric) coefficients A_{ij} , B_{ij} and C all are.

The corresponding conjugate momentum and Hamiltonian are:

$$\pi_i(x) = \frac{\partial \mathcal{L}_0}{\partial \dot{\phi}^i} = A_{ij} \dot{\phi}^j(x) \quad (1.57)$$

$$\begin{aligned} \text{so} \quad \mathcal{H}_0 &= \pi_i \dot{\phi}^i - \mathcal{L}_0 \\ &= +\frac{1}{2} \left[A_{ij} \dot{\phi}^i \dot{\phi}^j + A_{ij} \nabla \phi^i \cdot \nabla \phi^j + B_{ij} \phi^i \phi^j \right] + C \end{aligned} \quad (1.58)$$

This Hamiltonian is bounded below provided that the matrices A_{ij} and B_{ij} are non-negative definite. We assume in what follows that A_{ij} is strictly positive definite, since there would otherwise be a particle without any kinetic energy.

There are considerably more parameters appearing in the Lagrangian, Eq. (1.56), than appeared in the Hamiltonian, Eq. (1.50). This is because many of the constants in Eq. (1.56) may be absorbed into redefinitions of the field variables by putting \mathcal{L}_0 into *canonical form*. Only linear transformations

$$\phi^i = M^i_j \phi'^j \equiv (M \phi')^i \quad (1.59)$$

need be considered since these are the only ones that ensure that \mathcal{L}_0 remains quadratic when expressed in terms of the new variable, ϕ'^j . We use this freedom to put A_{ij} and B_{ij} into standard form.

Since A_{ij} is assumed positive definite, its eigenvalues a_1, \dots, a_N are all positive and its square root and inverse exist. If we define the new fields ϕ'^i as

$$\phi^i = (A^{-1/2} \phi')^i \quad (1.60)$$

then \mathcal{L}_0 becomes

$$\mathcal{L}_0 = -\frac{1}{2} \partial_\mu \phi'^i \partial^\mu \phi'^i - \frac{1}{2} B'_{ij} \phi'^i \phi'^j - C \quad (1.61)$$

where

$$B'_{ij} \equiv (A^{-1/2} B A^{-1/2})_{ij} \quad (1.62)$$

This does not exhaust the freedom (1.59) to redefine fields. Indeed, the transformation $\phi' = \mathcal{O} \varphi$ in which $\mathcal{O}^T \mathcal{O} = I$ preserves the form (1.61). Recall now that any real symmetric matrix can be diagonalized by an orthogonal

transformation

$$\mathcal{O}^T B' \mathcal{O} = \begin{pmatrix} b_1 & & \\ & b_2 & \\ & & b_N \end{pmatrix} \quad (1.63)$$

with $b_k \geq 0$ from stability. The redefinition $\phi' = \mathcal{O}\phi$ with this \mathcal{O} then diagonalizes the mass matrix, B'_{ij} , giving:

$$\mathcal{L}_0 = -\frac{1}{2}\partial_\mu\phi^i\partial^\mu\phi^i - \frac{1}{2}b_i\phi^i\phi^i - C \quad (1.64)$$

$$\text{and } \mathcal{H}_0 = \frac{1}{2}\dot{\phi}^i\dot{\phi}^i + \frac{1}{2}(\nabla\phi^i) \cdot (\nabla\phi^i) + \frac{1}{2}b_i\phi^i\phi^i + C \quad (1.65)$$

Unless some of the eigenvalues of the matrix B_{ij} are degenerate, this exhausts our freedom to linearly redefine fields. Equation (1.64) is then the standard form for \mathcal{L}_0 . The equations of motion are

$$(-\partial^\mu\partial_\mu + b_i)\phi^i = 0 \quad (1.66)$$

The parameters appearing in \mathcal{L}_0 may be related to the physical vacuum energy, E_0 , and masses, m_i , by expressing the total Hamiltonian, Eq. (1.65), in terms of $a_{\mathbf{p}i}$ and comparing to Eq. (1.50):

$$\begin{aligned} H_0 &= \int d^3x \mathcal{H}(x) \\ &= E_0 + \sum_{i=1}^N \int \frac{d^3\mathbf{p}}{2E_{\mathbf{p}}(2\pi)^3} E_{\mathbf{p}} a_{\mathbf{p}i}^* a_{\mathbf{p}i} \end{aligned} \quad (1.67)$$

$$\text{with } E_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + b_i} \quad (1.68)$$

$$\text{and } E_0 = C \int d^3x + \sum_i \frac{1}{2} \int \frac{d^3\mathbf{p}}{(2\pi)^3} E_{\mathbf{p}} (2\pi)^3 \delta^3(0) \quad (1.69)$$

Clearly the eigenvalues $b_i = m_i^2$ give the square of the particle masses. The vacuum energy is more delicate since it diverges at both long and short distances. The long-distance divergence may be regularized by putting the system within a space of finite, but large, volume Ω . The divergence of E_0 as $\Omega \rightarrow \infty$ merely indicates that the total energy is not the quantity of physical interest, since the total energy is by construction an extensive variable that grows with the size of the system. The well behaved quantity in this limit is the energy density, $\rho = E_0/\Omega$. Using

$$(2\pi)^3 \delta^3(0) = \int_{\Omega} d^3x e^{i(\mathbf{p}=0) \cdot \mathbf{x}} = \Omega \quad (1.70)$$

the energy density is

$$\begin{aligned}\frac{E_0}{\Omega} &= C + \sum_{i=1}^N \int_0^\Lambda \frac{1}{2} \frac{d^3 p}{(2\pi)^3} E_{\mathbf{p}} \\ &= C + \frac{1}{16\pi^2} \sum_{i=1}^N \left[\Lambda^4 + m_i^2 \Lambda^2 - \frac{1}{4} m_i^4 \log \left(\frac{\Lambda^2}{m_i^2} \right) + o \left(\frac{m_i^2}{\Lambda^2} \right) \right] \quad (1.71)\end{aligned}$$

The short distance divergence has been regulated by cutting off the integration at a maximum momentum, $|p| < \Lambda$. The Λ -dependence can then be renormalized by canceling it with a Λ -dependent constant C .

1.3.2 Spin-half particles

We assume familiarity with the Dirac equation and the Lorentz group in the following; readers unfamiliar with one or both may consult Appendix C.

Spin-half particle states are labeled by $|\mathbf{p}, \sigma\rangle$, in which the label $\sigma = \pm \frac{1}{2}$ represents the projection of intrinsic angular momentum along some axis. Representation theory of the Poincaré group implies that spin- $\frac{1}{2}$ particles are most easily represented by *spinor* fields. Four-component spinor fields transform as follows under Lorentz transformations:

$$U(\omega)\psi(x)U(\omega)^* = D(-\omega)\psi(\Lambda \cdot x) \quad (1.72)$$

in which $D(\omega)$ is the four-by-four matrix given explicitly by

$$D(\omega) = \exp \left[\frac{i}{2} \omega_{\mu\nu} \mathcal{J}^{\mu\nu} \right] \quad (1.73)$$

with the matrices $\mathcal{J}^{\mu\nu}$ given, in the chiral basis which will be used throughout this book, by

$$\mathcal{J}_k = \frac{1}{2} \epsilon_{klm} \mathcal{J}^{lm} = \begin{pmatrix} \frac{1}{2} \sigma_k & 0 \\ 0 & \frac{1}{2} \sigma_k \end{pmatrix} \quad (1.74)$$

$$\mathcal{K}_k = \mathcal{J}_{k0} = \begin{pmatrix} -\frac{i}{2} \sigma_k & 0 \\ 0 & \frac{i}{2} \sigma_k \end{pmatrix} \quad (1.75)$$

Here the two-by-two matrices, σ_k with $k = 1, 2, 3$, denote the usual Pauli spin matrices.

It is clear that this representation is block-diagonal and so is *reducible*. That is, the upper two components of a spinor field never “mix” with the lower two components under any Poincaré transformation. Therefore, it is consistent to consider quantum field theories in which only the upper or lower components of a spinor exist as fields of the theory. Though this does not happen for quantum electrodynamics—the electron can be represented

by a 4-component Dirac spinor – it turns out that it *does* happen for *every* spinor field in the standard model.

There are two equivalent choices of notation to handle such fields, which we will now list.

(i) *Weyl* spinors:

A Weyl spinor is one for which the upper two or lower two components are zero. That is, define *left-handed* and *right-handed* spinors by

$$\psi_L = \frac{1}{2}(1 + \gamma_5)\psi = P_L \psi = \begin{pmatrix} \xi \\ 0 \end{pmatrix} \quad (1.76)$$

$$\psi_R = \frac{1}{2}(1 - \gamma_5)\psi = P_R \psi = \begin{pmatrix} 0 \\ \chi \end{pmatrix} \quad (1.77)$$

in which ξ and χ are two-component objects and γ_5 is the following four-by-four matrix:

$$\gamma_5 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} (= -i\gamma^0\gamma^1\gamma^2\gamma^3) \quad (1.78)$$

I here denotes the two-by-two unit matrix, and γ^μ are defined below.

(ii) *Majorana* spinors:

Alternately, we may work in terms of 4-component spinors where the bottom two components are not independent but are determined by the upper two components. Specifically, first define a two-by-two, real antisymmetric matrix ϵ ,

$$\epsilon \equiv i\sigma_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (1.79)$$

Now note that if ξ is left-handed under Lorentz transformations, then $\chi = \epsilon\xi^*$ is right-handed. This follows from the property

$$\epsilon\sigma_i^* = -\sigma_i\epsilon \quad (1.80)$$

With this in mind, a Majorana spinor is then defined by

$$\psi_M = \begin{pmatrix} \xi \\ \epsilon\xi^* \end{pmatrix} \quad (1.81)$$

These two formulations of fermions with two independent components are equivalent, and the choice of which one to use to formulate a theory is a matter of taste. It is our preference in this book to work with the Majorana notation, mostly because it is simple to make contact with the 4 component γ -matrix algebra in which calculations are generally performed.

The relation between a Majorana spinor field and the creation and annihilation operators is

$$\psi(x) = \sum_{\sigma=\pm\frac{1}{2}} \int \frac{d^3\mathbf{p}}{2E_{\mathbf{p}}(2\pi)^3} \left[u(\mathbf{p}, \sigma) a_{\mathbf{p}\sigma} \left(e^{ipx} + v(\mathbf{p}, \sigma) a_{\mathbf{p}\sigma}^* \right) e^{-ipx} \right] \quad (1.82)$$

In this expression $\psi(x)$, $u(\mathbf{p}, \sigma)$, and $v(\mathbf{p}, \sigma)$ are all 4-component objects with $v(\mathbf{p}, \sigma)$ defined in terms of $u(\mathbf{p}, \sigma)$ by

$$v\left(\mathbf{p}, \sigma = \pm\frac{1}{2}\right) = \pm\gamma_5 u\left(\mathbf{p}, \sigma = \mp\frac{1}{2}\right) \quad (1.83)$$

It turns out that for this decomposition to be consistent with Lorentz invariance, the spinor u in the rest frame, $\mathbf{p} = 0$, must satisfy

$$m\beta u\left(\mathbf{p} = 0, \sigma = \pm\frac{1}{2}\right) = mu\left(\mathbf{p} = 0, \sigma = \pm\frac{1}{2}\right) \quad (1.84)$$

where β denotes the following matrix:

$$\beta = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \quad (= i\gamma^0) \quad (1.85)$$

The non-zero \mathbf{p} generalization of Eq. (1.84) can be found by applying a boost, using Eq. (1.73). The mass m on the right-hand side becomes the four-vector p_μ , which in the rest frame has a single component, $E = m$. The matrix $-i\beta$ is really the time component of a four-vector of matrices, the Dirac matrices γ^μ , so Eq. (1.84) in a general frame becomes

$$(i\not{p} + m)u(\mathbf{p}, \sigma) = 0 \quad (1.86)$$

with \not{p} defined by $\not{p} = \gamma_\mu p^\mu$ (and in general $\not{a} \equiv \gamma^\mu a_\mu$). Equation (1.73) uniquely determines the Dirac matrices:

$$\gamma^0 = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}, \quad \gamma_k = \begin{pmatrix} 0 & -i\sigma_k \\ i\sigma_k & 0 \end{pmatrix} \quad (1.87)$$

In position space, Eq. (1.86) is the *Dirac equation*:

$$(\not{\partial} + m)\psi = 0 \quad (1.88)$$

The Dirac or gamma matrices γ^μ used here differ by a factor of i from the form they would take if we adopted a $\eta_{\mu\nu} = \text{diag}[+---]$ Lorentz metric notation. The reader should be aware of this notation choice. This is discussed in some detail in Appendix E.

The matrix $\gamma^0 = -i\beta$ is anti-Hermitian, while the spatial γ matrices are Hermitian. Therefore the Dirac matrices transform differently under Hermitian conjugation. Similarly, the matrices \mathcal{J}_k which perform rotations are

Hermitian, while the matrices \mathcal{K}_k which perform boosts are anti-Hermitian; so $D^\dagger(\omega)$ does not equal $D^{-1}(\omega)$ in general. However, the matrix β satisfies

$$\beta = \beta^\dagger = \beta^T = \beta^{-1}, \quad \beta\gamma_\mu^\dagger = -\gamma_\mu\beta, \quad \beta\gamma_5 = -\gamma_5\beta \quad (1.89)$$

Also, since $\mathcal{J}^{\mu\nu} = -i[\gamma^\mu, \gamma^\nu]/4$, these imply that

$$\mathcal{K}_k^\dagger\beta = \beta\mathcal{K}_k, \quad \mathcal{J}_k^\dagger\beta = \beta\mathcal{J}_k \quad (1.90)$$

Because of these properties of β , it is convenient to define the *Dirac conjugate* of a spinor, $\bar{\psi}$, as

$$\bar{\psi} \equiv \psi^\dagger\beta \quad (1.91)$$

which transforms under Lorentz transformations as

$$U(\omega)\bar{\psi}(x)U(\omega)^* = \bar{\psi}(\Lambda \cdot x)D^{-1}(-\omega) \quad (1.92)$$

Therefore $\bar{\psi}\psi$ transforms as a Lorentz scalar. As can be readily checked, $D^{-1}(\omega)\gamma^\mu D(\omega) = \Lambda^\mu{}_\nu\gamma^\nu$, so $\bar{\psi}\gamma^\mu\psi$ transforms as a vector.

It is also convenient to introduce the *charge conjugation matrix* C , as the matrix which relates a Majorana spinor to its Dirac conjugate:

$$C = \begin{pmatrix} -\epsilon & 0 \\ 0 & \epsilon \end{pmatrix} (\equiv \gamma^2\beta), \quad \text{so} \quad \psi_M = C\bar{\psi}_M^T, \quad \text{and} \quad \psi_M^T = -\bar{\psi}_M C \quad (1.93)$$

Its properties are

$$-C = C^\dagger = C^{-1} = C^T, \quad \gamma_\mu^T C = -C\gamma_\mu, \quad C\beta = -\beta C, \quad C\gamma_5 = \gamma_5 C \quad (1.94)$$

Returning to Eq. (1.86), we can solve explicitly for the spinor $u(\mathbf{p}, \sigma)$, giving

$$u(\mathbf{p}, \sigma) = \frac{1}{\sqrt{2}} \begin{pmatrix} A_+ - A_- \sigma \cdot \hat{\mathbf{p}} & 0 \\ 0 & A_+ + A_- \sigma \cdot \hat{\mathbf{p}} \end{pmatrix} \begin{pmatrix} \chi(\sigma) \\ \chi(\sigma) \end{pmatrix} \quad (1.95)$$

where

$$\chi\left(\sigma = +\frac{1}{2}\right) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \chi\left(\sigma = -\frac{1}{2}\right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (1.96)$$

$\hat{\mathbf{p}}$ is the unit vector $\hat{\mathbf{p}} = \mathbf{p}/|\mathbf{p}|$, and the coefficients A_\pm are the following functions of the particle energy $E_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2}$:

$$A_\pm(\mathbf{p}) = \sqrt{E_{\mathbf{p}} \pm m} \quad (1.97)$$

As defined by Eq. (1.95), $u(\mathbf{p}, \sigma)$ satisfies the normalization condition:

$$\bar{u}(\mathbf{p}, \sigma') u(\mathbf{p}, \sigma) = 2m \delta_{\sigma\sigma'} \quad (1.98)$$

The dyadics $u\bar{u}$ and $v\bar{v}$ are often encountered in calculations. They can be thought of as matrices, with values

$$u(\mathbf{p}, \sigma) \bar{u}(\mathbf{p}, \sigma) = \frac{1}{2} (m - i\not{p})(1 + i\gamma_5 \not{s}) \quad (1.99)$$

and

$$v(\mathbf{p}, \sigma) \bar{v}(\mathbf{p}, \sigma) = -\frac{1}{2} (m + i\not{p})(1 + i\gamma_5 \not{s}) \quad (1.100)$$

In these expressions $s^\mu(\sigma)$ is the *spin four-vector*. It is defined in the following way. Suppose the spin projection, $\sigma = \pm \frac{1}{2}$, is measured along the direction defined by the unit vector \mathbf{e} in the particle rest frame. Define s^μ in this frame by $s^0 = 0$ and $\mathbf{s} = \pm \mathbf{e}$ in which the sign \pm denotes the sign of σ . The result in any other frame is found by performing the appropriate Lorentz boost. Notice that this definition implies the following invariant properties:

$$s^2 = s^\mu s_\mu = +1 \quad \text{and} \quad s \cdot p = s^\mu p_\mu = 0 \quad (1.101)$$

Now we repeat the exercise of showing that it is always possible to write a free theory of spin-half particles in a canonical form. Consider the Lagrangian description of a system of N non-interacting spin-half particles. Just as there is no loss in choosing our scalar fields to be real, we may always take our spinor fields to be Majorana. The Lagrangian must then be a Lorentz-invariant function of N Majorana spinors, ψ^m , that is at most quadratic in the fields and involves the fewest (nonzero) number of derivatives. The most general such Lagrangian is

$$\mathcal{L}_0 = -\frac{1}{2} A_{mn} \bar{\psi}^m \not{\partial} \psi^n - \frac{i}{2} B_{mn} \bar{\psi}^m \gamma_5 \not{\partial} \psi^n - \frac{1}{2} C_{mn} \bar{\psi}^m \psi^n - \frac{i}{2} D_{mn} \bar{\psi}^m \gamma_5 \psi^n - E \quad (1.102)$$

The Lagrangian must be Hermitian; together with the results of problem 1.1, this implies that A_{mn} , B_{mn} , C_{mn} , D_{mn} , and E must all be real. We may also take the matrices A , C , and D symmetric and B antisymmetric, since the operators multiplying them have the same property.

As usual, most of the parameters in this Lagrangian may be eliminated by performing field redefinitions. The purpose of the remainder of this section is to use this freedom to put the Lagrangian (Eq.(1.102)) into a standard form in which all parameters have an obvious physical significance. Consider then the following field redefinition:

$$\psi^m = V_n^m \psi'^n + i U_n^m \gamma_5 \psi'^n \quad (1.103)$$

with real matrices V and U . This is the most general transformation that preserves the Majorana character of the spinors and the quadratic form of the Lagrangian. It is convenient in what follows to handle the left- and right-handed parts of the fields separately. We therefore rewrite Eq. (1.102) and Eq. (1.103) as

$$P_L \psi^m = (V + iU)_n^m P_L \psi'^n \quad (1.104)$$

$$P_R \psi^m = (V - iU)_n^m P_R \psi'^n \quad (1.105)$$

$$\mathcal{L}_0 = -\frac{1}{2} \left[(A + iB)_{mn} \bar{\psi}^m P_L \not{\partial} \psi^n + \frac{1}{2} (C + iD)_{mn} \bar{\psi}^m P_L \psi^n \right] + \text{h.c.} - E \quad (1.106)$$

Define the complex matrices $\mathcal{A} = (A + iB)$, $\mathcal{C} = (C + iD)$, and $\mathcal{V} = (V + iU)$. The properties of A , B , C , and D then imply that \mathcal{A} is Hermitian and \mathcal{C} is symmetric. For stability we require that \mathcal{A} be positive definite. In terms of the new variables the Lagrangian is then:

$$\mathcal{L}_0 = -\frac{1}{2} \left[(\mathcal{V}^T \mathcal{A} \mathcal{V}^*)_{mn} \bar{\psi}'^m P_L \not{\partial} \psi'^n - \frac{1}{2} (\mathcal{V}^T \mathcal{C} \mathcal{V})_{mn} \bar{\psi}'^m P_L \psi'^n \right] + \text{h.c.} - E \quad (1.107)$$

In order to simplify \mathcal{L}_0 choose \mathcal{V} as follows:

$$\mathcal{V} = (\mathcal{A}^*)^{-\frac{1}{2}} \mathcal{M} \quad (1.108)$$

in which \mathcal{M} is the unitary matrix that satisfies the following property:

$$\mathcal{M}^T \mathcal{C}' \mathcal{M} = \begin{pmatrix} c_1 & & \\ & c_2 & \\ & & \ddots \\ & & & c_N \end{pmatrix} \quad (1.109)$$

\mathcal{C}' is the complex symmetric matrix $\mathcal{C}' = [\mathcal{A}^{-\frac{1}{2}} \mathcal{C} (\mathcal{A}^*)^{-\frac{1}{2}}]$. For any such matrix, a unitary matrix, \mathcal{M} , satisfying (1.109) always exists (see Problem 1.6). In fact, \mathcal{M} may always be chosen such that the numbers c_k , $k = 1, \dots, N$ are all real and non-negative. It must be emphasized that since Eq. (1.109) is *not* a similarity transformation, the c_k are *not* the eigenvalues of the matrix \mathcal{C} or \mathcal{C}' . Instead, c_k^2 turn out to be the eigenvalues of the Hermitian matrix $\mathcal{C}'^\dagger \mathcal{C}'$.

Having made this redefinition, the Lagrangian is in canonical form:

$$\mathcal{L}_0 = -\frac{1}{2} \bar{\psi}^m \not{\partial} \psi^m - \frac{1}{2} c_m \bar{\psi}^m \psi^m - E \quad (1.110)$$

The equation of motion for this action is

$$(\not{\partial} + c_m) \psi^m = 0 \quad (1.111)$$

which is recognized as the Dirac equation with mass c_m . To confirm this connection we compare the resulting free Hamiltonian with the general form (1.50):

$$\begin{aligned} H_0 &= \int d^3x \bar{\psi}^m (\gamma \cdot \nabla + c_m) \psi^m + E \\ &= E_0 + \sum_{m=1}^N \sum_{\sigma=\pm\frac{1}{2}} \int \frac{d^3\mathbf{p}}{2E_{\mathbf{p}}(2\pi)^3} E_{\mathbf{p}} a_{\mathbf{p}\sigma m}^* a_{\mathbf{p}\sigma m} \end{aligned} \quad (1.112)$$

$$\text{with } E_0 = E \int d^3x - \sum_m \sum_{\sigma} \frac{1}{2} \int \frac{d^3\mathbf{p}}{(2\pi)^3} E_{\mathbf{p}} (2\pi)^3 \delta^3(0) \quad (1.113)$$

The corresponding vacuum energy density is

$$\frac{E_0}{\Omega} = E - \frac{1}{8\pi^2} \sum_{i=1}^N \left[\Lambda^4 + m_i^2 \Lambda^2 - \frac{1}{4} m_i^4 \log \left(\frac{\Lambda^2}{m_i^2} \right) + o \left(\frac{m_i^2}{\Lambda^2} \right) \right] \quad (1.114)$$

Notice the relative factor of -2 between the zero-point energy, Eq. (1.114), of free spin-half Majorana fermions and that, Eq. (1.71), of free real scalars.

1.3.3 Spin-one particles

The fields that are most convenient for representing spin-one particles differ for massive and massless particles. This is as might have been expected given that massive and massless spin-one particles have differing numbers of spin states. The particle states are labeled by $|\mathbf{p}, \lambda\rangle$ in which $\lambda = \pm 1$ for massless particles and $\lambda = 0, \pm 1$ for massive ones.

1.3.3.1 Massive spin-one particles

Massive particles are most conveniently represented in terms of a four-vector field, V^μ . This transforms under a Lorentz transformation according to

$$U(\omega) V^\mu(x) U(\omega)^* = (\Lambda^{-1})^\mu{}_\nu V^\nu(\Lambda \cdot x) \quad (1.115)$$

The relation between such a field and the creation and annihilation operators for a massive spin-one particle is,

$$V^\mu(x) = \sum_{\lambda=-1}^1 \int \frac{d^3\mathbf{p}}{2E_{\mathbf{p}}(2\pi)^3} \left[\epsilon^\mu(\mathbf{p}, \lambda) a_{\mathbf{p}\lambda} e^{ipx} + \epsilon^{\mu*}(\mathbf{p}, \lambda) \bar{a}_{\mathbf{p}\lambda} e^{-ipx} \right] \quad (1.116)$$

Here the three four-vectors $\epsilon^\mu(\mathbf{p}, \lambda)$ denote the three linearly independent directions that correspond to each polarization λ . For example, for linearly polarized particles these would correspond to the three unit vectors \mathbf{e}_x , \mathbf{e}_y , and \mathbf{e}_z in the particle rest frame. For circularly polarized particles choose

instead the combinations \mathbf{e}_z and $\mathbf{e}_\pm = \frac{1}{\sqrt{2}}(\mathbf{e}_x \pm i\mathbf{e}_y)$. These polarization vectors are all characterized by the covariant constraint that is the analogue of Eq. (1.101):

$$p_\mu \epsilon^\mu(\mathbf{p}, \lambda) = 0 \quad (1.117)$$

They satisfy the normalization condition

$$\epsilon^{\mu*}(\mathbf{p}, \lambda) \epsilon_\mu(\mathbf{p}, \lambda') = \delta_{\lambda\lambda'} \quad (1.118)$$

and completeness relation

$$\sum_{\lambda=-1}^1 \epsilon_\mu(\mathbf{p}, \lambda) \epsilon_\nu^*(\mathbf{p}, \lambda) = \eta_{\mu\nu} + \frac{p_\mu p_\nu}{m^2} \quad (1.119)$$

Together with the condition $p^2 + m^2 = 0$, Eq. (1.115) implies that in position space $V^\mu(x)$ must satisfy

$$(-\partial^2 + m^2)V^\mu = 0 \quad \text{and} \quad \partial^\mu V_\mu = 0 \quad (1.120)$$

These are the conditions that V^μ must satisfy in order to represent massive spin-one particles.

Turn now to the Lagrangian formulation of a system of free massive spin-one particles. We must construct the most general quadratic, Lorentz-invariant etc. Lagrangian whose equations of motion imply Eq. (1.120). The new feature here is that the condition that the equations of motion be equivalent to Eq. (1.120) will be found to impose conditions on what form we may entertain for the Lagrangian. This is unlike what we encountered for spin-zero and spin-half particles, where the most general Lagrangian automatically implied the analogues of Eq. (1.120), i.e. the Klein-Gordon or Dirac equations. This new feature arises because, unlike for scalar or spinor fields, a four-vector may a priori represent particles of more than one spin. It may correspond to either spin zero or spin one. (Schematically, a vector represents a spin-zero particle when it is the gradient of a scalar.)

To see how this works consider the most general quadratic Lagrangian for a single vector field, given by

$$\mathcal{L}_0 = -\frac{1}{2}A\partial_\mu V_\nu \partial^\mu V^\nu - \frac{1}{2}B\partial_\mu V_\nu \partial^\nu V^\mu - \frac{1}{2}CV^\mu V_\mu - D \quad (1.121)$$

The constants A , B , C , and D must all be real. The equations of motion for such a Lagrangian are

$$A\Box V^\mu + B\partial^\mu \partial_\nu V^\nu - CV^\mu = 0 \quad (1.122)$$

Taking the divergence of Eq. (1.122) gives the further equation,

$$[(A+B)\Box - C]\partial^\mu V_\mu = 0 \quad (1.123)$$

These equations only imply that $\partial \cdot V = 0$ when $A + B = 0$ and $C \neq 0$. In this case they are equivalent to Eq. (1.120). We may also always rescale V^μ to ensure that $A = 1$. The Lagrangian must therefore be

$$\begin{aligned}\mathcal{L}_0 &= -\frac{1}{2}(\partial_\mu V_\nu \partial^\mu V^\nu - \partial_\mu V_\nu \partial^\nu V^\mu) - \frac{1}{2}C' V^\mu V_\mu - D \\ &= -\frac{1}{4}f_{\mu\nu}f^{\mu\nu} - \frac{1}{2}C' V^\mu V_\mu - D\end{aligned}\quad (1.124)$$

in which $f_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu$, which is called the *field strength*.

Comparison with Eq. (1.120) or the expression for the corresponding free Hamiltonian implies that $C' = C/A = m^2$ has the interpretation of the squared mass of the particle being described. The vacuum energy is similarly

$$\frac{E_0}{\Omega} = D + \frac{3}{16\pi^2} \sum_{i=1}^N \left[\Lambda^4 + m_i^2 \Lambda^2 - \frac{1}{4} m_i^4 \log \left(\frac{\Lambda^2}{m_i^2} \right) + O \left(\frac{m_i^2}{\Lambda^2} \right) \right] \quad (1.125)$$

For N massive spin-one particles the argument above, together with one that exactly parallels that given for scalar fields, implies that the most general Lagrangian,

$$\mathcal{L}_0 = -\frac{1}{2}A_{ab}\partial_\mu V_\nu^a \partial^\mu V^{b\nu} - \frac{1}{2}B_{ab}\partial_\mu V_\nu^a \partial^\nu V^{b\mu} - \frac{1}{2}C_{ab}V^{a\mu}V_\mu^b - D \quad (1.126)$$

may be rewritten as

$$\begin{aligned}\mathcal{L}_0 &= -\frac{1}{2}(\partial_\mu V_\nu^a \partial^\mu V^{a\nu} - \partial_\mu V_\nu^a \partial^\nu V^{a\mu}) - \frac{1}{2}C'_a V^{a\mu}V_\mu^a - D \\ &= -\frac{1}{4}f_{\mu\nu}^a f^{a\mu\nu} - \frac{1}{2}C'_a V^{a\mu}V_\mu^a - D\end{aligned}\quad (1.127)$$

1.3.3.2 Massless spin-one particles

Massless spin-one particles are, on the other hand, most conveniently represented in terms of an antisymmetric tensor field, $f_{\mu\nu}$. The relation between such a field and the creation and annihilation operators for a massless spin-one particle are:

$$f_{\mu\nu}(x) = \sum_{\lambda=\pm 1} \int \frac{d^3\mathbf{p}}{2E_{\mathbf{p}}(2\pi)^3} \left[(ip_\mu \epsilon_\nu(\mathbf{p}, \lambda) - ip_\nu \epsilon_\mu(\mathbf{p}, \lambda)) a_{\mathbf{p}\lambda} e^{ipx} + \text{h.c.} \right] \quad (1.128)$$

Here the two quantities $\epsilon^\mu(\mathbf{p}, \lambda)$ denote the linearly independent directions that correspond to each polarization λ . For particles moving along the Z axis, linearly polarized particles correspond to the choice of the unit vectors \mathbf{e}_x and \mathbf{e}_y perpendicular to the particle motion. The alternative combinations $\mathbf{e}_\pm = \frac{1}{\sqrt{2}}(\mathbf{e}_x \pm i\mathbf{e}_y)$ correspond instead to circularly polarized particles.

Notice that Eq. (1.128) only determines the polarization vector, ϵ^μ , up to the *gauge* freedom

$$\epsilon^\mu(\mathbf{p}, \lambda) \rightarrow \epsilon^\mu(\mathbf{p}, \lambda) + p^\mu \quad (1.129)$$

This freedom may be used to ensure that ϵ^μ satisfies the following Lorentz-covariant properties:

$$\bar{p}_\mu \epsilon^\mu(\mathbf{p}, \lambda) = p_\mu \epsilon^\mu(\mathbf{p}, \lambda) = 0 \quad (1.130)$$

in which \bar{p}_μ is a null vector $\bar{p}_\mu \bar{p}^\mu = p_\mu p^\mu = 0$ satisfying $p_\mu \bar{p}^\mu = -1$. The normalization and completeness relations satisfied by such polarization vectors are

$$\epsilon^{\mu*}(\mathbf{p}, \lambda) \epsilon_\mu(\mathbf{p}, \lambda') = \delta_{\lambda\lambda'} \quad (1.131)$$

and

$$\sum_{\lambda=\pm 1} \epsilon_\mu(\mathbf{p}, \lambda) \epsilon_\nu^*(\mathbf{p}, \lambda) = \eta_{\mu\nu} + p_\mu \bar{p}_\nu + p_\nu \bar{p}_\mu \quad (1.132)$$

Note that the null vector \bar{p} is not unique; indeed, the substitution

$$\bar{p}^\mu \rightarrow \bar{p}^\mu + a \epsilon^\mu + \frac{a^2}{2} p^\mu \quad (1.133)$$

for any spacelike ϵ^μ satisfying $\epsilon \cdot \bar{p} = 0 = \epsilon \cdot p$ and $\epsilon_\mu^* \epsilon^\mu = 1$, yields a new vector satisfying the required properties for \bar{p} . However, if we choose particular polarization vectors $\epsilon_\mu(\lambda)$ and require $\bar{p} \cdot \epsilon(\lambda) = 0$ for each λ , then the choice is made unique.

In position space, the conditions, Eq. (1.128) through Eq. (1.130), imply that

$$f_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (1.134)$$

$$A_\mu(x) = \sum_{\lambda=\pm 1} \int \frac{d^3 \mathbf{p}}{2E_{\mathbf{p}}(2\pi)^3} [\epsilon_\mu(\mathbf{p}, \lambda) a_{\mathbf{p}\lambda} e^{ipx} + \dots] \quad (1.135)$$

in which the *gauge potential*, $A^\mu(x)$, is only defined up to the freedom, Eq. (1.129)

$$A_\mu \rightarrow A_\mu + \partial_\mu \omega(x) \quad (1.136)$$

where $\omega(x)$ is an arbitrary function. The mass-shell condition $p^2 = 0$ then becomes

$$\partial_\mu f^{\mu\nu} = 0 \quad (1.137)$$

or, using Eq. (1.136) to impose the *gauge condition* $\partial^\mu A_\mu = 0$, equivalently:

$$\square A^\mu = 0 \quad (1.138)$$

The corresponding free Lagrangian then is

$$\mathcal{L}_0 = -\frac{1}{4} \sum_{a=1}^N f_{\mu\nu}^a f^{a\mu\nu} \quad (1.139)$$

It is crucial to realize that, whereas the field-strength $f_{\mu\nu}$ defined in this way is a tensor under Lorentz transformations, the gauge potential, A_μ , is *not* a four-vector. Rather, it transforms as:

$$U(\omega)A^\mu(x)U(\omega)^* = \Lambda_\nu^\mu A^\nu(\lambda \cdot x) + \partial^\mu \omega(x) \quad (1.140)$$

for some scalar field $\omega(x)$. That is, A^μ transforms as a four-vector only up to a gauge transformation. This is a crucial observation because if we wish to write down interactions that do not vanish in the zero-momentum limit between massless spin-one particles and other particles (such as, for example, the Coulomb interaction in electromagnetism) then we must build our Lagrangian from the field A^μ rather than $f^{\mu\nu}$. Since A^μ is only a Lorentz four-vector up to gauge transformations, we see that Lorentz invariance of the Lagrangian requires that the interactions be invariant under the gauge transformations of Eq. (1.136). In this way we see gauge invariance emerge as a consequence of Lorentz invariance for massless particles of high spin. (A similar argument may be made for massless particles with spin-3/2 or -2, leading to supersymmetry or general covariance.)

1.4 Implications of symmetries

We pause here for a short aside on the general symmetry features that may arise in a Lagrangian. There are two motivations for this aside, corresponding to the two roles played by symmetries in what follows. First, symmetries are useful because they often allow us to make exact statements, even without a detailed understanding of a theory's dynamics. Namely, they can provide general conservation laws and spectral degeneracies familiar from quantum mechanics. Second, symmetries play a crucial role in the couplings of massless (or light) spin-one particles, by virtue of the requirement of gauge invariance that must be imposed. In this section we address the first of these roles in the first two subsections and return to the issue of gauge invariance in the last subsection.

1.4.1 Symmetries and conservation laws

Perhaps the simplest example of the connection between symmetry and a conservation law is given by the example of a discrete symmetry. For

example, suppose the Hamiltonian of a system has a symmetry, in the sense that it remains unchanged after the replacement $\phi(x) \rightarrow -\phi(x)$; i.e.

$$H(-\phi, -\partial_\mu \phi) = H(\phi, \partial_\mu \phi) \quad (1.141)$$

identically for any field configuration $\phi(x)$. This ensures that there is a conservation law, inasmuch as it is possible to define a unitary operator, \mathcal{X} , which represents this replacement in the following sense:

$$\mathcal{X}\phi(x)\mathcal{X}^* = -\phi(x) \quad (1.142)$$

and so

$$\mathcal{X}a_p\mathcal{X}^* = -a_p, \quad (1.143)$$

Such an operator necessarily satisfies the symmetry property of a quantum symmetry: $\mathcal{X}H = H\mathcal{X}$.

If any *Hermitian* operator, \mathcal{X} , satisfies the condition $[\mathcal{X}, H] = 0$, it defines a conservation law. (For instance, in the example being discussed the condition $\mathcal{X}^2 = I$ together with the unitarity of \mathcal{X} automatically ensures \mathcal{X} is Hermitian.) It defines a conservation law because the fact that \mathcal{X} commutes with H ensures that energy eigenstates may be labeled consistently by the eigenvalues of \mathcal{X} : $\mathcal{X}|E, x\rangle = x|E, x\rangle$. Furthermore, this label is conserved because it cannot change under time evolution:

$$\mathcal{X}|E, x; t\rangle = \mathcal{X}e^{-iHt}|E, x\rangle = e^{-iHt}\mathcal{X}|E, x\rangle = x|E, x; t\rangle \quad (1.144)$$

If it is true that $\mathcal{X}^2 = I$, then the eigenvalues satisfy $x = \pm 1$.

It bears emphasis that this conservation is an exact statement, provided only that \mathcal{X} commutes with the exact Hamiltonian of the system, and so can have very powerful consequences. It implies, for example, that the lowest-energy state having eigenvalue $x = -1$ must be absolutely stable. It must be stable since it cannot decay into lower energy states, since energy conservation requires that any decay products have lower energy and yet they must also share the eigenvalue $x = -1$. Since no states satisfy both requirements, the decay cannot occur.

1.4.2 Local conservation laws: continuous symmetries

A particularly important class of conservation laws arises in the case when the theory has a continuous symmetry: $U(g)H = HU(g)$, where $U(g)$ is a unitary operator and g is any element $g \in G$ of a continuous group (whose properties are reviewed in appendix B). Since any element of the group g can be written as $g = \exp(i\epsilon_a t_a)$ with t_a the Lie algebra elements of the group,

the unitary operator can be written $U(g) = \exp(i\epsilon_a Q_a)$. The operators Q defined in this way satisfy both $[Q, H] = 0$ and $Q^* = Q$, with the latter condition following as a consequence of the unitarity of $U(g)$.

This connection between a conserved charge, Q , and a symmetry holds equally well regardless of whether one is interested in classical mechanics, “ordinary” quantum mechanics of a few degrees of freedom, or field theory. For example, the symmetries of time translation, spatial translation or spatial rotations imply the conservation of energy, linear, and angular momentum respectively.

A new feature which appears in field theories having continuous symmetries is that the resulting conservation law holds *locally* through the existence of a spacetime-dependent conserved current, according to *Noether’s theorem*. This local conservation may be seen as follows.

Suppose the Lagrangian density, $\mathcal{L}[\phi, \partial_\mu \phi]$, is invariant with respect to a local transformation of the field variables, $\phi^i(x)$

$$\delta\phi^i(x) = \epsilon^a F_a^i[\phi, \partial_\mu \phi; x] \quad (1.145)$$

in which ϵ^a represent a set, $a = 1, \dots, N$ of *spacetime-independent* infinitesimal parameters and F_a^i indicates a local functional of the fields. The invariance of the action may be expressed as

$$\begin{aligned} 0 &\equiv \delta\mathcal{L} \\ &= \frac{\partial\mathcal{L}}{\partial\phi^i(x)}\epsilon^a F_a^i + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}\epsilon^a \partial_\mu F_a^i \\ &= \left[\frac{\partial\mathcal{L}}{\partial\phi^i(x)} - \partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))} \right) \right] \epsilon^a F_a^i + \partial_\mu \left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))} F_a^i \right] \epsilon^a \end{aligned} \quad (1.146)$$

The first term in the final line of Eq. (1.146) vanishes once the equations of motion for ϕ^i are used. The final line then shows that the equations of motion imply that the four-vector *Noether current*,

$$j_a^\mu(x) \equiv -\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))} F_a^i \quad (1.147)$$

is conserved; $\partial_\mu j_a^\mu(x) = 0$ for each a . (The overall minus sign is conventional.) This last equation expresses conservation because it implies that the *charge*, Q , defined by

$$Q_a(t) \equiv \int d^3x j_a^0(\mathbf{x}, t) \quad (1.148)$$

is time-independent:

$$\frac{dQ_a}{dt} = \int d^3x \times \frac{\partial j_a^0}{\partial t} = - \int d^3x \times \nabla \cdot \vec{j} = \oint d^2x \times \vec{n} \cdot \vec{j} = 0.$$

We assume here that there is no net flux going out of the boundary at infinity.

A symmetry for which the Lagrangian density is invariant as in Eq. (1.146) is known as an *internal symmetry*. This is to distinguish it from *spacetime symmetries* such as Poincaré transformations. In general, symmetries that act on spacetime coordinates as well as the fields cannot leave the Lagrangian density invariant because the Lagrangian density is not constant throughout spacetime. In this case a slightly more general form for Noether's theorem is necessary.

Suppose, then, that under the transformations

$$\begin{aligned}\delta\phi^i(x) &= \epsilon^a F_a^i[\phi, \partial_\mu\phi; x] \\ \delta x^\mu &= \epsilon^a \xi_a^\mu(x)\end{aligned}\tag{1.149}$$

the Lagrangian density transforms into a total derivative (so the action $\int \mathcal{L} d^4x$ is invariant)

$$\delta\mathcal{L} \equiv \epsilon^a \partial_\mu V_a^\mu\tag{1.150}$$

for some Lorentz-vector fields, $V_a^\mu[\phi^i, \partial\psi]$, that are local functionals of $\phi^i(x)$. Repeating the arguments leading to Eq. (1.146) again implies conserved currents, $\partial_\mu j_a^\mu(x) = 0$, with $j_a^\mu(x)$ given by

$$j_a^\mu(x) = -\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}F_a^i + V_a^\mu(x)\tag{1.151}$$

Conservation laws such as these are significant because they are exact results, and so allow conclusions even in the absence of a detailed understanding of the dynamics of a particular system. In a quantum theory the conserved charges, Q_a , are of particular interest since they are Hermitian and commute with the system Hamiltonian (since they are conserved!). They are therefore ideal operators for labeling the individual particle states. With particle states labeled in this way, conservation laws imply general selection rules concerning how quantum numbers must be related before and after collision processes.

It is often true that the symmetry transformation law given in the first line of Eq. (1.149) is more general than is necessary for a particular physical situation. It is often sufficient to consider symmetry transformations that are linear in the field variables

$$\delta\phi^i(x) = i\epsilon^a (T_a)_j^i \phi^j(x)\tag{1.152}$$

1.4.3 Spectral relations

The second major conclusion that may be drawn from symmetry properties of the Lagrangian of a system concerns the system's energy spectrum. The general statement is that states that are related by a symmetry transformation must have the same energy. This is a simple consequence of the fact that the conserved charge, Q_a , commutes with the system Hamiltonian. If, for instance, two energy eigenstates are related by $|\psi\rangle = Q_a|\chi\rangle$, then

$$\begin{aligned} H|\psi\rangle &= HQ_a|\chi\rangle \\ &= Q_aH|\chi\rangle \\ &= E_\chi Q_a|\chi\rangle \\ &= E_\chi|\psi\rangle \end{aligned} \tag{1.153}$$

It follows that $|\psi\rangle$ and $|\chi\rangle$ have the same energy eigenvalue, $E_\psi = E_\chi$, or are *degenerate*.

In general, states in the Hilbert space fall into unitary representations of the symmetry and all of the elements of a given representation must have the same energy.

Now, in a field theory we would like to apply this reasoning to the single-particle states in order to derive relations among the particle masses. This can be done subject to a single *caveat*: the ground state of the theory must be invariant under the symmetry transformations. That is to say, if the symmetry transformations are represented in the Hilbert space by the unitary transformations $U(\epsilon) = \exp(i\epsilon^a Q_a)$, then the invariance of the ground state, $|0\rangle$, is expressed by: $U(\epsilon)|0\rangle = |0\rangle$ or, equivalently, $Q_a|0\rangle = 0$.

The connection between the invariance of the vacuum and symmetry relations among particle masses arises because symmetry transformations in field theory are usually defined as acting on the fields representing the various particles. If the fields representing a particular two particles are related by a symmetry transformation, it does not necessarily follow that the corresponding particle states are related by this same symmetry. It is this link between the fields and the particles that relies on the invariance of the ground state.

To see this in some detail, suppose that the fields, $\phi_1(x)$ and $\phi_2(x)$, corresponding to particle types “1” and “2”, are related by the action of some symmetry:

$$\phi_1(x) = i[Q, \phi_2(x)] \tag{1.154}$$

where $Q^* = Q$ is Hermitian. Then the same is true for the corresponding

creation and annihilation operators:

$$a_1 = i[Q, a_2] \quad (1.155)$$

The particle states are therefore related as follows:

$$\begin{aligned} |1\rangle &= a_1^*|0\rangle \\ &= i[Q, a_2^*]|0\rangle \\ &= iQa_2^*|0\rangle - ia_2^*Q|0\rangle \\ &= iQ|2\rangle - ia_2^*Q|0\rangle \end{aligned} \quad (1.156)$$

The particle states therefore satisfy $|1\rangle = iQ|2\rangle$ if the no-particle state is invariant: $Q|0\rangle = 0$. Once it is known that the particle states are related in this way, the arguments leading to Eq. (1.153) may be used to infer that they have equal masses.

To summarize, the general quantum-mechanical result, which implies that states that are related by symmetry transformations must be degenerate, applies equally well within the field-theoretical context. It does *not* follow, however, that particles whose representative fields are related by symmetry transformations must be degenerate (i.e. have equal masses). This last implication does hold, though, if the ground state of the system is invariant under the action of the symmetry. It is a general feature of field theories that the ground state need not be invariant with respect to symmetry transformations. If the ground state is not invariant, the symmetry is said to be *spontaneously broken*. For spontaneously broken symmetries it is generic that naive symmetry relationships among masses fail.

The conserved currents discussed in the previous section, however, exist regardless of whether a symmetry is spontaneously broken or not, because Noether's theorem only uses the invariance of the action. It is true, however, that spontaneous breaking of a symmetry makes it impossible to use the corresponding charge to define conserved quantum numbers for particle states.

1.5 Renormalizable interactions

We now turn to the construction of general interactions involving particles with spin-zero, -half, or -one. The goal is to construct the most general form for these interactions that is consistent with the five principles outlined in Section 1.2. In this section the general form for renormalizable interactions involving particles of spins zero through one is summarized, largely without proof. The purpose is to outline the general features of these interactions.

1.5.1 Spin-zero and spin-half particles

In order to get started, consider first the most general renormalizable interactions allowed for N interacting spin-zero particles. As outlined in Subsection 1.3.1, we may, without loss of generality, represent these particles with N real scalar fields, $\phi^i(x)$, $i = 1, \dots, N$.

We are instructed to write down a Lorentz-invariant Lagrangian density,

$$\mathcal{L}_s = \mathcal{L}_0 + \mathcal{L}_{\text{int}} \quad (1.157)$$

where \mathcal{L}_0 is the free Lagrangian of Subsection 1.3.1 and \mathcal{L}_{int} is the interaction term that is by definition not quadratic or linear in the fields. \mathcal{L}_{int} is to be constructed solely from $\phi^i(x)$ and $\partial_\mu \phi^i(x)$ subject to the requirement (renormalizability) that it involves interactions of at most dimension four in powers of mass. In order to do so it is necessary to compute the mass dimension of the fields, $\phi^i(x)$, themselves. This is easily done once the free Lagrangian is put into canonical form.

Comparing with standard form, Eq. (1.64) shows that the scalar field must have dimensions of M^1 (when $\hbar = c = 1$) if \mathcal{L}_0 is to have dimension M^4 . This may then be used to infer the restrictions imposed on \mathcal{L}_{int} by renormalizability. It is easy to now show that the most general renormalizable interactions possible among N spin-zero particles are:

$$\begin{aligned} \mathcal{L}_s &= \mathcal{L}_0 - V(\phi) \\ &= -\frac{1}{2}\partial_\mu \phi^i \partial^\mu \phi^i - \rho - v_i \phi^i - \frac{1}{2}\mu_i^2 \phi^i \phi^i - \frac{1}{3!}\xi_{ijk} \phi^i \phi^j \phi^k \\ &\quad - \frac{1}{4!}\lambda_{ijkl} \phi^i \phi^j \phi^k \phi^l \end{aligned} \quad (1.158)$$

The generalization to include also spin-half particles is again straightforward. Inspection of the canonically normalized kinetic term, Eq. (1.110), implies that a spinor field carries dimension $M^{\frac{3}{2}}$. This implies that the most general renormalizable Lagrangian involving spins zero and half must be:

$$\mathcal{L}_m = \mathcal{L}_s - \frac{1}{2}\bar{\psi}^n \not{\partial} \psi^n - \frac{1}{2}m_n \bar{\psi}^n \psi^n - g_{mni} \bar{\psi}^m \psi^n \phi^i - i h_{mni} \bar{\psi}^m \gamma_5 \psi^n \phi^i \quad (1.159)$$

Here \mathcal{L}_s is as in Eq. (1.158) and the new spin-half/spin-zero interaction terms are known generically as *Yukawa couplings*.

1.5.2 Spin-one couplings: gauge invariance

We would like to write down a general set of renormalizable couplings involving particles from spins zero through one. It turns out not to be possible

to do so for the massive spin-one particle (apart from one exception that is a special case of the general situation considered below). We turn therefore directly to the case of massless spin-one particles.

The straightforward thing to try is to couple massless spin-one particles to other particles by writing down interactions that involve the field-strength, $f_{\mu\nu}$. Dimension counting again shows that this is impossible because the free Lagrangian, Eq. (1.139), implies that $f_{\mu\nu}$ has dimensions of M^2 . The lowest-dimension interaction possible would then be something like $\bar{\psi}\gamma^{\mu\nu}\psi f_{\mu\nu}$ which has dimension M^5 and so is not renormalizable.

The only remaining possibility then is to build couplings directly from the gauge potential, $A_\mu(x)$. This is somewhat delicate, because as we have seen, $A_\mu(x)$ does *not* transform as a four-vector – it is only a four-vector up to a gauge transformation: $A_\mu \rightarrow A_\mu + \partial_\mu\omega$. It follows that the interaction Lagrangian itself will only be Lorentz invariant provided that the interactions are required to be gauge invariant.

It is beyond the scope of this book to work out the requirements of gauge invariance in all of their detail. We content ourselves here with simply motivating the construction and then quoting the final results.

Suppose, then, that we write down an interaction term

$$\mathcal{L}_{\text{int}} = A_\mu(x)J^\mu[\phi] \quad (1.160)$$

with $J^\mu[\phi]$ some four-vector function of the other fields and possibly their derivatives. Under a gauge transformation, $\delta A_\mu(x) = \partial_\mu\omega(x)$, if $\delta\phi^i = 0$, this interaction Lagrangian transforms to

$$\delta\mathcal{L}_{\text{int}}(x) = \partial_\mu\omega(x)J^\mu[\phi(x)] \quad (1.161)$$

We need to cancel Eq. (1.161) with the contribution from another term in the Lagrangian. One can imagine doing so in one of two ways. Extra interaction terms can be added, and/or the transformation rules can be altered. The first of these options must fail in the present instance because the required term would have to be linear in the gauge potential in order to produce a variation like Eq. (1.161), and Eq. (1.160) is already the most general such Lagrangian.

The required transformation rule may be most easily seen by repeating the steps leading to Eq. (1.146) in the proof of Noether's theorem, with one alteration. In the previous section Noether's theorem was derived subject to the condition that the transformation parameter, ϵ^a , be independent of spacetime position, x^μ . In the present case, however, the transformation parameter, ω , cannot be spacetime independent because the gauge potential transforms into its gradient. Consider, then, the variation of the Lagrangian

under a transformation as in Eq. (1.145)

$$\delta\phi^i(x) = \epsilon^a(x)F_a^i[\phi, \partial_\mu\phi; x] \quad (1.162)$$

but with the transformation parameter a function of x^μ . Suppose further that the Lagrangian would be invariant if ϵ^a had been chosen as constant. The Lagrangian in this case fails to be invariant with spacetime-dependent ϵ^a only because of its dependence on the derivatives, $\partial_\mu\phi^i$, of the fields. The variation of the Lagrangian therefore becomes

$$\begin{aligned} \delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi^i(x)}\epsilon^a(x)F_a^i[\phi, \partial_\mu\phi; x] + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}\partial_\mu(\epsilon^a(x)F_a^i[\phi, \partial_\mu\phi; x]) \\ &= \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}F_a^i[\phi, \partial_\mu\phi; x]\partial_\mu\epsilon^a(x) \\ &= -j_a^\mu(x)\partial_\mu\epsilon^a(x) \end{aligned} \quad (1.163)$$

Comparing Eq. (1.161) with Eq. (1.163) shows that the gauge variation of the spin-one coupling can cancel against the variation of the spin-zero and spin-half “matter” Lagrangian if

- (i) the coefficient function, $J_a^\mu[\phi]$, is identified with the conserved current,

$$J_a^\mu[\phi] = j_a^\mu(x) \quad (1.164)$$

associated with a symmetry of this matter Lagrangian, and

- (ii) the gauge transformations are enlarged to include the transformation of the matter fields with respect to this symmetry with a spacetime-dependent parameter:

$$\delta A_\mu(x) = \partial_\mu\omega(x) \quad (1.165)$$

$$\delta\phi^i(x) = \omega(x)F^i[\phi, \partial_\mu\phi; x] \quad (1.166)$$

This promotion of a spacetime-independent symmetry of the matter Lagrangian to a spacetime-dependent symmetry of the matter/spin-one Lagrangian is called the *gauging* of the symmetry. The corresponding spin-one particles are known as *gauge bosons*.

More generally, if there are more than one spin-one fields, and if the symmetries involved transform one spin-one particle into another, then the conserved current, $j_a^\mu(x)$, will itself depend on the $A_\mu^a(x)$ s. This leads to self-couplings of the gauge bosons amongst themselves. Such a symmetry is called a *non-abelian* symmetry, and will require a generalization of the above discussion. We here summarize the results of such a generalization.

Consider a (renormalizable) Lagrangian, $\mathcal{L}_m[\phi]$, depending on a collection of spin-zero and spin-half “matter” fields. Suppose that \mathcal{L}_m is invariant with respect to the following *global* (i.e. spacetime-independent) symmetry transformations:

$$\delta\phi^i(x) = i\omega^a(T_a)^i_j\phi^j(x) \quad (1.167)$$

In general, repetition of several symmetry transformations produces further symmetries so the transformations, Eq. (1.167), form a Lie algebra and the matrices $(T_a)^i_j$ necessarily satisfy the *commutation relations* (see Appendix B):

$$[T_a, T_b] = if_{ab}^c T_c \quad (1.168)$$

where the coefficients f_{ab}^c are a set of numbers that are characteristic of the algebra involved. The good news is that all of the algebras of this type that are of physical interest have been found and are cataloged once and for all.

The most general renormalizable way to couple this Lagrangian to a bunch of spin-one particles is given by the following prescription.

- (i) Associate each spin-one particle, $A_\mu^a(x)$, with one of the generators, (T_a) , of the symmetry algebra.
- (ii) Replace ordinary spacetime derivatives everywhere in \mathcal{L}_m with the following *covariant derivatives*:

$$D_\mu\phi^i(x) \equiv \partial_\mu\phi^i(x) - iA_\mu^a(x)(T_a)^i_j\phi^j(x) \quad (1.169)$$

- (iii) Add the following gauge-boson Lagrangian

$$\mathcal{L}_g \equiv -\frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} \quad (1.170)$$

with the covariant field strength, $F_{\mu\nu}^a(x)$, defined by

$$F_{\mu\nu}^a \equiv \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + f_{bc}^a A_\mu^b A_\nu^c \quad (1.171)$$

The total Lagrangian is then given by the sum: $\mathcal{L} = \mathcal{L}_m[\phi, D_\mu\phi] + \mathcal{L}_g$. It is invariant (in fact \mathcal{L}_m and \mathcal{L}_g are separately invariant) under the *local* or gauged generalization of transformation, Eq. (1.167):

$$\delta A_\mu^a(x) = \partial_\mu\omega^a(x) - f_{bc}^a\omega^b(x)A_\mu^c(x), \quad (1.172)$$

$$\delta\phi^i(x) = i\omega^a(x)(T_a)^i_j\phi^j(x) \quad (1.173)$$

1.6 Some illustrative examples

Before proceeding it is useful to consider a few illustrative examples.

1.6.1 Quantum electrodynamics: an abelian gauge theory

Consider, first, the theory describing physics at scales below the mass of the muon, $m_\mu = 106$ MeV. The elementary particles in this energy range are the electron and the neutrinos, represented by a Dirac spinor field, $e(x)$, and three Majorana spinor fields, $\nu_i(x)$; and the photon, represented by the gauge potential, $A_\mu(x)$. We wish to write down the most general renormalizable interactions of these particles, which should furnish a reasonable description of their behavior at energies much less than $2m_\mu$.

From the previous discussion, the coupling of the photon must be to some conserved current – in this case electric charge. The current is

$$J_{\text{em}}^\mu(x) = -ie\bar{e}\gamma^\mu e(x) \quad (1.174)$$

(where unfortunately the electric coupling and the electron field have the same symbol e and must be told apart by context), and the corresponding local symmetry transformation is therefore

$$\begin{aligned} \delta e(x) &= -ie\omega(x)e(x) \\ \delta \nu_i(x) &= 0 \\ \delta A_\mu(x) &= \partial_\mu \omega(x) \end{aligned} \quad (1.175)$$

The most general renormalizable interaction must therefore be

$$\mathcal{L} = -\bar{e}(\not{D} + m_e)e - \bar{\nu}_i(\not{\partial} + m_{\nu_i})\nu_i - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (1.176)$$

in which

$$\begin{aligned} F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu \\ \text{and} \quad D_\mu e(x) &= \partial_\mu e(x) + ieA_\mu(x)e(x) \end{aligned} \quad (1.177)$$

Equation (1.176) has two features that are worth remarking on here. The first is that the Lagrangian has broken up into the sum of two terms: $\mathcal{L} = \mathcal{L}_{\text{QED}} + \mathcal{L}_\nu$ in which \mathcal{L}_{QED} is independent of the neutrino fields and \mathcal{L}_ν depends only on the neutrino fields. Since \mathcal{L}_ν is quadratic this implies that the neutrinos cannot interact at all with the other particles through renormalizable interactions. This is the major part of the present understanding of why it is that neutrinos couple so feebly to the rest of matter. The other observation is that the part of the Lagrangian, \mathcal{L}_{QED} , that depends on electrons and photons is precisely the standard Lagrangian for quantum electrodynamics (QED). This Lagrangian is indeed known to give an extremely precise description of the interactions of electrons and photons. We here have the beginnings of an explanation of why it must have the form that it does. To the extent that any theory at higher energies has the observed

spectrum of particles and preserves the conservation of electric charge, it must reproduce QED at energies, E , well below the mass of the muon, up to non-renormalizable corrections that are suppressed by powers of (E/m_μ) .

1.6.2 Scalar electrodynamics: spontaneous symmetry breaking

The gauge-invariant Lagrangian of the previous sections appears to have the serious drawback that it can only describe the interactions of massless spin-one particles. This turns out not to be true in general, as we shall demonstrate using a less orthodox example, called the abelian Higgs model. The theory consists of a single charged spinless particle, with complex field $\phi(x) = (\phi_{re} + i\phi_{im})/\sqrt{2}$, coupled to electromagnetism, $A_\mu(x)$.

The most general renormalizable matter Lagrangian that is invariant under the global rephasing (or $U(1)$) symmetry $\phi \rightarrow e^{i\omega}\phi$ (and is analytic in ϕ) is

$$\mathcal{L}_\phi = -\partial_\mu \phi^* \partial^\mu \phi - a(\phi^* \phi)^2 - b(\phi^* \phi) - c \quad (1.178)$$

Gauging this symmetry and coupling to the photon gives the Lagrangian,

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - D_\mu \phi^* D^\mu \phi - a(\phi^* \phi)^2 - b(\phi^* \phi) - c \quad (1.179)$$

in which

$$D_\mu \phi = \partial_\mu \phi - ie A_\mu \phi \quad (1.180)$$

and the field strength is as in Eq. (1.177). Although stability implies that the real constant $a = \lambda^2$ must be non-negative, the sign of b is arbitrary.

We wish to extract the spectrum of this theory for weak couplings $e \ll 1$ and $\lambda^2 \ll 1$. There are two qualitatively different possibilities, depending on the sign of b . If $b = \mu^2$ is positive, then the unperturbed Lagrangian simply consists of those terms that are quadratic in the fields. The spectrum for this unperturbed theory was worked out in the previous sections and consists of a massless spin-one photon and a charged, spinless particle with mass $m_\phi^2 = \mu^2$ (see the sentence following Eq. (1.69)).

Things are different if it should happen that $b = -\mu^2$ were negative. In this case a naive repetition of the steps outlined earlier would have us identify the quadratic part of Eq. (1.179) as the unperturbed Lagrangian. One sign that this cannot be quite right is that the mass of the spinless particle in this unperturbed theory would then be imaginary: $m_\phi^2 = -\mu^2$. A tachyonic mass such as this is the sign that the assumed ground-state field configuration – in this case $\phi = 0$ – is unstable, since a negative squared-mass implies that the

field modes with $|\mathbf{p}| < \mu$ have a complex energy: $E = \sqrt{\mathbf{p}^2 - \mu^2} = E_r - iE_i$, and so have a runaway time dependence: $\exp(-iEt) = \exp[+E_i t - iE_r t]$.

More properly, since we are interested in the energies of the lowest excitations about the ground state, i.e. the vacuum, we must first check that we have properly identified the ground state. The weak-coupling limit we are interested in may be used to justify doing so semiclassically. In the semiclassical limit the ground state is just described by its classical field configuration. Being a ground state, this configuration must by definition minimize the energy. Furthermore, the energy of the configuration is semiclassically dominated by the classical energy which is easily computable from the system's Lagrangian. In the present instance the energy density is

$$\mathcal{H} = \frac{\partial \phi^*}{\partial t} \frac{\partial \phi}{\partial t} + D\phi^* \cdot D\phi + \lambda^2(\phi^* \phi)^2 + b \phi^* \phi + c + \frac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) \quad (1.181)$$

Here $E_{-i} = F_{-io}$ and $B_{-i} = \frac{1}{2} - \epsilon_{ijk} F^{jk}$. Since this is a sum of non-negative terms, it is minimized by minimizing each term separately. The electromagnetic field energy is minimized at zero field, $\mathbf{B} = \mathbf{E} = 0$, and the gradient terms in the scalar energy are smallest for constant fields, $\partial \phi / \partial t = \nabla \phi = 0$. If $b \geq 0$ then the potential energy is also minimized by zero field, $\phi = 0$, as was implicitly assumed above. If $b = -\mu^2$, however, then the scalar-field energy is minimized when $\phi^* \phi \equiv v^* v = \mu^2 / (2\lambda^2)$. This value v which the scalar field takes in vacuum is called its vacuum expectation value, or *v.e.v.*

The low-energy excitations are found semiclassically by perturbing about this stable field configuration. The unperturbed system consists of all terms that are quadratic or less in the fluctuations about the minimum-energy field configuration. Since the ground-state constructed in this way is by construction stable, tachyonic modes never appear in such an expansion. When $b \geq 0$ and the ground-state configuration is zero, this agrees with the naive treatment outlined earlier.

For $b < 0$ we must expand instead in powers of the difference: $\varphi \equiv \phi - v$. Doing so with the Lagrangian of Eq. (1.179) gives the following unperturbed result:

$$\begin{aligned} \mathcal{L}_0 = & -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \partial_\mu \varphi^* \partial^\mu \varphi + ie A_\mu (v \partial^\mu \varphi^* - v^* \partial^\mu \varphi) - e^2 v^* v A_\mu A^\mu \\ & - V_0 - \lambda^2 (v^* \varphi + v \varphi^*)^2 \end{aligned} \quad (1.182)$$

The constant V_0 contains all of the φ -independent terms and so represents the ground-state energy density.

Unfortunately, because of the terms that mix the vector with scalar fields, we cannot directly use the results of the previous sections to read off the particle spectrum. Happily enough, gauge invariance now comes to our aid.

Recall that the Lagrangian, and so all of the physics, is unchanged by the gauge transformation

$$\begin{aligned} A_\mu(x) &\rightarrow A_\mu(x) + \partial_\mu \omega(x) \\ \phi(x) &\rightarrow \exp[ie\omega(x)]\phi(x) \end{aligned} \quad (1.183)$$

We may therefore use this freedom to redefine fields to put the Lagrangian into a particularly convenient form. A useful choice for the present purposes is to use the transformations of Eq. (1.183) to make the scalar field everywhere real, $\phi^*(x) = \phi(x)$ for all x . The utility of this choice arises from the observation that the $A_\mu \partial^\mu \varphi$ cross terms then vanish.

The spectrum may now be directly read off as before. The quadratic terms in the electromagnetic potential describe a spin-one particle with mass $M_A^2 = 2e^2 v^2$. The photon is no longer massless! The spin-zero sector now consists of a single real scalar of mass $m_\varphi^2 = 4\lambda^2 v^2 = 2\mu^2$. Since the gauge condition completely eliminates the imaginary part of the scalar field, an entire scalar degree of freedom has been “removed” from the spectrum. This degree of freedom has re-emerged as the longitudinal spin state of the massive spin-one particle. This process, in which a vector field “eats” a scalar one in the process of becoming massive, is known as the *Higgs mechanism*.

The process of using the gauge freedom to impose conditions on the fields is known as “choosing a gauge.” The choice made here is known as “unitary” or “physical” gauge since it makes the spectrum of the theory easy to identify.

The lesson to be learned is that a gauge symmetry need not imply that the corresponding spin-one gauge particle need be massless. This is the second time we have encountered an exception to a general symmetry consequence for the particle spectrum. The circumstances here are similar to those described in Subsection 1.4.3. In both cases the root cause lies in the fact that the ground state is not invariant under the symmetry in question, and it is this non-invariance that ruins the symmetry predictions for the spectrum of fluctuations about that ground state. This is again the phenomenon of spontaneous symmetry breaking.

To see that the ground state indeed breaks the relevant symmetry in the present example, notice that any ground state field configuration $\phi = v$ is not invariant under the transformations of Eq. (1.183). This condition is intimately related to what was our working definition of spontaneous symmetry breaking in Subsection 1.4.3. There we defined it by the condition that the conserved charge, Q , not annihilate the ground state, $Q|\Omega\rangle \neq 0$. The one condition is a consequence of the other, since $\langle\Omega|\phi|\Omega\rangle = v \neq 0$ implies

that the commutator $\langle \Omega | [Q, \phi] | \Omega \rangle$ cannot be zero as would be required if $Q| \Omega \rangle = 0$.

1.6.3 QCD: an $SU(3)$ gauge theory

To a good approximation, the theory of nuclei and their constituents is quantum chromodynamics (QCD), a gauge theory with group $SU_c(3)$. We review it here in some detail, because it is a good lesson in how non-abelian gauge theories work, as well as being directly a component of the standard model.

The theory of QCD contains several types of Dirac fermions called *quarks*, labeled u, d, s, \dots for up, down, strange, \dots (There are six altogether, named u, d, s, c, b , and t , but only u, d, s are light.) However, when we say there is “a” quark u , we really mean there are three quark fields, written u_r, u_g , and u_b (rgb for “red,” “green,” and “blue”), which have exactly the same mass; similarly, d, s, \dots are replicated in triplicate, also labeled r, g, b. It is convenient to group these three fields in a column vector, $[u_r, u_g, u_b]^T$, or u_a in index notation. It is customary when possible to suppress this index (matrix notation), and it is important to appreciate that the index a is not the spinorial index we have already met – each u_r, u_g, u_b has four spinor components. When one writes $\bar{u}u$, it really means $\bar{u}_a u_a$ with the a sum implicit and where the spinor indices are summed over for each color separately (spinorial and color indices are independent). The free Lagrangian for the up quarks is,

$$\mathcal{L}_{0,u} = -\bar{u}_a(\not{\partial} + m_u)u_a \equiv \bar{u}(\not{\partial} + m_u)u \quad (1.184)$$

and the Lagrangians for the d, s quarks are similar.

At the free theory level, nothing would change if we made the replacement, $[u_r, u_g, u_b]^T \rightarrow [u_g, -u_r, u_b]^T$, exchanging the role of red and green quarks. More generally, nothing is changed by making an arbitrary unitary rotation $u_a \rightarrow \tilde{U}_{ab}u_b$, $\tilde{U}^\dagger = \tilde{U}^{-1}$, under which the free Lagrangian changes to

$$-\bar{u}_a(\not{\partial} + m_u)u_a \rightarrow -\bar{u}_b\tilde{U}_{ba}^\dagger(\not{\partial} + m_u)\tilde{U}_{ac}u_c = -\bar{u}_a(\not{\partial} + m_u)u_a \quad (1.185)$$

At the free level, the theory has a symmetry under $U(3)$ (3×3 unitary matrix) rotations between the u quarks – and separately under independent $U(3)$ rotations of each other quark type. This $U(3)$ matrix can be decomposed as $\tilde{U} = e^{i\theta}U$, with $U \in SU(3)$ a *special unitary matrix*, that is, unitary matrix of determinant 1. Any $SU(3)$ matrix can be exponentiated as $U = \exp(iM)$, with M a traceless, 3×3 Hermitian matrix. An $N \times N$

complex Hermitian matrix has N^2 independent entries, and the tracelessness condition removes one, so there are eight independent parameters to describe M . Such matrices can always be written in terms of a standard basis of traceless Hermitian matrices, $U = \exp(i\omega_\alpha \lambda_\alpha/2)$, with ω_α some coefficients and λ_α the Gell-Mann matrices, explicitly,

$$\begin{aligned} \lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ \lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} & \lambda_5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} & \lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\ \lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} & \lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \end{aligned} \quad (1.186)$$

chosen to satisfy $\text{tr } \lambda_\alpha \lambda_\beta = 2\delta_{\alpha\beta}$. (Do not confuse the index α with the indices a, b earlier: the α index runs over the eight such independent matrices, while a, b are row and column indices for these matrices and run over three values.) The Gell-Mann matrices satisfy an algebra,

$$\left[\frac{\lambda_\alpha}{2}, \frac{\lambda_\beta}{2} \right] = i f_{\alpha\beta}^\gamma \frac{\lambda_\gamma}{2} \quad (1.187)$$

where $f_{\alpha\beta}^\gamma$, the *structure constants* of the group $SU(3)$, are real and anti-symmetric in all three indices.

QCD is defined as the interacting theory for which the rotations in which u, d, s, \dots are each rotated by the same $SU(3)$ matrix are gauged. (These gauge interactions break all the remaining symmetries except each $U(1)$ symmetry associated with separate phase rotations for each quark species. We return to this issue in Section 2.5.) Now let us see in this example why the conditions, Eq. 1.169–1.171, are necessary. For the u field kinetic term to be invariant under symmetry transformations, it must involve the covariant derivative,

$$\mathcal{L}_u = -\bar{u}(\not{D} + m_u)u, \quad D_\mu = \partial_\mu - ig_3 G_\mu^\alpha \frac{\lambda_\alpha}{2} \quad (1.188)$$

Here G_μ^α are eight spin-1 gauge fields, called *gluon fields*, with the sum on α implicit and g_3 a coupling constant analogous to the electric charge of QED, called the *strong coupling* (frequently written as g_s). (Remember that we suppress matrix indices; λ is a 3×3 matrix multiplying the column vector

u , so $\lambda_\alpha u$ means $(\lambda_\alpha)_{ab} u_b$.) However, the invariance of this expression also requires a specific transformation rule for the field G_μ^α . Under an infinitesimal gauge transformation,

$$u \rightarrow \left(1 + ig_3 \omega^\alpha \frac{\lambda_\alpha}{2}\right) u \quad (1.189)$$

and taking G_μ^α to change to $G_\mu^\alpha + \delta G_\mu^\alpha$ under gauge transformations, this Lagrangian changes to

$$\begin{aligned} \mathcal{L}_u &\rightarrow \bar{u} \left(1 - ig_3 \omega^\alpha \frac{\lambda_\alpha^\dagger}{2}\right) \left[m_u + \gamma^\mu \left(\partial_\mu - ig_3 (G_\mu^\beta + \delta G_\mu^\beta) \frac{\lambda_\beta}{2} \right) \right] \\ &\quad \times \left(1 + ig_3 \omega^\gamma \frac{\lambda_\gamma}{2}\right) u \\ &= \bar{u} \left[m_u + \gamma^\mu \left(\partial_\mu - ig_3 G_\mu^\alpha \frac{\lambda_\alpha}{2} - ig_3 \delta G_\mu^\alpha \frac{\lambda_\alpha}{2} \right. \right. \\ &\quad \left. \left. + ig_3 (\partial_\mu \omega^\alpha) \frac{\lambda_\alpha}{2} + ig_3^2 f_{\beta\gamma}^\alpha G_\mu^\beta \omega^\gamma \frac{\lambda_\alpha}{2} \right) \right] \quad (1.190) \end{aligned}$$

(at linear order in infinitesimal ω), which is unchanged only if we identify the change under gauge transformations of the field G as

$$\delta G_\mu^\alpha = \partial_\mu \omega^\alpha - gf_{\beta\gamma}^\alpha \omega^\beta G_\mu^\gamma \quad (1.191)$$

reproducing Eq. (1.172). The combination $\partial_\mu G_\nu^\alpha - \partial_\nu G_\mu^\alpha$ transforms quite non-trivially under this gauge transformation rule, and is not the correct object to identify as a field strength. However, the combination (Eq. (1.171)) compare with

$$G_{\mu\nu}^\alpha \equiv \partial_\mu G_\nu^\alpha - \partial_\nu G_\mu^\alpha + gf_{\beta\gamma}^\alpha G_\mu^\beta G_\nu^\gamma \quad (1.192)$$

transforms as

$$G_{\mu\nu}^\alpha \rightarrow G_{\mu\nu}^\alpha - f_{\beta\gamma}^\alpha \omega^\beta G_{\mu\nu}^\gamma \quad (1.193)$$

and therefore the combination $G_{\mu\nu}^\alpha G^{\alpha\mu\nu}$ is invariant, and may appear in the Lagrangian. The full Lagrangian of QCD is therefore

$$\mathcal{L}_{QCD} = -\sum_q \bar{q}(\not{D} + m)q - \frac{1}{4} G_{\mu\nu}^\alpha G^{\alpha\mu\nu} \quad (1.194)$$

where $q = u, d, s$.

The physics of this theory is quite non-trivial and occupies Chapter 8 and Chapter 9.

1.7 Problems

[1.1] Identities for Majorana spinors

Prove the following useful relations for Majorana spinors ψ_1, ψ_2 ,

$$\begin{aligned}
 \bar{\psi}_1 \psi_2 &= +\bar{\psi}_2 \psi_1 \\
 \bar{\psi}_1 \gamma_5 \psi_2 &= +\bar{\psi}_2 \gamma_5 \psi_1 \\
 \bar{\psi}_1 \gamma^\mu \psi_2 &= -\bar{\psi}_2 \gamma^\mu \psi_1 \\
 \bar{\psi}_1 \gamma^\mu \gamma_5 \psi_2 &= +\bar{\psi}_2 \gamma^\mu \gamma_5 \psi_1 \\
 \bar{\psi}_1 [\gamma^\mu, \gamma^\nu] \psi_2 &= -\bar{\psi}_2 [\gamma^\mu, \gamma^\nu] \psi_1
 \end{aligned}$$

Hint: It is possible to invert the order of a series of matrices which contract a column vector on the right and row vector on the left, $cM_1 M_2 v = v^T M_2^T M_1^T c^T$, for instance. However, since the operators ψ_1, ψ_2 are anti-commuting objects, there is a factor of -1 when doing so here; so $\bar{\psi}_1 \psi_2 = -\psi_2^T \bar{\psi}_1^T$. Use this manipulation, and the identities in Eq. (1.93) and Eq. (1.94).

Next, show that for any spinors, Hermitian conjugation takes the form,

$$\begin{aligned}
 (\bar{\psi}_1 \psi_2)^* &= +\bar{\psi}_2 \psi_1 \\
 (\bar{\psi}_1 \gamma_5 \psi_2)^* &= -\bar{\psi}_2 \gamma_5 \psi_1 \\
 (\bar{\psi}_1 \gamma^\mu \psi_2)^* &= -\bar{\psi}_2 \gamma^\mu \psi_1 \\
 (\bar{\psi}_1 \gamma^\mu \gamma_5 \psi_2)^* &= -\bar{\psi}_2 \gamma^\mu \gamma_5 \psi_1 \\
 (\bar{\psi}_1 [\gamma^\mu, \gamma^\nu] \psi_2)^* &= -\bar{\psi}_2 [\gamma^\mu, \gamma^\nu] \psi_1
 \end{aligned}$$

by using repeatedly Eq. (1.90). Note that Hermitian conjugation involves a reversal of the order of operators, so $(\psi_1^\dagger \psi_2)^\dagger = \psi_2^\dagger \psi_1$ without a minus sign.

Combine these to get the following relations for Majorana spinors:

$$\begin{aligned}
 (\bar{\psi}_1 \psi_2)^* &= +\bar{\psi}_1 \psi_2 \\
 (\bar{\psi}_1 \gamma_5 \psi_2)^* &= -\bar{\psi}_1 \gamma_5 \psi_2 \\
 (\bar{\psi}_1 \gamma^\mu \psi_2)^* &= +\bar{\psi}_1 \gamma^\mu \psi_2 \\
 (\bar{\psi}_1 \gamma^\mu \gamma_5 \psi_2)^* &= -\bar{\psi}_1 \gamma^\mu \gamma_5 \psi_2 \\
 (\bar{\psi}_1 [\gamma^\mu, \gamma^\nu] \psi_2)^* &= +\bar{\psi}_1 [\gamma^\mu, \gamma^\nu] \psi_2
 \end{aligned}$$

Use these to justify the requirements on the coefficients A, B, C, D , and E mentioned under Eq. (1.102).

[1.2] $\mathcal{O}(N)$ scalar theories

The kinetic term $\frac{1}{2}\partial_\mu\varphi_i\partial^\mu\varphi_i$ for N real scalar fields is invariant under a symmetry $\varphi_i \rightarrow \mathcal{O}_{ij}\varphi_j$, where $\mathcal{O}^T\mathcal{O} = 1$, $i, j = 1, \dots, N$. These form the group of $N \times N$ real orthogonal matrices $\mathcal{O}(N)$. When N is even, $\mathcal{O}(N)$ contains as a subgroup the group of $(N/2) \times (N/2)$ complex unitary matrices, $U(N/2)$. When the interactions respect only this subgroup rather than the full $\mathcal{O}(N)$ group, it is often convenient to use complex fields.

[1.2.1] Example 1: $N = 2$.

- (i) Write down the most general renormalizable Lagrangian for two real scalar fields, φ_1 and φ_2 , subject to the discrete symmetries $\varphi_1 \rightarrow -\varphi_1$, $\varphi_2 \rightarrow \varphi_2$ and $\varphi_1 \rightarrow \varphi_1$, $\varphi_2 \rightarrow -\varphi_2$.
- (ii) Re-express this Lagrangian in terms of the complex variables $\psi = \frac{1}{\sqrt{2}}(\varphi_1 + i\varphi_2)$ and $\psi^* = \frac{1}{\sqrt{2}}(\varphi_1 - i\varphi_2)$.
- (iii) In this case the groups $\mathcal{O}(2)$ and $U(1)$ are equivalent to one another. If the $\mathcal{O}(2)$ transformations are written

$$\mathcal{O}(\theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

find the transformation rules for ψ and ψ^* .

- (iv) What further restrictions are placed on the Lagrangian by requiring that it be $\mathcal{O}(2)$ invariant (including interaction terms)? Write the resulting Lagrangian in terms of both the variables (φ_1, φ_2) and (ψ, ψ^*) .
- (v) Assuming the coupling to be weak, allowing a semiclassical approximation, what is the ground state (i.e. background value for the fields) and spectrum (i.e. masses) of this $\mathcal{O}(2)$ -symmetric model if the coefficient of the quadratic term of the potential is positive? What are the ground states and spectrum if the coefficient of the quadratic term is negative? Which field is massless (such a massless field is called a Goldstone boson)?

[1.2.2] Example 2: $N = 4$.

- (i) What is the most general form for a renormalizable theory of four real scalars, (assuming as above invariance under separate reflections of each field)?
- (ii) In this case the maximal symmetry group is $\mathcal{O}(4)$ which consists of 4×4 real orthogonal matrices. These by definition are the group that leaves $\phi^T\phi = (\phi_1)^2 + (\phi_2)^2 + (\phi_3)^2 + (\phi_4)^2$ invariant. As a group $\mathcal{O}(4)$ is equivalent to $SU(2) \times SU(2)$. This can be seen as follows: Define the complex fields $\varphi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$ and

$\psi = \frac{1}{\sqrt{2}}(\phi_3 + i\phi_4)$ together with their complex conjugates and construct the 2×2 matrix whose columns are $\chi \equiv \begin{pmatrix} \varphi \\ \psi \end{pmatrix}$ and $\bar{\chi} \equiv \epsilon\chi^* = \begin{pmatrix} \psi^* \\ -\varphi^* \end{pmatrix}$, i.e.

$$\Phi = \begin{pmatrix} \varphi & \psi^* \\ \psi & -\varphi^* \end{pmatrix}$$

Then Φ satisfies

$$\bar{\Phi} \equiv \epsilon\Phi^*\epsilon = \Phi \quad (1.195)$$

$$\det \Phi = -(\varphi^*\varphi + \psi^*\psi) = -\frac{1}{2} \sum_{i=1}^4 \phi_i^2 = -\frac{1}{2} \phi^T \phi \quad (1.196)$$

The group $\mathcal{O}(4)$ can therefore be described as those linear transformations of Φ that preserve Eq. (1.195) and Eq. (1.196). Show that these conditions are satisfied by

- (a) $\Phi \rightarrow U\Phi$; or
- (b) $\Phi \rightarrow \Phi V$

for U and V arbitrary 2×2 complex unitary matrices with unit determinant. Transformations (A) and (B) each form an $SU(2)$ group and $\mathcal{O}(4) \approx SU(2) \times SU(2)$.

- (iii) The complex variable $\chi = \begin{pmatrix} \varphi \\ \psi \end{pmatrix}$ is convenient if invariance under only one of the $SU(2)$ s is required. Choosing this to be the $SU(2)$ formed by multiplication on the left, χ and $\bar{\chi}$ transform as doublets: $\chi \rightarrow U\chi, \bar{\chi} \rightarrow U\bar{\chi}$. Construct the most general renormalizable Lagrangian consistent with invariance under a single $SU(2)$. Did you include the invariant term

$$\sum_{a=1}^3 (\chi^+ \tau^a \chi)(\chi^+ \tau^a \chi)$$

with τ^a being the Pauli matrices? Should you? Which terms, if any, are not invariant under the “other” $SU(2)$?

- (iv) For the $SU(2)$ -invariant model, give the ground state and spectrum in the semiclassical approximation for both choices of sign for the coefficient of the quadratic term of the potential. When

the background field is non-zero, what subgroup of the original invariance group leaves the background fields invariant? What is the dimension of this subgroup? What is the dimension of the original symmetry group? How many massless states are there?

[1.3] Vacuum energies

Consider the model consisting of one free Majorana fermion and one complex scalar field:

$$\mathcal{L} = -\frac{1}{2}\bar{\psi}(\not{\partial} + m)\psi - (\partial_\mu\varphi)^*(\partial^\mu\varphi) - \mu^2\varphi^*\varphi$$

The Hamiltonian density for this model is (defining $\dot{\varphi} = \partial_t\varphi$)

$$\mathcal{H}(x) = \dot{\varphi}^*\dot{\varphi} + (\nabla\varphi)^*(\nabla\varphi) + \mu^2\varphi^*\varphi + \frac{1}{2}\bar{\psi}(\gamma\cdot\nabla + m)\psi$$

Use the mode expansions

$$\varphi(x) = \frac{1}{\sqrt{2}} \left[\varphi^{(1)} + i\varphi^{(2)} \right]$$

$$\varphi^{(i)}(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_{\mathbf{p}}} \left[e^{ipx} a_{\mathbf{p}}^{(i)} + e^{-ipx} a_{\mathbf{p}}^{*(i)} \right]$$

$$\psi(x) = \sum_{\sigma=\pm\frac{1}{2}} \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_{\mathbf{p}}} \left[u_{\mathbf{p},\sigma} e^{ipx} b_{\mathbf{p},\sigma} + v_{\mathbf{p},\sigma} e^{-ipx} b_{\mathbf{p},\sigma}^* \right]$$

to express the total energy, H , in terms of the creation and annihilation operators $a_{\mathbf{p}}^{(i)}$ and $b_{\mathbf{p},\sigma}$. What is the zero-point energy in this theory? What is the zero-point energy when $\mu = m$? Assume the standard ordering convention: $AB \rightarrow \frac{1}{2}(AB + BA) \equiv \frac{1}{2}\{A, B\}$ when quantizing. Also assume the standard relations:

$$[a_{\mathbf{p}}^{(i)}, a_{\mathbf{p}'}^{*(j)}] = 2E_{\mathbf{p}}(2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}') \delta^{ij}$$

$$\{b_{\mathbf{p},\sigma}, b_{\mathbf{p}',\sigma'}^*\} = \delta_{\sigma,\sigma'} 2E_{\mathbf{p}}(2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}')$$

and

$$[a_{\mathbf{p}}^{(i)}, b_{\mathbf{p}',\sigma}] = 0 \text{ etc.}$$

[1.4] Symmetries and Yukawa interactions

Consider a theory with one Majorana fermion, and two real scalar fields φ, χ subject to the symmetry:

$$\begin{aligned}\delta\psi &= i\omega\gamma_5\psi \\ \delta\varphi &= 2\omega\chi \\ \delta\chi &= -2\omega\varphi\end{aligned}$$

for ω an infinitesimal, spatially constant parameter.

[1.4.1] Write down the most general renormalizable Lagrangian coupling the scalars to each other and to the fermion. Identify the vacuum field configuration and mass spectrum both in the broken and unbroken phases (i.e. for both choices of sign for the coefficient of the quadratic term of the potential).

[1.4.2] Couple a spin-one particle to this symmetry; i.e., write down covariant derivatives for the fields ψ , χ , and φ and construct an action invariant with respect to these transformations with $\partial_\mu\omega \neq 0$. Again identify the spectrum in both broken and unbroken phases.

[1.5] Spinor identities

Derive the following formulae concerning the spin-half wave function:

$$\begin{aligned}u(\mathbf{p}, \sigma)\bar{u}(\mathbf{p}, \sigma) &= \frac{1}{2}(m - i\not{p})(1 + i\gamma_5\not{s}) \\ \text{and } \sum_{\sigma=\pm\frac{1}{2}} u(\mathbf{p}, \sigma)\bar{u}(\mathbf{p}, \sigma) &= (m - i\not{p})\end{aligned}$$

in which p^μ is the particle four-momentum and s^μ is a four-vector whose components in the rest frame are $s^0 = 0$, $\mathbf{s} = 2\sigma\mathbf{e}$ where \mathbf{e} is the unit vector in the direction along which the spin components, $\sigma = \pm\frac{1}{2}$, are measured. (Choose \mathbf{e} to lie along the positive x^3 -axis.) Notice these imply the frame-independent conditions: $s^\mu s_\mu = +1$ and $s^\mu p_\mu = 0$. Recall, also that $(i\not{p} + m)u(\mathbf{p}, \sigma) = 0$ and $p^\mu p_\mu = -m^2$.

Hint: Since $u(\mathbf{p}, \sigma)\bar{u}(\mathbf{p}, \sigma)$ is a 4×4 matrix, expand it in terms of the basis matrices S, P, V, A , and T , defined as $S = \mathbf{1}$, $P = \gamma_5$, $V = \gamma^\mu$, $A = \gamma^\mu\gamma^5$, and $T = [\gamma^\mu, \gamma^\nu]$. Since $E_{\mathbf{p}}u\bar{u}$ transforms covariantly under Lorentz transformations, the coefficients of these matrices are scalars, pseudoscalars, vectors, etc. Evaluate the coefficients by taking traces after multiplying by an appropriate matrix. It may prove convenient to evaluate those coefficients that transform as vectors and tensors in the rest frame of the particle.

Is the resulting expression well behaved in the zero-mass limit?

[1.6] Fermion mass matrix diagonalization

Prove the theorem that, for any complex, symmetric matrix, A , there exists a unitary matrix, U , for which

$$U^T A U = M$$

is real, diagonal and non-negative. (Recall we used this theorem to show that the spin-half mass matrix could always be put into standard form.)

Hint: Notice that this would be trivial if $[A, A^\dagger] = 0$ because then if we break A into its real and imaginary parts, $A = R + iS$ for R, S real and symmetric, we see that $[A, A^\dagger] = 2i[S, R] = 0$. Since S and R are both real and symmetric and commute, they can both be diagonalized by the same real orthogonal matrix, \mathcal{O} . This implies that $\mathcal{O}^T A \mathcal{O} = \text{diag}$ and we could define $U = \mathcal{O} D$ with D being a diagonal matrix whose elements are phases that can be chosen to make the entries of M real and non-negative. In the general case when $[A, A^\dagger] \neq 0$, we know that $A^\dagger A$ is Hermitian and so is diagonalizable by a unitary matrix, V . Define the new matrix $B \equiv V^T A V$ and show that $B = B^T$ and $[B, B^\dagger] = 0$.

[1.7] A Dirac matrix identity

Prove the identity which shows that $\gamma_{\mu\nu}\gamma_5$ is not linearly independent of $\epsilon_{\mu\nu\lambda\rho}\gamma^{\lambda\rho}$

$$\epsilon_{\mu\nu\lambda\rho}\gamma^{\lambda\rho} = 2i\gamma_{\mu\nu}\gamma_5 \quad (1.197)$$

Here $\gamma^{\mu\nu} = \frac{1}{2}[\gamma^\mu, \gamma^\nu]$.

[1.8] More useful identities: Prove the following identities:

[1.8.1]

$$\begin{aligned} \gamma^\mu\gamma^\nu\gamma^\lambda\gamma^\rho P_R &= (\eta^{\mu\nu}\eta^{\lambda\rho} - \eta^{\mu\lambda}\eta^{\nu\rho} + \eta^{\mu\rho}\eta^{\lambda\nu} - i\epsilon^{\mu\nu\lambda\rho})P_R \\ &+ (\eta^{\mu\nu}\gamma^{\lambda\rho} - \eta^{\mu\lambda}\gamma^{\nu\rho} + \eta^{\mu\rho}\gamma^{\nu\lambda} - \eta^{\nu\lambda}\gamma^{\rho\mu} + \eta^{\nu\rho}\gamma^{\lambda\mu} - \eta^{\lambda\rho}\gamma^{\nu\mu})P_R \end{aligned} \quad (1.198)$$

Here $P_R = \frac{1}{2}(1 - \gamma_5)$ is the usual projection matrix onto right-handed spinors and $\gamma^{\mu\nu} = \frac{1}{2}[\gamma^\mu, \gamma^\nu]$ is half of the commutator of two Dirac matrices.

[1.8.2] For X and Y any product of an *odd* number of gamma matrices prove the following trace formula:

$$\text{tr}[XY P_R] = \frac{1}{2} \text{tr}[X\gamma^\mu P_R] \text{tr}[Y\gamma_\mu P_L] \quad (1.199)$$

P_R is as before and $P_L = \frac{1}{2}(1 + \gamma_5)$.

[1.9] **Fiertz rearrangements**

The sixteen Dirac matrices I , γ_5 , γ^μ , $\gamma^\mu\gamma_5$, and $\gamma^{\mu\nu} = \frac{1}{2}[\gamma^\mu, \gamma^\nu]$ provide a basis in terms of which any 4×4 complex matrix can be expressed (prove this). Given this property, show that this provides the following useful way to rewrite a dyadic product of two anticommuting spinors:

$$\begin{aligned}
 P_L[\psi_1\bar{\psi}_2]P_R &= -\frac{1}{2}[\bar{\psi}_2\gamma^\mu P_L\psi_1]\gamma_\mu P_R \\
 P_R[\psi_1\bar{\psi}_2]P_L &= -\frac{1}{2}[\bar{\psi}_2\gamma^\mu P_R\psi_1]\gamma_\mu P_L \\
 P_L[\psi_1\bar{\psi}_2]P_L &= -\frac{1}{2}[\bar{\psi}_2 P_L\psi_1]P_L - \frac{1}{8}[\bar{\psi}_2\gamma^{\mu\nu} P_L\psi_1]\gamma_{\mu\nu}P_L \\
 P_R[\psi_1\bar{\psi}_2]P_R &= -\frac{1}{2}[\bar{\psi}_2 P_R\psi_1]P_R - \frac{1}{8}[\bar{\psi}_2\gamma^{\mu\nu} P_R\psi_1]\gamma_{\mu\nu}P_R \quad (1.200)
 \end{aligned}$$

2

The standard model: general features

The last chapter developed the general principles for writing down a relativistic quantum field theory. It showed what types of fields are possible, and explained that spin-one fields can only appear in an interacting, renormalizable theory if they are coupled via the gauge principle.

In this chapter, we write down specifically what the field content of the standard model is. The interactions will then follow as the most general set of renormalizable interactions, compatible with that field content. We then explore what the vacuum and the particle content are, and write down the complete interaction Hamiltonian in the particle basis.

We will not attempt to motivate theoretically why the particle content of the standard model is what it is. We have no deep understanding of why the gauge group is $SU_c(3) \times SU_L(2) \times U_Y(1)$, for instance. We just take the field content as observed fact, and present it. The exception is the Higgs boson, which has not been observed. This is the weakest part of our understanding of the standard model. Note however that the field content of the standard model is not completely arbitrary; once the gauge group is known, the fermionic field content is somewhat constrained by the requirement of anomaly cancellation, which we discuss at the end of the chapter.

2.1 Particle content

The strong, weak, and electromagnetic interactions are understood as arising due to the exchange of various spin-one bosons amongst the spin-half particles that make up matter. The gauged symmetry group of the standard model is $SU_c(3) \times SU_L(2) \times U_Y(1)$. The specific gauge bosons associated with

the generators of the algebra of the group are:

$$\begin{array}{ccccc}
 SU_c(3) & \times & SU_L(2) & \times & U_Y(1) \\
 \downarrow & & \downarrow & & \downarrow \\
 8 \ G_\mu^\alpha & & 3 \ W_\mu^a & & B_\mu \\
 \alpha = 1, \dots, 8 & & a = 1, 2, 3 & &
 \end{array} \tag{2.1}$$

The eight spin-one particles, $G_\mu^\alpha(x)$, associated with the factor $SU_c(3)$ are called *gluons* and the associated subscript “c” is meant to denote “color.” Gluons are thought to be massless. Any particle that transforms with respect to this factor of the gauge group, and so which couples to the gluons, is said to be colored or to carry color. This interaction is also called the “strong interaction,” and any particle which couples to the gluons is said to be “strongly interacting.” Three spin-one particles, $W_\mu^a(x)$, are associated with the factor $SU_L(2)$, and one, $B_\mu(x)$, with the factor $U_Y(1)$. The subscript “L” is meant to indicate that only the left-handed fermions turn out to carry this quantum number. The subscript “Y” is meant to distinguish the group associated with the quantum number (defined below) of *weak hypercharge*, denoted Y , from the group associated with ordinary electric charge, denoted Q . The electromagnetic group will be written as $U_{em}(1)$. The four spin-one bosons associated with the factors $SU_L(2) \times U_Y(1)$ are related to the physical bosons that mediate the weak interactions, W^\pm and Z^0 , and the familiar photon from QED, in a way we will explain in Section 2.3.

Apart from spin-one particles we are aware of a number of fundamental spin-half particles. Our knowledge to date about the character of the interactions of these fermions may be compactly summarized by giving their transformation properties with respect to the gauge group $SU_c(3) \times SU_L(2) \times U_Y(1)$. The fermions transform in a relatively complicated way with respect to this symmetry group. There are three copies (or families) of particles, each copy of which couples identically to all spin-one particles.

Leptons are, by definition, those spin-half particles which do not take part in the strong interactions. Six leptons are known to date. They are denoted individually by $e, \mu, \tau, \nu_e, \nu_\mu$, and ν_τ , and collectively by ℓ .

Hadrons, on the other hand, are defined as those particles that do take part in the strong interactions. The spectrum of known hadrons is rich and varied but, as we shall see, appears to be accounted for as the bound states of six quarks u, d, s, c, b , and t , denoted collectively as q .

Because of the relatively large number of spin-half fields involved, a few words on notation may be appropriate. Spinors written in capital letters L, E, D, U, Q , or script letters $\mathcal{E}, \mathcal{U}, \mathcal{D}$, and neutrinos ν_i are taken as

Majorana spinors. The left- and right-handed components of these spinors are denoted by subscripts L, R. Spinors written in lower case Roman letters $l_i, u_i, d_i, e, u, c, t, d, s, b$, or by μ, τ are Dirac spinors, which we will introduce in turn.

For example, the electron field is represented in quantum electrodynamics by the Dirac spinor, $e(x)$. Denote the left- and right-handed components of this spinor by e_L and e_R respectively:

$$e = \begin{pmatrix} e_L \\ e_R \end{pmatrix} \quad (2.2)$$

In the standard model, however, the electron is represented by two Majorana fields, $\mathcal{E}(x)$ and $E(x)$, that are defined to contain the left- and right-handed parts of $e(x)$ respectively. That is,

$$\mathcal{E} = \begin{pmatrix} e_L \\ \epsilon e_L^* \end{pmatrix}, \quad E = \begin{pmatrix} -\epsilon e_R^* \\ e_R \end{pmatrix} \quad (2.3)$$

where the 2×2 matrix ϵ is defined in Eq. (1.79). The Dirac spinor, e , is therefore related to the Majorana fields, \mathcal{E} and E , by projecting onto the left- or right-handed part:

$$e = P_L \mathcal{E} + P_R E \quad (2.4)$$

The “left-handed” electron field, \mathcal{E} , itself appears within an $SU_L(2)$ -doublet with the field, ν , whose left-handed part contains the left-handed electron-neutrino. This doublet is denoted $L(x)$:

$$P_L L = \begin{pmatrix} P_L \nu \\ P_L \mathcal{E} \end{pmatrix} \quad (2.5)$$

The notation here is somewhat confusing; the matrix structure shown for L above does not show spinorial matrix structure, but shows matrix structure under the group $SU_L(2)$; each component, ν and \mathcal{E} , is a 4-component Majorana spinor. Generally, when possible spinorial structure is suppressed in what follows.

Members of successive generations are denoted by a generation index, m , that runs from 1 to 3. The generations are numbered in increasing order with respect to the mass of the corresponding charged lepton:

$$\begin{aligned} \nu_m &\text{ denotes } \nu_1 = \nu_e, \quad \nu_2 = \nu_\mu, \quad \nu_3 = \nu_\tau \\ e_m &\text{ denotes } e_1 = e, \quad e_2 = \mu, \quad e_3 = \tau \\ u_m &\text{ denotes } u_1 = u, \quad u_2 = c, \quad u_3 = t \\ \text{and } d_m &\text{ denotes } d_1 = d, \quad d_2 = s, \quad \text{and } d_3 = b \end{aligned} \quad (2.6)$$

The transformation properties of the fermions are summarized by giving the representation of the gauge group in which they transform. A standard way to label the representations of $SU_L(2)$ and $SU_c(3)$ is with their dimension. So the two-dimensional spinor representation of $SU_L(2)$ is written **2** (familiar from the physics of spin as the spin-half representation) and the two three-dimensional representations of $SU_c(3)$ would be **3** or $\bar{\mathbf{3}}$. The trivial (invariant) representation is written as **1**. The transformation properties of the known fermions with respect to $U_Y(1)$ may be specified by giving the corresponding eigenvalue of the generator, Y , called the *weak hypercharge*. Y is normalized so that the action of $U_Y(1)$ on a field with eigenvalue y is given by $\psi \rightarrow \exp[i\omega(x)y]\psi$.

With these conventions the particle content of the standard model may be summarized as follows:

$$\begin{array}{ll}
 P_L L_m = \begin{pmatrix} P_L \nu_m \\ P_L \mathcal{E}_m \end{pmatrix} & \text{transforms as } \left(\mathbf{1}, \mathbf{2}, -\frac{1}{2} \right) \\
 P_R E_m & \left(\mathbf{1}, \mathbf{1}, -1 \right) \\
 P_L Q_m = \begin{pmatrix} P_L \mathcal{U}_m \\ P_L \mathcal{D}_m \end{pmatrix} & \left(\mathbf{3}, \mathbf{2}, +\frac{1}{6} \right) \\
 P_R U_m & \left(\mathbf{3}, \mathbf{1}, +\frac{2}{3} \right) \\
 P_R D_m & \left(\mathbf{3}, \mathbf{1}, -\frac{1}{3} \right) \quad (2.7)
 \end{array}$$

Here the first number represents the $SU_c(3)$ representation, the second number is the $SU_L(2)$ representation and the final number is the eigenvalue of the weak hypercharge, Y . In the case of $SU_L(2)$ doublets, we have named their upper and lower $SU_L(2)$ components, $L_m = (P_L \nu_m P_L \mathcal{E}_m)^T$ and $Q_m = (P_L \mathcal{U}_m P_L \mathcal{D}_m)^T$. We could in principle do this for the three separate colors of the Q , U , and D fields; but it turns out to be useful to do so for the $SU_L(2)$ content but not for the $SU_c(3)$ content.

Since the left- and right-handed pieces of a Majorana spinor are the complex conjugates of one another, they must transform in complex-conjugate representations. It follows then that

$$\begin{array}{ll}
 P_R L_m = \begin{pmatrix} P_R \nu_m \\ P_R \mathcal{E}_m \end{pmatrix} & \text{transforms as } \left(\mathbf{1}, \mathbf{2}, +\frac{1}{2} \right) \\
 P_L E_m & \left(\mathbf{1}, \mathbf{1}, +1 \right)
 \end{array}$$

$$\begin{aligned}
P_R Q_m &= \begin{pmatrix} P_R \mathcal{U}_m \\ P_R \mathcal{D}_m \end{pmatrix} & \left(\bar{\mathbf{3}}, \mathbf{2}, -\frac{1}{6} \right) \\
P_L U_m & & \left(\bar{\mathbf{3}}, \mathbf{1}, -\frac{2}{3} \right) \\
P_L D_m & & \left(\bar{\mathbf{3}}, \mathbf{1}, +\frac{1}{3} \right)
\end{aligned} \tag{2.8}$$

We note in passing that if the standard model were to be supplemented to include a right-handed neutrino field, N_m , this field would be a singlet,

$$P_R N_m \quad \text{transforms as:} \quad (\mathbf{1}, \mathbf{1}, 0) \tag{2.9}$$

with respect to the gauge group $SU_c(3) \times SU_L(2) \times U_Y(1)$. We will discuss such a singlet some more in Chapter 10, see also Problem 2.3.

Apart from fermions, the Lagrangian must also involve the fields representing the spin-one gauge bosons. These fields and their transformation rules are denoted as follows:

$$\begin{aligned}
G_\mu^\alpha & \quad \text{transforms as:} \quad (\mathbf{8}, \mathbf{1}, 0) \\
W_\mu^a & \quad (\mathbf{1}, \mathbf{3}, 0) \\
B_\mu & \quad (\mathbf{1}, \mathbf{1}, 0)
\end{aligned} \tag{2.10}$$

This representation content is merely a short form for the invariance of the Lagrangian under the following symmetries:

$$\begin{aligned}
\delta L_m &= \left[\left(-\frac{i}{2} \omega_1(x) + \frac{i}{2} \omega_2^a(x) \tau_a \right) P_L + \left(\frac{i}{2} \omega_1(x) - \frac{i}{2} \omega_2^a(x) \tau_a^* \right) P_R \right] L_m \\
\delta E_m &= [i \omega_1(x) P_L - i \omega_1(x) P_R] E_m \\
\delta Q_m &= \left[\left(\frac{i}{6} \omega_1(x) + \frac{i}{2} \omega_2^a(x) \tau_a + \frac{i}{2} \omega_3^\alpha(x) \lambda_\alpha \right) P_L + \right. \\
&\quad \left. + \left(-\frac{i}{6} \omega_1(x) - \frac{i}{2} \omega_2^a(x) \tau_a^* - \frac{i}{2} \omega_3^\alpha(x) \lambda_\alpha^* \right) P_R \right] Q_m \\
\delta U_m &= \left[\left(-\frac{2i}{3} \omega_1(x) - \frac{i}{2} \omega_3^\alpha(x) \lambda_\alpha^* \right) P_L + \left(\frac{2i}{3} \omega_1(x) + \frac{i}{2} \omega_3^\alpha(x) \lambda_\alpha \right) P_R \right] U_m \\
\delta D_m &= \left[\left(\frac{i}{3} \omega_1(x) - \frac{i}{2} \omega_3^\alpha(x) \lambda_\alpha^* \right) P_L + \left(-\frac{i}{3} \omega_1(x) + \frac{i}{2} \omega_3^\alpha(x) \lambda_\alpha \right) P_R \right] D_m \\
\delta G_\mu^\alpha &= \partial_\mu \omega_3^\alpha(x) - f_{\beta\gamma}^\alpha \omega_3^\beta(x) G_\mu^\gamma \\
\delta W_\mu^a &= \partial_\mu \omega_2^a(x) - \epsilon^{abc} \omega_2^b(x) W_\mu^c \\
\delta B_\mu &= \partial_\mu \omega_1(x)
\end{aligned} \tag{2.11}$$

In these expressions the generators of $SU_L(2)$ have been explicitly written as $T_a = \frac{1}{2} \tau_a$ where $\tau_a, a = 1, 2, 3$ denotes the 2×2 *Pauli* matrices that act

on the $SU_L(2)$ -indices

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2.12)$$

(The same matrices appeared in discussing the spin structure of fermions in Section 1.3. We use the notation τ_i when they act on $SU_L(2)$ indices and σ_i when they act on spinorial indices.) Similarly, the generators of $SU_c(\mathbf{3})$ (when acting on the $\mathbf{3}$ representation) are given explicitly by $T_\alpha = \frac{1}{2}\lambda_\alpha$ where $\lambda_\alpha, \alpha = 1, \dots, 8$ denote the 3×3 Gell-Mann matrices given in Eq. (1.186).

The electric charge Q of a field is defined in terms of the hypercharge Y and the $SU_L(2)$ charge's T_3 component, according to $Q = T_3 + Y$. Note that the electromagnetic group is *not* directly the $U_Y(1)$ component of the standard model gauge group, and electric charge Q is *not* one of the basic charges particles carry under $SU_c(3) \times SU_L(2) \times U_Y(1)$; rather it is a derived quantity.

2.2 The Lagrangian

The most general renormalizable Lagrangian involving these fields is

$$\begin{aligned} \mathcal{L}_{fg} = & -\frac{1}{4}G_{\mu\nu}^{\alpha}G^{\alpha\mu\nu} - \frac{1}{4}W^{a\mu\nu}W_{\mu\nu}^a - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{g_3^2\Theta_3}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}G^{\alpha\mu\nu}G^{\alpha\lambda\rho} \\ & - \frac{g_2^2\Theta_2}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}W^{a\mu\nu}W^{a\lambda\rho} - \frac{g_1^2\Theta_1}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}B^{\mu\nu}B^{\lambda\rho} - \frac{1}{2}\bar{L}_m\not{D}L_m \\ & - \frac{1}{2}\bar{E}_m\not{D}E_m - \frac{1}{2}\bar{Q}_m\not{D}Q_m - \frac{1}{2}\bar{U}_m\not{D}U_m - \frac{1}{2}\bar{D}_m\not{D}D_m \end{aligned} \quad (2.13)$$

in which the gauge field-strengths are given by

$$G_{\mu\nu}^{\alpha} = \partial_{\mu}G_{\nu}^{\alpha} - \partial_{\nu}G_{\mu}^{\alpha} + g_3f^{\alpha}_{\beta\gamma}G_{\mu}^{\beta}G_{\nu}^{\gamma} \quad (2.14)$$

$$W_{\mu\nu}^a = \partial_{\mu}W_{\nu}^a - \partial_{\nu}W_{\mu}^a + g_2\epsilon_{abc}W_{\mu}^bW_{\nu}^c \quad (2.15)$$

$$B_{\mu\nu} = \partial_{\mu}B_{\nu} - \partial_{\nu}B_{\mu} \quad (2.16)$$

The gauge-covariant derivatives are

$$\begin{aligned} D_{\mu}L_m = & \partial_{\mu}L_m + \left[\frac{i}{2}g_1B_{\mu} - \frac{i}{2}g_2W_{\mu}^a\tau_a \right] P_L L_m \\ & + \left[-\frac{i}{2}g_1B_{\mu} + \frac{i}{2}g_2W_{\mu}^a\tau_a^* \right] P_R L_m \end{aligned} \quad (2.17)$$

$$D_{\mu}E_m = \partial_{\mu}E_m + ig_1B_{\mu}(P_R E_m) - ig_1B_{\mu}(P_L E_m) \quad (2.18)$$

$$D_{\mu}Q_m = \partial_{\mu}Q_m + \left[-\frac{i}{2}g_3G_{\mu}^{\alpha}\lambda_{\alpha} - \frac{i}{2}g_2W_{\mu}^a\tau_a - \frac{i}{6}g_1B_{\mu} \right] P_L Q_m$$

$$+ \left[\frac{i}{2} g_3 G_\mu^\alpha \lambda_\alpha^* + \frac{i}{2} g_2 W_\mu^a \tau_a^* + \frac{i}{6} g_1 B_\mu \right] P_R Q_m \quad (2.19)$$

$$D_\mu U_m = \partial_\mu U_m + \left[-\frac{i}{2} g_3 G_\mu^\alpha \lambda_\alpha - \frac{2i}{3} g_1 B_\mu \right] P_R U_m + \left[\frac{i}{2} g_3 G_\mu^\alpha \lambda_\alpha^* + \frac{2i}{3} g_1 B_\mu \right] P_L U_m \quad (2.20)$$

$$D_\mu D_m = \partial_\mu D_m + \left[-\frac{i}{2} g_3 G_\mu^\alpha \lambda_\alpha + \frac{i}{3} g_1 B_\mu \right] P_R D_m + \left[+\frac{i}{2} g_3 G_\mu^\alpha \lambda_\alpha^* - \frac{i}{3} g_1 B_\mu \right] P_L D_m \quad (2.21)$$

It is worth emphasizing at this point why certain terms do *not* appear in \mathcal{L}_{fg} , particularly mass terms for the fermionic fields. The reason is that only terms which are singlets under $SU_c(3) \times SU_L(2) \times U_Y(1)$ can appear in the Lagrangian – otherwise it would not respect gauge invariance, that is, it would change under a gauge transformation. The rules for telling if a combination of fields is a singlet under $SU_c(3)$ or $SU_L(2)$ are summarized in appendix B; basically the rule is that all color and $SU_L(2)$ indices must “tie off” against each other. The rule for $U_Y(1)$ is even easier; the charges of the fields must add to zero.

Consider for instance the would-be mass term for the E field,

$$\mathcal{L}_{\text{would-be}} = -\frac{m_{mn}}{2} \bar{E}_m E_n$$

Write $\bar{E}_m E_n = \bar{E}_m P_L E_n + \bar{E}_m P_R E_n$, and just consider the P_R term. $P_R E$ has hypercharge -1 . The hypercharge of $\bar{E} P_R$ is also -1 . To see this, note that

$$\bar{E} P_R = E^\dagger \beta P_R = E^\dagger P_L \beta \quad (2.22)$$

is actually the conjugate field of $P_L E$, and has the opposite charge as $P_L E$. Therefore, the combination $\bar{E} P_R E$ is hypercharge -2 and is not a gauge singlet. The combination $\bar{E} P_L E$ is hypercharge -2 and is also not allowed. One can quickly check that no combination of two spinor fields is hypercharge neutral, so no such mass is permitted. The kinetic terms *are* invariant because $P_L \gamma^\mu = \gamma^\mu P_R$; so the left-handed component of a field couples to the Hermitian conjugate of the left-handed component and the gauge dependence does cancel.

The spectrum of this theory may be analyzed by perturbing in the gauge couplings, g_i , $i = 1, 2, 3$. (We return to the accuracy of this approximation in more detail later.) The unperturbed part of the Lagrangian becomes in this case those terms that are quadratic in the fields. The spectrum of this unperturbed theory is therefore that of a system of free spin-half and spin-one

particles, as was described in the previous chapter. Following the discussion leading up to Eq. (1.110)–Eq. (1.125), their masses may be read off from the Lagrangian and are zero!

Since the perturbative semiclassical analysis should apply to at least the electroweak part of the theory, the Lagrangian, Eq. (2.13), cannot be the whole story. In fact, as is clear from the discussion of Subsection 1.6.2, the vanishing of the masses is a consequence of the $SU_c(3) \times SU_L(2) \times U_Y(1)$ invariance of the theory and can be evaded only if this symmetry is spontaneously broken by the ground state. We must couple the known particles to some hitherto undiscovered sector whose ground state is not $SU_L(2) \times U_Y(1)$ invariant.

The simplest way to do so is to add a weakly-coupled spin-zero particle to the theory with a potential that is minimized at a non-zero field value. The transformation properties of this scalar field are largely determined if we require that no new spin-half fields are to be included. Since the scalar field is supposed to produce a mass for the fermions after it develops a *v.e.v.* it must have Yukawa couplings with the fermions. But all of the fermions are either singlets or doublets under $SU_L(2)$ so the new scalar field must itself be either a doublet or a triplet if it is to combine with two fermions into a gauge invariant Yukawa interaction. It turns out that a scalar triplet cannot by itself couple in a way that can generate masses for all the known massive fermions, but a scalar doublet can. The simplest choice is therefore to add a single (complex) scalar doublet, called the Higgs field:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (2.23)$$

transforming as $(1, \mathbf{2}, 1/2)$. Its complex conjugate,

$$\tilde{\phi} \equiv \begin{pmatrix} \phi^{0*} \\ -\phi^{+*} \end{pmatrix} = \epsilon \phi^* \quad (2.24)$$

then transforms as $(1, \mathbf{2}, -1/2)$. It must be emphasized that apart from the conditions that the appropriate masses be generated, there is precious little known about this symmetry-breaking sector. Although the Higgs-doublet model we are using is the simplest, its foundations are much less firm than are those of the rest of the theory. This is explored further in the problem section of this chapter.

The new terms that may then appear in \mathcal{L} are

$$\begin{aligned} \mathcal{L}_{\text{Higgs}} = & -(D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi^\dagger \phi) \\ & -(f_{mn} \bar{L}_m P_R E_n \phi + h_{mn} \bar{Q}_m P_R D_n \phi + g_{mn} \bar{Q}_m P_R U_n \tilde{\phi} + \text{h.c.}) \end{aligned} \quad (2.25)$$

in which

$$\begin{aligned} V(\phi^\dagger\phi) &= \lambda \left[\phi^\dagger\phi - \mu^2/2\lambda \right]^2 \\ &= \lambda(\phi^\dagger\phi)^2 - \mu^2\phi^\dagger\phi + \mu^4/4\lambda \end{aligned} \quad (2.26)$$

Unitarity requires that the constants λ and μ^2 be real and stability demands that λ be positive. In order to ensure that the ground state not be $SU_L(2) \times U_Y(1)$ invariant we further require that μ^2 be positive.

The gauge transformation rules for ϕ are explicitly

$$\delta\phi = \frac{i}{2}\omega_2^a\tau_a\phi + \frac{i}{2}\omega_1\phi \quad (2.27)$$

so its covariant derivative must be

$$D_\mu\phi = \partial_\mu\phi - \frac{i}{2}g_2W_\mu^a\tau_a\phi - \frac{i}{2}g_1B_\mu\phi \quad (2.28)$$

The complete standard model Lagrangian then becomes

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{fg} + \mathcal{L}_{\text{Higgs}} \quad (2.29)$$

The following general features of \mathcal{L}_{SM} bear special mention.

- (i) \mathcal{L}_{fg} , $\mathcal{L}_{\text{Higgs}}$ and \mathcal{L}_{SM} are the most general Lagrangian consistent with the given particle content and invariance under $SU_c(3) \times SU_L(2) \times U_Y(1)$. If the predictions made from such an \mathcal{L} are wrong, then either the particle-content or renormalizability or the gauge group is wrong.
- (ii) Because of $SU_c(3) \times SU_L(2) \times U_Y(1)$ invariance, all masses vanish in the absence of $\mathcal{L}_{\text{Higgs}}$.
- (iii) There are six parameters in \mathcal{L}_{fg} of which only four enter into physical predictions (since Θ_1 and Θ_2 turn out to have no physical effects, for reasons we will not discuss). $\mathcal{L}_{\text{Higgs}}$, on the other hand, contains no less than 15 parameters (as we shall see these may be taken to be the ten masses, the Higgs self-coupling, and the four Kobayashi–Maskawa angles). In this sense $\mathcal{L}_{\text{Higgs}}$ parameterizes most of our ignorance and is the part of the theory that is the least understood. All of the couplings also turn out to be small (modulo some restrictions to which we return for g_3), allowing the use of perturbation theory to calculate the predictions of \mathcal{L} .

2.3 The perturbative spectrum

The first step in analyzing the consequences of the standard model is to find its spectrum. We do so semiclassically, following the procedure of Subsection 1.6.2. For these purposes it is convenient here, as it was there, to use

the gauge freedom to transform to unitary gauge. In the present context unitary gauge is defined by the following condition:

$$\phi = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + H(x)) \end{pmatrix} \quad (2.30)$$

where $H(x)$ is a real field and v is a real constant that minimizes the scalar potential. It may be shown that it is always possible to reach Eq. (2.30) from an arbitrary initial field configuration via a gauge transformation. The motivation for this gauge choice is that it ensures that no vector-scalar cross terms survive in the quadratic terms once we expand about the ground state. It is worth noting in passing that the gauge, Eq. (2.30), does not fix those gauge invariances that leave the Higgs *v.e.v.* invariant. In the present context, as is shown later in this section, this means that the electromagnetic gauge invariance still remains to be fixed.

v is determined by minimizing the potential in Eq. (2.26) and satisfies

$$v^2 = \mu^2/\lambda \quad (2.31)$$

In order to read off the particle masses we must identify the unperturbed Lagrangian, \mathcal{L}_0 . This is equal to that part of \mathcal{L}_{SM} that is quadratic in the fluctuations. The expansion of \mathcal{L}_{fg} is trivial and just contributes the spin-half and spin-one kinetic terms to \mathcal{L}_0 . Everything else comes from the expansion of $\mathcal{L}_{\text{Higgs}}$. Using the following result,

$$D_\mu \phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \partial_\mu H \end{pmatrix} - \frac{i}{2\sqrt{2}} \begin{pmatrix} g_2 W_\mu^3 + g_1 B_\mu & g_2 W_\mu^1 - i g_2 W_\mu^2 \\ g_2 W_\mu^1 + i g_2 W_\mu^2 & -g_2 W_\mu^3 + g_1 B_\mu \end{pmatrix} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \quad (2.32)$$

the expansion of the scalar-field kinetic term becomes:

$$\begin{aligned} -(D_\mu \phi)^\dagger (D^\mu \phi) &= -\frac{1}{2} \partial_\mu H \partial^\mu H - \frac{1}{8} (v + H)^2 g_2^2 (W_\mu^1 - i W_\mu^2)(W^{1\mu} + i W^{2\mu}) \\ &\quad - \frac{1}{8} (v + H)^2 (-g_2 W^{3\mu} + g_1 B^\mu)(-g_2 W_{3\mu} + g_1 B_\mu) \end{aligned} \quad (2.33)$$

The scalar potential term contributes

$$\begin{aligned} V &= \frac{\lambda}{4} [(v + H)^2 - \mu^2/\lambda]^2 \\ &= \frac{\lambda}{4} (2vH + H^2)^2 \\ &= \lambda v^2 H^2 + \lambda v H^3 + \frac{\lambda}{4} H^4 \end{aligned} \quad (2.34)$$

The Yukawa couplings may be expanded in an identical way:

$$\begin{aligned}\bar{L}_m P_R E_n \phi &= \frac{1}{\sqrt{2}} \begin{pmatrix} \bar{\nu}_m \\ \bar{\mathcal{E}}_m \end{pmatrix}^T P_R E_n \begin{pmatrix} 0 \\ v + H \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} (v + H) \bar{\mathcal{E}}_m P_R E_n\end{aligned}\quad (2.35)$$

and similarly for Q , d , and D , and

$$\begin{aligned}\bar{Q}_m P_R U_n \tilde{\phi} &= \frac{1}{\sqrt{2}} \begin{pmatrix} \bar{\mathcal{U}}_m \\ \bar{\mathcal{D}}_m \end{pmatrix} P_R U_n \begin{pmatrix} v + H \\ 0 \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} (v + H) \bar{\mathcal{U}}_m P_R U_n\end{aligned}\quad (2.36)$$

Combining all of these results gives the expansion of $\mathcal{L}_{\text{Higgs}}$ to be

$$\begin{aligned}\mathcal{L}_{\text{Higgs}} &= -\frac{1}{2} \partial_\mu H \partial^\mu H - \lambda v^2 H^2 - \lambda v H^3 - \frac{\lambda}{4} H^4 \\ &\quad - \frac{1}{8} g_2^2 (v + H)^2 |W_\mu^1 - iW_\mu^2|^2 \\ &\quad - \frac{1}{8} (v + H)^2 (-g_2 W_\mu^3 + g_1 B_\mu)^2 \\ &\quad - \frac{1}{\sqrt{2}} (v + H) [f_{mn} \bar{\mathcal{E}}_m P_R E_n + \text{h.c.}] \\ &\quad - \frac{1}{\sqrt{2}} (v + H) [g_{mn} \bar{\mathcal{U}}_m P_R U_n + \text{h.c.}] \\ &\quad - \frac{1}{\sqrt{2}} (v + H) [h_{mn} \bar{\mathcal{D}}_m P_R D_n + \text{h.c.}]\end{aligned}\quad (2.37)$$

2.3.1 Boson masses

$\mathcal{L}_{\text{Higgs}}$ contains all of the mass terms, although some of these are not diagonal. They are, in more detail

2.3.1.1 Spin-zero particles

Comparing the H^2 term of $\mathcal{L}_{\text{Higgs}}$ with the standard form, $-\frac{1}{2} m_H^2 H^2$, gives

$$m_H^2 = 2\lambda v^2 = 2\mu^2 \quad (2.38)$$

2.3.1.2 Spin-one particles

The relevant terms in this case are:

$$-\frac{1}{8} g_2^2 v^2 |W_\mu^1 - iW_\mu^2|^2 - \frac{1}{8} v^2 (-g_2 W_\mu^3 + g_1 B_\mu)^2 \quad (2.39)$$

The fields W_μ^1 and W_μ^2 only appear in the combination $W_1^\mu W_{1\mu} + W_2^\mu W_{2\mu}$ and do not mix with any other fields. Their masses can therefore be read by inspection. Comparing this term to

$$-\frac{1}{2}M_1^2 W_\mu^1 W^{1\mu} - \frac{1}{2}M_2^2 W_\mu^2 W^{2\mu} \quad (2.40)$$

gives the masses

$$M_1^2 = M_2^2 = \frac{g_2^2 v^2}{4} \quad (2.41)$$

It is not an accident that these masses are equal. They are equal because the particles W_1 and W_2 are related by a symmetry that is not spontaneously broken, even when $v \neq 0$. To see this, consider performing a constant gauge transformation, $\partial_\mu \omega^a = 0$. The ground-state scalar field configuration then transforms as

$$\begin{aligned} \delta \begin{pmatrix} 0 \\ v \end{pmatrix} &= \frac{i}{2} \omega_2^a \tau_a \begin{pmatrix} 0 \\ v \end{pmatrix} + \frac{i}{2} \omega_1 \begin{pmatrix} 0 \\ v \end{pmatrix} \\ &= \frac{i}{2} \begin{pmatrix} [\omega_2^1 - i\omega_2^2]v \\ [\omega_1 - \omega_2^3]v \end{pmatrix} \end{aligned} \quad (2.42)$$

which vanishes provided that $\omega_2^1 = \omega_2^2 = 0$ and $\omega_1 = \omega_2^3 \equiv \omega$. This particular combination of $SU_L(2) \times U_Y(1)$ -transformations is therefore a symmetry of the ground state.

Under this symmetry the W fields transform according to Eq. (2.11):

$$\delta W_\mu^a = -\epsilon^{abc} \omega_2^b W_\mu^c, \quad \text{or,} \quad \delta \begin{pmatrix} W_\mu^1 \\ W_\mu^2 \end{pmatrix} = \omega \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} W_\mu^1 \\ W_\mu^2 \end{pmatrix} \quad (2.43)$$

This shows that W_μ^1 and W_μ^2 transform into one another under this symmetry. The condition $\omega_2^3 = \omega_1$ implies that the generator of this unbroken symmetry is $T_3 + Y$. Now, we saw earlier that the electric charge, Q , of a field is related to the $SU_L(2) \times U_Y(1)$ -generators by $Q = T_3 + Y$. It is precisely the electromagnetic gauge invariance, $U_{\text{em}}(1)$, which is unbroken by the vacuum. W_μ^1 and W_μ^2 must therefore correspond to the two degrees of freedom associated with the distinct particle and antiparticle states required for an electrically charged particle. It is convenient in these cases to deal with fields that diagonalize the generator of electric charge. This corresponds, in the present case, to writing W_1 and W_2 as the real and imaginary parts of a complex, charged field:

$$W_\mu^\pm \equiv \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \quad (2.44)$$

which satisfies $\delta W_\mu^\pm = \pm i\omega W_\mu^\pm$ under electromagnetic gauge transformations, Eq. (2.43).

The mass term appropriate to such a charged field is $-M_W^2 W_\mu^+ W^{-\mu}$. Comparing with the Lagrangian, Eq. (1.121), therefore gives the W^\pm mass to be

$$M_W = M_1 = M_2 = \frac{g_2 v}{2} \quad (2.45)$$

The remaining vector fields that appear in the mass term are W_μ^3 and B_μ . They also only appear in one particular combination, $g_1 B_\mu - g_2 W_\mu^3$. We may normalize this combination (in order not to alter the standard form for the kinetic terms) to define the mass eigenstate:

$$\begin{aligned} Z_\mu &\equiv \frac{-g_1 B_\mu + g_2 W_\mu^3}{\sqrt{g_1^2 + g_2^2}} \\ &\equiv W_\mu^3 \cos \theta_W - B_\mu \sin \theta_W \end{aligned} \quad (2.46)$$

This last equation defines the weak-mixing angle or Weinberg angle, θ_W , given by

$$\begin{aligned} \cos \theta_W &= \frac{g_2}{\sqrt{g_1^2 + g_2^2}} \\ \sin \theta_W &= \frac{g_1}{\sqrt{g_1^2 + g_2^2}} \end{aligned} \quad (2.47)$$

In terms of this field the mass term, Eq. (1.124), is

$$-\frac{1}{8}v^2(g_1^2 + g_2^2)Z_\mu Z^\mu \quad (2.48)$$

from which the mass may be read off:

$$M_Z^2 = \frac{1}{4}(g_1^2 + g_2^2)v^2 \quad (2.49)$$

The final mass eigenstate is the combination of W_μ^3 and B_μ that is orthogonal to Z_μ :

$$A_\mu = W_\mu^3 \sin \theta_W + B_\mu \cos \theta_W = \frac{g_1 W_\mu^3 + g_2 B_\mu}{\sqrt{g_1^2 + g_2^2}} \quad (2.50)$$

This is massless, as are the gluons, G_μ^α , that gauge $SU_c(3)$. The masslessness of A_μ corresponds to the fact that the linear combination $Q = T_3 + Y$ is unbroken even when $v \neq 0$. A_μ is the corresponding massless gauge boson required for this unbroken symmetry. Since Q is the electric charge, we expect A_μ to have the couplings of the usual photon.

To summarize the relations between field bases, writing $c_W \equiv \cos \theta_W$ and $s_W \equiv \sin \theta_W$,

$$\begin{aligned}
W_\mu^3 &= c_W Z_\mu + s_W A_\mu & Z_\mu &= c_W W_\mu^3 - s_W B_\mu \\
B_\mu &= -s_W Z_\mu + c_W A_\mu & A_\mu &= s_W W_\mu^3 + c_W B_\mu \\
\sqrt{2}W_\mu^+ &= W_\mu^1 - iW_\mu^2 & \sqrt{2}W_\mu^1 &= W_\mu^+ + W_\mu^- \\
\sqrt{2}W_\mu^- &= W_\mu^1 + iW_\mu^2 & \sqrt{2}W_\mu^2 &= iW_\mu^+ - iW_\mu^- \\
\sqrt{g_2^2 + g_1^2}W_\mu^3 &= g_2 Z_\mu + g_1 A_\mu & \sqrt{g_2^2 + g_1^2}Z_\mu &= g_2 W_\mu^3 - g_1 B_\mu \\
\sqrt{g_2^2 + g_1^2}B_\mu &= -g_1 Z_\mu + g_2 A_\mu & \sqrt{g_2^2 + g_1^2}A_\mu &= g_1 W_\mu^3 + g_2 B_\mu
\end{aligned} \tag{2.51}$$

2.3.2 The custodial $SU(2)$

Notice that there is a relation amongst the three quantities M_W , M_Z , and θ_W

$$\frac{M_W}{M_Z} = \frac{g_2}{\sqrt{g_1^2 + g_2^2}} = \cos \theta_W \tag{2.52}$$

It is natural to ask how much this relation depends on the details of how the symmetry $SU_L(2) \times U_Y(1)$ is broken, since any information that can restrict the arbitrariness in the symmetry breaking sector is welcome. Consider therefore the most general form for the spin-one mass matrix that is consistent with the symmetry-breaking pattern $SU_L(2) \times U_Y(1) \rightarrow U_{\text{em}}(1)$:

$$\begin{pmatrix} M_W^2 & & & \\ & M_W^2 & & \\ & & M_3^2 & m^2 \\ & & m^2 & M_0^2 \end{pmatrix} \tag{2.53}$$

This form has a simple explanation. As we saw above, unbroken electromagnetic gauge invariance dictates that the upper left 2×2 block of the matrix be proportional to the unit matrix: $M_W^2 I_{2 \times 2}$. It similarly implies that the upper-right and the lower-left blocks must vanish. The lower-right 2×2 block is a priori an arbitrary symmetric matrix, subject to the one constraint that one of its eigenvalues must vanish. The vanishing of one of the eigenvalues corresponds to the masslessness of the photon, and is a general consequence of the fact that the electromagnetic gauge invariance is unbroken.

The requirement that one eigenvalue be zero is equivalent to the vanishing of the determinant:

$$\det \begin{pmatrix} M_3^2 & m^2 \\ m^2 & M_0^2 \end{pmatrix} = M_3^2 M_0^2 - m^4 = 0, \quad (2.54)$$

implying the condition $m^2 = |M_0 M_3|$. The corresponding zero eigenvector may be written as

$$\begin{pmatrix} -\sin \theta_w \\ \cos \theta_w \end{pmatrix} \quad (2.55)$$

Equation (2.55) defines the mixing angle, θ_w , in the general case. We may now eliminate M_0^2 in favor of θ_w . The required relation is

$$\tan \theta_w = \frac{m^2}{M_3^2} = \left| \frac{M_0}{M_3} \right| \quad (2.56)$$

The non-zero eigenvalue, M_Z , is then given in terms of M_3 and θ_w by

$$\begin{aligned} M_Z^2 &= \text{tr} \begin{pmatrix} M_3^2 & m^2 \\ m^2 & M_0^2 \end{pmatrix} \\ &= M_0^2 + M_3^2 = M_3^2 (1 + \tan^2 \theta_w) = M_3^2 \sec^2 \theta_w \end{aligned} \quad (2.57)$$

The mass relation implied by the symmetry breaking pattern $SU_L(2) \times U_Y(1) \rightarrow U_{\text{em}}(1)$ is therefore $M_3 = M_Z \cos \theta_w$. An alternative way of expressing the mass formula, Eq. (2.52), is therefore $M_1 = M_2 = M_3 = M_W$.

The equality of M_3 and M_W within the standard model is a consequence of using a scalar $SU_L(2)$ -doublet, ϕ , to break $SU_L(2) \times U_Y(1)$. The connection arises because of an *accidental symmetry* of the scalar self-couplings that determine the symmetry-breaking pattern that in turn determines the gauge boson mass matrix. The Higgs doublet, ϕ , may be thought of as four real scalar fields, corresponding to the real and imaginary parts of ϕ^0 and ϕ^+ in Eq. (2.23). An alternative way to write these four real fields would be as a column vector:

$$\Phi = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \end{pmatrix} \quad (2.58)$$

As we saw in Subsection 1.3.1, the kinetic terms for four real scalar fields can be written as $\partial_\mu \Phi^T \partial^\mu \Phi$ and so is always invariant under the multiplication of Φ by an arbitrary 4×4 orthogonal matrix, $\mathcal{O} \in O(4)$. Now, in general

the interaction terms of the Lagrangian break this symmetry completely. However, for the standard model, the two requirements of gauge invariance and renormalizability imply that the only possible scalar self-couplings are of the form $V = V(\phi^\dagger \phi) = V(\Phi^T \Phi)$. Even though it was not required to be so, this potential is therefore also invariant under these general $O(4)$ transformations. Any such global symmetry that appears as a simple consequence of gauge invariance and renormalizability is known as an accidental symmetry.

Once Φ develops a *v.e.v.*,

$$\langle \Phi \rangle = \begin{pmatrix} v \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (2.59)$$

this $O(4)$ -invariance gets broken to the 3×3 orthogonal, $O(3)$, transformations that shuffle the lower three components amongst themselves. Since this $O(3)$ is unbroken, it constrains the form that the mass matrix may take. The ϕ gauge couplings that ultimately produce the gauge boson mass matrix are also invariant under these $O(3)$ transformations if the W_μ^a s transform as a three-dimensional vector. Invariance of the mass matrix under this 3×3 transformation therefore implies that the upper-left 3×3 block of the spin-one mass matrix, Eq. (2.53), must be proportional to the unit matrix, implying $M_3 = M_1 = M_2 = M_W$ as required.

Since the group $O(3)$ is locally isomorphic to the group $SU(2)$, it is said that the symmetry-breaking sector has an accidental *custodial* $SU(2)$ invariance that is responsible for the mass formula, Eq. (2.52).

The utility of having such a symmetry understanding of this mass formula is that it points to the circumstances under which it might be altered and to how big the corrections might be. In fact, some of the interactions in the standard model, like the $\phi - B_\mu$ coupling and the Yukawa couplings, do *not* respect this custodial symmetry. We may expect, then, that radiative (quantum) corrections that involve these interactions can alter the mass relation. This is discussed in Section 7.5. Experimental verification of this relation is clearly of great importance since deviations point to detailed effects within the standard model, and potentially to indications of new physics.

2.3.3 Fermion masses

The terms quadratic in the fermion fields come from the Yukawa couplings after the shifting of the scalar field by v . The relevant terms are

$$\mathcal{L} = -\frac{v}{\sqrt{2}} [f_{mn} \bar{\mathcal{E}}_m P_R E_n + g_{mn} \bar{\mathcal{U}}_m P_R U_n + h_{mn} \bar{\mathcal{D}}_m P_R D_n + \text{h.c.}] \quad (2.60)$$

(It now becomes clear why it was convenient to label separately the different $SU_L(2)$ components of the fields L and Q ; the fact that the $v.e.v.$ of the Higgs field breaks $SU_L(2)$ symmetry means that a Yukawa coupling introduces a mass which picks out one or the other component.)

The mass terms induced by the Yukawa couplings of fermions to the Higgs $v.e.v.$ are in general not diagonal in the generation indices, m and n . They may be diagonalized following the procedure outlined in Subsection 1.3.2. To this end, redefine the spin-half fields as follows:

$$\begin{aligned} P_L \mathcal{E}_m &= U_{mn}^{(e)} P_L \mathcal{E}'_n & P_R E_m &= V_{mn}^{(e)} P_R E'_n \\ P_L \mathcal{U}_m &= U_{mn}^{(u)} P_L \mathcal{U}'_n & P_R U_m &= V_{mn}^{(u)} P_R U'_n \\ P_L \mathcal{D}_m &= U_{mn}^{(d)} P_L \mathcal{D}'_n & P_R D_m &= V_{mn}^{(d)} P_R D'_n \end{aligned} \quad (2.61)$$

where the matrices $U^{(e)}, U^{(u)}, U^{(d)}, V^{(e)}, V^{(u)}, V^{(d)}$ act on the generation indices (e.g. connect e to μ to τ) and must be unitary in order to preserve the canonical form for the kinetic terms.

As argued in Subsection 1.3.2, it is always possible to choose $U^{(e)} = V^{(e)*}, U^{(u)} = V^{(u)*}, U^{(d)} = V^{(d)*}$, and then choose $U^{(e)}$ to ensure that the new mass matrices are diagonal:

$$U^{(e)\dagger} f V^{(e)} = V^{(e)T} f V^{(e)} = \text{diag}(f_e, f_u, f_\tau) \quad (2.62)$$

with f_e, f_μ, f_τ real and non-negative. The same may be done for $V^{(u)T} g V^{(u)}$ and $V^{(d)T} h V^{(d)}$. The resulting mass terms then become (dropping the primes on the new fields)

$$\mathcal{L} = -\frac{1}{\sqrt{2}} v [f_m \bar{\mathcal{E}}_m P_R E_m + g_m \bar{\mathcal{U}}_m P_R U_m + h_m \bar{\mathcal{D}}_m P_R D_m + \text{h.c.}] \quad (2.63)$$

This has a simple expression in terms of the Dirac spinors, e_m, d_m , and u_m , defined as

$$\begin{aligned} e_m &\equiv P_L \mathcal{E}_m + P_R E_m \\ d_m &\equiv P_L \mathcal{D}_m + P_R D_m \\ u_m &\equiv P_L \mathcal{U}_m + P_R U_m \end{aligned} \quad (2.64)$$

To see this, use

$$\begin{aligned}
\bar{\mathcal{E}}_m P_R E_m + \text{h.c.} &= \bar{\mathcal{E}}_m P_R E_m + \bar{\mathcal{E}}_m P_L E_m \\
&= \bar{\mathcal{E}}_m P_R E_m + \bar{E}_m P_L \mathcal{E}_m \\
&= \bar{e}_m P_R e_m + \bar{e}_m P_L e_m \\
&= \bar{e}_m e_m
\end{aligned} \tag{2.65}$$

(The derivation of the identities used here was the subject of Problem 1 of Chapter 1.)

In terms of these Dirac spinors, the final form for the mass terms is

$$\mathcal{L} = -\frac{1}{\sqrt{2}}v(f_m \bar{e}_m e_m + g_m \bar{u}_m u_m + h_m \bar{d}_m d_m) \tag{2.66}$$

which, when compared to the standard mass term, $-m\bar{\psi}\psi$, gives the fermion masses as

$$m_n^{(e)} = \frac{1}{\sqrt{2}}f_n v, \quad m_n^{(u)} = \frac{1}{\sqrt{2}}g_n v, \quad m_n^{(d)} = \frac{1}{\sqrt{2}}h_n v \tag{2.67}$$

Notice that there is a separate Yukawa parameter, f_n , for every independent mass, m_n , so there are no mass formulae along the lines of Eq. (2.52) for the fermions. The numerical values of these fermion masses are presented in Appendix A.

Note that no mass term for the neutrinos is generated. If only renormalizable interactions and the minimal field content of the standard model are included, then this is exactly true, not just at the semiclassical level. A neutrino mass could appear if we extended the theory to include right-handed neutrinos N_m , because this would allow another Yukawa matrix between L and N . However, nothing forbids a mass term $m_m \bar{N}_m N_m$ for such right-handed neutrinos. One interpretation of the recent evidence for neutrino masses is that such right-handed neutrinos exist but their mass is very heavy. This is discussed in more detail in Chapter 10 and in Problem 2.3.

2.3.4 Hadrons

What we have just presented is the *perturbative* spectrum, that is, the spectrum assuming all interactions are weak. As we will discuss in Section 7.4, this is a valid approximation except for the $SU_c(3)$ (“strong”) interactions, which become strong at scales of order 500 MeV. The result is that quarks and gluons do not appear as actual particles of the spectrum. Rather, the particles we observe are bound states of quarks and gluons, in appropriate

combinations to be color singlets. Such bound states are called *hadrons*. This is discussed in much more detail in Chapter 8. Here, we will just briefly explain the results and the nomenclature.

There are three ways to form a colorless combination of quarks and gluons. One is to have a bound state made purely of two or more gluons, called a “glueball.” It is believed that such states should be heavy and highly unstable, making their identification difficult. The next way is to have a bound state made up of a quark and an antiquark, $q\bar{q}$ (possibly together with gluons and more $q\bar{q}$ pairs). Such bound states exist and are called *mesons*; the lightest meson is the pion, made up of a $u\bar{d}$ (π^+), a $d\bar{u}$ (π^-), or $(u\bar{u} - d\bar{d})/\sqrt{2}$ (π^0). The final way is to have a bound state of three quarks (possibly together with gluons and more $q\bar{q}$ pairs). Such a three-quark state is called a *baryon*, and its antiparticle, with three antiquarks, is an antibaryon. The lightest two baryons are the familiar proton and neutron, made up of uud and udd respectively. There is no straightforward way to relate the masses of the hadrons to the masses of the constituent quarks and gluons, because the binding energies involved are of order 500 MeV. In the case of the b - and c -containing hadrons, however, the mass is dominated by the mass of the heavy quark, making possible simpler relations between hadron and quark masses.

When energies are large compared to the hadronic binding energy, the language of quarks and gluons can be appropriate – within limits. For instance, in computing Z boson decays in Chapter 4, we will see that the total rate of decay into hadrons is given, up to small corrections, by the rate of decay into quarks; how the quarks stick together into hadrons determines what the actual final state is, but not the likelihood for the Z boson to create the quarks. Similarly, when a hadron is one of the particles participating in a collision, then at high energies we can often describe the collision in terms of the quarks and gluons residing within the hadron, as discussed in Chapter 9.

2.4 Interactions

We have determined the particle masses in terms of the various parameters of the Lagrangian. The predictive nature of the theory only appears once we identify how these parameters determine the strengths of particle interactions and compare the interactions we see with those that are predicted.

This section is largely bookkeeping. The most important parts to understand are the charged and neutral current interactions and the necessity

of the Kobayashi–Maskawa matrix. Most of the content of this section is summarized by the Feynman rules presented in Section 5.4.

2.4.1 Higgs couplings

The couplings of the Higgs boson are found in the expansion of the Higgs Lagrangian, $\mathcal{L}_{\text{Higgs}}$, of Eq. (2.37):

$$\begin{aligned}\mathcal{L}_{\text{Higgs}} = & -\frac{1}{2}\partial_\mu H \partial^\mu H - \lambda v^2 H^2 - \lambda v H^3 - \frac{1}{4}\lambda H^4 \\ & -\frac{1}{8}g_2^2(v+H)^2|W_\mu^1 - iW_\mu^2|^2 \\ & -\frac{1}{8}(v+H)^2(-g_2 W_\mu^3 + g_1 B_\mu)^2 \\ & -\frac{1}{\sqrt{2}}(v+H)[f_{mn}\bar{\mathcal{E}}_m P_R E_n + \text{h.c.}] \\ & -\frac{1}{\sqrt{2}}(v+H)[g_{mn}\bar{\mathcal{U}}_m P_R U_n + \text{h.c.}] \\ & -\frac{1}{\sqrt{2}}(v+H)[h_{mn}\bar{\mathcal{D}}_m P_R D_n + \text{h.c.}]\end{aligned}$$

This Lagrangian completely specifies the Higgs couplings to other particles.

2.4.1.1 Higgs self-couplings

The couplings of the Higgs to itself are easily read from the potential in Eq. (2.37):

$$\begin{aligned}\mathcal{L}_{\text{H-H}} &= -\lambda v H^3 - \frac{1}{4}\lambda H^4 \\ &= -\frac{m_{\text{H}}^2}{2v}H^3 - \frac{m_{\text{H}}^2}{8v^2}H^4\end{aligned}\tag{2.68}$$

2.4.1.2 Higgs–gauge-boson couplings

The Higgs–gauge boson couplings are similarly given by

$$\begin{aligned}\mathcal{L}_{\text{H-g}} &= -\frac{1}{8}g_2^2(2vH + H^2)|W_\mu^1 - iW_\mu^2|^2 - \frac{1}{8}(2vH + H^2)(-g_2 W_\mu^3 + g_1 B_\mu)^2 \\ &= -\left(\frac{H}{v} + \frac{H^2}{2v^2}\right)\left(2M_W^2 W_\mu^+ W^{-\mu} + M_Z^2 Z_\mu Z^\mu\right)\end{aligned}\tag{2.69}$$

2.4.1.3 Higgs-fermion couplings

The final Higgs interactions consist of Yukawa couplings between the Higgs scalar and the various fermions:

$$\begin{aligned}\mathcal{L}_{\text{H-f}} &= -\frac{1}{\sqrt{2}}H(f_m\bar{e}_m e_m + g_m\bar{u}_m u_m + h_m\bar{d}_m d_m) \\ &= -\sum_f \frac{m_f}{v}\bar{f}fH\end{aligned}\tag{2.70}$$

Here and in the following we use f (for fermion) to run over the nine Dirac and three Majorana species labels e_i, u_i, d_i, ν_i ; but the m_ν are zero.

Several points about these couplings are worth noting.

- (i) Notice first that all other particles couple to the Higgs boson with strength m/v , in which m is the mass of the particle in question and v (which turns out to equal 246 GeV) is the symmetry-breaking vacuum expectation value. This ratio is small provided that $m \ll v$, which is true for all known particles, though only marginally so for the top quark, t . H must therefore couple weakly to all of the particles that have been discovered to date, and must furthermore couple preferentially to the heavier particles.
- (ii) The Higgs-fermion couplings are automatically flavor-diagonal when expressed in terms of mass eigenstates. That is to say, the act of emission of a Higgs particle by a fermion does not convert one type (or “flavor”) of fermion into another. This is an important property of the model since there are very strong limits on the existence of any transitions of this type. The only known interactions that can change fermion flavor are the W^\pm interactions we meet later. The strongest limits on these types of flavor-changing couplings arise for those that involve the strange quark, $H\bar{s}d$ for example. Such an interaction would contribute to the extremely well measured mass difference, $m_{K_L} - m_{K_S} = (3.490 \pm 0.006) \times 10^{-12}$ MeV, between the two neutral kaons, K_L and K_S , or to *flavor-changing neutral-current* processes such as the decay $K_L \rightarrow \mu^+ e^-$, which has never been observed to occur. More quantitatively, this last process is known to happen less frequently than once in every 5×10^{12} K_L decays.

This property need not be preserved if the symmetry-breaking sector is more complicated, so it serves as a useful clue to the unknown nature of this sector.

- (iii) As will be shown in Section 2.5, these Higgs couplings also conserve the discrete symmetries of charge conjugation, C, parity, P, and time

reversal, T. This property is also not a general feature of more complicated symmetry-breaking sectors.

- (iv) The Higgs boson itself has not been seen (as of 2006), so its mass is undetermined. Since the Higgs self-coupling is $2\lambda = (m_{\text{H}}/v)^2$ and the accuracy of the perturbative series is governed by the combination $\lambda/(16\pi^2)$, we expect perturbation theory to fail unless $m_{\text{H}} \ll \sqrt{8\pi}v \approx 1.2 \text{ TeV}$. This is often expressed more precisely as a unitarity limit: lowest-order perturbation theory in λ violates unitarity if $m_{\text{H}} \gtrsim 500 \text{ GeV}$.

More generally, numerical results concerning the triviality of scalar field theories indicate that, if a Higgs particle exists, it must have a mass $\lesssim 1 \text{ TeV}$, even for strong couplings. At the time of writing, the lower limit on the Higgs boson mass (assuming that the standard model is correct) is 114 GeV, and the upper bound inferred from radiative effects of the Higgs boson on electroweak observables (and assuming that the standard model is correct) is about 200 GeV.

2.4.2 Strong interactions

The strong interactions are by definition those that involve the spin-one gluons. The relevant terms in \mathcal{L} are

$$\begin{aligned} \mathcal{L}_{\text{strong}} = & -\frac{1}{4}G_{\mu\nu}^{\alpha}G^{\alpha\mu\nu} - \frac{g_3^2\Theta_3}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}G^{\alpha\mu\nu}G^{\alpha\lambda\rho} \\ & -\frac{1}{2}\bar{Q}_m\not{D}Q_m - \frac{1}{2}\bar{U}_m\not{D}U_m - \frac{1}{2}\bar{D}_m\not{D}D_m \end{aligned} \quad (2.71)$$

2.4.2.1 Gluon self-couplings

The $G_{\mu\nu}^{\alpha}G^{\alpha\mu\nu}$ term describes the couplings of the gluons among themselves:

$$\mathcal{L}_{\text{gl-gl}} = -\frac{1}{4}\mathcal{G}_{\mu\nu}^{\alpha}\mathcal{G}^{\alpha\mu\nu} - \frac{g_3}{2}f_{\beta\gamma}^{\alpha}\mathcal{G}_{\mu\nu}^{\alpha}G^{\beta\mu}G^{\gamma\nu} - \frac{g_3^2}{4}f_{\beta\gamma}^{\alpha}f_{\delta\epsilon}^{\alpha}G_{\mu}^{\beta}G_{\nu}^{\gamma}G^{\delta\mu}G^{\epsilon\nu} \quad (2.72)$$

plus the Θ_3 term which we have not written out. Here, $\mathcal{G}_{\mu\nu}^{\alpha}$ denotes the linearized field strength, $\partial_{\mu}G_{\nu}^{\alpha} - \partial_{\nu}G_{\mu}^{\alpha}$. The Θ_3 term has almost no impact in the following, because it has no effect on any perturbative calculation, and because Θ_3 is numerically almost exactly zero. This is a mystery, discussed in Subsection 11.4.2.

2.4.2.2 Gluon-fermion couplings

The couplings between gluons and fermions may be read from Eq. (2.71),

$$\mathcal{L}_{\text{gl-f}} = +\frac{ig_3}{2}\sum_q G_{\mu}^{\alpha}\bar{q}\gamma^{\mu}\lambda_{\alpha}q \quad (2.73)$$

in which the sum is over the six Dirac spinors representing the different flavors of quarks, $q = u_m, d_m$.

The emission of a gluon by a fermion causes a transition in the fermion's color quantum numbers. We return to these couplings in more detail later. In the meantime some features of these couplings bear comment.

- (i) Because the standard model gauge group, $SU_c(3) \times SU_L(2) \times U_Y(1)$, is the *product* of a strong-interaction factor, $SU_c(3)$, with an electroweak factor, $SU_L(2) \times U_Y(1)$, all of the particles of the theory can be divided into two classes according to whether or not they carry strong-interaction quantum numbers. Quarks and gluons do and electrons, neutrinos, the Higgs particle, and the electroweak gauge bosons, W, Z, A , do not. This is the origin of the classification of elementary particles as *hadrons* or *leptons*. Hadrons involve the quarks and gluons and so participate in the strong interactions. For historical reasons only the spin-half particles that do not interact strongly are called leptons, and these therefore consist of the electron-type and neutrino-type fermions.
- (ii) Gluon interactions are called “strong,” as will be pursued in more detail in subsequent chapters, because unlike the electroweak interactions, the spectrum of strongly interacting particles cannot be described perturbatively in the gluon coupling, g_3 . The observed hadrons consist of bound states of the more elementary quarks and gluons. This greatly complicates the interpretation of interactions that involve hadrons as initial or final particles. As we shall see, it turns out that it is nevertheless possible to accurately describe some carefully chosen observables in hadron collisions at sufficiently high energies within perturbation theory.
- (iii) Just as was the case for the Higgs–fermion couplings, the emission of a gluon by a fermion can never change the flavor of the fermion. This may be seen from the above expressions, since the gluon–fermion interactions always have the form $G\bar{q}q$ and never involve two different types of quark, such as $G\bar{q}q'$. As a result, flavor type is conserved by the strong interactions. This has important consequences for the interactions and spectrum of all strongly-interacting particles, which will be explored in more detail later.
- (iv) Apart from the Θ_3 term, the strong interactions as given above are invariant with respect to all three of the discrete symmetries, C, P, and T. (This conclusion is justified in more detail in Section 2.5.) The present evidence for the invariance of the strong interactions under

these discrete symmetries (principally the current upper bound on the neutron's intrinsic electric dipole moment) implies that the *strong-CP parameter*, $|\Theta_3|$, must be smaller than $\approx 10^{-9}$. The potential significance of this CP-violating parameter is taken up in more detail in Subsection 11.4.2.

- (v) The strength of all strong interactions is governed by a single coupling constant, g_3 , so the strong interactions have a *universal* strength that is independent of the particle type that is participating in the interaction. This is an important experimental fact that is explained here as the natural consequence of the observation that the gluons are gauge bosons, and that all of the strongly-interacting fermions fall into the same representation (in this case triplets or antitriplets) of the gauge group $SU_c(3)$.

2.4.3 Electroweak interactions

We next turn to the couplings that involve the electroweak gauge bosons – those spin-one particles that correspond to the $SU_L(2) \times U_Y(1)$ factor of the gauge group. These come in two basic types. There are self-couplings that arise due to the non-linear terms in the gauge potentials within the $SU_L(2) \times U_Y(1)$ field strengths, and there are couplings with other particles that arise due to the use of gauge covariant derivatives in the kinetic-energy terms of the Lagrangian. We consider each of these in turn.

2.4.3.1 Electroweak boson self-interactions

There are both cubic and quartic self-couplings of the spin-one electroweak gauge bosons. Both arise from the non-linear terms in the $SU_L(2)$ gauge boson field strength

$$\mathcal{L} = -\frac{1}{4}W_{\mu\nu}^a W^{a\mu\nu} \quad (2.74)$$

The cubic terms are

$$\mathcal{L}_{\text{cubic}} = -\frac{1}{2}g_2\epsilon_{abc}\mathcal{W}_{\mu\nu}^a W^{b\mu}W^{c\nu} = \mathcal{L}_{WW\gamma} + \mathcal{L}_{WWZ} \quad (2.75)$$

with the W -photon and W - Z trilinear couplings given in terms of the mass eigenstates, $W_\mu^1 = \frac{1}{\sqrt{2}}(W_\mu^+ + W_\mu^-)$, $W_\mu^2 = \frac{-i}{\sqrt{2}}(W_\mu^- - W_\mu^+)$, and $W_\mu^3 = Z_\mu \cos\theta_W + A_\mu \sin\theta_W$, by

$$\mathcal{L}_{WW\gamma} = ig_2 \sin\theta_W \left[W_{\mu\nu}^+ W^{-\mu} A^\nu - W_{\mu\nu}^- W^{+\mu} A^\nu + W_\mu^+ W_\nu^- F^{\mu\nu} \right] \quad (2.76)$$

$$\mathcal{L}_{WWZ} = ig_2 \cos\theta_W \left[W_{\mu\nu}^+ W^{-\mu} Z^\nu - W_{\mu\nu}^- W^{+\mu} Z^\nu + W_\mu^+ W_\nu^- Z^{\mu\nu} \right] \quad (2.77)$$

In these expressions, $\mathcal{W}_{\mu\nu}^a$, $W_{\mu\nu}^\pm$, $Z_{\mu\nu}$ and $F_{\mu\nu}$ are the linear curls of the gauge potentials W_μ^a , W_μ^\pm , Z_μ and A_μ respectively, eg, $W_{\mu\nu}^\pm = \partial_\mu W_\nu^\pm - \partial_\nu W_\mu^\pm$.

The interaction terms that are quartic in these fields are

$$\begin{aligned}\mathcal{L}_{\text{quartic}} &= -\frac{1}{4}g_2^2\epsilon_{abc}\epsilon_{ade}W_\mu^bW_\nu^cW^{d\mu}W^{e\nu} \\ &= -\frac{1}{4}g_2^2\left[(W_\mu^aW_a^\mu)^2 - W_{a\mu}W_\nu^aW_b^\mu W^{b\nu}\right] \\ &= \mathcal{L}_{WWWW} + \mathcal{L}_{WWZZ} + \mathcal{L}_{WW\gamma\gamma} + \mathcal{L}_{WZ\gamma\gamma}\end{aligned}\quad (2.78)$$

which, using the relation $W_\mu^aW_\nu^a = W_\mu^-W_\nu^+ + W_\mu^+W_\nu^- + W_\mu^3W_\nu^3$ with $W_\mu^3 = Z_\mu \cos \theta_W + A_\mu \sin \theta_W$, gives,

$$\begin{aligned}\mathcal{L}_{\text{quartic}} &= -\frac{1}{2}g_2^2\left[(W_\mu^+W^{-\mu})^2 - (W_\mu^+W^{+\mu})(W_\nu^-W^{-\nu})\right] \\ &\quad -g_2^2\left[(W_\mu^+W^{-\mu})X_\nu X^\nu - (W_\mu^+X^\mu)(W_\nu^-X^\nu)\right]\end{aligned}\quad (2.79)$$

so

$$\mathcal{L}_{WWWW} = -\frac{1}{2}g_2^2\left[(W_\mu^+W^{-\mu})^2 - (W_\mu^+W^{+\mu})(W_\nu^-W^{-\nu})\right]\quad (2.80)$$

$$\mathcal{L}_{WWZZ} = -g_2^2\cos^2\theta_W\left[(W_\mu^+W^{-\mu})Z_\nu Z^\nu - (W_\mu^+Z^\mu)(W_\nu^-Z^\nu)\right]\quad (2.81)$$

$$\mathcal{L}_{WW\gamma\gamma} = -g_2^2\sin^2\theta_W\left[(W_\mu^+W^{-\mu})(A_\nu A^\nu) - (W_\mu^+A^\mu)(W_\nu^-A^\nu)\right]\quad (2.82)$$

$$\begin{aligned}\mathcal{L}_{WZ\gamma\gamma} &= -g_2^2\sin\theta_W\cos\theta_W\left[2(W_\mu^+W^{-\mu})(Z_\nu A^\nu) - (W_\mu^+A^\mu)(W_\nu^-Z^\nu)\right. \\ &\quad \left. - (W_\mu^+Z^\mu)(W_\nu^-A^\nu)\right]\end{aligned}\quad (2.83)$$

Some brief comments.

- (i) These self-interactions have been probed by the LEP-II experiments at the 2–3% level. However, compared to the precision with which the electroweak interactions of the fermions have been probed, these measurements are comparatively poor.
- (ii) The interactions of the W particles with the massless A boson only involve the particular combination of couplings $g_2\sin\theta_W$. As will become clear once the remainder of the A couplings are presented, this combination has the interpretation of being the electromagnetic coupling constant, e , as is appropriate for the interactions of the photon, A , with a particle of electric charge 1.
- (iii) These interactions preserve C, P, and T (see Section 2.5).

2.4.3.2 “Charged-current” fermion interactions

The only other electroweak interactions in the theory are the couplings between the electroweak bosons and spin-half and spin-zero particles. Since the couplings with the Higgs boson are given in Subsection 2.4.1, they need not be reconsidered again here.

The W_μ^a and B_μ -fermion couplings arise from the following kinetic terms,

$$\mathcal{L} = -\frac{1}{2}\bar{L}_m \not{D} L_m - \frac{1}{2}\bar{E}_m \not{D} E_m - \frac{1}{2}\bar{Q}_m \not{D} Q_m - \frac{1}{2}\bar{U}_m \not{D} U_m - \frac{1}{2}\bar{D}_m \not{D} D_m \quad (2.84)$$

Expanding each field in terms of the mass eigenstates gives

$$\begin{aligned} \mathcal{L}_{\text{ew}} = & +\frac{i}{4}\begin{pmatrix} \bar{\nu}_m \\ \bar{\mathcal{E}}_m \end{pmatrix}^T \gamma_\mu P_L \begin{pmatrix} -g_1 B_\mu + g_2 W_\mu^3 & g_2(W_\mu^1 - iW_\mu^2) \\ g_2(W_\mu^1 + iW_\mu^2) & -g_1 B_\mu - g_2 W_\mu^3 \end{pmatrix} \begin{pmatrix} \nu_m \\ \mathcal{E}_m \end{pmatrix} \\ & +\frac{i}{4}\begin{pmatrix} \bar{U}_m \\ \bar{\mathcal{D}}_m \end{pmatrix}^T \gamma_\mu P_L \begin{pmatrix} \frac{1}{3}g_1 B_\mu + g_2 W_\mu^3 & g_2(W_\mu^1 - iW_\mu^2) \\ g_2(W_\mu^1 + iW_\mu^2) & \frac{1}{3}g_1 B_\mu - g_2 W_\mu^3 \end{pmatrix} \begin{pmatrix} U_m \\ \mathcal{D}_m \end{pmatrix} \\ & +\frac{i}{3}g_1 B_\mu \bar{U}_m \gamma^\mu P_R U_m - \frac{i}{6}g_1 B_\mu \bar{D}_m \gamma^\mu P_R D_m \\ & -\frac{i}{2}g_1 B_\mu \bar{E}_m \gamma^\mu P_R E_m + \text{h.c.} \end{aligned} \quad (2.85)$$

The couplings between fermions and the charged spin-one particle, W_μ^+ , are called the *charged-current* interactions. Because these interactions always involve projection operators P_L or P_R , we may replace the Majorana fermions $\mathcal{U}, \mathcal{D}, \mathcal{E}$ with the Dirac fermions u, d, e (since the additional U, D, E fields introduced in this substitution are removed by the projection operator), giving

$$\mathcal{L}_{\text{cc}} = \frac{ig_2}{\sqrt{2}} \left[W_\mu^+ (\bar{\nu}_m \gamma^\mu P_L e_m + \bar{u}_m \gamma^\mu P_L d_m) + W_\mu^- (\bar{e}_m \gamma^\mu P_L \nu_m + \bar{d}_m \gamma^\mu P_L u_m) \right] \quad (2.86)$$

Unfortunately, as written this expression is correct in the generation basis we had before making the field redefinitions described in Subsection 2.3.3. To learn what the interactions are in terms of the mass basis, we must perform the same transformations, $e_m = U_{mn}^{(e)} e'_n$, $u_m = U_{mn}^{(u)} u'_n$, and $d_m = U_{mn}^{(d)} d'_n$, on this expression. Since there is no mass term for neutrinos, we are free to also redefine the neutrino field by $\nu_m = U_{mn}^{(e)} \nu'_n$, since this does not alter their mass or kinetic terms (see, however, Chapter 10). Defining

$$V_{mn} = (U^{(u)\dagger} U^{(d)})_{mn} \quad (2.87)$$

and introducing $e_W \equiv g_2/2\sqrt{2}$, gives the following expression:

$$\begin{aligned} \mathcal{L}_{cc} = & ie_W \left[W_\mu^+ (\bar{\nu}'_m \gamma^\mu (1+\gamma_5) e'_m + V_{mn} \bar{u}'_m \gamma^\mu (1+\gamma_5) d'_n) \right. \\ & \left. + W_\mu^- (\bar{e}'_m \gamma^\mu (1+\gamma_5) \nu'_{1n} + (V^\dagger)_{mn} \bar{d}'_m \gamma^\mu (1+\gamma_5) u'_n) \right] \end{aligned} \quad (2.88)$$

V_{mn} is a 3×3 unitary matrix called the *Kobayashi–Maskawa* (KM) – or sometimes the *Cabbibo–Kobayashi–Maskawa* (CKM)–matrix. It arises due to the necessity to perform different field redefinitions for up- and down-type quarks when diagonalizing masses. Since the matrix V_{mn} is 3×3 and unitary, it is described by nine real parameters. Not all of these nine parameters can be physically significant, however, because they may be changed by performing a field redefinition which has no other effects on the standard model Lagrangian. The only field redefinitions which can alter V_{mn} but which do not affect any other terms in the Lagrangian consist of multiplication of the various quark fields, u'_n and d'_n by a phase. Notice that since an overall rotation of all quarks by a common phase is a symmetry of the entire Lagrangian, and so leaves V_{mn} unchanged, this freedom to redefine fields allows the removal of at most five phases from V_{mn} . This would leave only four parameters of potential physical significance.

The choice of how to use these phase redefinitions to rotate the KM matrix is somewhat arbitrary. Partly for this reason, there are several different conventional ways in which to parameterize the KM matrix. The principal three are listed here for convenience. The parameterization advocated by the Particle Data Group is:

$$\begin{aligned} V &= \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \end{aligned} \quad (2.89)$$

$$(2.90)$$

in which c_{ij} and s_{ij} are shorthand for $\cos \theta_{ij}$ and $\sin \theta_{ij}$ respectively, and the *mixing angles*, θ_{ij} , are experimentally known to satisfy $\theta_{13} \ll \theta_{23} \ll \theta_{12} \ll 1$. This implies that (for unknown reasons) charged-current interactions that link fermions of differing generation are highly suppressed in the standard model and so in particular V_{mn} is very close to being a unit matrix. We return to the experimental constraints on the matrix V_{mn} shortly.

There are two other parameterizations of the KM matrix that are commonly used in the literature. Many of the older sources parameterize the KM matrix in terms of the Euler angles of an $O(3)$ rotation together with one phase:

$$\begin{aligned}
 V &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_2 & -s_2 \\ 0 & s_2 & c_2 \end{pmatrix} \begin{pmatrix} c_1 & s_1 & 0 \\ -s_1 & c_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{i\delta} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_3 & s_3 \\ 0 & -s_3 & c_3 \end{pmatrix} \\
 &= \begin{pmatrix} c_1 & s_1 c_3 & s_1 s_3 \\ -c_2 s_1 & c_1 c_2 c_3 + s_2 s_3 e^{i\delta} & c_1 c_2 s_3 - c_3 s_2 e^{i\delta} \\ -s_1 s_2 & c_1 s_2 c_3 - s_3 c_2 e^{i\delta} & c_1 s_2 s_3 + c_2 c_3 e^{i\delta} \end{pmatrix} \quad (2.91)
 \end{aligned}$$

Again $c_i (= \cos \theta_i)$ and $s_i (= \sin \theta_i)$ denote trigonometric functions of the Euler angles.

The third common parameterization is the *Wolfenstein* parameterization, which indicates the size of each matrix element in a particularly simple way. It is given, up to fourth order in the small quantity λ , by:

$$V = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} \quad (2.92)$$

The utility of this parameterization is that, since λ is found experimentally to be a small quantity, $\lambda \approx 0.2$, and A and $\rho^2 + \eta^2$ are $O(1)$, Eq. (2.92) summarizes the small size and hierarchy of the off-diagonal elements of V_{mn} .

It turns out (see Subsection 2.5.1) that these interactions preserve time reversal symmetry, T (or equivalently, CP) if the KM matrix can be made real by suitably redefining fields. Hence, it is interesting to know under which circumstances this is possible. In the generic case in which V_{mn} does not take any special form this can be decided by comparing the number of parameters available in a real versus a complex unitary matrix.

It is instructive to make the argument for the case of N generations of fermions. The counting goes as follows. The KM matrix is an $N \times N$ unitary matrix and so generically contains N^2 real parameters. If the KM matrix were real then it would be an orthogonal matrix, which can be described in terms of $\frac{1}{2}N(N-1)$ real parameters. The difference between these numbers, $N^2 - \frac{1}{2}N(N-1) = \frac{1}{2}N(N+1)$, is therefore the number of complex “phases” contained in V_{mn} . Not all of these phases, however, are physically significant, since some may be removed by absorbing phases into the various quark fields. Since such a redefinition does not affect any other term in the Lagrangian, any phase that can be removed in this way cannot cause any physical effects. Even though there are $2N$ species of quark fields, only $2N - 1$ phases may

be removed in this way, since the overall multiplication of all quark fields by a common phase is a symmetry of the Lagrangian and does not change V_{mn} . The number of remaining physical phases is therefore

$$\begin{aligned} P &= \left[N^2 - \frac{1}{2}N(N-1) \right] - (2N-1) \\ &= \frac{1}{2}(N-1)(N-2) \end{aligned} \quad (2.93)$$

Notice that if there were only two generations, then $P = 0$ and so the KM matrix could be chosen to be a real 2×2 orthogonal matrix:

$$V_{mn} = \begin{pmatrix} \cos \theta_c & \sin \theta_c \\ -\sin \theta_c & \cos \theta_c \end{pmatrix} \quad (2.94)$$

It happens that the experimental values for the angles in the full KM matrix are such that those parts of it that mix the first two generations are very close to being of the form of Eq. (2.94). For historical reasons the first few components of the KM matrix are therefore sometimes written in this way. The *Cabbibo angle* is accordingly defined by: $\cos \theta_c = V_{ud}$ and $\sin \theta_c = V_{us}$.

Some comments.

- (i) The charged-current interactions are the only ones within the model that connect fermions with differing flavors. In the absence of these charged-current interactions, the lightest species of fermion of any flavor would be absolutely stable, since flavor would be conserved. As a result, the charged-current interactions are the ones responsible for the majority of particle decays that have been observed.
- (ii) Since there is no Kobayashi–Maskawa matrix in the leptonic component of the charged-current interactions, all leptons participate in these interactions with equal strength, determined by g_2 . Just as was the case with the strong interactions, this result follows theoretically from the spin-one and hence gauge nature of the W boson, and the fact that all leptons that couple to the W boson are in doublets of $SU_L(2)$. The experimentally observed property that all leptons participate in charged-current weak interactions with equal strength is called *weak universality*.
- (iii) Weak universality does not hold for charged-current interactions involving quarks, because of the appearance there of the Kobayashi–Maskawa matrix, although there will be relationships amongst various hadronic charged-current interactions that follow from the unitarity of the KM matrix.
- (iv) As is shown in Section 2.5, the charged-current interactions violate both C and P, since they involve only the left-handed components of

the various fermion fields. They can only violate T if the KM matrix cannot be made real by a suitable choice of fields. It follows that all charged-current lepton interactions must preserve T and that the hadronic charged-current interactions can violate T only in a very specific way and only if there are at least three generations. At this time (2006), this source of T-violation is consistent with all of the experimental evidence.

- (v) Although the lepton sector of the standard model does not involve a KM matrix and so cannot violate CP, this would not be so if the model were enlarged in such a way as to generate a neutrino mass matrix. As discussed in Chapter 10, very small neutrino masses are in fact observed. These suggest that CP violation in the neutrino sector may be observable. The observation of CP violation is a major experimental goal of modern neutrino physics.

2.4.3.3 “Neutral-current” fermion interactions

It remains to write out the couplings of the two neutral gauge bosons, A_μ, Z_μ , of the electroweak gauge group, $SU_L(2) \times U_Y(1)$. Using the expressions

$$\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} \quad (2.95)$$

we see that these couplings are flavor-diagonal and of the form

$$\mathcal{L}_{\text{nc}} = \sum_f -Z_\mu \left[\bar{f} \gamma^\mu P_L (-ig_2 W_\mu^3 T_3 - ig_1 B_\mu Y_L) f + \bar{f} \gamma^\mu P_R (-ig_1 B_\mu Y_R) f \right] \quad (2.96)$$

where Y_L is the hypercharge of the left-handed fermion and Y_R is that of the right-handed one, e.g., $Y_L = -1/2$ for $P_L e_m = P_L \mathcal{E}_m$ and $Y_R = -1$ for $P_R e_m = P_R \mathcal{E}_m$ etc. Notice that Y_R agrees with the electric charge, Q , since all right-handed fields are singlets under $SU_L(2)$ and so have $T_3 = 0$. This then implies $Q = T_3 + Y_L = Y_R$.

Now, define the combination of Dirac matrices, gauge potentials and group generators, T_3 and $Y_{L,R}$, that appear in Eq. (2.96) above as M_μ . It may be reexpressed in the following form:

$$\begin{aligned} M_\mu &\equiv P_L g_2 A_\mu^3 T_3 + P_L g_1 B_\mu Y_L + P_R g_1 B_\mu Y_R \\ &= P_L g_2 A_\mu^3 T_3 + P_L g_1 B_\mu (Q - T_3) + P_R g_1 B_\mu Q \\ &= T_3 P_L (g_2 A_\mu^3 - g_1 B_\mu) + g_1 B_\mu Q \\ &= T_3 P_L [g_2 (Z_\mu \cos \theta_W + A_\mu \sin \theta_W) - g_1 (A_\mu \cos \theta_W - Z_\mu \sin \theta_W)] \\ &\quad + g_1 (A_\mu \cos \theta_W - Z_\mu \sin \theta_W) Q \end{aligned} \quad (2.97)$$

This simplifies further if we use the following relations among the coupling constants

$$g_2 = \cos \theta_W \sqrt{g_1^2 + g_2^2} \quad \text{and} \quad g_1 = \sin \theta_W \sqrt{g_1^2 + g_2^2}$$

so

$$g_1 \cos \theta_W = g_2 \sin \theta_W \equiv e$$

and

$$g_2 \cos \theta_W + g_1 \sin \theta_W = \sqrt{g_1^2 + g_2^2} = \frac{e}{\sin \theta_W \cos \theta_W}$$

Therefore,

$$M_\mu = \frac{e}{\sin \theta_W \cos \theta_W} \left[T_3 P_L - Q \sin^2 \theta_W \right] Z_\mu + e Q A_\mu \quad (2.98)$$

It is easily verified that the form of these interactions are not changed by the process of rotating to a basis of mass eigenstates for the fermion fields.

We may read from this the fermion couplings with the Z -boson and the massless photon, A . The photon-fermion coupling is

$$\mathcal{L}_{\text{em}} = \sum_f i e A_\mu \bar{f} \gamma^\mu Q f \quad (2.99)$$

in which the sum is over all fermion types, $f = e_m, \nu_m, d_m, u_m$, weighted by their electric charge, Q . Since the neutrino is electrically neutral it does not appear in the electromagnetic interactions. Comparing the interaction of Eq. (2.99) with that of QED in Eq. (1.176), we see that it is the combination $e = g_1 \cos \theta_W = g_2 \sin \theta_W = \sin \theta_W \cos \theta_W \sqrt{g_1^2 + g_2^2}$ that plays the role of the electromagnetic coupling constant – i.e. the absolute value of the electron charge – in this theory.

The Z_μ – or *neutral-current* – couplings are similarly given by

$$\begin{aligned} \mathcal{L}_{\text{nc}} &= \frac{i e}{\sin \theta_W \cos \theta_W} \sum_f Z_\mu \bar{f} \gamma^\mu \left[P_L T_3 - Q \sin^2 \theta_W \right] f \\ &= \frac{i e}{\sin \theta_W \cos \theta_W} \sum_f Z_\mu \bar{f} \gamma^\mu (g_V + \gamma_5 g_A) f \end{aligned} \quad (2.100)$$

in which $g_V = \frac{1}{2} T_3 - Q \sin^2 \theta_W$ and $g_A = \frac{1}{2} T_3$. Here T_3 refers to the charge, under the third generator of $SU_L(2)$, of the left-handed constituent of f , that is, \mathcal{E} , \mathcal{D} , or \mathcal{U} . The values of the charges g_V , g_A are given in Table 2.1.

These interactions share several noteworthy properties.

- (i) The couplings of the massless spin-one particle are precisely those of quantum electrodynamics, justifying its identification with the

Table 2.1 *Neutral-current charges of the fermions*

Fermion type	T_3	Q	g_V	g_A
ν_e, ν_μ, ν_τ	$+\frac{1}{2}$	0	+0.25	+0.25
e, μ, τ	$-\frac{1}{2}$	-1	-0.0189	-0.25
u, c, t	$+\frac{1}{2}$	$+\frac{2}{3}$	+0.0959	+0.25
d, s, b	$-\frac{1}{2}$	$-\frac{1}{3}$	-0.1730	-0.25

photon. This is not an accident, but follows as a result of the requirement that the symmetry-breaking order parameter not break the gauge symmetry generated by the electric charge, Q .

- (ii) The neutral-current interactions that couple fermions to Z -bosons never involve fermions of more than one flavor at a time and so cannot change flavor. As was indicated earlier for the Higgs and strong interactions, the experimental absence of such flavor-changing neutral currents was a strong clue to the structure of the standard model and was even used to predict the existence of the fourth type of quark, c !
- (iii) Electromagnetic interactions all conserve P, C, and T separately.
- (iv) The neutral-current interactions, on the other hand, violate both P and C but do not break T (see Section 2.5 for details).

This concludes the tabulation of the interactions that are contained in the standard model Lagrangian.

2.5 Symmetry properties*

When exploring the consequences for experiment of any potential theoretical model, it is always necessary to make use of various approximation schemes. It is therefore of crucial importance to understand which of the predictions of the model are of general validity, and which depend on more details of the approximation scheme used. For this reason, the first step to take in exploring any model is to identify the symmetries that it predicts, since these can be used to draw exact conclusions concerning the existence of conservation laws and of systematics (such as degeneracies) in the spectrum of particles. Therefore, we will now discuss at some length the symmetries of the standard model, and what exact conservation laws they predict.

* This section, while good for your teeth and bones, is not necessary for most of the development of this book, and can be skipped in whole or in part if necessary.

One of the most beautiful features of the standard model is its success in reproducing precisely the conservation laws and symmetries that had been distilled from experiment over the several decades before the discovery of the model. This accomplishment is all the more remarkable in light of the fact that the standard model is the most general theory consistent with a few very general principles, together with the given particle content and the requirement of renormalizability. As a result, none of the properties to be discussed in this section are built into the model as assumptions, and so they may be understood as general consequences of the basic principles of Section 1.2, together with the explicit particle content of the model.

Symmetries such as these, that are simply consequences of gauge invariance, particle content and renormalizability, are known as *accidental* symmetries. One example that has already been encountered is the custodial $SU(2)$ of the symmetry-breaking sector of Subsection 2.3.2.

2.5.1 Discrete symmetries

There are three discrete transformations that naturally arise within the quantum mechanics of any relativistic system. Two of these – parity, P , and time reversal, T – are related to (i.e. *automorphisms* of) the Lorentz group itself. The third discrete transformation – charge conjugation, C – consists of the interchange of every particle with its antiparticle.

It turns out that *none* of these are symmetries of the standard model, although the combined symmetry CPT is (and, in fact, is a symmetry of any quantum field theory which satisfies the basic principles laid out in Section 1.2). Nevertheless, we will take some time to discuss them. The reasons for doing so are, first, that the combined symmetry CP (or equivalently T) is *almost* a symmetry of the standard model, broken by very small subtle effects; and, second, that while C and P are very far from being symmetries of the standard model, at low energies $E \ll M_W$ they turn out to be accidental symmetries, as we will discuss in Section 7.3.

2.5.1.1 Definitions: P and T

The existence of the operations of parity and time reversal is related to the connectedness of the Lorentz group itself. The Lorentz group is reviewed in Appendix C. We show there that not all coordinate transformations permitted in special relativity can be built infinitesimally from the identity. In particular, two transformations of coordinates cannot: the parity

transformation,

$$x^\mu \rightarrow P^\mu_\nu x^\nu, \quad P^\mu_\nu = \begin{pmatrix} +1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \quad (2.101)$$

which reflects each space coordinate, and the time reversal transformation,

$$x^\mu \rightarrow T^\mu_\nu, \quad T^\mu_\nu = \begin{pmatrix} -1 & & & \\ & +1 & & \\ & & +1 & \\ & & & +1 \end{pmatrix} \quad (2.102)$$

which reverses the sign of time (see Appendix C).

Transformations P and T need not be symmetries of a given theory. If they are symmetries, and if their representations in the theory's Hilbert space are denoted by \mathcal{P} and \mathcal{T} respectively, then \mathcal{P} can always be chosen to be a unitary operator and although \mathcal{T} cannot be made unitary, it may always be chosen to be anti-unitary (that is, an operator which flips the sign of i). The reason \mathcal{T} is antiunitary is that H must transform under the symmetry into an operator which still has a positive spectrum; this will be satisfied if $\mathcal{P}H\mathcal{P}^* = H$ and $\mathcal{T}H\mathcal{T}^* = H$. On the other hand, time evolution by a positive amount of time t , e^{-iHt} , should be carried under time reversal to time evolution by a negative amount of time $-t$, $\mathcal{T}e^{-iHt}\mathcal{T}^* = e^{iHt}$. The only way that both of these can be true is if \mathcal{T} is an anti-unitary operator, reversing the sign of i .

2.5.1.2 Definition: \mathcal{C}

Charge conjugation is defined as the interchange of every particle with its antiparticle. The unitary operator that represents this interchange in the Hilbert space will be denoted by \mathcal{C} .

Notice that the condition that a theory be charge-conjugation invariant is stronger than the condition of crossing symmetry discussed in Section 1.2. Crossing symmetry is a general consequence of relativistic quantum mechanics; it states that particles and antiparticles must appear in the action only in the schematic combination $(a + \bar{a}^*)$. This ensures that particles and antiparticles appear in all interactions with the same strength but does *not* imply that all interactions must be invariant with respect to interchange of a with \bar{a} .

It is a theorem, though, that the combined action of all three of these discrete transformations, CPT, must be a symmetry in any Lorentz invariant, local field theory.

2.5.1.3 Transformation rules

The action of \mathcal{P} , \mathcal{T} , and \mathcal{C} on particle states and on fields is determined (up to a conventionally fixed freedom to redefine fields) by their transformation properties under Lorentz transformations. Their action on a state, $|\mathbf{p}, \sigma\rangle$, that describes a particle of three-momentum \mathbf{p} , total spin j , and third component of angular momentum σ , may be chosen to be

$$\begin{aligned}\mathcal{P}|\mathbf{p}, \sigma\rangle &= \alpha_p |-\mathbf{p}, \sigma\rangle \\ \mathcal{T}|\mathbf{p}, \sigma\rangle &= \alpha_t (-)^{j-\sigma} |-\mathbf{p}, -\sigma\rangle \\ \mathcal{C}|\mathbf{p}, \sigma\rangle &= \alpha_c \overline{|\mathbf{p}, \sigma\rangle}\end{aligned}\tag{2.103}$$

In these expressions, α_p , α_t , and α_c are phases that are characteristic of each particle type, and the state $\overline{|\cdots\rangle}$ denotes the antiparticle for the state $|\cdots\rangle$. The transformation properties of the corresponding creation and annihilation operators are determined by those of the particle states

$$\begin{aligned}\mathcal{P}a_{\mathbf{p},\sigma}^*\mathcal{P}^* &= \alpha_p a_{-\mathbf{p},\sigma}^* \\ \mathcal{T}a_{\mathbf{p},\sigma}^*\mathcal{T}^* &= \alpha_t (-)^{j-\sigma} a_{-\mathbf{p},-\sigma}^* \\ \mathcal{C}a_{\mathbf{p},\sigma}^*\mathcal{C}^* &= \alpha_c \bar{a}_{\mathbf{p},\sigma}^*\end{aligned}\tag{2.104}$$

The transformation rules for the fields are then determined by their expansions in terms of creation and annihilation operators. Since these have the generic form

$$\phi \sim \sum_{\mathbf{p},\sigma} [u(\mathbf{p}, \sigma) a_{\mathbf{p},\sigma} + v(\mathbf{p}, \sigma) \bar{a}_{\mathbf{p},\sigma}^*]\tag{2.105}$$

the transformation rules for fields representing spin-zero particles become

$$\begin{aligned}\mathcal{P}\phi(x)\mathcal{P}^* &= \alpha_p^* \phi(x_p) \\ \mathcal{C}\phi(x)\mathcal{C}^* &= \alpha_c^* \phi^*(x)\end{aligned}\tag{2.106}$$

in which $x_p = (-\mathbf{x}, t)$ denotes the image of $x = (\mathbf{x}, t)$ under parity. (Since invariance of the theory under the combination CPT is guaranteed on general grounds, T-invariance is equivalent to CP-invariance. For this reason it suffices to have explicit expressions for the transformation rules under \mathcal{C} and \mathcal{P} in order to determine its symmetry properties.)

For spinor fields we have instead,

$$\begin{aligned}\mathcal{P}\psi(x)\mathcal{P}^* &= \alpha_p^* \beta \psi(x_p) \\ \mathcal{C}\psi(x)\mathcal{C}^* &= \alpha_c^* C \bar{\psi}^T(x)\end{aligned}\tag{2.107}$$

in which β and C are the matrices defined in Eq. (1.85) and Eq. (1.93)

respectively. (The factor β exchanges left- and right-handed components and is necessary because parity flips handedness.)

Finally, for spin-one gauge potentials, V_a^μ , that correspond to the gauge generator, t_a , we have (up to gauge transformations)

$$\begin{aligned}\mathcal{P}[t_a V_a^\mu(x)]\mathcal{P}^* &= P^\mu{}_\nu [t_a V_a^\nu(x_p)] \\ \mathcal{C}[t_a V_a^\mu(x)]\mathcal{C}^* &= -[t_a V_a^\mu(x)]^*\end{aligned}\quad (2.108)$$

The phase in the transformation rule for the gauge potentials is fixed by the requirement that the covariant derivative, $D = \partial - iT_a V_a$, transform properly.

2.5.1.4 Invariance of the model

Using these transformation rules, we can test the standard model interactions of the previous section for invariance under the three independent symmetries of C, P, and CP.

The typical interaction Lagrangian density is the sum of several local operators, $\mathcal{O}_n(x)$, with some constant coefficients, c_n : $\mathcal{L}_{\text{int}} = \sum_n c_n \mathcal{O}_n(x)$. The transformation properties of the operators, $\mathcal{O}_n(x)$, can be inferred in terms of those of the various fields of the theory in terms of which they are expressed. The resulting transformation rule for the interaction Lagrangian is

$$\begin{aligned}\mathcal{P}\mathcal{L}_{\text{int}}\mathcal{P}^* &= \sum_n (\alpha_n)_p c_n \mathcal{O}_n(x_p) \\ \mathcal{C}\mathcal{L}_{\text{int}}\mathcal{C}^* &= \sum_n (\alpha_n)_c c_n \mathcal{O}_n^*(x) \\ (\mathcal{CP})\mathcal{L}_{\text{int}}(\mathcal{CP})^* &= \sum_n (\alpha_n)_p (\alpha_n)_c c_n \mathcal{O}_n^*(x_p)\end{aligned}\quad (2.109)$$

The phases $(\alpha_n)_p$ and $(\alpha_n)_c$ are products of the phases associated with the transformation of each field.

Since the action is given by the integral of $\mathcal{L}(x)$ over spacetime, the condition $\mathcal{P}\mathcal{L}(x)\mathcal{P}^* = \mathcal{L}(x_p)$ suffices to ensure that the action is invariant. The condition for parity invariance is therefore that there exist a choice of phases, α_{ps} , for each of the fields for which

$$(\alpha_n)_p = 1 \quad \text{for all } n \quad (2.110)$$

This is a nontrivial condition because there can be more interactions, \mathcal{O}_n , than there are fields appearing within them.

The Lagrangian is also required by unitarity to be Hermitian, so the following relation among the operators is also true: $\sum_n c_n^* \mathcal{O}_n^* = \sum_n c_n \mathcal{O}_n$. The

action is therefore charge-conjugation invariant provided that there exists a choice of charge-conjugation phases, α_{cs} , for each of the fields for which the coefficient of \mathcal{O}_n^* is unchanged:

$$(\alpha_n)_c c_n = c_n^* \quad \text{for all } n \quad (2.111)$$

CP-invariance is similarly ensured if phases can be chosen such that

$$(\alpha_n)_c (\alpha_n)_p c_n = c_n^* \quad \text{for all } n \quad (2.112)$$

If we apply this formalism to the standard model Lagrangian then we find the results quoted in Section 2.4. The Higgs interactions, gluon interactions, and electromagnetic interactions all respect each of the three discrete symmetries, C, P, and CP. The neutral current couplings of the fermions to the neutral Z boson break both C and P but in such a way that the combination CP is unbroken. Finally, the charged-current coupling of the fermions to the W boson not only violates C and P, but can also violate CP, provided that there is not sufficient freedom to make the Kobayashi–Maskawa matrix real. As an illustration we show the manipulations for the charged-current quark interactions,

$$\mathcal{L} = \frac{ig_2}{2\sqrt{2}} \left[V_{mn} W_\mu^+ \bar{u}_m \gamma^\mu (1+\gamma_5) d_n + (V^\dagger)_{mn} W_\mu^- \bar{d}_m \gamma^\mu (1+\gamma_5) u_n \right] \quad (2.113)$$

In this case the transformation rules for the spin-one fields become $\mathcal{C} W_\mu^\pm \mathcal{C}^* = -W_\mu^\mp$ and $\mathcal{P} W_\mu^\pm \mathcal{P}^* = P^\nu{}_\mu W_\nu^\pm$. Then, under charge conjugation, we have

$$\begin{aligned} \mathcal{C} \mathcal{L} \mathcal{C}^* = \frac{ig_2}{2\sqrt{2}} \bigg\{ & (\alpha_{u_m})_c (\alpha_{d_n})_c^* V_{mn} W_\mu^- [\bar{d}_n \gamma^\mu (1-\gamma_5) u_m]^* \\ & + (\alpha_{u_n})_c^* (\alpha_{d_m})_c (V^\dagger)_{mn} W_\mu^+ [\bar{u}_n \gamma^\mu (1-\gamma_5) d_m]^* \bigg\} \end{aligned} \quad (2.114)$$

and under parity transformations we get

$$\begin{aligned} \mathcal{P} \mathcal{L} \mathcal{P}^* = \frac{ig_2}{2\sqrt{2}} \bigg[& (\alpha_{u_m})_p (\alpha_{d_n})_p^* V_{mn} W_\mu^+ \bar{u}_m \gamma^\mu (1-\gamma_5) d_n \\ & + (\alpha_{d_m})_p^* (\alpha_{u_n})_p (V^\dagger)_{mn} W_\mu^- \bar{d}_m \gamma^\mu (1-\gamma_5) u_n \bigg] \end{aligned} \quad (2.115)$$

It is clear that there is no choice of phases for which the Lagrangian is parity or charge-conjugation invariant, because any choice that would make the term involving γ^μ invariant would make the $\gamma_5 \gamma^\mu$ term not invariant (and vice versa). The point is that each operation replaces the projector $P_L = (1+\gamma_5)/2$ with the projector $P_R = (1-\gamma_5)/2$.

Combining both transformations, however, gives the following result:

$$\begin{aligned}
 (\mathcal{CP})\mathcal{L}(\mathcal{CP})^* &= \frac{ig_2}{2\sqrt{2}} \\
 &\times \left\{ (\alpha_{u_m})_c (\alpha_{d_n})_c^* (\alpha_{u_m})_p (\alpha_{d_n})_p^* V_{mn} W_\mu^- [\bar{d}_n \gamma^\mu (1+\gamma_5) u_m]^* \right. \\
 &\quad \left. + (\alpha_{u_n})_c^* (\alpha_{d_m})_c (\alpha_{u_n})_p^* (\alpha_{d_m})_p (V^\dagger)_{mn} W_\mu^+ [\bar{u}_n \gamma^\mu (1+\gamma_5) d_m]^* \right\} \quad (2.116)
 \end{aligned}$$

If the phases can be chosen to satisfy $(\alpha_{u_m})_c (\alpha_{d_n})_c^* (\alpha_{u_m})_p (\alpha_{d_n})_p^* = 1$, and the KM matrix can be simultaneously chosen to be real, then this last equation would be precisely the complex conjugate of the original Lagrangian. Inspection of the other terms in the Lagrangian confirms that the phase choice can be made provided that the KM matrix may be chosen to be real. Therefore, as claimed, the standard model fails to conserve CP invariance only in that the KM matrix cannot be made purely real.

2.5.2 Continuous symmetries

It is of considerable interest to determine the continuous global symmetries of the standard model Lagrangian. The purpose of this section is to identify the exact, and some approximate, symmetries of this Lagrangian.

The starting point is the class of symmetries of the Lagrangian in the absence of all interactions or mass terms. This will give the maximum possible symmetry group which could exist, given the particle content of the model. The interactions of the theory will not respect all of this symmetry. We will consider each interaction in turn and see how it cuts down the size of the actual symmetry group, until we find what symmetries remain.

As is discussed in Chapter 1, when the basis of fields is chosen to be real (or Majorana), this class consists of a general independent orthogonal rotation among all of the bosonic fields of a given spin, as well as a unitary rotation amongst the left-handed fermions. For the standard model the group of all such transformations is $G_{\max} = O(4) \times O(12) \times U(45)$, corresponding to the four real scalar fields, 12 gauge potentials and three generations of fermions each containing 15 different species of fermion (one E , two from L , three each from U and D , and six from Q). We will write this group as $G_{\max} = G_0 \times G_{\frac{1}{2}} \times G_1$, with $G_0 = O(4)$ the group of scalar transformations, $G_{\frac{1}{2}} = U(45)$ the group of fermionic transformations, and $G_1 = O(12)$ the group of gauge-field transformations.

We wish to determine what subgroup of this group of transformations is preserved once the interactions are turned on. One immediate subgroup of

this type is the group of gauge transformations themselves: $G_g \equiv SU_c(3) \times SU_L(2) \times U_Y(1) \subset G$.

2.5.2.1 Gauge self-interactions

We next describe conditions G must satisfy if it is not to be broken by the gauge interactions.

Consider first the self-interactions of the twelve gauge bosons. As is discussed in more detail in Chapter 1, the free kinetic terms for these fields are invariant under the replacement of each field by an arbitrary linear combination of the fields, $\delta V_\mu^a = M_b^a V_\mu^b$, provided that the 12×12 matrix M_b^a is antisymmetric (and so its exponential, $[\exp(M)]_b^a$, is orthogonal). The group formed by these transformations is the group $G_1 = O(12)$. We wish to determine what subgroup of these transformations are also symmetries of the gauge boson self-interactions. In order to be an invariance of these interactions, a candidate symmetry transformation must preserve the structure constants of the gauge group

$$M_a^b c_{bd}^c + M_d^b c_{ab}^c = M_b^c c_{ad}^b \quad (2.117)$$

The algebra of infinitesimal symmetry transformations of the gauge boson self interactions is given by that subalgebra of G_1 that satisfies Eq. (2.117). This subalgebra must include the Lie algebra of the gauge group itself, because infinitesimal gauge rotations, $\delta V_\mu^a = \epsilon^b c_{bc}^a V_\mu^c$, automatically satisfy Eq. (2.117) by virtue of the Jacobi identity that is satisfied by the structure constants, c_{bc}^a .

An immediate consequence of Eq. (2.117) is that if the gauge group consists of several mutually commuting factors, $G_g = H_1 \times H_2 \times \cdots$ (as is the case for the standard model), then $M_a^b = 0$ unless both a and b correspond to generators that are in the same factor of G_g . It is a theorem of the theory of compact semisimple Lie groups that the only Lie subgroup of G_1 that satisfies Eq. (2.117) is the gauge subgroup itself (i.e. $G_1 \sim G_g$ consists of the group of *inner* automorphisms of G_g). As a result, there are no accidental global symmetries within the gauge boson sector of the theory.

2.5.2.2 Scalar-gauge and scalar self-couplings

The next simplest case is the scalar sector of the model. The Higgs doublet consists of four real scalar fields, $\phi_i = \phi_i^*$, and so the free kinetic terms of these fields are invariant under arbitrary $G_0 = O(4)$ rotations, $\delta \phi_i = iR_j^i \phi^j$ with $R + R^T = 0$, of these fields into one another. As discussed in Subsection 2.5.2, this symmetry is not broken by the scalar self-interactions as described by the scalar potential. We wish to know which subgroup of

G_0 is also a symmetry of the scalar-gauge interactions. Our answer to this question is not specific to the example $O(4)$ but applies more generally for larger symmetry groups, G_0 .

Consider a group G_R of symmetry transformations, with group generators we will designate as R . If the generators of the gauge transformations are t_a , then the condition for the group of symmetry transformations to be unbroken by the gauge transformations is

$$[t_a, R] = N_a^b t_b \quad (2.118)$$

for each R and each t_a of the gauge group. The coefficients N_a^b represent a rotation among the gauge potentials of the theory that might be necessary to compensate for the effects of the scalar rotation, R . For our application, we are interested in the case where G_R is a subgroup of G_0 .

Note that R and t_a are all generators of the group G_0 ; so Eq. (2.118) is a special case of the Lie algebra of G_0 . Choose a basis for the generators of G_0 such that the structure constants f^A_{BC} are totally antisymmetric. Then $[t_a, R] = f_{aR}^B g_B$, with g_B one of the generators of G_0 . For Eq. (2.118) to hold, either f_{aR}^B vanishes, or g_B must be one of the t_b . But that would imply that $[t_a, t_b] \propto R$, which cannot be – the t_a must be a subgroup of G_0 , so their algebra should be closed. Therefore, R must either be a generator of the group of gauge transformations, or it must commute with all of the generators of the gauge group.

Since the solution in which R is a gauge transformation generator does not represent a new, accidental, symmetry, we focus on the alternative for which R commutes with all of the gauge transformations in G_g . By Schur's lemma, this implies that the transformations, R , cannot mix fields that transform in different irreducible representations of the gauge group. The resulting symmetry of the gauge interactions then becomes a product of orthogonal groups, $O(N_1) \times O(N_2) \times \cdots$ in which each factor describes the rotations of the N_i fields that all transform in the common representation, r_i , of the gauge group.

Since only a single irreducible representation of scalar fields appears in the standard model, and since there is no other subgroup of $O(4)$ which commutes with the $SU_L(2) \times U_Y(1)$ subgroup, there are no accidental global symmetries of the scalar gauge couplings. It is purely the $U_Y(1)$ gauge couplings that break the potential $O(4)$ symmetry of the scalar sector. One way to see this is to notice that the Lie algebra of $O(4)$ is isomorphic to that of the algebra $SU(2) \times SU(2)$, of which one of the $SU(2)$ factors may be taken to be the gauge group $SU_L(2)$. In the absence of the $U_Y(1)$ gauge couplings to the scalars, there would therefore be an entire $SU(2)$ subgroup of G_0 that

commutes with the gauge group. This is the origin of the custodial $SU(2)$ symmetry of Subsection 2.3.2.

Although the standard model is not invariant under the full $O(4)$ invariance, conclusions based on this symmetry do become correct in the limit that the $U_Y(1)$ gauge coupling—and, as we shall see, the Yukawa couplings—vanish. Since this coupling is known to be experimentally small, it follows that the $O(4)$ symmetry is a good *approximate symmetry* of the standard model. Such approximate symmetries can be almost as useful as exact symmetries if the non-invariant couplings are sufficiently small.

2.5.2.3 Fermion–gauge couplings

The only place left to look for accidental global symmetries is inside the group $G_{\frac{1}{2}} = U(45)$ of transformations between the 45 species of left-handed fermions. (The number 45 arises as three generations times one E , two L , three U , three D , and six Q fields per generation. A quark species counts for three because of its three colors, L and Q count double because of the two flavors in each.)

If we work with a basis of fermions which are in definite representations of the gauge group—as opposed to being mass eigenstates—the condition that the symmetry transformations be preserved by the fermion gauge interactions is a direct analog of Eq. (2.118). It follows that a subgroup of $G_{\frac{1}{2}}$ preserves the fermion–gauge interactions if it is either the subgroup of the gauge transformations themselves, or it commutes with this gauge subgroup. Since the 15 fermion species of a given generation transform under the gauge group

$$\begin{aligned} SU_c(3) \times SU_L(2) \times U_Y(1) \text{ as } & \left(\mathbf{3}, \mathbf{2}, +\frac{1}{6} \right) \oplus \left(\bar{\mathbf{3}}, \mathbf{1}, -\frac{2}{3} \right) \oplus \left(\bar{\mathbf{3}}, \mathbf{1}, +\frac{1}{3} \right) \\ & \oplus \left(\mathbf{1}, \mathbf{2}, -\frac{1}{2} \right) \oplus (\mathbf{1}, \mathbf{1}, +1) \end{aligned}$$

and since none of these irreducible representations is big enough to admit an internal potential symmetry that commutes with the gauge group, there are no accidental symmetry transformations relating the fermions within a single generation.

The accidental symmetries of the fermion–gauge couplings are therefore

$$G_f \equiv U_Q(3) \times U_U(3) \times U_D(3) \times U_L(3) \times U_E(3) \subset G_{\frac{1}{2}} \quad (2.119)$$

Each factor of this group corresponds to a unitary rotation in generation space of the five types of irreducible $SU_c(3) \times SU_L(2) \times U_Y(1)$ representations of the model's fermion content.

2.5.2.4 Yukawa interactions

From the previous paragraphs, the only exact non-gauge symmetries of the gauge interactions of the standard model are $G_f = [U(3)]^5$, representing independent transformations, in generation space, of each type of fermion fields. The final issue is to determine which of these potential symmetries also preserves the Yukawa interactions of the theory.

The conditions that must be satisfied in order for these transformations to preserve the form of the Yukawa couplings of Eq. (2.25) are

$$\begin{aligned} (U_L^T f U_E)_{mn} &= f_{mn} \\ (U_Q^T g U_U)_{mn} &= g_{mn} \\ (U_Q^T h U_D)_{mn} &= h_{mn} \end{aligned} \quad (2.120)$$

These equations imply that the potential symmetry transformations must also satisfy the following additional conditions, which each involve only left-handed or only right-handed unitary transformations:

$$\begin{aligned} (U_E^\dagger f^\dagger f U_E)_{mn} &= (f^\dagger f)_{mn} \\ (U_L^T f f^\dagger U_L^*)_{mn} &= (f f^\dagger)_{mn} \\ (U_U^\dagger g^\dagger g U_U)_{mn} &= (g^\dagger g)_{mn} \\ (U_Q^T g g^\dagger U_Q)_{mn} &= (g g^\dagger)_{mn} \\ (U_D^\dagger h^\dagger h U_D)_{mn} &= (h^\dagger h)_{mn} \\ (U_Q^T h h^\dagger U_Q)_{mn} &= (h h^\dagger)_{mn} \end{aligned} \quad (2.121)$$

In order to analyze the implications of these conditions, it is convenient to work with a basis of fields for which the fermion mass matrix, and so also the Yukawa coupling matrices, are real and diagonal. (Since the transformation to this basis introduces the Kobayashi–Maskawa matrix into the charged-current fermion gauge couplings, these couplings must be re-examined for invariance at the end.)

In this basis, and taking the experimental information that none of the eigenvalues of the Yukawa coupling matrices f_{mn} , g_{mn} , and h_{mn} vanish or are degenerate, Eq. (2.121) implies that each of the unitary matrices must be diagonal with phases along their diagonals. This reduces the candidate symmetry group for the fermions to the multiplication of the left- and right-handed parts of each mass eigenstate by an independent $U(1)$ phase.

Using this form for the unitary transformations in the original condition of Eq. (2.120) implies that the left- and right-handed transformations must be equal for each type of fermion; that is $U_Q = U_U^* = U_D^*$ and $U_L = U_E^*$.

For leptons this is the end of the story, implying that the accidental symmetry of the lepton sector is $U_e(1) \times U_\mu(1) \times U_\tau(1)$:

$$U_L = U_E^* = \begin{pmatrix} e^{i\theta_e} & & \\ & e^{i\theta_\mu} & \\ & & e^{i\theta_\tau} \end{pmatrix} \quad (2.122)$$

For quarks, we must also check that these phase transformations preserve the form for the charged-current gauge interactions when written in terms of mass eigenstates as in Eq. (2.115). To be invariant, the candidate transformation must therefore commute with the KM matrix. For a generic unitary KM matrix the only combination of such transformations are those that are proportional to the unit matrix in generation space, and so which rotate all quarks by a common phase. Therefore, there is only a single $U(1)$ transformation left:

$$U_Q = U_U^* = U_D^* = \begin{pmatrix} e^{i\theta_B/3} & & \\ & e^{i\theta_B/3} & \\ & & e^{i\theta_B/3} \end{pmatrix} \quad (2.123)$$

The corresponding group is $U_B(1)$. The factor of $1/3$ is chosen so that the charge of a quark under this $U(1)$ is $1/3$. Since bound states of quarks always contain a multiple of 3 quarks (see Chapter 8), they have integer charge (0 or ± 1) under this symmetry.

The accidental global symmetry group of the standard model is therefore

$$G = U_e(1) \times U_\mu(1) \times U_\tau(1) \times U_B(1) \quad (2.124)$$

Each of the four generators of this symmetry group corresponds to a quantum number that appears to be experimentally conserved. They are:

- (i) *electron number*: $L_e(e^-) = L(\nu_e) = +1$, $L_e(e^+) = L(\bar{\nu}_e) = -1$,
 $L_e = 0$ for all others;
- (ii) *muon number*: $L_\mu(\mu^-) = L(\nu_\mu) = +1$, $L_\mu(\mu^+) = L(\bar{\nu}_\mu) = -1$, $L_\mu = 0$
for all others;
- (iii) *tau number*: $L_\tau(\tau^-) = L(\nu_\tau) = +1$, $L_\tau(\tau^+) = L(\bar{\nu}_\tau) = -1$, $L_\tau = 0$
for all others;
- (iv) *baryon number*: $B(q) = \frac{1}{3}$ for all quarks, q , $B(\bar{q}) = -\frac{1}{3}$ for antiquarks,
and $B = 0$ for all others.

The sum $L = L_e + L_\mu + L_\tau$ is also known as *lepton number*. It is one of the triumphs of the standard model that its accidental symmetries

correspond exactly with those conserved quantum numbers that had been experimentally observed.

Conservation of these quantum numbers immediately implies the stability of the lightest particles that carry nonzero values for them. Given that the neutrinos are massless and the charged leptons are not, we conclude that all neutrino types are absolutely stable in this theory. Similarly, the lightest baryon, which turns out to be the proton, is also predicted never to decay. The electron is similarly stable because it is the lightest particle in the theory that carries electric charge.

These conservation laws similarly forbid processes such as the reaction $\mu \rightarrow e\gamma$, since these do not conserve L_e or L_μ . This agrees with the current experimental upper bound on this decay, which at present indicates that it must occur less frequently than once in every 10^{11} μ decays.

In fact, there is now evidence that the separate lepton numbers are *not* conserved, and that neutrinos are not perfectly massless – though the effects which violate lepton number are tiny and are of no bearing in most particle physics experiments. Chapter 10 discusses the evidence for this violation, together with some of its implications. At present, experiments do not provide evidence for $\mathcal{L} = \mathcal{L}_e + \mathcal{L}_\mu + \mathcal{L}_\tau$ violation.

As it happens, one of the puzzling features of the standard model is the small size of the Yukawa couplings for almost all of the fermions of the theory. An equivalent way to phrase the same puzzle is to ask why the fermion masses (apart from that of the top quark) are all so small in comparison to, say, the masses of the W and the Z . To the extent that these Yukawa couplings can be ignored, there is a larger approximate flavor symmetry, $[U_L(3) \times U_E(3)]$ for leptons and $[U_Q(2) \times U_U(2) \times U_D(3)]$ for quarks. A related $\mathcal{U}_L(3) \times \mathcal{U}_R(3)$ approximate symmetry emerges when electroweak interactions are turned off, and is very useful for analyzing the low-energy properties of the strongly interacting quark sector in which the implications of such a *chiral* $U_L(3) \times U_R(3)$ symmetry provides otherwise unobtainable information about the spectrum of the light strongly interacting particles. These approximate symmetries are considered in much more detail in Chapter 8.

2.5.3 Anomalies

The discussion of the previous sections has dealt exclusively with the symmetries of the *classical* action of the model and has neglected quantum considerations. We devote this section to a discussion of the potential complications that arise when considering symmetries within a quantum, as opposed to classical, field theory.

In order to outline the issue at stake, recall that there are several uses to which symmetries are applied. The most important place is in the coupling of light spin-one particles. Here it was argued that these interactions could only be Lorentz invariant and unitary if they were also invariant under local gauge transformations. Another application was to use the existence of global (or local) symmetries to infer the existence of local conservation laws and symmetry relations amongst the energy eigenvalues of the system concerned.

The logic used in all of these applications has been: (i) The invariance of the classical action under a particular symmetry transformation ensures, by Noether's theorem, the existence of a set of currents, j_a^μ , whose conservation, $\partial_\mu j_a^\mu = 0$, follows from the equations of motion for the fields; (ii) these conserved currents may be used to construct conserved charges, $Q_a = \int j_a^0 d^3x$, for which the equations of motion for the fields imply $[H, Q_a] = 0$.

Unfortunately, such classical arguments do not always hold in a quantum theory. The process of quantizing a given classical theory introduces ambiguities associated with the ordering of operators in the quantum theory. In a field theory this operator-ordering ambiguity is intimately related with the divergences at short distances, since operators only fail to commute when their spacelike separations tend to zero. Since different operator orderings for the system Hamiltonian give rise to different equations of motion, and since the conservation of the Noether current depends on these equations of motion, the form taken by the conserved current will in general depend on how these operator-ordering issues are resolved.

It could potentially happen that there is no operator ordering under which all would-be currents are conserved, even if they should be conserved at the classical level. That is to say, it might happen that the existence of a symmetry of the classical action might not be sufficient for the existence of a conserved quantum charge operator. Should this occur, we would lose the exact results we hoped to derive from the existence of the symmetry. The discovery that classical symmetries can fail in this way was so surprising when it was discovered that this failure of a symmetry to survive quantization was termed an *anomaly*. The purpose of the remainder of this section is to summarize under what circumstances a symmetry is "anomalous" in this way.

Precisely such an anomaly can indeed occur for a current if the symmetry at issue involves transformations on Majorana fermions. Since the distinction between right- and left-handed fields is essential here, the anomaly is termed the *chiral anomaly*. While it is beyond the scope of this book to derive how such an anomaly arises, the condition for the absence of a chiral

anomaly may be fairly simply stated. Suppose that the generators of a classical symmetry acting on left-handed spinor fields are denoted by T_a . Then, as is discussed in Section 2.1, the action of the symmetry on a Majorana spinor becomes $\delta\psi^m = i\epsilon^a[(T_a)_n^m P_L - (T_a^*)_n^m P_R]\psi^n$. The classical symmetry survives quantization, and so is called *anomaly free*, if the *anomaly coefficients*, A_{abc} , vanish for all a, b , and c . These coefficients are completely symmetric under permutations of the indices a, b , and c , and are defined by

$$A_{abc} = \text{tr}(T_a\{T_b, T_c\}) \quad (2.125)$$

The curly brackets in this equation denote the anticommutator, $\{T_b, T_c\} \equiv T_b T_c + T_c T_b$, and the trace means that a sum is to be taken over all types of fermions, *e.g.* every color of every flavor of quark and every lepton, in each generation, with T denoting the action of the symmetry on that particular particle type (so if T_a represents the action of one of the color generators, it is $\lambda_a/2$ in color space when acting on a quark, and 0 when acting on a lepton, since leptons are colorless and do not change under a color rotation).

In particular, when the anomaly coefficient A_{abc} does not vanish and the indices b, c correspond to gauge symmetries, then the conservation of the current J_a^μ is violated by

$$\partial_\lambda J_a^\lambda = \frac{A_{abc}}{64\pi^2} \epsilon^{\mu\nu\alpha\beta} g F_{\mu\nu}^b g F_{\alpha\beta}^c \quad (2.126)$$

with F the field strength corresponding to symmetry b and g the associated gauge coupling.

A consequence of the structure of Eq. (2.125) is that there are no anomalies for *real* (or pseudoreal) fermion representations. A (pseudo-) real representation is defined to be one for which the generators iT_a are real up to a similarity transformation: $T_a^* = -S T_a S^{-1}$ for some invertible matrix S . To see that this ensures freedom from anomalies, notice that since the generators T_a are Hermitian it follows that $T_a^T = T_a^*$. Then

$$\begin{aligned} A_{abc} &= \text{tr}(T_a\{T_b, T_c\}) \\ &= \text{tr}[(T_a\{T_b, T_c\})^T] \\ &= \text{tr}(\{T_c^T, T_b^T\}T_a^T) \\ &= \text{tr}(\{T_c^*, T_b^*\}T_a^*) \\ &= -\text{tr}(S\{T_c, T_b\}T_a S^{-1}) \\ &= -\text{tr}(\{T_c, T_b\}T_a) \\ &= -A_{abc} = 0 \end{aligned} \quad (2.127)$$

This will make the calculation of several anomaly coefficients much easier.

An important special case of this last result occurs when fermion number is conserved and when the left- and right-handed fermions (as opposed to antifermions) transform in the same representation, t_a say, of the group of interest. In this case the generator of this group acting on *all* of the left-handed spinors (for fermions *and* antifermions) may be written in the block-diagonal form

$$T_a = \begin{pmatrix} t_a & 0 \\ 0 & -t_a^* \end{pmatrix} \quad (2.128)$$

This is manifestly pseudoreal since $T_a^* = -ST_aS^{-1}$. It follows that any symmetry that is left-right symmetric in this way must be anomaly-free.

Because of the central role symmetries play in field theory, we must check two things.

- (i) First, since the gauge symmetries of the standard model are chiral in the sense just described, we must verify that they are anomaly-free, that is, that all anomalies involving three gauge symmetries vanish. Otherwise, the gauge fields will not couple to conserved currents, and the gauge interactions will not be simultaneously Lorentz-invariant and unitary. Since these are both basic principles of quantum field theory, a theory with anomalous gauge symmetries *does not exist* (is not a valid theory).
- (ii) Next, we must see whether the exact and approximate global “accidental” symmetries of the standard model have anomalies or not. No issues of consistency need arise if they do have anomalies, since these symmetries are not associated with the couplings of any spin-one particles. It is nevertheless important to understand which are anomalous, since anomalies negate the argument that would allow these classical symmetries to imply the existence of exact conservation laws or spectral relations.

These two issues are the topics of the following two sections.

2.5.3.1 Cancellation of gauge anomalies

Let us verify that the anomaly coefficient, A_{abc} , vanishes in the standard model when all of the indices, a , b , and c , correspond to gauge group generators. As we shall see, this *anomaly cancellation* relies on the detailed quantum numbers of the standard model fermions and requires all of the members of a complete generation in order to work.

We consider each combination of generators in turn. We will use the notation “A(3, 3, 3)” for the anomaly coefficient involving three $SU_c(3)$

generators etc. We demonstrate that the contribution to the anomaly coefficient from each generation separately vanishes.

- (i) $A(3, 3, 3)$: The $SU_c(3)$ representations are all left-right symmetric. This anomaly coefficient must therefore vanish for the general reasons given above.
- (ii) $A(3, 3, 2)$: These coefficients are all proportional to the trace of the Pauli matrices since these furnish the two-dimensional $SU_L(2)$ representations. Since the Pauli matrices are all traceless this anomaly coefficient must vanish.
- (iii) $A(3, 3, 1)$: The three-dimensional $SU_c(3)$ generators are given by the Gell-Mann matrices, $\lambda_\alpha/2$, of Eq. (1.186). These are all tracefree and satisfy the following property:

$$\{\lambda_\alpha, \lambda_\beta\} = \frac{4}{3}\delta_{\alpha\beta} + 2d_{\alpha\beta\gamma}\lambda_\gamma$$

The trace over colors of $\delta_{\alpha\beta}$ will give 3, while the trace over $d_{\alpha\beta\gamma}\lambda_\gamma$ gives zero; so $A(3,3,1)$ is therefore proportional to the trace over all left-handed colored fields (i.e. quarks) of the $U_Y(1)$ generator—weak hypercharge, Y . The anomaly coefficient therefore is

$$\begin{aligned} A(3, 3, 1) = \sum_{\text{quarks}} Y &= 3(2y_{Q_L} + y_{U_L} + y_{D_L}) \\ &= 3 \left[2 \left(\frac{1}{6} \right) + \left(-\frac{2}{3} \right) + \left(\frac{1}{3} \right) \right] \\ &= 0 \end{aligned} \tag{2.129}$$

The overall factor of 3 is the number of generations. The factor of 2 on y_{Q_L} is because of the two $SU_L(2)$ flavors.

- (iv) $A(3, X, Y)$: This coefficient vanishes for X and Y equal to either 2 or 1 since it is proportional to the trace of a Gell-Mann matrix, which vanishes.
- (v) $A(2, 2, 2)$: As observed above, the only nontrivial $SU_L(2)$ representations that appear within the standard model are doublets, and so the generators are represented by the Pauli matrices. Since all three Pauli matrices satisfy the following identity, $\tau_a^* = -\tau_2\tau_a\tau_2$, it follows that this representation is pseudoreal, and so the anomaly coefficient must vanish by the general argument of Eq. (2.125).
- (vi) $A(2, 2, 1)$: The Pauli matrices satisfy an identity similar to that satisfied by the Gell-Mann matrices: $\{\tau_a/2, \tau_b/2\} = \delta_{ab}/2$, which is doubled when summed over a doublet. This anomaly coefficient is

therefore the sum over $SU_L(2)$ doublets of the weak hypercharge, Y :

$$\begin{aligned}
 A(2, 2, 1) &= \sum_{\text{doublets}} Y = 3(y_{L_L} + 3y_{Q_L}) \\
 &= 3 \left[\left(-\frac{1}{2}\right) + 3 \left(\frac{1}{6}\right) \right] \\
 &= 0
 \end{aligned} \tag{2.130}$$

The factor of 3 on the Q contribution arises from the trace on colors.

- (vii) $A(2, 1, 1)$: This coefficient vanishes simply because it is proportional to the trace of a single Pauli matrix, which is zero.
- (viii) $A(1, 1, 1)$: This coefficient is proportional to the sum over all left-handed fermions of the cube of the weak hypercharge:

$$\begin{aligned}
 A(1, 1, 1) &= 2 \sum_{\text{all}} Y^3 = 6(2y_{L_L}^3 + y_{E_L}^3 + 6y_{Q_L}^3 + 3y_{U_L}^3 + 3y_{D_L}^3) \\
 &= 6 \left(2 \left(-\frac{1}{2}\right)^3 + (+1)^3 + 6 \left(\frac{1}{6}\right)^3 + 3 \left(-\frac{2}{3}\right)^3 + 3 \left(\frac{1}{3}\right)^3 \right) \\
 &= 0
 \end{aligned} \tag{2.131}$$

It is clear that anomaly cancellations in the standard model require non-trivial relationships between the number of species of and the quantum numbers for the quarks and leptons. It is also clear that the values of the hypercharges of the different species are not accidental. The relations $Y_{E_L} + Y_{L_L} = 1/2$, $Y_{D_L} + Y_{Q_L} = 1/2$, and $Y_{U_L} + Y_{Q_L} = -1/2$ are enforced by the requirement that the Yukawa interaction terms be hypercharge-invariant. However, until now, the fact that $Y_{E_L} = 1$ and not, say, $1 + \epsilon$, has been a mystery. This is important; if it were $1 + \epsilon$, the neutrinos would possess electric charges of $-\epsilon$. Similarly, Y_{D_L} could be $1/3 + \delta$ rather than $1/3$, in which case the neutron would be charged, and the electron and proton charges would differ. (The proton charge is $2Q_u + Q_d$.) In fact, limits on neutrino and neutron charges and on proton–electron charge differences are very strong; for instance, the electron and proton charges differ in absolute value by no more than a part in 10^{21} . The reason is that Eq. (2.130) and Eq. (2.131) only sum to zero if $\epsilon = \delta = 0$. Therefore the equality of the proton charge and the electron charge, and the vanishing of the neutrino and neutron charges, are exact identities within the standard model.

We next consider the potential anomalies that could arise in the Lorentz algebra. The Lorentz group has been treated here as a global rather than a gauge symmetry and so might be treated in the following section. However, the introduction of gravitational interactions requires it to be gauged, so

if gauge-Lorentz anomalies exist, then the theory of gravitation would be inconsistent. Therefore we consider it here.

The only standard model particles that are in complex representations of the Lorentz group are the fermions. Since the Lorentz generators on fermions (c.f. Subsection 1.3.2) are essentially equivalent to $SU(2)$ transformations, the anomaly cancellation arguments are similar to those for an $SU(2)$ gauge group. It follows that the only anomaly coefficient that does not vanish immediately due to the properties of the Pauli matrices is $A(J, J, 1)$, in which J generically denotes the Lorentz generators. The condition that this anomaly coefficient be zero is that the trace of the weak hypercharge over all left-handed fermions vanish:

$$\begin{aligned}
 \text{tr}_{\text{all}} Y &= 3(2y_{L_L} + y_{E_L} + 6y_{Q_L} + 3y_{U_L} + 3y_{D_L}) \\
 &= 3 \left[2 \left(-\frac{1}{2} \right) + (+1) + 6 \left(\frac{1}{6} \right) + 3 \left(-\frac{2}{3} \right) + 3 \left(\frac{1}{3} \right) \right] \\
 &= 0
 \end{aligned} \tag{2.132}$$

2.5.3.2 Anomalies in global symmetries

We next compute the anomalies for the accidental global symmetries and for some of the approximate global symmetries that were identified in the previous sections.

For baryon number, B , the anomaly coefficients are:

$$\begin{aligned}
 A(3, 3, B) &= \sum_{\text{quarks}} B = 18 \left(\frac{1}{3} \right) + 9 \left(-\frac{1}{3} \right) + 9 \left(-\frac{1}{3} \right) = 0 \\
 A(2, 2, B) &= \sum_{\text{doublets}} B = 9 \left(\frac{1}{3} \right) = 3 \\
 A(1, 1, B) &= \sum_{\text{all}} 2Y^2 B \\
 &= 36 \left(\frac{1}{6} \right)^2 \left(\frac{1}{3} \right) + 18 \left(-\frac{2}{3} \right)^2 \left(-\frac{1}{3} \right) + 18 \left(\frac{1}{3} \right)^2 \left(-\frac{1}{3} \right) = -3; \\
 A(1, B, B) &= \sum_{\text{all}} 2Y B^2 \\
 &= 36 \left(\frac{1}{6} \right) \left(\frac{1}{3} \right)^2 + 18 \left(-\frac{2}{3} \right) \left(-\frac{1}{3} \right)^2 + 18 \left(\frac{1}{3} \right) \left(-\frac{1}{3} \right)^2 = 0; \\
 A(B, B, B) &= \sum_{\text{all}} B^3 = (36 - 18 - 18)/27 = 0; \\
 A(J, J, B) &= \sum_{\text{all}} B = (12 - 6 - 6) = 0
 \end{aligned} \tag{2.133}$$

For lepton numbers, L_e , L_μ , L_τ , each of these charges gets contributions only from its own generation, so the factors of 3 from the generation sum in

the baryon results will be absent. It suffices to compute the anomalies for one of them since the results are identical for the others. Anomalies between lepton symmetries vanish.

$$\begin{aligned}
A(2, 2, L_e) &= \sum_{\text{doublets}} L_e = 1 \\
A(1, 1, L_e) &= \sum_{\text{all}} 2Y^2 L_e = 4 \left(-\frac{1}{2}\right)^2 (+1) + 2(+1)^2 (-1) = -1 \\
A(1, L_e, L_e) &= \sum_{\text{all}} 2Y L_e^2 = 4 \left(-\frac{1}{2}\right) (+1)^2 + 2(+1) (-1)^2 = 0; \\
A(L_e, L_e, L_e) &= \sum_{\text{all}} 2L_e^3 = 4(+1)^3 + 2(-1)^3 = 2 \\
A(J, J, L_e) &= \sum_{\text{all}} L_e = 2(+1) + 1(-1) = 1
\end{aligned} \tag{2.134}$$

Chiral $U(3)$ is an approximate symmetry under which the three lightest left- and right-handed quarks get shuffled amongst one another, $U_{qL}(3) \times U_{qR}(3)$. It will be of interest in Chapter 8, where we will need to know how much of this approximate symmetry group is anomaly-free. We consider here only the quark sector since this is the case that is of most direct interest in subsequent chapters. For brevity, we consider only the left-handed case explicitly here. Denote a general $U_{qL}(3)$ generator by T_a and denote its specific 3×3 representation by t_a . Then

$$\begin{aligned}
A(3, 3, T_a) &\propto \text{tr } t_a \\
A(2, 2, T_a) &\propto \text{tr } t_a \\
A(1, 1, T_a) &\propto \text{tr } t_a \\
A(1, T_a, T_b) &\propto \text{tr}(t_a t_b) \propto \delta_{ab} \\
A(T_a, T_b, T_c) &\propto \text{tr}(t_a \{t_b, t_c\}) \\
A(J, J, T_a) &\propto \text{tr } t_a
\end{aligned} \tag{2.135}$$

Some comments.

- (i) Perhaps the most basic observation about these anomaly coefficients is that they are not zero. It follows that the naive conclusions that are based on the corresponding symmetries can break down and so must be treated with caution. It turns out, however, that for physics at temperatures low compared to the W boson mass, any violation of the corresponding conservation laws due to quantum effects are proportional to $\exp(-8\pi^2/g^2)$ and so are negligibly small for weak couplings ($g \ll 1$). The same arguments indicate that those global symmetries that have anomalies due to any strong interactions are

strongly broken, and so should not provide good approximations to the dynamics of the full quantum theory.

As a result, all of the consequences of the exact global symmetries are expected to hold for the standard model to an extremely good approximation. However, those symmetries having $SU_c(3)$ anomalies are expected to be strongly broken.

- (ii) The only anomaly-free global symmetries of the standard model are found by taking appropriate linear combinations of the anomalous symmetries given above. The symmetries free of all anomalies, including gravitational anomalies, are $L_e - L_\mu$, $L_e - L_\tau$, and $L_\mu - L_\tau$ (which is linearly dependent on the first two).
- (iii) Notice that all of the $SU_L(3) \times SU_L(2) \times U_Y(1)$ anomalies are the same for baryon number, B , as they are for the total lepton number, $L = L_e + L_\mu + L_\tau$. The Lorentz, B^3 , and L^3 anomalies would also agree if the model were to be supplemented by a right-handed neutrino field per generation. This suggests that the combination $B-L$ would be anomaly free, including gravitational effects, in the presence of right-handed neutrinos.
- (iv) It is clear from Eq. (2.135) that all of the chiral $U(3)$ transformations have anomalies of one type or another. Only those with a non-vanishing trace receive $SU_c(3)$ anomalies, however, so the traceless ones would be bona fide symmetries to the extent that the electroweak interactions are negligible. Now, since the group $U(3)$ is generated by arbitrary 3×3 Hermitian matrices, and since any such matrix may always be decomposed as a linear combination of traceless Gell-Mann matrices and the unit matrix, it follows that the Lie algebra for $U(3)$ is equivalent to that of the product $SU(3) \times U(1)$. Since only the $U(1)$ generator has a non-vanishing trace, only it suffers from an $SU_c(3)$ anomaly. As a result, the strong interactions break the approximate symmetry $U_{qL}(3) \times U_{qR}(3)$ down to its subgroup $SU_{qL}(3) \times SU_{qR}(3) \times U_B(1)$. The unbroken $U_B(1)$ is that combination of the $U(1)$ s that acts equally on left- and right-handed quark fields, and so may be recognized simply as baryon number.

2.6 Problems

[2.1] Anomaly cancellation and charge assignments

Complete the proof that anomaly cancellation fixes the charges of the standard model fermions.

First, take the hypercharge of the Higgs field ϕ to be exactly $-1/2$. This can be considered as the definition of the normalization of g_1 . Then, write the hypercharges of P_L and $P_L Q$ as $q_L \equiv -1/2 - \epsilon$ and $q_Q \equiv 1/6 - \delta$.

Show that the hypercharges of the E , U , and D fields are fixed by the requirement that the Yukawa interactions be gauge invariant, and find expressions for q_E , q_D , and q_U , the hypercharges of $P_L E$, $P_L D$, and $P_L U$.

Then find expressions for the two anomaly conditions, Eq. (2.130) and Eq. (2.131), in terms of δ and ϵ . Show that the only simultaneous solution to both equations is $\epsilon = \delta = 0$.

[2.2] Muon decay

The muon μ decays via the reaction

$$\mu^- \rightarrow e^- \nu_\mu \bar{\nu}_e$$

However, the decay

$$\mu^- \rightarrow e^- \gamma$$

with γ a photon has never been observed. Explain in terms of symmetries why there is no obstacle in principle to the first decay, but the second decay is forbidden and is expected to have a rate in the standard model of zero.

[2.3] Right-handed neutrinos

Suppose a right-handed neutrino for each generation (invariant under $SU_c(3) \times SU_L(2) \times U_Y(1)$) is added to the standard model.

[2.3.1] Show that the only new renormalizable terms that can appear in the Lagrangian are (also rewriting the kinetic term for the left handed leptons):

$$\mathcal{L} = -\frac{1}{2} \bar{L}_m \not{D} L_m - \frac{1}{2} \bar{N}_m \not{D} N_m - \frac{1}{2} M_m \bar{N}_m N_m - (k_{mn} \bar{L}_m P_R N_n \tilde{\phi} + \text{h.c.})$$

where N_m is the Majorana spinor whose right-handed piece is the right-handed neutrino and L_m is the usual lepton doublet. M_m is a real mass parameter and k_{mn} are Yukawa coupling constants.

[2.3.2] Do any combinations of electron-number, muon-number and tau-number remain conserved in the presence of these terms?

[2.3.3] Argue that these new terms induce a neutrino mass. Specializing to the case of one generation for simplicity, write down the neutrino mass matrix and identify the basis of fields in which it is diagonal and positive.

[2.3.4] Express the lepton-Higgs and lepton-gauge-boson interactions in terms of these mass eigenstates. (It is most convenient to keep using

Majorana spinors here because the mass matrix does not take a simple form in terms of Dirac spinors.)

[2.4] Two Higgs doublet models

Suppose the Higgs doublet of the standard model is supplemented by a second complex doublet, ψ , transforming as $(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$ under $SU_c(3) \times SU_L(2) \times U_Y(1)$.

[2.4.1] If ψ is written $\psi = \begin{pmatrix} \chi \\ \xi \end{pmatrix}$, what are the electric charges of the component fields χ and ξ ?

[2.4.2] Write out the covariant derivative $D_\mu \psi$ explicitly in terms of the gauge fields G_μ^α , W_μ^a and B_μ .

[2.4.3] Assuming the potential must be a function of the invariants $a = \phi^\dagger \phi$, $b = \psi^\dagger \psi$, and $c = \phi^T \epsilon \psi$, where ϕ is the usual Higgs doublet, what is the most general renormalizable form? How many independent real parameters does it contain? Need the parameters appearing in the potential be real? Is the combination $d = \phi^\dagger \psi$ $SU_L(2) \times U_Y(1)$ invariant?

[2.4.4] Suppose the parameters of the potential are such that it is minimized when

$$\phi = \phi_{\min} = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}$$

$$\psi = \psi_{\min} = \begin{pmatrix} \frac{1}{\sqrt{2}}(u + iw) \\ 0 \end{pmatrix}$$

u, v, w all real. Do these values break the electromagnetic group $U_{\text{em}}(1)$ generated by the electric charge $Q = T_3 + Y$? Identify the terms in the Lagrangian that are quadratic in the gauge fields and find their masses in terms of u, v , and w . Call the mass eigenstates $W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2)$, $Z_\mu = W_\mu^3 \cos \theta - B_\mu \sin \theta$, and $A_\mu = B_\mu \cos \theta + W_\mu^3 \sin \theta$. Express $\cos \theta$ in terms of the gauge couplings g_1 and g_2 . Is the standard model mass relation $M_W = M_Z \cos \theta$ also true for this model?

[2.4.5] What are the possible Yukawa couplings of the spin zero fields, ϕ and ψ , to the fermions? Suppose the Lagrangian is required to be invariant under the symmetry:

$$P_R E_m \rightarrow e^{i\theta} P_R E_m, \quad P_R U_m \rightarrow e^{i\theta} P_R U_m, \quad P_R D_m \rightarrow e^{i\theta} P_R D_m$$

$$\phi \rightarrow e^{-i\theta} \phi \quad \text{and} \quad \psi \rightarrow e^{-i\theta} \psi$$

with θ a real constant and all other fields being invariant. What are

the resulting restrictions on the Yukawa couplings and Higgs potential, $V(\phi, \psi)$?

[2.5] **Adjoint Higgs fields**

Suppose that the standard model is supplemented by a second complex Higgs field that transforms as a triplet of $SU_L(2)$ rather than as a doublet; i.e.

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix}$$

and

$$\delta_2 \psi = i\omega_2^a t_a \psi$$

with

$$t_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad t_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad t_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

(You can verify that t_1, t_2, t_3 satisfy the algebra of $SU_L(2)$ generators.) Suppose also that the hypercharge, Y , of the field ψ is zero.

[2.5.1] What is the electric charge of each component field, ψ_1, ψ_2 , and ψ_3 ?

[2.5.2] Suppose the potential for ψ and the usual Higgs field, ϕ , is minimized when

$$\phi = \phi_{\min} = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}$$

$$\psi = \psi_{\min} = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(u + iw) \\ 0 \end{pmatrix}$$

Do these values respect the electromagnetic gauge group $U_{\text{em}}(1)$ generated by the electric charge $Q = T_3 + Y$?

[2.5.3] Find the masses of the spin-one fields W_μ^\pm, Z_μ , and A_μ , where, as usual, $Z_\mu = W_\mu^3 \cos \theta - B_\mu \sin \theta$ and $A_\mu = B_\mu \cos \theta + W_\mu^3 \sin \theta$. What is $\cos \theta$ in terms of the gauge couplings? Is the mass relation $M_W = M_Z \cos \theta$ still valid?

[2.6] **Gauged B - L coupling**

Suppose the standard model is extended to contain an extra $U(1)$ symmetry $U(1)'$, with gauge boson F_μ and gauge coupling g_4 . Suppose that

the Higgs boson has charge 0 under this gauge boson, but the left-handed lepton doublet $P_L L$ has charge -1.

Also assume a complex scalar field χ , of charge +1 under the new symmetry but uncharged under hypercharge, is added to the Lagrangian. Write its effective potential as

$$V(\chi) = \lambda_\chi \left(\chi^* \chi - \frac{\mu^2}{2} \right)^2 \quad (2.136)$$

so that when $\mu^2 > 0$, it develops a vacuum expectation value. (There can also be an interaction term between the Higgs boson and χ , but assume that such a term is absent.)

[2.6.1] Revisit Problem 2.3, where a right-handed neutrino N is added to the standard model. What is the charge of $P_R N$ under $U(1)'$, and is the Majorana neutrino mass $M \bar{N} N$ still allowed?

[2.6.2] Based on the requirement that the Yukawa couplings preserve $U(1)'$ symmetry, and that all gauge anomalies cancel (in particular, the $(3, 3, 1')$, $(2, 2, 1')$, $(1, 1, 1')$, $(1, 1', 1')$, and $(1', 1', 1')$ anomaly coefficients are non-trivial), what must be the charges of the standard model fermions? Show that anomaly cancellation actually *demand*s that the theory possess an N field.

[2.6.3] What linear combinations of baryon number and the three lepton numbers remain conserved? Are there any Yukawa couplings involving the χ field?

[2.6.4] Argue that if $\mu^2 < 0$ so the χ field has no condensate, the F field is massless. In analogy with the Coulomb interaction mediated by the electromagnetic A field between charged particles, argue that there will be a Coulomb-like interaction between the electron and the neutron. Is it attractive or repulsive? How might it be observed or (very tightly!) constrained?

[2.6.5] Suppose that $\mu^2 > 0$. What is the spectrum of bosons? Does the normal relation between W and Z boson masses hold? Is there any mixing between F_μ and Z_μ , A_μ ?

[2.7] Colored scalar fields

Suppose the standard model is extended to include a complex scalar field \tilde{U} , transforming under the $(\mathbf{3}, \mathbf{1}, +\frac{2}{3})$ representation of $SU_c(3) \times SU_L(2) \times U_Y(1)$;

$$\delta \tilde{U} = \left(\frac{2ig_1}{3} \omega_1 + \frac{ig_3}{2} \lambda_\alpha \omega_3^\alpha \right) \tilde{U}, \quad D_\mu \tilde{U} = \left(\partial_\mu - \frac{ig_3}{2} G_\mu^\alpha \lambda_\alpha - \frac{2ig_1}{3} B_\mu \right) \tilde{U}$$

(This is the same as the transformation property of $P_R U$.)

[2.7.1] Show that \tilde{U}^* transforms under the $(\bar{\mathbf{3}}, \mathbf{1}, -\frac{2}{3})$ representation of $SU_c(3) \times SU_L(2) \times U_Y(1)$ (which is the same as the transformation rule for $P_L U$, see Eq. (2.11) and Eq. (2.20)), and that $\tilde{U}^\dagger \tilde{U}$ (with the contraction over the color indices implicit; the \dagger means that \tilde{U}^* is written as a row vector) is an $SU_c(3) \times SU_L(2) \times U_Y(1)$ invariant.

[2.7.2] Show that the following renormalizable interactions are allowed for the \tilde{U} field: a kinetic and gauge interaction term,

$$-(D_\mu \tilde{U})^\dagger (D^\mu \tilde{U})$$

a mass term,

$$-M_{\tilde{U}}^2 \tilde{U}^\dagger \tilde{U}$$

the following scalar interaction terms,

$$-\lambda'(\tilde{U}^\dagger \tilde{U})^2 - \lambda'' \phi^\dagger \phi \tilde{U}^\dagger \tilde{U}$$

and the following new Yukawa interactions:

$$-x_{mn} \bar{Q}_m P_R L_n \tilde{U} - y_{mn} \bar{D}_m P_L E_n \tilde{U} - z_{mn} \epsilon_{rst} \tilde{U}^r \bar{D}_m^s P_R D_n^t + \text{h.c.}$$

in which r, s, t are color indices, ϵ_{rst} is the totally antisymmetric tensor on color indices, color indices are implicitly summed in the other two terms, and m, n are generation indices. Note that z_{mn} must be odd in m, n .

Argue that there are no other renormalizable interactions which are gauge invariant and satisfy all of the basic principles.

[2.7.3] What is the mass squared of \tilde{U} , including both the explicit effects of its mass term and the effects of v the *v.e.v.* of the Higgs boson? Is the mass of \tilde{U} determined by its coupling to the Higgs boson, or is it an independent free parameter of the model?

[2.7.4] Argue that there is *no* assignment of lepton or baryon number to the \tilde{U} field which leaves either B or L symmetry unbroken. Hence, the addition of such a scalar field generically leads to the violation of B and L symmetries.

Show, however, that if the Lagrangian is required to be invariant under a discrete symmetry, $\tilde{U} \rightarrow -\tilde{U}$ with all other fields unaffected, then none of the Yukawa couplings are permitted and conserved baryon and lepton numbers can again be defined. Further, show that in this case there is a new global $U(1)$ symmetry $\tilde{U} \rightarrow e^{i\theta} \tilde{U}$ which ensures that the number of \tilde{U} particles is conserved.

[2.8] **Adjoint representation fermions**

Suppose that two Majorana fermions were added to the standard model; \tilde{W} , a triplet under $SU_L(2)$, transforming as $(\mathbf{1}, \mathbf{3}, 0)$, and \tilde{G} , an octet under $SU_c(3)$, transforming as $(\mathbf{8}, \mathbf{1}, 0)$. That is, the transformation properties are,

$$\delta P_L \tilde{W}^a = -\epsilon_{abc} \omega_2^b P_L \tilde{W}^c, \quad D_\mu P_L \tilde{W}^a = \left(\partial_\mu \delta_{ac} + g_2 \epsilon_{abc} W_\mu^b \right) P_L \tilde{W}^c,$$

and

$$\delta P_L \tilde{G}^\alpha = -f_{\alpha\beta\gamma} \omega_3^\beta P_L \tilde{G}^\gamma, \quad D_\mu P_L \tilde{G}^\alpha = \left(\partial_\mu \delta_{\alpha\gamma} + g_3 f_{\alpha\beta\gamma} G_\mu^\beta \right) P_L \tilde{G}^\gamma$$

[2.8.1] Show that the reality of ϵ_{abc} and $f_{\alpha\beta\gamma}$ cause $P_R \tilde{W}$ and $P_R \tilde{G}$ to have the same transformation properties as $P_L \tilde{W}$ and $P_L \tilde{G}$.

[2.8.2] Show that, contrary to what happened with the fermions of the standard model, the new fields \tilde{W} and \tilde{G} *do* have $SU_c(3) \times SU_L(2) \times U_Y(1)$ invariant mass terms,

$$-\frac{m_{\tilde{W}}}{2} \overline{\tilde{W}} \tilde{W} - \frac{m_{\tilde{G}}}{2} \overline{\tilde{G}} \tilde{G}$$

Therefore, these particles may possess masses independent of their coupling to the Higgs boson.

[2.8.3] Show that the only new Yukawa interaction is

$$y_m \bar{L}_m \tau_a \tilde{\phi} P_L \tilde{W}_a + \text{h.c.}$$

3

Cross sections and lifetimes

Most of the applications of the standard model to experimental situations are concerned with processes in which almost free particles interact briefly and over short distances. These processes could be the collisions of various elementary particles within an accelerator (Chapter 6 and Chapter 9) or they could be the decay of an unstable elementary particle in flight (Chapter 4 and Chapter 5). Scattering (S -matrix) theory is the formalism that has been devised to study these systems.

This chapter presents a whirlwind review of the quantum theory of scattering. The purpose is to gather into one place all of the results that are required in order to use the Lagrangian of Chapter 2 to predict the outcomes of experiments. The first section sets up the notion of scattering states, which are meant to represent in a precise way the idea that the particles involved do not interact except for a short time interval. This is followed by a review of the calculation of scattering amplitudes using time-dependent perturbation theory.

In later chapters this formalism is finally used to compute the Feynman rules that describe the interactions contained within the standard model Lagrangian.

Readers in a hurry, or who find themselves bogged down in this section, should try to understand Section 3.2 and will need to learn the results at the end of Section 3.3, particularly Eq. (3.40) and Eq. (3.43).

3.1 Scattering states and the S -matrix

In a real scattering (or decay) process, the particles involved only interact briefly because they physically move apart from one another. For instance, in a scattering experiment, the initial particles are initially well separated from one another, but moving with velocities which bring them into mutual

contact. From the perspective of quantum mechanics, this means that these initial states cannot be exact momentum eigenstates, since such states are not spatially localized at all. Similarly, they cannot be exact energy eigenstates to the extent that their profiles in position space change with time (as opposed to simply being multiplied by an overall phase e^{-iEt}). Instead, the initial particles are usually given by wave packets which are somewhat localized in both position and momentum (in a way which is consistent with the uncertainty relations), with the packets describing the relative approach of initially well-separated particles.

To the extent that the initially colliding particles are not correlated with one another and that the reactions do not depend on the environment within which they occur, one expects the probability of any given reaction to factorize into the product of the probability for the particles to meet, times the probability for the reaction to occur given that the meeting has taken place. Of these, the first factor can be expected to depend on the details of the wave packets which describe the initial state, since this controls things like how many particles are present and how quickly they approach one another. The second factor, however, might be expected to be independent of the details of the initial state and instead be more of an intrinsic property of the interactions involved. Indeed, these expectations are borne out in practice for collisions, and motivate the definition of initial-state-independent quantities, like *cross sections*, which describe the part of the reaction which does not depend on the details of how a particular reaction has been set up.

It is the inference of quantities like cross sections from experimental measurements which is of practical interest, since these directly bear the information about the underlying interactions like those described in earlier chapters. Because they are largely insensitive to the details of the wave packets describing the initial states, it turns out to be possible to compute quantities like cross sections directly in the limit that these initial states become energy and momentum eigenstates, even though this is not the limit within which real experiments take place. The idealized energy eigenstates to which one is led in this way are called *scattering states*, and their definition is the topic of this section.

Suppose, then, that the complete Hamiltonian, H , can be broken into two pieces, $H = H_0 + V$, in such a way that H_0 describes the evolution of the initial and final wave packets before and after the scattering. In the simplest instance H_0 might describe just the kinetic energy of moving free particles, with all of the interactions being put into V . But more complicated divisions of H are also possible, such as by including the strong and/or electromagnetic interactions in H_0 while placing the weak interactions into V .

In general the Hilbert space, \mathcal{H} , for the full system divides into two parts,

$$\mathcal{H} = \mathcal{B} \oplus \mathcal{S} \quad (3.1)$$

for which \mathcal{S} contains those states of the full system whose evolution in time using H is well approximated at late or early times by evolution using H_0 . That is, \mathcal{S} are the states (particles) of the theory with Hamiltonian H_0 . It is useful to define the origin of time so that the initial and final wave packets of the interacting particles are sufficiently widely separated that H_0 evolution suffices outside of a region $-T < t < T$, for some appropriately large and positive T . Not all states need reside in \mathcal{S} , and those which do not live in \mathcal{B} , which we loosely call bound states. For example, if our system consisted of electrons and protons interacting electromagnetically, then \mathcal{S} might contain freely-moving electrons and protons, but \mathcal{B} might contain bound hydrogen atoms.

Let us denote the eigenstates of H_0 by $|\alpha\rangle$, with α collectively denoting all of the labels which are required to describe single- and many-particle states and $H_0|\alpha\rangle = E_\alpha|\alpha\rangle$. We write a wave packet of such states as

$$|\phi_g\rangle \equiv \int d\alpha g(\alpha)|\alpha\rangle \quad (3.2)$$

where $g(\alpha)$ defines an appropriately normalizable packet. The label α here is treated as a continuous variable because we envisage it to include (possibly among other labels) the momenta of the various particles included in the state. We assume that H_0 has the same spectrum on \mathcal{H} as H does on \mathcal{S} , so the same labels, α , and energies, E_α , may be used to describe the eigenstates of the full system, $H|\alpha\rangle\rangle = E_\alpha|\alpha\rangle\rangle$ (where the double angle $\rangle\rangle$ is used to denote an eigenstate of H).

To describe scattering processes we work within the Schrödinger picture, where the burden of time evolution is carried by the state of the system. In a scattering problem we imagine that the time evolution of states prepared in appropriate wave packets, $|\phi_g\rangle$, have essentially the same evolution in the remote past and the remote future, $|t| \gg T$, using either H or H_0 . That is, we require that there must exist an *out* state, $|\phi_g\rangle\rangle_o$, which at late times evolves under H in the same way as does any properly normalizable packet $|\phi_g\rangle$ under H_0 :

$$\lim_{t \gg T} e^{-iHt} |\phi_g\rangle\rangle_o = \lim_{t \gg T} e^{-iH_0 t} |\phi_g\rangle \quad (3.3)$$

There must similarly exist an *in* state, $|\phi_g\rangle\rangle_i$, – in general different than $|\phi_g\rangle\rangle_o$ – whose evolution under H agrees with the evolution of a packet $|\phi_g\rangle$

under H_0 in the remote past:

$$\lim_{t \ll -T} e^{-iHt} |\phi_g\rangle_i = \lim_{t \ll -T} e^{-iH_0 t} |\phi_g\rangle \quad (3.4)$$

By choosing the limiting case of appropriately peaked wave packets, $g(\alpha)$, we may also formally define in this way idealized scattering eigenstates of the full Hamiltonian, $|\alpha\rangle_{o,i}$, which satisfy

$$\lim_{t \gg T} e^{-iHt} |\alpha\rangle_o = \lim_{t \gg T} e^{-iH_0 t} |\alpha\rangle \quad \text{and} \quad \lim_{t \ll -T} e^{-iHt} |\alpha\rangle_i = \lim_{t \ll -T} e^{-iH_0 t} |\alpha\rangle \quad (3.5)$$

In terms of these states a scattering event corresponds to the transition from a state resembling a packet $|\phi_g\rangle$ at asymptotically early times to one resembling a different packet $|\phi_f\rangle$ at asymptotically late times. From the above definitions the amplitude for a such a process is given by the overlap

$${}_o\langle\langle\phi_f|\phi_g\rangle\rangle_i \quad (3.6)$$

Any such scattering event may therefore be found from the limiting amplitude for the ideal process where the initial and final state are approximately energy eigenstates, and the matrix of all possible such amplitudes,

$$S_{\beta\alpha} := {}_o\langle\langle\beta|\alpha\rangle\rangle_i \quad (3.7)$$

therefore plays an important role, and is called the S -matrix. It is also convenient to define the operator, S , whose matrix elements between H_0 eigenstates, $|\alpha\rangle$, reproduce these transition amplitudes:

$$\langle\beta|S|\alpha\rangle := S_{\beta\alpha} \quad (3.8)$$

Our goal is to provide an explicit expression for S in terms of the known operators H_0 and V . A step towards this end is the definition of the Møller wave operators

$$\Omega(t) := e^{iHt} e^{-iH_0 t} \quad (3.9)$$

in terms of which we have

$$|\alpha\rangle_o = \lim_{t \gg T} \Omega(t) |\alpha\rangle \quad \text{and} \quad |\alpha\rangle_i = \lim_{t \ll -T} \Omega(t) |\alpha\rangle \quad (3.10)$$

Since $|\alpha\rangle$ and $|\alpha\rangle_{o,i}$ are normalized, $\Omega^\pm = \lim_{t \rightarrow \pm\infty} \Omega(t)$ are isometric operators. Notice, however, that the states $|\alpha\rangle_{o,i}$ only span \mathcal{S} , while $|\alpha\rangle$ span \mathcal{H} , so Ω^\pm can only be unitary if $\mathcal{B} = \emptyset$ (i.e. there are no bound states).

These operators are useful because the S -matrix can be constructed from them using

$$S = \lim_{t \rightarrow \infty} \lim_{t' \rightarrow -\infty} \Omega^*(t) \Omega(t') = (\Omega^+)^* \Omega^- \quad (3.11)$$

The limit $t \rightarrow \mp\infty$ must of course be defined with some care, using appropriately normalized wave packets. This complication is ignored here with the understanding that a more careful treatment justifies the formal manipulations we present.

3.2 Time-dependent perturbation theory

We now derive an approximate expression for S as powers of the interaction V . In order to express S in a form that lends itself to such a perturbative approximation, we rewrite the operator $\Omega^*(t)\Omega(t')$ by re-expressing it as a solution to a first-order differential equation in the variable t . That is, $\Omega^*(t)\Omega(t')$ satisfies

$$\begin{aligned}\Omega^*(t)\Omega(t') &= e^{iH_0t} e^{-iHt} e^{iHt'} e^{-iH_0t'} \\ &= e^{iH_0t} e^{-iH(t-t')} e^{-iH_0t'}\end{aligned}\quad (3.12)$$

Evidently,

$$\begin{aligned}i \frac{d}{dt} [\Omega^*(t)\Omega(t')] &= e^{iH_0t} (H - H_0) e^{-iH(t-t')} e^{-iH_0t'} \\ &= (e^{iH_0t} V e^{-iH_0t}) \Omega^*(t)\Omega(t') \\ &= V(t) \Omega^*(t)\Omega(t')\end{aligned}\quad (3.13)$$

where this last equality defines the interaction picture V operator at time t , $V(t) := e^{iH_0t} V e^{-iH_0t}$.

Solutions of this differential equation, together with the initial condition $\Omega^*(t')\Omega(t') = 1$, are equivalent to solutions of the integral equation

$$\Omega^*(t)\Omega(t') = 1 - i \int_{t'}^t d\tau V(\tau) \Omega^*(\tau)\Omega(t') \quad (3.14)$$

This has the obvious iterative solution

$$\Omega^*(t)\Omega(t') = \sum_{n=0}^{\infty} (-i)^n \int_{t'}^t d\tau_1 \int_{t'}^{\tau_1} d\tau_2 \cdots \int_{t'}^{\tau_{n-1}} d\tau_n V(\tau_1) V(\tau_2) \cdots V(\tau_n) \quad (3.15)$$

The S -matrix becomes

$$\begin{aligned}S &= \lim_{\substack{t \rightarrow +\infty \\ t' \rightarrow -\infty}} \Omega^*(t)\Omega(t') \\ &= \sum_{n=0}^{\infty} (-i)^n \int_{-\infty}^{\infty} d\tau_1 \int_{-\infty}^{\tau_1} d\tau_2 \cdots \int_{-\infty}^{\tau_{n-1}} d\tau_n V(\tau_1) V(\tau_2) \cdots V(\tau_n)\end{aligned}\quad (3.16)$$

One of our goals is to make Lorentz invariance as manifest as possible, so to this end it is desirable to rewrite this expression in a form where the

temporal integration is over the same range as any spatial integrations, i.e. from $-\infty$ to ∞ . This can be done via the following trick. Define the *time-ordering* operation by

$$\begin{aligned} T[V(t_1) \cdots V(t_n)] &\equiv V(t_{\text{latest}}) \cdots V(t_{\text{earliest}}) \\ &= \sum_{P_n} V(t_{P_1}) \cdots V(t_{P_n}) \theta(t_{P_1} - t_{P_2}) \cdots \theta(t_{P_{n-1}} - t_{P_n}) \end{aligned} \quad (3.17)$$

The sum here is over all permutations of the n times t_1, \dots, t_n , and the Heaviside step function,

$$\theta(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases} \quad (3.18)$$

ensures that only the permutation in which $t_{P_1} > t_{P_2} > \cdots > t_{P_n}$ contributes. Consider, then, the integral,

$$\begin{aligned} I &\equiv \int_{-\infty}^{\infty} d\tau_1 \cdots \int_{-\infty}^{\infty} d\tau_n T[V(\tau_1) \cdots V(\tau_n)] \\ &= \sum_{P_n} \int_{-\infty}^{\infty} d\tau_{P_1} \cdots \int_{-\infty}^{\infty} d\tau_{P_n} V(\tau_{P_1}) \cdots V(\tau_{P_n}) \theta(\tau_{P_1} - \tau_{P_2}) \cdots \theta(\tau_{P_{n-1}} - \tau_{P_n}) \\ &= n! \int_{-\infty}^{\infty} d\tau_1 \cdots \int_{-\infty}^{\tau_{n-1}} d\tau_n V(\tau_1) \cdots V(\tau_n) \end{aligned} \quad (3.19)$$

Comparing the last line with the iterative expression for S , given above, implies that

$$S = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} d\tau_1 \cdots d\tau_n T[V(\tau_1) \cdots V(\tau_n)] \quad (3.20)$$

This will be the final form for the perturbative expansion of the S -matrix in time-dependent perturbation theory.

Equation (3.20) has a particularly pretty form if the interaction Hamiltonian is given as an integral over a local Hamiltonian density,

$$V(t) = \int d^3\mathbf{x} \mathcal{H}_I(\mathbf{x}, t) \quad (3.21)$$

since in this case the S -matrix becomes

$$S = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} d^4x_1 \cdots d^4x_n T[\mathcal{H}_I(x_1) \cdots \mathcal{H}_I(x_n)] \quad (3.22)$$

This last equation is one of the main results of this chapter.

If we use energy and momentum eigenstates it is convenient to use the identity:

$$\langle \beta | \mathcal{O}(x) | \alpha \rangle = \langle \beta | e^{-iP \cdot x} \mathcal{O}(x=0) e^{iP \cdot x} | \alpha \rangle = e^{i(p_\alpha - p_\beta) \cdot x} \langle \beta | \mathcal{O}(x=0) | \alpha \rangle \quad (3.23)$$

to factor an overall energy-momentum conserving factor out of the S -matrix:

$$S_{\beta\alpha} = \delta_{\beta\alpha} - i\mathcal{M}_{\beta\alpha}(2\pi)^4\delta^4(p_\beta - p_\alpha) \quad (3.24)$$

The quantity $\mathcal{M}_{\beta\alpha}$ is called the *matrix element* for the transition from state α to state β . It is also conventional to define the *T-matrix element*, in which only the energy conserving delta function is factored out:

$$S_{\beta\alpha} = \delta_{\beta\alpha} - iT_{\beta\alpha}2\pi\delta(p_\beta^0 - p_\alpha^0) \quad (3.25)$$

We can read off the first few terms in the expansion of \mathcal{M} directly from Eq. (3.22):

$$\mathcal{M}_{\beta\alpha} = \langle\beta|\mathcal{H}_I(x=0)|\alpha\rangle + \frac{(-i)^2}{2!} \int d^4x \langle\beta|T[\mathcal{H}_I(x)\mathcal{H}_I(x=0)]|\alpha\rangle + \cdots \quad (3.26)$$

This is an important result because it gives the S -matrix in terms of quantities that we know, namely the matrix elements of the interaction Hamiltonian density.

Equation (3.22) or Eq. (3.26) do not quite appear Lorentz-invariant, for two reasons. One reason is the appearance of the time-ordering operation, which leads to the functions $\theta(t_i - t_j)$ whose values may differ in different frames. (Recall that different Lorentz observers can disagree on the ordering in time of spacelike separated events.) This turns out not to be important because the operator ordering is only relevant for operators which do not commute. Locality ensures that commutators vanish for spacelike separated points; it is only for timelike or lightlike separated operators that the time ordering operation is important, and for such operators the time ordering is the same in all frames. Therefore, the time-ordering operation on a product of local operators is Lorentz invariant in a local theory, and this is not an obstacle to the Lorentz invariance of Eq. (3.22).

The other reason to doubt the Lorentz invariance of the S -matrix is because the integral of the Hamiltonian density need not be Lorentz-invariant. Note, however, that it is only the interaction part of the Hamiltonian density which appears in the above formulae, and to the extent that this does not involve derivatives of the fields it is typically related to the interaction part of the Lagrangian density by $\mathcal{H} = -\mathcal{L}$. When this is so we see that the Lorentz invariance of the S -matrix and of \mathcal{M} is manifest, since we know that $\int d^4x \mathcal{L}$ is Lorentz invariant by construction. As we see in later chapters Lorentz invariance also holds for interactions involving derivatives of fields, although this invariance arises in a more subtle way.

We shall use these equations – Eq. (3.23), Eq. (3.24), and Eq. (3.26) – extensively throughout what follows.

3.3 Decay rates and cross sections

The expressions obtained above for the S -matrix are proportional to an energy-conserving (and possibly to a momentum-conserving) delta function when expressed in terms of energy eigenstates rather than wave packets. This means that the square of S -matrix elements – the transition probabilities – are proportional to $\delta(0)$ and so must diverge. Physically, this divergence reflects the fact discussed earlier that scattering processes necessarily involve wave packets and *cannot* involve energy eigenstates. (It is also related to the difficulty, in infinite volume, of correctly normalizing an energy eigenstate.) If the initial and final states are energy and momentum eigenstates then their interactions never really turn on and off, because their wave functions spread throughout all of space, which prevents their influence on one another from changing over time. As a result, if we insist on using such eigenstates to compute the S -matrix (as we shall for convenience of calculation), we must more carefully sort out the relationship between physical quantities and the S -matrix elements we find. This is the purpose of the present section.

3.3.1 Wave packets

If the initial state is described by a wave packet, $|\phi_g\rangle_i = \int d\alpha g(\alpha)|\alpha\rangle_i$, then the probability of finding the system in the final state labeled by β becomes

$$P_g(\beta) = |{}_o\langle\langle\beta|\phi_g\rangle_i|^2 = \int d\alpha d\alpha' g^*(\alpha')g(\alpha) {}_o\langle\langle\beta|\alpha\rangle_i {}_i\langle\langle\alpha'|\beta\rangle_o \quad (3.27)$$

In most cases of practical interest, the initial state is prepared in such a way that the function $g(\alpha)$ is peaked about some value $\bar{\alpha}$, and the width of the wave packet is classical in the sense that the resolution of initial position and momentum measurements are much too large to push the limits of the uncertainty relations. It is also usually true that support of the initial wave packet is chosen to be over a region of α , over which $S_{\beta\alpha}$ depends only weakly on α . For instance, the energy width of a wave packet is usually small compared to the energy dependence of the scattering cross section or particle decay width. (Otherwise the experiment does a poor job in measuring the S -matrix, because it uses an inadequately resolved initial state.)

Under these circumstances (and assuming β is distinguishable from all of the α in the support of $g(\alpha)$, so we may write $S_{\beta\alpha} = -iT_{\beta\alpha}2\pi\delta(E_\beta - E_\alpha)$), then Eq. (3.27) is approximately given by

$$P_g(\beta) \approx |T_{\beta\bar{\alpha}}|^2 \int d\hat{\alpha} d\hat{\alpha}' g^*(\alpha')g(\alpha) \quad (3.28)$$

In this expression $d\alpha 2\pi\delta(E_\alpha - E_\beta) = d\hat{\alpha}$, and we use the fact that $T_{\beta\alpha}$ is approximately independent of α within the domain of support of $g(\alpha)$ to bring it outside of the integral. Notice that the energy-conserving delta functions are no longer a problem since they are used to perform part of the integration over α and α' .

We see that the probability in this case factorizes into a reaction dependent factor ($|T_{\beta\alpha}|^2$) and a factor depending on the details of the experimental set-up. Our interest in the remainder of this section is in precisely identifying a convenient quantity which captures the initial-condition-independent factor.

3.3.2 The finite-volume trick

For the present purposes the important consequence of the previous section is Eq. (3.28), which expresses how reaction probabilities factorize in the situations of common practical interest. Since our interest is in finding a convenient way to identify the $|T_{\beta\alpha}|^2$ factor in a calculation of $S_{\beta\alpha}$ based on energy and momentum eigenstates, we may feel free to use any old specification of the initial state, provided it captures this factorization (involves narrow ranges of energy and momentum). Obviously we should choose one which makes the calculations convenient.

A particularly simple way of specifying states, and seeing how to handle the subtleties associated with the delta functions in $S_{\beta\alpha}$, is to imagine the system being inside a box having large but finite volume Ω , and allowing the interactions to last only over a large but finite time interval, T . In this case we may simply use energy and momentum eigenstates, with the knowledge that the divergences associated with squaring delta functions are regularized by T and Ω . Once the regularization dependence cancels in the final physical quantities of interest, we may drop the temporary theoretical contrivance of the box.

In a finite-volume box we use particles in momentum states, $|\mathbf{p}\rangle$, that are normalized to 1 in the box,

$$[\mathbf{p}|\mathbf{p}'] = \delta_{\mathbf{p},\mathbf{p}'} \quad (3.29)$$

which satisfy the completeness relation

$$\sum_{\mathbf{p}} |\mathbf{p}\rangle[\mathbf{p}| = 1 \quad (3.30)$$

This is to be distinguished from the continuum normalization we use in the infinite-volume limit,

$$\langle \mathbf{p}|\mathbf{p}' \rangle = 2E_{\mathbf{p}}(2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}') \quad (3.31)$$

for which completeness is expressed by

$$\int \frac{d^3\mathbf{p}}{2E_{\mathbf{p}}(2\pi)^3} |\mathbf{p}\rangle \langle \mathbf{p}| = 1 \quad (3.32)$$

For a cubic box of volume Ω , subject to periodic boundary conditions on the walls, momentum eigenvalues take discrete values. There is one state for each cube of volume $(2\pi)^3/\Omega$ in momentum space. In the limit $\Omega \rightarrow \infty$, the spacing between momentum levels goes to zero and sums over momenta go to integrals according to

$$\frac{1}{\Omega} \sum_{\mathbf{p}} f(\mathbf{p}) \rightarrow \int \frac{d^3\mathbf{p}}{(2\pi)^3} f(\mathbf{p}) \quad (3.33)$$

Here $f(\mathbf{p})$ represents an arbitrary function that satisfies the boundary conditions at the edge of the box. Comparison with the completeness relations shows that the states $|\mathbf{p}\rangle = (2E\Omega)^{1/2} |\mathbf{p}]$ are the ones which have the desired normalization for large Ω .

For a state, $|\alpha\rangle$, involving N_α particles this implies $|\alpha\rangle = (2E\Omega)^{N_\alpha/2} |\alpha]$. The box-normalized matrix element $S_{\beta\alpha}^\square \equiv [\beta|S|\alpha]$ is therefore related to the continuum-normalized $S_{\beta\alpha} = \langle\beta|S|\alpha\rangle$ by

$$S_{\beta\alpha} = (2E\Omega)^{(N_\alpha+N_\beta)/2} S_{\beta\alpha}^\square \quad (3.34)$$

When particle energies differ, $(2E)^{(N_\alpha+N_\beta)/2}$ is to be interpreted as the square root of the product of the energies of the particles in the in and out states.

At finite volume, in translationally invariant theories, the T -matrix is

$$T_{\beta\alpha} \equiv \mathcal{M}_{\beta\alpha} (2\pi)^3 \delta_\Omega^3(\mathbf{p}_\beta - \mathbf{p}_\alpha) \quad (3.35)$$

so the S -matrix is given by

$$S_{\beta\alpha} = \delta(\beta - \alpha) - i(2\pi)^4 \delta_{\Omega T}^4(p_\beta - p_\alpha) \mathcal{M}_{\beta\alpha} \quad (3.36)$$

The delta functions express energy and momentum conservation and appear in the form,

$$(2\pi)^3 \delta_\Omega^3(\mathbf{p}_\alpha - \mathbf{p}_\beta) = \frac{1}{\Omega} \int_\Omega d^3x e^{i(\mathbf{p}_\alpha - \mathbf{p}_\beta) \cdot \mathbf{x}} \quad (3.37)$$

$$(2\pi)^4 \delta_{\Omega T}^4(p_\alpha - p_\beta) = \frac{1}{\Omega T} \int_{\Omega T} d^4x e^{i(p_\alpha - p_\beta) \cdot x} \quad (3.38)$$

The spatial integration is over the volume, Ω , and the temporal integration is from $-T/2$ to $+T/2$ respectively. As $\Omega T \rightarrow \infty$, $\delta_{\Omega T}$ goes to the standard delta-function but for finite T and Ω , $(2\pi)^4 \delta_{\Omega T}(0) = \Omega T$.

In a time-translationally invariant theory it is the transition probability per unit time, or the transition rate, which is independent of time and so is well behaved as $T \rightarrow \infty$. Similarly, as $\Omega \rightarrow \infty$ the number of states in any finite momentum range diverges, making the probability of a transition to a specific state go to zero. It is therefore the rate, $d\Gamma$, for the state $|\alpha\rangle$ to make a transition into any state in a small number, $\Delta\beta$, of states in the vicinity of $|\beta\rangle$ that is well behaved as $\Omega T \rightarrow \infty$. Since the density of states in momentum space is $\Omega/(2\pi)^3$, the number of states in an interval $d\beta$ for an N_β -particle state is $\Delta\beta = (2E\Omega)^{N_\beta} d\beta$. Here we have absorbed the powers of 2π into the measure on $d\beta$, so that $d\beta \equiv \prod d^3k/[(2\pi)^3 2E_{\mathbf{k}}]$. With this notational convention, the rate becomes

$$\begin{aligned}
 d\Gamma(\alpha \rightarrow \beta) &= \frac{dP(\alpha \rightarrow \beta)}{T} \\
 &= \frac{|S_{\beta\alpha}^\square|^2}{T} \Delta\beta \\
 &= \left[\frac{|S_{\beta\alpha}|^2}{T} \left(\frac{1}{2E\Omega} \right)^{(N_\alpha + N_\beta)} \right] \Delta\beta \\
 &= \frac{1}{T} (2\pi)^4 \delta_{\Omega T}^4(p_\beta - p_\alpha) (2\pi)^4 \delta_{\Omega T}^4(0) |\mathcal{M}_{\beta\alpha}|^2 \left(\frac{1}{2E\Omega} \right)^{N_\alpha + N_\beta} \Delta\beta \\
 &= \Omega (2\pi)^4 \delta^4(p_\alpha - p_\beta) \frac{1}{(2E\Omega)^{N_\alpha}} |\mathcal{M}_{\beta\alpha}^2| d\beta \\
 &= \Omega^{1-N_\alpha} \left[\prod_{i \in \alpha} \frac{1}{2E_i} \right] |\mathcal{M}_{\beta\alpha}|^2 (2\pi)^4 \delta_{\Omega T}^4(P_\alpha - P_\beta) d\beta \quad (3.39)
 \end{aligned}$$

where the product means a product over the particles in the initial state. Notice that the δ -function ensures that the final integral over β runs over a *finite* range of integration and so can never diverge unless $\mathcal{M}_{\beta\alpha}$ is singular for some momenta.

Consider now the cases of most present interest, with $N_\alpha = 1$, $N_\alpha = 2$, and $N_\alpha > 2$.

3.3.2.1 Decay processes: $N_\alpha = 1$

In the limit $\Omega \rightarrow \infty$ and $T \rightarrow \infty$ the decay rate for a single particle is explicitly independent of Ω and T , and is given by

$$\begin{aligned}
 d\Gamma(\alpha \rightarrow \beta) &= \frac{1}{2E_\alpha} |\mathcal{M}_{\beta\alpha}|^2 (2\pi)^4 \delta^4(P_\alpha - P_\beta) d\beta, \\
 d\beta &\equiv \prod_{f \in \beta} \frac{d^3\mathbf{k}_f}{2E_{\mathbf{k}_f} (2\pi)^3} \quad (3.40)
 \end{aligned}$$

This result is not quite Lorentz-invariant, because of the $1/(2E_\alpha)$ in front. But indeed, it should not be Lorentz-invariant, since a fast-moving particle's lifetime should be extended by time dilation; the $1/(2E_\alpha)$ factor precisely generates this time dilation effect.

3.3.2.2 Two-body scattering: $N_\alpha = 2$

When $N_\alpha = 2$, $d\Gamma$ is proportional to Ω^{-1} . Since the single-particle states are normalized with $\int_\Omega d^3x |\psi(x)|^2 = 1$, the number density of particles in the box as seen by an incident particle is $n = \Omega^{-1}$. The fact that $d\Gamma$ is inversely proportional to the volume reflects the property that in the absence of initial-state coherence the reaction rate is proportional to the number density of target particles.

It is convenient and conventional to remove this dependence on the number of particles by dividing out a factor proportional to the incident flux of particles. Define, then, the cross section, $d\sigma$, by

$$d\sigma(\alpha \rightarrow \beta) = \frac{d\Gamma}{F}(\alpha \rightarrow \beta) \quad (3.41)$$

In this expression the denominator, F , is fixed by requiring that (a) $d\sigma$ be Lorentz invariant; and (b) F , when evaluated in the rest-frame of either of the particles, equals the particle flux: $nv_{\text{rel}} = v_{\text{rel}}/\Omega$.

Our next task is to find the function, F , determined by these conditions. Condition (a) implies that F must transform the same way as $d\Gamma$ does under Lorentz transformations. Because of our choice of state normalization and integration measure $d\beta$, the final-state factors are already Lorentz-invariant. Invariance of the cross section is therefore ensured if $F = f/(4E_1E_2\Omega)$, where E_k denotes the energy of the particles in the initial two-particle state, $|\alpha\rangle$, and f is a Lorentz-invariant function chosen to satisfy condition (b). Since the relative velocity of two particles,

$$v_{\text{rel}} = \sqrt{1 - \frac{m_1^2 m_2^2}{(p_1 \cdot p_2)^2}} \quad (3.42)$$

is Lorentz-invariant and the scalar $-p_1 \cdot p_2$ equals $E_1 E_2$ in the particle rest frame the solution is $f = -4V_{\text{rel}}(p_1 \cdot p_2)$.

We are led in this way to the following expression for the two-body cross section:

$$d\sigma(\alpha \rightarrow \beta) = \frac{|\mathcal{M}_{\beta\alpha}|^2}{f} (2\pi)^4 \delta^4(p_\alpha - p_\beta) d\beta \quad (3.43)$$

$$\text{with } f = (-4p_1 \cdot p_2)v_{\text{rel}} = 4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2} \quad (3.44)$$

To be completely explicit, and for later convenience, we pause here to calculate the factor $(2\pi)^4 \delta^4(p_\alpha - p_\beta) d\beta$ for a two-body final state, $N_\beta = 2$, in the center-of-mass frame of the two bodies. Denote the final-state quantum numbers by primes. In an arbitrary frame $\delta^4(p_\alpha - p_\beta) d\beta$ is

$$\begin{aligned}
 (2\pi)^4 \delta^4(p_\alpha - p_\beta) d\beta &= (2\pi)^4 \delta^4(p_\alpha - p'_1 - p'_2) \frac{d^3 \mathbf{p}'_1 d^3 \mathbf{p}'_2}{(2\pi)^6 4E'_1 E'_2} \\
 &= 2\pi \delta(E_\alpha - E'_1 - E'_2) \frac{d^3 \mathbf{p}'_1}{(2\pi)^3 4E'_1 E'_2} \Big|_{\mathbf{p}'_2 = \mathbf{p}_\alpha - \mathbf{p}'_1} \\
 &= \frac{p'^2_1 d^2 \Omega'_1}{(2\pi)^2 4E'_1 E'_2 |d(E'_1 + E'_2)/dp'_1|} \\
 &= \frac{p'^3_1 d^2 \Omega'_1}{16\pi^2 (E'_2 \mathbf{p}'_1 - E'_1 \mathbf{p}'_2) \cdot \mathbf{p}'_1} \quad (3.45)
 \end{aligned}$$

$d^2 \Omega'_1 = \sin \theta' d\theta' d\phi'$ is the element of solid angle where θ' and ϕ' give the direction of the vector \mathbf{p}'_1 . In the center-of-mass frame, $\mathbf{p}'_1 = -\mathbf{p}'_2$ and $E'_1 + E'_2 = E_\alpha$, so

$$(2\pi)^4 \delta^4(p_\alpha - p_\beta) d\beta = \frac{p'_1 d^2 \Omega'_1}{16\pi^2 E_\alpha} \quad (\text{c.m.}) \quad (3.46)$$

In this case, the final-state integral consists of the sum over the direction of one of the two final-state particles.

3.3.2.3 Many-body collisions: $N_\alpha > 2$

The reaction rate per unit volume, $d\Gamma/\Omega$, is proportional to Ω^{-N_α} . For N_α distinct particles in the initial state this again represents the incident-particle density that is expected for incoherent scattering:

$$\Omega^{-N_\alpha} = \prod_{i=1}^{N_\alpha} n_i \quad (3.47)$$

In this case the reaction rate per unit volume becomes

$$\frac{d\Gamma(\alpha \rightarrow \beta)}{\Omega} = \prod_{i=1}^{N_\alpha} \left[\frac{n_i}{2E_i} \right] |\mathcal{M}_{\beta\alpha}|^2 (2\pi)^4 \delta^4(p_\alpha - p_\beta) d\beta \quad (3.48)$$

Part II

Applications: leptons

4

Elementary boson decays

We wish to put the formalism of the previous chapters to use to describe the properties of the standard-model particles. Since many of the properties of the theory are simpler at higher energies we choose to do this by starting with the properties of the heavy bosons of the theory and then working our way down in energy towards more familiar particles. We also choose to focus here on the properties of the elementary bosons since these furnish among the simplest examples of the scattering formalism of the previous chapter.

Among the most basic particle properties are their masses and lifetimes. The masses of the gauge bosons of the theory have been dealt with in previous (and in subsequent) chapters, so we concentrate here on their lifetimes.

4.1 Z^0 decay

4.1.1 Z^0 decay: *preliminaries*

We wish to compute within the standard-model the decay lifetime of the neutral electroweak gauge boson, Z^0 , as a function of the parameters of the model. We do so using the perturbative framework of Chapter 3. The basic result of that chapter, for the present purposes, is given by Eq. (3.24) and Eq. (3.26),

$$\begin{aligned} S_{\beta\alpha} &= \delta_{\beta\alpha} - i(2\pi)^4 \delta^4(p_\beta - p_\alpha) \mathcal{M}_{\beta\alpha}, \quad \text{with} \\ \mathcal{M}_{\beta\alpha} &= \langle \beta | \mathcal{H}_I(0) | \alpha \rangle + \frac{(-i)^2}{2!} \int d^4x \langle \beta | T [\mathcal{H}_I(x) \mathcal{H}_I(0)] | \alpha \rangle + \cdots \end{aligned} \quad (4.1)$$

We see that, in the absence of other effects, the dominant contribution to Z^0 decay will come from any interactions of the model for which the matrix element

$$\langle \beta | \mathcal{H}_I(0) | Z^0 \rangle \neq 0 \quad (4.2)$$

for some final state $|\beta\rangle$ into which the Z^0 may kinematically decay. If there is no such final state or interaction then the dominant contribution must instead be second order, i.e.,

$$\frac{(-i)^2}{2!} \int d^4x \langle \beta | T[\mathcal{H}_I(x) \mathcal{H}_I(x=0)] | Z^0 \rangle \neq 0 \quad (4.3)$$

We must continue in this way until a nonzero result is eventually obtained.

If the Z^0 boson is to decay, it cannot appear in the final state. It follows that, in order to contribute to the matrix element of Eq. (4.2), any candidate interaction must be strictly linear in the field $Z_\mu(x)$. Inspection of the Z^0 couplings of Section 2.4 shows that there are only a few candidate interactions of this type. The candidates are \mathcal{L}_{WWZ} of Eq. (2.77), $\mathcal{L}_{WWZ\gamma}$ of Eq. (2.83), and \mathcal{L}_{nc} of Eq. (2.100). These would respectively describe the processes $Z^0 \rightarrow W^+W^-$, $Z^0 \rightarrow W^+W^-\gamma$, and $Z \rightarrow f\bar{f}$. Conservation of four-momentum implies that the sum of the masses in any candidate final state, $|\beta\rangle$, must be less than the mass of the Z^0 . This rules out the first two processes, leaving only the decay of the Z^0 into a fermion–antifermion pair through a neutral-current weak interaction.

We now compute the resulting Z^0 decay rate. We do so in some detail in this section in order to develop some of the computational tools that are useful for general calculations of this sort. The first step is to identify the interaction Hamiltonian that corresponds to \mathcal{L}_{nc} . Since this term of the Lagrangian does not involve any time derivatives it is tempting to conclude that $\mathcal{H}_{nc} = -\mathcal{L}_{nc}$. This is not quite true in the present instance, however, because of the appearance of the time component of the gauge potential, $Z_0(x)$. The additional terms in \mathcal{H}_{nc} that arise from this source are the analogs of the contact Coulomb interaction of quantum electrodynamics and are not even Lorentz invariant. At this point one might sensibly worry that they could potentially ruin the Lorentz invariance of the S -matrix being computed. Happily, their effect turns out to precisely cancel another source of Lorentz non-invariance that is encountered in Section 5.2. The upshot is that the naive relation, $\mathcal{H}_I = -\mathcal{L}_I$, may be used after all, so these terms are therefore ignored in all of what follows.

The interaction Hamiltonian density therefore is

$$\mathcal{H}_I = -\mathcal{L}_{nc} = -ie_Z Z_\mu \bar{f} \gamma^\mu (g_V + g_A \gamma_5) f \quad (4.4)$$

in which the coupling constant is $e_Z = e/(\sin \theta_W \cos \theta_W)$. The desired matrix

element then becomes

$$\begin{aligned}
 \mathcal{M}(Z \rightarrow f\bar{f}) &= \langle f(\mathbf{p}, \sigma); \bar{f}(\mathbf{q}, \zeta) | \mathcal{H}_I(0) | Z(\mathbf{k}, \lambda) \rangle \\
 &= -ie_Z \langle f(\mathbf{p}, \sigma); \bar{f}(\mathbf{q}, \zeta) | \bar{f}\gamma^\mu (g_V + g_A\gamma_5) f Z_\mu | Z(\mathbf{k}, \lambda) \rangle \\
 &= -ie_Z \langle 0 | b_{p,\sigma} \bar{b}_{q,\zeta} \bar{f}\gamma^\mu (g_V + g_A\gamma_5) f Z_\mu a_{k,\lambda}^* | 0 \rangle \quad (4.5)
 \end{aligned}$$

This matrix element may be evaluated once the fields appearing within the interaction Hamiltonian are expressed in terms of creation and annihilation operators. These are given in Chapter 1 by Eq. (1.116) and Eq. (1.82):

$$Z_\mu(x) = \sum_{\lambda'=-1}^1 \int \frac{d^3k'}{2E_{\mathbf{k}'}(2\pi)^3} \left[\epsilon_\mu(\mathbf{k}', \lambda') a_{\mathbf{k}', \lambda'} e^{ik'x} + \text{h.c.} \right] \quad (4.6)$$

$$\psi(x) = \sum_{\sigma'=\pm\frac{1}{2}} \int \frac{d^3p'}{2E_{\mathbf{p}'}(2\pi)^3} \left[u(\mathbf{p}', \sigma') b_{\mathbf{p}', \sigma'} e^{ip'x} + v(\mathbf{p}', \sigma') \bar{b}_{\mathbf{p}', \sigma'}^* e^{-ip'x} \right] \quad (4.7)$$

The matrix element, Eq. (4.5), clearly gets contributions only from those terms in the expansion of the fields, Eq. (4.6) and Eq. (4.7), in which the destruction operator, a , appearing in $Z_\mu(x)$ destroys the incoming Z^0 boson, and the creation operators, b^* from $\bar{f}(x)$ and \bar{b}^* from $f(x)$, create the fermion–antifermion pair. The matrix element then is

$$\mathcal{M}(Z \rightarrow f\bar{f}) = -ie_Z \epsilon_\mu(\mathbf{k}, \lambda) \bar{u}(\mathbf{p}, \sigma) \gamma^\mu (g_V + g_A\gamma_5) v(\mathbf{q}, \zeta) \quad (4.8)$$

The differential decay rate is related to this result by Eq. (3.40):

$$\begin{aligned}
 d\Gamma[Z(\mathbf{k}, \lambda) \rightarrow f\bar{f}] &= |\mathcal{M}(Z \rightarrow f\bar{f})|^2 (2\pi)^4 \delta^4(k-p-q) \frac{d^3p \, d^3q}{4E_{\mathbf{p}} E_{\mathbf{q}} (2\pi)^6} \\
 &= e_Z^2 |\epsilon_\mu \bar{u} \gamma^\mu (g_V + g_A\gamma_5) v|^2 \\
 &\quad \times (2\pi)^4 \delta^4(k-p-q) \frac{d^3p \, d^3q}{4E_{\mathbf{p}} E_{\mathbf{q}} (2\pi)^6} \quad (4.9)
 \end{aligned}$$

The next step we must take is to evaluate the square of the matrix elements, $|\epsilon_\mu \bar{u} \gamma^\mu (g_V + g_A\gamma_5) v|^2$, that arise in this last expression. The evaluation proceeds differently depending on whether the particles involved are polarized or unpolarized. We consider the two cases of polarized and unpolarized initial Z^0 bosons separately.

4.1.2 Unpolarized Z^0 decay

Consider the decay of a sample of Z^0 s that have no net polarization. We take the initial density matrix in the 3×3 spin space of the Z^0 meson to be

the unit matrix:

$$\rho = \frac{1}{3} \sum_{\lambda=-1}^1 |Z(\mathbf{k}, \lambda)\rangle \langle Z(\mathbf{k}, \lambda)| \quad (4.10)$$

In order to proceed we need to generalize the S -matrix formalism slightly to include the case for which the initial state is not a pure state, $|\alpha\rangle$, but is rather described by a density matrix, ρ . In this case the probability of there being a transition to a final state, $|\beta\rangle$, is given by the trace

$$p(\beta) = \text{tr}(\rho P_\beta) \quad (4.11)$$

in which $P_\beta = |\beta\rangle \langle \beta|$ is the projection operator onto the subspace of Hilbert space that is spanned by $|\beta\rangle$. In the special case where the initial state is a pure state, $\rho = |\alpha\rangle \langle \alpha|$, this reduces to the squared amplitude $|\langle \beta | \alpha \rangle|^2$. More generally, if the initial system could be in state $|i\rangle$ with probability P_i , then $\rho = \sum_i P_i |i\rangle \langle i|$ and $p(\beta) = \sum_i P_i |\langle \beta | i \rangle|^2$.

Using this expression, the differential decay rate for a sample of Z^0 s that is described by the density matrix of Eq. (4.10) is then given by averaging the result of Eq. (4.9) over the initial Z^0 spin, λ . If, as is usually the case, the spins of the final fermions are not measured in the detector, then we must also sum over all possible final-state polarizations:

$$d\Gamma[Z(\mathbf{k}) \rightarrow f\bar{f}] = \frac{1}{3} \sum_{\lambda=-1}^1 \sum_{\sigma=\pm\frac{1}{2}} \sum_{\zeta=\pm\frac{1}{2}} d\Gamma[Z(\mathbf{k}, \lambda) \rightarrow f\bar{f}] \quad (4.12)$$

The spin sums may be evaluated using the polarization vector identity given by Eq. (1.119) and the spinor identities given in Eq. (1.99) and Eq. (1.100).

That part of the squared amplitude which involves the Z^0 polarization then becomes

$$\begin{aligned} & \sum_{\lambda=-1}^1 |\epsilon_\mu \bar{u} \gamma^\mu (g_V + g_A \gamma_5) v|^2 \\ &= \sum_{\lambda=-1}^1 \epsilon_\mu(\mathbf{k}, \lambda) \epsilon_\nu^*(\mathbf{k}, \lambda) [\bar{u} \gamma^\mu (g_V + g_A \gamma_5) v] [\bar{u} \gamma^\nu (g_V + g_A \gamma_5) v]^* \\ &= \left[\eta_{\mu\nu} + \frac{k_\mu k_\nu}{M_Z^2} \right] [\bar{u} \gamma^\mu (g_V + g_A \gamma_5) v] [\bar{u} \gamma^\nu (g_V + g_A \gamma_5) v]^* \end{aligned} \quad (4.13)$$

A similar manipulation may be performed for the fermion spinors, u and v , once the trick of rewriting the spinor product as a trace over Dirac matrices is used:

$$\bar{u} M u = \sum_{ij} \bar{u}_i M_{ij} u_j = \text{tr}[M(u\bar{u})] \quad (4.14)$$

In this last expression, $(u\bar{u})$ denotes the dyadic matrix whose matrix elements are given by $(u\bar{u})_{ij} = u_i\bar{u}_j$. Using this trick gives

$$\begin{aligned} & [\bar{u}\gamma^\mu(g_V + g_A\gamma_5)v][\bar{u}\gamma^\nu(g_V + g_A\gamma_5)v]^* \\ &= -[\bar{u}\gamma^\mu(g_V + g_A\gamma_5)v][\bar{v}\gamma^\nu(g_V + g_A\gamma_5)u] \\ &= -\text{tr}[\gamma^\mu(g_V + g_A\gamma_5)v\bar{v}\gamma^\nu(g_V + g_A\gamma_5)u\bar{u}] \end{aligned} \quad (4.15)$$

The utility of this way of writing things is that the dyadics $u\bar{u}$ and $v\bar{v}$ have simple expressions, given by Eq. (1.99) and Eq. (1.100) respectively, when both of the spinors in the dyadic refer to the same particle. Performing the fermion spin sums using these expressions gives

$$\sum_{\sigma=\pm\frac{1}{2}} u(\mathbf{p}, \sigma)\bar{u}(\mathbf{p}, \sigma) = (m_f - i\not{p}) \quad (4.16)$$

$$\sum_{\sigma=\pm\frac{1}{2}} v(\mathbf{q}, \zeta)\bar{v}(\mathbf{q}, \zeta) = (-m_f - i\not{q}) \quad (4.17)$$

so summing the result of Eq. (4.15) over the fermion spins then gives

$$\begin{aligned} & \sum_{\sigma, \zeta=\pm\frac{1}{2}} [\bar{u}\gamma^\mu(g_V + g_A\gamma_5)v][\bar{u}\gamma^\nu(g_V + g_A\gamma_5)v]^* \\ &= \text{tr}[\gamma^\mu(g_V + g_A\gamma_5)(m_f + i\not{q})\gamma^\nu(g_V + g_A\gamma_5)(m_f - i\not{p})] \end{aligned} \quad (4.18)$$

4.1.3 Evaluating Dirac traces

Further progress requires the evaluation of various traces over Dirac matrices, of form $\text{tr}[\gamma_{\mu_1} \dots \gamma_{\mu_n}]$ or $\text{tr}[\gamma_5 \gamma_{\mu_1} \dots \gamma_{\mu_n}]$. (Traces involving multiple γ_5 can always be handled by anti-commuting a γ_5 across the γ_μ which separate it from another, and using $\gamma_5\gamma_5 = \mathbf{1}$.)

There are two procedures for evaluating such traces. One procedure is to use repeatedly the identity, Eq. (C.56) from Appendix C and the cyclicity of the trace. Here we will present an alternative, in some respects more powerful, approach. Namely, we take advantage of their transformation properties under the (improper) Lorentz group.

The key observation is that the Dirac gamma-matrices, γ_μ , satisfy the following property:

$$D^{-1}(\Lambda)\gamma^\mu D(\Lambda) = \Lambda^\mu_\nu \gamma^\nu \quad (4.19)$$

in which Λ^μ_ν is an arbitrary Lorentz transformation whose representation on spinor fields – c.f. Eq. (1.72) – is denoted $D(\Lambda)$. This implies that a trace over n gamma matrices is an *invariant tensor* of the Lorentz group. That is,

$$\Lambda^{\mu_1}_{\nu_1} \dots \Lambda^{\mu_n}_{\nu_n} \text{tr}[\gamma^{\nu_1} \dots \gamma^{\nu_n}] = \text{tr}[\gamma^{\mu_1} \dots \gamma^{\mu_n}] \quad (4.20)$$

for all Lorentz transformations. A trace that includes a factor of the matrix γ_5 is similarly an invariant Lorentz *pseudotensor*:

$$\Lambda^{\mu_1}_{\nu_1} \cdots \Lambda^{\mu_n}_{\nu_n} \text{tr}[\gamma_5 \gamma^{\nu_1} \cdots \gamma^{\nu_n}] = \det(\Lambda) \text{tr}[\gamma_5 \gamma^{\mu_1} \cdots \gamma^{\mu_n}] \quad (4.21)$$

Now comes the main point: *any* such invariant tensor of the Lorentz group may be constructed from products of the invariant metric tensor, $\eta_{\mu\nu}$. Similarly, any invariant pseudotensor may be constructed from products of the metric tensor and an odd power of the completely antisymmetric *Levi–Civita symbol*, $\epsilon^{\mu\nu\lambda\rho}$. This last tensor is an invariant pseudotensor by virtue of the following identity that is satisfied by any 4×4 matrix:

$$\Lambda^{\mu_1}_{\nu_1} \Lambda^{\mu_2}_{\nu_2} \Lambda^{\mu_3}_{\nu_3} \Lambda^{\mu_4}_{\nu_4} \epsilon^{\nu_1 \nu_2 \nu_3 \nu_4} = \det(\Lambda) \epsilon^{\mu_1 \mu_2 \mu_3 \mu_4} \quad (4.22)$$

The traces may therefore be evaluated up to an overall multiplicative factor by writing down the most general combinations of metric and Levi–Civita tensors that has the same number and symmetry of indices. The multiplicative factor may then be chosen by evaluating the trace for a particularly simple choice of indices. This procedure may be illustrated as follows.

(i)

$$\text{tr}[\gamma^{\mu_1} \cdots \gamma^{\mu_n}] = 0 \quad \text{if } n \text{ is odd.} \quad (4.23)$$

This is so because the result must be expressed as a combination of metrics and Levi–Civita symbols. However, each of these has an even number of indices. They cannot be combined into an object with an odd number of indices, so the result must vanish.

(ii)

$$\text{tr}[\gamma_5 \gamma^{\mu_1} \cdots \gamma^{\mu_n}] = 0 \quad \text{if } n \text{ is odd} \quad (4.24)$$

This result is an immediate consequence of the previous one since $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$ involves an even number of gamma matrices.

(iii)

$$\text{tr}[\gamma^\mu \gamma^\nu] = 4\eta^{\mu\nu} \quad (4.25)$$

There is only one invariant second-rank symmetric tensor: the metric itself, $\eta^{\mu\nu}$. This establishes Eq. (4.25) up to the value of the proportionality constant. To fix this constant, choose the special case where $\mu = \nu = 1$, for which $\text{tr}[(\gamma^1)^2] = \text{tr}[1] = 4 = 4\eta^{11}$.

(iv)

$$\text{tr}[\gamma_5 \gamma^\mu \gamma^\nu] = 0 \quad (4.26)$$

To see this, note that $\gamma_5 \gamma^\mu = -\gamma^\mu \gamma_5$. The γ^μ may then be moved

to the end by cyclicity of the trace, proving that the result must be antisymmetric in μ, ν . But the only second-rank invariant tensor is symmetric, so the answer must be zero.

(v)

$$\text{tr}[\gamma^\mu \gamma^\nu \gamma^\lambda \gamma^\rho] = 4(\eta^{\mu\nu} \eta^{\lambda\rho} - \eta^{\mu\lambda} \eta^{\nu\rho} + \eta^{\mu\rho} \eta^{\nu\lambda}) \quad (4.27)$$

A fourth-rank invariant tensor (as opposed to pseudotensor) must be constructed from a sum of pairs of metric tensors. The three distinct pairs that are possible are those that appear on the right-hand side of Eq. (4.27). The coefficient of each of these terms is most easily determined by evaluating both sides with a simple choice for the indices. For example, the coefficient of the first term is determined to be 4 by the choice $\mu = \nu = 0$ and $\lambda = \rho = 1$. With this choice only the first term on the right-hand side is nonzero since the metric is diagonal, and the left-hand side becomes $\text{tr}[(\gamma^0)^2 (\gamma^1)^2] = \text{tr}[-1] = -4 = 4\eta^{00}\eta^{11}$.

(vi)

$$\text{tr}[\gamma_5 \gamma^\mu \gamma^\nu \gamma^\lambda \gamma^\rho] = 4i\epsilon^{\mu\nu\lambda\rho} \quad (4.28)$$

The right-hand side of this result is again the unique fourth-rank invariant pseudotensor. Its coefficient is easily determined by the evaluating the choice $\mu = 0, \nu = 1, \lambda = 2$ and $\rho = 3$ for which the right-hand side is $4i\epsilon^{0123} = 4i$ (c.f. Eq. (1.33)) and the left-hand side is $\text{tr}[\gamma_5 \gamma^0 \gamma^1 \gamma^2 \gamma^3] = \text{tr}[i(\gamma_5)^2] = 4i$.

These results suffice for the present purposes. Traces involving more than four gamma matrices may be evaluated in a similar fashion.

4.1.4 Z^0 decay: formulae

With these results, we can evaluate the traces that arise in Eq. (4.18):

$$\begin{aligned} & \text{tr}[\gamma^\mu (g_V + g_A \gamma_5)(m_f + i\not{q})\gamma^\nu (g_V + g_A \gamma_5)(m_f - i\not{p})] \\ = & m_f^2 \text{tr}[\gamma^\mu (g_V + g_A \gamma_5)\gamma^\nu (g_V + g_A \gamma_5)] \\ & + \text{tr}[\gamma^\mu (g_V + g_A \gamma_5)\not{q}\gamma^\nu (g_V + g_A \gamma_5)\not{p}] \\ = & m_f^2 \text{tr}[\gamma^\mu (g_V^2 - g_A^2)\gamma^\nu] \\ & + \text{tr}[\gamma^\mu (g_V^2 + g_A^2 + 2g_V g_A \gamma_5)\not{q}\gamma^\nu \not{p}] \\ = & 4m_f^2 (g_V^2 - g_A^2) \eta^{\mu\nu} \\ & + 4(g_V^2 + g_A^2)(q^\mu p^\nu + p^\mu q^\nu - \eta^{\mu\nu} p \cdot q) + 8ig_V g_A \epsilon^{\mu\nu\alpha\beta} p_\alpha q_\beta \quad (4.29) \end{aligned}$$

In going from the first to second expressions, we have dropped terms linear in m_f because they involve an odd number of gamma matrices, and therefore vanish in the trace. Between the second and third expressions, we have moved $(g_V + g_A \gamma_5)$ across either 1 or 2 intervening gamma matrices; as γ_5 anticommutes with each γ_α , its sign flips once for each intervening gamma matrix. The last step uses the trace identities numbered 3, 5, and 6 above.

Contracting against $[\eta_{\mu\nu} + k_\mu k_\nu / M_Z^2]$ from Eq. (4.13), and using (due to the δ function) $k_\mu = p_\mu + q_\mu$, the averaged matrix element squared $\bar{\mu}^2 = \frac{1}{3} \sum_{\gamma\sigma\sigma'} |\mu|^2$ becomes

$$\bar{\mathcal{M}}^2 = \frac{4e_Z^2}{3} \left[-2(g_V^2 + g_A^2)p \cdot q + 4m_f^2(g_V^2 - g_A^2) + \frac{4m_f^2}{M_Z^2} g_A^2(m_f^2 - p \cdot q) \right] \quad (4.30)$$

This should be combined with Eq. (4.9) to give the polarization averaged differential decay rate,

$$d\bar{\Gamma}[Z(\mathbf{k}) \rightarrow f\bar{f}] = \frac{1}{2k^0} \bar{\mathcal{M}}^2 (2\pi)^4 \delta^4(k - p - q) \frac{d^3p d^3q}{2p^0 2q^0 (2\pi)^6} \quad (4.31)$$

Notice that this displays the proper Lorentz-transformation properties appropriate to a decay rate. All of the factors in Eq. (4.31) are manifestly Lorentz invariant except for the $1/2k^0$ prefactor. Since the Z^0 -boson energy, k^0 , is related to its rest mass, M_Z , and speed, v , by $k^0 = M_Z / \sqrt{1 - v^2}$, it follows that in a general frame $d\Gamma = d\Gamma_{\text{rest}} \sqrt{1 - v^2}$, implying the correct time dilation for the lifetime $\tau = 1/\Gamma$.

The decay rate in the Z^0 rest frame is found by making the substitution $k^\mu = (M_Z, \mathbf{0})$, which implies that $(2\pi)^4 \delta^4(p + q - k) = 2\pi \delta(p^0 + q^0 - M_Z) (2\pi)^3 \delta^3(\mathbf{p} + \mathbf{q})$. It follows that the outgoing fermion and antifermion have a specific energy in the Z^0 rest frame. In this case, because the fermion and antifermion have equal masses, the outgoing fermion energies and momenta are

$$\begin{aligned} p^0 &= q^0 = M_Z/2 \\ |\mathbf{p}| &= |\mathbf{q}| = \sqrt{(p^0)^2 - m_f^2} = \frac{1}{2} \sqrt{M_Z^2 - 4m_f^2} \end{aligned} \quad (4.32)$$

This kind of delta-function distribution of outgoing-particle energies is characteristic of a two-body decay process.

The rest-frame differential decay rate may be simplified by using the delta functions to perform the integrals over \mathbf{q} and $p = |\mathbf{p}|$. Suppose θ and ϕ are the polar angles that give the direction of the outgoing fermion in the Z^0 rest frame. Then the result $p \cdot q = -p^0 q^0 + \mathbf{p} \cdot \mathbf{q} = -(p^0)^2 - \mathbf{p}^2 = m_f^2 - M_Z^2/2$ implies that the differential decay rate, $d\Gamma$, for the decay of *unpolarized* Z^0 s is independent of θ and ϕ . This is not surprising, as the initial state is rotationally symmetric.

The total and differential decay rate in the Z^0 rest frame is therefore,

$$\begin{aligned}\Gamma(Z \rightarrow f\bar{f}) &= 4\pi \frac{d\Gamma}{\sin\theta d\theta d\phi}(Z \rightarrow f\bar{f}) \\ &= \frac{e_Z^2}{12\pi} M_Z \left[(g_V^2 + g_A^2) + 2(g_V^2 - 2g_A^2) \frac{m_f^2}{M_Z^2} \right] \sqrt{1 - \frac{4m_f^2}{M_Z^2}} \quad (4.33)\end{aligned}$$

Before turning to the implications of this expression, a short aside is in order to compute the same quantity for a perfectly polarized sample of Z^0 s.

4.1.5 Polarized Z^0 decay

The differential decay rate for polarized Z^0 s is found using the same techniques. Assuming that the spin of the outgoing fermion and antifermion are not observed, the main difference is that there is in this case no sum over the initial Z^0 spin, and so the identity used in Eq. (4.13) is no longer available.

The differential decay rate is therefore still given by Eq. (4.31), with the difference that in this case

$$\begin{aligned}\overline{\mathcal{M}}^2 &\rightarrow \overline{\mathcal{M}}_{\text{pol}}^2 \equiv e_Z^2 \sum_{\sigma, \zeta} |\epsilon_\mu \bar{u} \gamma^\mu (g_V + g_A \gamma_5) v|^2 \\ &= e_Z^2 \text{tr} [\not{\epsilon} (g_V + g_A \gamma_5) (m_f + i \not{\mathbf{q}}) \not{\epsilon}^* (g_V + g_A \gamma_5) (m_f - i \not{\mathbf{p}})] \quad (4.34)\end{aligned}$$

This trace may be evaluated using the techniques of Subsection 4.1.2. If the initial Z^0 is linearly polarized so that $\epsilon_\mu = \epsilon_\mu^*$ and $\epsilon \cdot \epsilon = 1$, then the result is

$$\overline{\mathcal{M}}_{\text{pol}}^2 = 4\{m_f^2(g_V^2 - g_A^2) - (g_V^2 + g_A^2)[p \cdot q - 2(\epsilon \cdot p)(\epsilon \cdot q)]\} \quad (4.35)$$

In the Z^0 rest frame, choose the direction of the Z^0 spin, ϵ_μ , to define the z -axis. Then taking the polar angles of the direction of the outgoing fermion to be (θ, ϕ) , we have $\epsilon \cdot p = -\epsilon \cdot q = |\mathbf{p}| \cos \theta$. The resulting differential cross section is independent of ϕ , as is expected due to the axial symmetry of the initial state, but does depend on θ in the following way:

$$\begin{aligned}\frac{d\Gamma}{\sin\theta d\theta} &= 2\pi \frac{d\Gamma}{\sin\theta d\theta d\phi} \\ &= \frac{e_Z^2 M_Z}{16\pi} \sqrt{1 - \frac{4m_f^2}{M_Z^2}} \\ &\quad \times \left[g_V^2 \left(1 - \cos^2 \theta + \frac{4m_f^2}{M_Z^2} \cos^2 \theta \right) + g_A^2 \left(1 - \frac{4m_f^2}{M_Z^2} \right) (1 - \cos^2 \theta) \right] \quad (4.36)\end{aligned}$$

As a check, notice that the integral of Eq. (4.36) over the interval $0 < \theta < \pi$ reproduces the same total decay rate as does the unpolarized result of Eq. (4.33), as it must.

4.1.5.1 The massless limit

Eq. (4.36) has a particularly simple physical interpretation in the limit of vanishing fermion mass, $m_f \rightarrow 0$. In this limit the differential decay rate becomes

$$\frac{d\Gamma}{\sin\theta d\theta} \approx \frac{e_Z^2}{16\pi} M_Z (g_V^2 + g_A^2) (1 - \cos^2\theta) \quad (4.37)$$

This result vanishes when the outgoing fermion comes out parallel or antiparallel to the initial Z^0 boson's polarization vector, ϵ_μ . This has a simple explanation in terms of the interplay between conservation of angular momentum and conservation of helicity (which is conserved in the limit of massless fermions).

The neutral-current interaction of Eq. (4.4) that is responsible for the Z^0 decay always pairs up fermions of definite helicity. That is, since this interaction Hamiltonian always involves the field combination $\bar{f}_L \gamma^\mu f_L$, it must always create a left-handed fermion together with the antiparticle to a left-handed fermion, which is a right-handed antifermion. The term which involves $\bar{f}_R \gamma^\mu f_R$ must similarly create a right-handed fermion and a left-handed antifermion. When the fermion or antifermion comes out along the direction of the initial Z^0 boson's polarization vector, then the total component of angular momentum along this direction is $J_z = \pm 1$. The angular momentum of the initial state along this direction is zero, however, so this decay configuration must be forbidden by conservation of angular momentum.

We return now to the main line of argument and explore the implications of Eq. (4.33) for the Z^0 decay width in the Z^0 rest frame.

4.1.6 Z^0 decay: applications

The rate for a Z^0 to decay into a particular species of fermion–antifermion pair, $f\bar{f}$, is given by Eq. (4.33):

$$\begin{aligned} \Gamma(Z \rightarrow f\bar{f}) &= \frac{e_Z^2}{12\pi} M_Z \left[(g_V^2 + g_A^2) + 2(g_V^2 - 2g_A^2) \frac{m_f^2}{M_Z^2} \right] \sqrt{1 - \frac{4m_f^2}{M_Z^2}} \\ &\approx \frac{e_Z^2}{12\pi} (g_V^2 + g_A^2) M_Z \quad \text{for } m_f^2 \ll M_Z^2 \end{aligned} \quad (4.38)$$

Table 4.1 *Fermion neutral-current coupling constants*

Fermion Type	T_3	Q	g_V	g_A	$(g_V^2 + g_A^2)$
ν_e, ν_μ, ν_τ	$+\frac{1}{2}$	0	0.25	0.25	0.125
e, μ, τ	$-\frac{1}{2}$	-1	-0.0189	-0.25	0.0629
u, c, t	$+\frac{1}{2}$	$+\frac{2}{3}$	0.0959	0.25	0.0717
d, s, b	$-\frac{1}{2}$	$-\frac{1}{3}$	-0.1730	-0.25	0.0924

The last line gives the approximate form for the decay rate to the extent that the mass ratio, m_f^2/M_Z^2 , is negligible. This is a very good approximation for all of the fermions of the standard model except the top quark, which is anyway too heavy to appear as a decay product for the Z^0 . The heaviest allowed decay product is the b quark, for which this mass ratio is $m_b^2/M_Z^2 \approx (5/90)^2 \approx 3 \times 10^{-3}$.

Given this formula for the Z^0 decay rate into differing fermion species, we may sum the contributions of all of the species of fermions in the standard model that are kinematically allowed to contribute, and thereby compute the total lifetime of the Z^0 within the standard model.

The coupling constants g_V and g_A in the standard model are given in terms of the third component of weak isospin, T_3 , and electric charge, Q , by $g_V = \frac{1}{2}T_3 - Q \sin^2 \theta_W$ and $g_A = \frac{1}{2}T_3$. The corresponding couplings of the standard model fermions are tabulated in Table 4.1 (using $\sin^2 \theta_W = 0.2311$, see Appendix A.)

From this table it is straightforward to compute the total Z^0 lifetime within the standard model.

Rather than computing the decay rate for each species of fermion in the model, it is convenient to compute the total decay rate, Γ_{tot} , and the fraction of Z^0 decays – or *branching fraction*, $B_f = \Gamma(Z \rightarrow f\bar{f})/\Gamma_{\text{tot}}$ – that go into each particular fermion species. The reason for quoting results in this way is that the branching fraction is more reliably computable since it just depends on the numbers g_V and g_A and so is less subject to errors in the values of the experimentally determined couplings. The branching fractions are also much easier to measure experimentally.

Using the numerical values for parameters given in Appendix A, we find

$$\begin{aligned}
 \Gamma(Z \rightarrow f\bar{f}) &= \frac{\alpha}{3 \sin^2 \theta_W \cos^2 \theta_W} M_Z (g_V^2 + g_A^2) N_c \\
 &= (1.336 \text{ GeV}) \cdot (g_V^2 + g_A^2) N_c
 \end{aligned} \tag{4.39}$$

Table 4.2 *Computed and measured Z^0 branching fractions*

Fermion type	Computed	Measured
$\nu_e \bar{\nu}_e + \nu_\mu \bar{\nu}_\mu + \nu_\tau \bar{\nu}_\tau$	20.5%	$f^1(20.00 \pm 0.06)\%$
$e^+ e^-$	3.45%	$(3.363 \pm 0.004)\%$
$\mu^+ \mu^-$	3.45%	$(3.366 \pm 0.007)\%$
$\tau^+ \tau^-$	3.45%	$(3.370 \pm 0.008)\%$
$b\bar{b}$	15.18%	$f^2(15.14 \pm 0.05)\%$
$u\bar{u} + d\bar{d} + s\bar{s} + c\bar{c}$	54%	$f^3(54.76 \pm 0.06)\%$
Total width	2.44 GeV	(2.4952 ± 0.0023) GeV

f^1 : i.e. $Z \rightarrow$ unobserved final state.

f^2 : i.e. $Z \rightarrow B\bar{B}$.

f^3 : i.e. $Z \rightarrow$ non- $B\bar{B}$ hadrons.

The constant N_c here represents the number of colors that is appropriate to fermion type f . $N_c = 1$ must therefore be chosen when f is a lepton and $N_c = 3$ when f is a quark. $\alpha = e^2/(4\pi)$ denotes the electromagnetic fine-structure constant whose value we take at $\mu = M_Z$ to be $\alpha = 1/127.9$. The total Z^0 width then becomes

$$\begin{aligned}\Gamma_{\text{tot}} &= (1.336 \text{ GeV})[3 \cdot (0.125) + 3 \cdot (0.0629) + 9 \cdot (0.0924) + 6 \cdot (0.0717)] \\ &= 2.44 \text{ GeV}\end{aligned}\tag{4.40}$$

The corresponding Z^0 lifetime is therefore

$$\tau(Z) = \frac{1}{\Gamma_{\text{tot}}} = 2.69 \times 10^{-25} \text{ s}\tag{4.41}$$

Since even an ultra-relativistic particle can only travel around 10^{-18} m in this time, Z^0 particles decay well before they are seen, and so must be reconstructed in a detector from their decay products.

Some of the branching fractions are listed in Table 4.2.

There are several points to be made about these results.

- (i) The factor $M_Z \sqrt{1 - (4m_f^2/M_Z^2)}$ in the decay rate has its origin in the integration over *phase space*. That is, it arises from the integration over the final-state momenta $\int d^3q d^3p$. For $m_f \ll M_Z$ this factor is $O(M_Z)$ since this is the typical size of the momentum available to the final-state particles. Since the total rate for a process is given by an integral over all of the final states that can take part, it is a rule of thumb that if two processes have equal-size couplings then the one with more available phase space (i.e. the one with more available final states) will have the larger rate.

The phase-space factor is proportional to the momentum available to the final fermions, and so tends to zero as m_f approaches $M_Z/2$, as is required by four-momentum conservation. In the event that m_f should be close to $M_Z/2$ this *phase-space suppression* can make the decay rate into fermion species f much smaller than might otherwise be expected.

- (ii) The overall order of magnitude of the Z^0 decay rate can be estimated reasonably well without performing the entire detailed calculation. This may be done by keeping track of factors of coupling constants and the volume of phase space appropriate to the process of interest. Since factors of 2π are ubiquitous in these calculations, and since their omission can appreciably affect the size of the result, it is important also to keep track of these factors. There is a factor of $(2\pi)^4$ from the momentum-conserving delta function, a $(2\pi)^{-3}$ from each final state particle's momentum integration, and a (2π) from the $d\Omega$ angular integral for all but one of the final-state particles. (As will be seen later there can also be an additional factor of $(4\pi)^{-2}$ for each loop in the relevant Feynman diagram if such loops arise.)

The matrix element for Z^0 decay is clearly proportional to the coupling constants, $e_Z g_V$ and $e_Z g_A$, of the neutral-current interaction term in the Lagrangian. Since, for massless fermions, the total rate is found by *incoherently* adding the rate due to left-handed fermions to that for right-handed fermions, these two couplings must appear in the combination $e_Z^2(g_V^2 + g_A^2)$ when fermion masses are neglected. The momentum integrals and squared matrix element therefore provide $\Gamma \sim [e_Z^2(g_V^2 + g_A^2)/2\pi]X$. Here the phase-space volume, X , represents the result obtained by integrating over all final-state momenta, and whose value can be estimated by dimensional analysis. In the present example the volume of phase space is $O(M_Z)$ if m_f is not too close to M_Z , since M_Z is the typical energy available in the decay. Since Γ has dimensions of mass (in units with $\hbar = c = 1$), we get $\Gamma \sim [e_Z^2(g_V^2 + g_A^2)/2\pi]M_Z$. Comparing this estimate with the full calculation, Eq. (4.38), shows that the estimate has only missed the purely numerical factor $1/6$. This is typical of the accuracy of this type of simple order-of-magnitude estimate for two-body decays (see also Subsection 5.1.1).

- (iii) The next feature of this result that bears remarking is that the decay width, Γ , is much smaller than the mass, M_Z , since $\Gamma/M_Z \approx e_Z^2/(12\pi) \sim 10^{-2}$. This implies that the Z^0 is reasonably stable for a particle of its mass. As we shall see, Z^0 s have been observed as

a resonance in e^+e^- annihilation in high-energy electron-positron storage rings at CERN and at SLAC. The small size of the width of the Z^0 translates into the narrowness of the resulting resonance (see Subsection 6.4.1).

- (iv) Inspection of the coupling constants, g_V and g_A , of the table shows that the neutrinos couple to the Z^0 with the largest strength of the fermions of the standard model. The vector coupling, g_V , of the remaining fermions is smaller due to a partial cancellation between $\frac{1}{2}T_3$ and $Q \sin^2 \theta_w$. This cancellation is most complete for the charged leptons, e , μ , and τ , and would be perfect if $\sin^2 \theta_w$ were exactly 0.25. As a result, the charged-lepton neutral-current couplings may be considered to first approximation as being purely axial in nature.
- (v) Although the data measures the decay rate into hadrons, the decay rate we have computed is really the decay rate into a quark-antiquark pair. Since the observed hadrons are really bound states of the quarks and since no isolated quark has ever been directly detected, it is not immediately clear that the rate for producing quark-antiquark pairs should be related to the rate for Z^0 decays into hadrons.

The argument is discussed in more detail in later chapters, but the main point can be made schematically here. The key observation that makes this connection relies on the fact that the rate we have computed is an *inclusive* rate in the sense that only the total rate for producing hadrons is considered without trying to distinguish one type of hadron from another. The observable therefore has the form of a sum over all possible final hadronic states:

$$d\Gamma(Z \rightarrow \text{hadrons}) \propto \text{tr}[\rho(Z)P(\beta)] \quad (4.42)$$

in which $\rho(Z)$ denotes the density matrix that describes the sample of initial Z^0 bosons and $P(\beta)$ is the projection matrix within the Hilbert space onto the subspace spanned by the observables labeling the final state. For the total rate for producing hadrons this projection matrix is the projector, P_H , onto the entire subspace of strongly interacting particles: $P_H = \sum_h |h\rangle\langle h|$ for some basis of hadronic states, $|h\rangle$. Now an equally good basis for the subspace of hadronic states is formed by the set of color-neutral many-quark and -gluon states, $|q, g\rangle$, even though no particular hadron-mass eigenstate may be well approximated by any particular multi-quark and -gluon state. The projector that appears in Eq. (4.41) may therefore be written $P_H = \sum_{q,g} |q, g\rangle\langle q, g|$. Once the projector is expressed in

terms of a sum on quark and gluon states the calculation simplifies dramatically. This is because the strong coupling constant is small, $\alpha_3 = g_3^2/(4\pi) \approx 0.12$, when it is evaluated at the scale, $\mu \approx M_Z$, appropriate to a Z^0 decay. It follows that the contribution of each of the quark basis states to Eq. (4.41) is well approximated at these energies by perturbation theory. To lowest order the dominant quark states that contribute are precisely the quark–antiquark pairs for which we have performed the calculation.

This is the general pattern. Although rates that involve identifying specific strongly interacting final-state particles cannot be computed without detailed knowledge of the wavefunctions of these particles, inclusive quantities that simply involve a sum over all possible hadronic states (possibly with some prescribed value for a quantum number such as B that is conserved by the strong interactions) may be reliably calculated (at high energies) within perturbation theory.

- (vi) The above table allows a comparison between the computed and observed widths for Z^0 decays into various final states. Since measurements of Z^0 properties have been made with great precision, this comparison furnishes a significant test of the standard model's accuracy. This is all the more true given the success of the model in describing other neutral-current phenomena (to be described in later chapters) using the same set of model parameters.

Before performing this comparison, however, we need to have an idea of the size of the potential corrections to the computed result. These corrections arise from processes that involve more than one power of the interaction Hamiltonian in Eq. (4.1). Corrections to the leading result can be expected to be suppressed in size by additional powers of the relevant coupling constants. For processes involving strongly interacting particles the typical size of a correction from additional strong interactions is $O(\alpha_3/(4\pi)) \approx 1\%$. All other things being equal, electroweak interactions can be expected to be even smaller since they are instead proportional to $O(\alpha/(4\pi \sin^2 \theta_w)) \approx 3 \times 10^{-3}$.

This counting turns out to be modified somewhat in the case when the correction involves the exchange of a massless particle such as a photon or a gluon. Then the appearance of infrared mass singularities can introduce factors of the logarithm of a large mass ratio which can increase the size of the correction. For strongly interacting particles this kind of effect would increase the above estimate to $O\{[\alpha_3/(4\pi)] \log(M_Z^2/\Lambda_{\text{QCD}}^2)\} \approx 5 \times 10^{-2}$. $\Lambda_{\text{QCD}} \approx 150$ MeV is a

scale that is typical of the strong interactions, and which is discussed at length in Chapter 8. The analogous estimate for the size of an electromagnetic correction is $O\{\alpha/(4\pi)\log(M_Z^2/m_f^2)\}$, which can be as large as 7×10^{-3} when f is an electron.

To summarize, we expect the uncertainty in the theoretical prediction to be in the neighborhood of around 5% for decays that involve strongly interacting quarks in the final state. Electromagnetic corrections should be the largest for Z^0 decays into electrons, for which they could be in the neighborhood of 1% of the lowest-order result. Electromagnetic corrections to decays into other final states should be smaller still. Since the neutrino does not interact strongly or electromagnetically, the prediction for the branching fraction into neutrino pairs should be accurate to within fractions of a percent.

These estimates would indicate that the uncertainty in the prediction, Eq. (4.42), for the total rate should be accurate to the level of roughly 0.13 GeV. The calculations of the hadronic branching fractions could also be in error at the few percent level. A real calculation of the size of these corrections is required in order to use the accuracy of the experiment to make a better test of the model. To date, such more precise comparisons between experiment and theory have been spectacularly successful. For instance, one of the best current experimental determinations of α_3 arises from the corrections it generates, in the width Γ_Z .

4.2 W^\pm decays

The calculation of the decay properties of the charged electroweak boson, W^\pm , follows the same lines as did that for the Z^0 . The total decay rate (but not necessarily the partial widths) of the W^+ and W^- are guaranteed to be equal to one another by the fact that CPT is a symmetry of the theory. The partial width $W^+ \rightarrow \beta$ must also be equal to the conjugate process, $W^- \rightarrow \bar{\beta}$, to the extent that the relevant interactions preserve CP. Since our interest in this section is restricted to the dominant decays of the W^\pm which are well-described within the Born approximation, which is CP perserving, it suffices to focus here on, say, the W^+ .

4.2.1 W^\pm decays: formulae

The first step is to identify the standard-model interactions for which the matrix element $\langle \beta | \mathcal{H}_I | W \rangle \neq 0$, since these can directly mediate the decay.

An argument that is identical to the one used for the Z^0 shows that the only such interaction is the charged-current fermion coupling of Eq. (2.88),

$$\begin{aligned}\mathcal{H}_I &= -\mathcal{L}_{cc} \\ &= -ie_W \left[W_\mu^+ (\bar{\nu}_m \gamma^\mu (1 + \gamma_5) e_m + V_{nm} \bar{u}_m \gamma^\mu (1 + \gamma_5) d_n) \right. \\ &\quad \left. + W_\mu^- (\bar{e}_m \gamma^\mu (1 + \gamma_5) \nu_{1n} + (V^\dagger)_{mn} \bar{d}_m \gamma^\mu (1 + \gamma_5) u_n) \right] \quad (4.43)\end{aligned}$$

As before, e_W is the coupling constant $e_W \equiv \frac{g_2}{2\sqrt{2}} = \frac{e}{2\sqrt{2}\sin\theta_W}$. The dominant W^\pm decays are therefore predicted to be into fermion–antifermion pairs, like $W^+ \rightarrow e^+ \nu_e$, $W^+ \rightarrow \bar{s}u$, etc.

4.2.1.1 Neglect of fermion masses

To the extent that all fermion masses may be neglected compared to M_W – an excellent approximation for the standard model given that the t quark is too heavy to allow the decay $W^+ \rightarrow t\bar{d}, t\bar{s}$, or $t\bar{b}$ – no additional calculation is necessary to determine the differential rate for W^+ decays. This is because the differential decay rate for the process $W^+ \rightarrow \bar{f}_m f_n$ may be directly lifted from the results of the previous section using the following translation table:

$$\begin{aligned}g_V, g_A &\rightarrow 1 \\ M_Z &\rightarrow M_W \\ e_Z &\rightarrow e_W U_{nm} \\ \text{with } U_{nm} &= \begin{cases} \text{unit matrix, } \delta_{mn} & \text{if } f_m, f_n \text{ are leptons} \\ \text{KM matrix, } V_{nm} & \text{if } f_m, f_n \text{ are quarks} \end{cases} \quad (4.44)\end{aligned}$$

The differential decay rate for the decay of a linearly polarized W^\pm boson into a fermion–antifermion pair, $\bar{f}_m f_n$, (with the final fermion spins unmeasured) therefore is

$$\frac{d\Gamma}{\sin\theta d\theta} [W^+ \rightarrow \bar{f}_m f_n] \approx \frac{e_W^2}{8\pi} |U_{nm}|^2 M_W N_c (1 - \cos^2\theta); \quad m_m^2, m_n^2 \ll M_W^2 \quad (4.45)$$

The notation is the same as in the previous section. θ denotes the polar angle of the outgoing fermion in the rest frame of the decaying W^+ with the initial polarization direction chosen to define the z -axis.

The total (unpolarized) decay rate is similarly

$$\Gamma(W^+ \rightarrow \bar{f}_m f_n) \approx \frac{e_W^2}{6\pi} |U_{nm}|^2 N_c M_W; \quad m_m^2, m_n^2 \ll M_W^2 \quad (4.46)$$

4.2.1.2 Non-vanishing fermion masses

Before exploring the implications of these expressions, we pause to generalize the above results to the case where the fermion masses are not neglected.

These generalizations may be straightforwardly proven using the techniques of the previous section.

The full expression for the differential decay rate for polarized W^+ bosons into an unpolarized fermion–antifermion pair, $\bar{f}_m f_n$, is

$$\begin{aligned} \frac{d\Gamma}{\sin\theta d\theta}[W^+ \rightarrow \bar{f}_m f_n] &= \frac{e_W^2}{8\pi} |U_{nm}|^2 N_c M_W \sqrt{\left(1 - \frac{\bar{m}^2}{M_W^2}\right)^2 - \frac{4m_m^2 m_n^2}{M_W^4}} \\ &\times \left[1 - \frac{\bar{m}^2}{M_W^2} - \left[\left(1 - \frac{\bar{m}^2}{M_W^2}\right)^2 - \frac{4m_m^2 m_n^2}{M_W^4} \right] \cos^2 \theta \right] \end{aligned} \quad (4.47)$$

in which $\bar{m}^2 = m_m^2 + m_n^2$ is the sum of the squared masses of the final spin-half particles.

Of particular interest is a special case of this last expression for which the mass of only one of the fermions is negligible. This is the result that is appropriate if f_n is a neutrino and the rate is desired as a function of the charged lepton mass. The results are

$$\begin{aligned} \frac{d\Gamma}{\sin\theta d\theta}[W \rightarrow \bar{f}_m f_n] &= \frac{e_W^2}{8\pi} |U_{nm}|^2 M_W N_c \left[1 - \left(1 - \frac{m_m^2}{M_W^2}\right) \cos^2 \theta \right] \\ &\times \left(1 - \frac{m_m^2}{M_W^2}\right)^2; \quad m_n^2 \ll m_m^2, M_W^2 \end{aligned} \quad (4.48)$$

The total decay rate in this last case becomes

$$\Gamma(W \rightarrow \bar{f}_m f_n) = \frac{e_W^2}{6\pi} |U_{nm}|^2 M_W N_c \left(1 + \frac{m_m^2}{M_W^2}\right) \left(1 - \frac{m_m^2}{M_W^2}\right)^2 \quad (4.49)$$

Since this is a two-body decay, four-momentum conservation implies that the spectrum of outgoing fermions has a delta-function distribution as a function of the energy of the outgoing fermions: $d\Gamma/dE \propto \delta(E - E_0)$. Since the fermions have equal and opposite three-momenta in the W^\pm rest frame, it follows that their energies in this frame are given in terms of the particle masses by

$$\begin{aligned} p_m^0 &= \frac{M_W}{2} + \frac{m_m^2 - m_n^2}{2M_W} \\ p_n^0 &= \frac{M_W}{2} - \frac{m_m^2 - m_n^2}{2M_W} \end{aligned} \quad (4.50)$$

This process is clearly kinematically allowed provided only that the sum of fermion and antifermion masses is smaller than M_W .

4.2.2 W^\pm decays: applications

The total decay rate and the branching fractions of the W^\pm boson into fermion–antifermion pairs may now be computed within the standard model by applying these formulae.

It is convenient to normalize the total decay rate by the partial rate for the decay of a W^+ into an positron–neutrino pair:

$$\begin{aligned}\Gamma(W^+ \rightarrow e^+ \nu_e) &= \frac{\alpha}{12 \sin^2 \theta_W} M_W \\ &= (226 \text{ MeV})\end{aligned}\quad (4.51)$$

In terms of this partial rate the total W^\pm decay width therefore is

$$\begin{aligned}\Gamma_{\text{tot}} &= \Gamma(W^+ \rightarrow e^+ \nu_e) \left[3 + 3 \sum_{n=1}^2 \sum_{m=1}^3 |V_{nm}|^2 \right] \\ &= 9\Gamma(W^+ \rightarrow e^+ \nu_e) \\ &= 2.04 \text{ GeV}\end{aligned}\quad (4.52)$$

The first factor of 3 in the square bracket in the first equation corresponds to the three families of leptons. The second 3 represents the three colors of each quark. The sum over ‘up-type’ quarks only runs over the first two generations because the top quark is too heavy to be a W^+ decay product. The second equality in Eq. (4.52) uses the unitarity of the Kobayashi–Maskawa matrix $\sum_{n=1}^2 \sum_{m=1}^3 |V_{nm}|^2 = \sum_{n=1}^2 [VV^\dagger]_{nn} = 2$.

The decay width of Eq. (4.52) corresponds to a W^\pm boson lifetime of

$$\tau(W^\pm) = \frac{1}{\Gamma_{\text{tot}}} = 3.22 \times 10^{-25} \text{ s} \quad (4.53)$$

These particles clearly decay well before they may themselves be directly detected.

Many branching fractions are once again calculable as pure numbers, independent of model parameters. A few branching fractions are presented in Table 4.3.

The W^\pm boson is again very long lived on the scale of its mass, and decays into leptons 33% of the time, and hadrons the rest of the time.

Since the strengths of the W^\pm couplings to fermions do not depend on the fermions’ electric charges or other such quantum numbers, to this approximation the only difference in the branching fractions into different species

Table 4.3 *Computed and measured W^+ branching fractions*

Fermion type	Computed	Measured
$e^+\nu_e$	11.1%	$(10.72 \pm 0.16) \%$
$\mu^+\nu_\mu$	11.1%	$(10.57 \pm 0.22) \%$
$\tau^+\nu_\tau$	11.1%	$(10.74 \pm 0.27) \%$
Hadrons	66.7%	$(67.96 \pm 0.35) \%$
Total width	2.04 GeV	$(2.118 \pm 0.042) \text{ GeV}$

of particles is due to the existence of the Kobayashi–Maskawa matrix. As a result, the model predicts absolutely no difference among the decay rates into lepton pairs until masses and radiative (loop) corrections are included. The size of these corrections are expected to be roughly the same size as for Z^0 decays – around a percent for electrons and less for μ s and τ s.

Like the Z^0 width, the W^\pm width is larger than our theoretical estimate. This is because of (computable) positive $O(\alpha_3)$ corrections; the width is in good agreement with a more detailed calculation.

4.3 Higgs decays

The last massive elementary boson of the model to be considered here is the spinless Higgs particle. We compute the decay rate of this particle, not to compare with experiment, but to predict the observations of future experiments of this (as yet undetected) particle.

The interaction terms in the Lagrangian that are linear in the Higgs scalar, which can potentially mediate Higgs decay through a matrix element of the form $\langle \beta | \mathcal{H}_I | H \rangle \neq 0$, are of two types: the Higgs-fermion Yukawa couplings of Eq. (2.70), and the Higgs-electroweak boson interactions of Eq. (2.69):

$$\mathcal{H}_f = -\mathcal{L}_{\text{Hf}} = \sum_f \frac{m_f}{v} \bar{f} f H \quad (4.54)$$

and

$$\mathcal{H}_g = -\mathcal{L}_{\text{H-g}} = \frac{H}{v} \left(2M_W^2 W_\mu^+ W^{-\mu} + M_Z^2 Z_\mu Z^\mu \right) \quad (4.55)$$

The first of these can mediate the potential decay $H \rightarrow f\bar{f}$, and the second can mediate $H \rightarrow W^+W^-$ and $H \rightarrow Z^0Z^0$. Since the coupling strength in both cases is proportional to the mass of the decay products, the dominant decay is expected to be into the heaviest particle that is still light enough for the decay to be kinematically allowed. It should be noted, however, that

a longitudinal W or Z boson counts for these purposes as a Higgs particle (see Subsection 4.3.2). The condition for the decay to be allowed is that the mass of the decay product must be less than half the mass of the Higgs boson.

Which decay dominates therefore depends on what the mass of the Higgs boson is. There are three possibilities, depending on the size of m_H .

(i) $m_H < 2M_W$:

In this case the heaviest available final state is a $b\bar{b}$ pair. Eq. (4.54) is therefore the relevant interaction.

(ii) $2M_W < m_H < 2m_t$:

In this scenario the top quark is too heavy to be produced, and so the Higgs prefers to decay into gauge-boson pairs. Here Eq. (4.55) describes the most relevant interaction.

(iii) $2M_W < 2m_t < m_H$:

Here the Higgs loves to decay into a $t\bar{t}$ pair. Eq. (4.54) is again the interaction of most interest. Note however that for a very heavy Higgs boson, it will turn out that decays into W bosons again dominate.

The next step is to compute the Higgs decay rate into fermion–antifermion pairs and into gauge-boson pairs.

4.3.1 Decays into fermions

The matrix element for this process may be evaluated using the expansion of the fields in terms of creation and annihilation operators, as in Subsection 4.1.3. The result is

$$\begin{aligned}
 \mathcal{M}(H \rightarrow f\bar{f}) &= \langle f(\mathbf{p}, \sigma); \bar{f}(\mathbf{q}, \zeta) | \mathcal{H}_f(0) | H(\mathbf{k}) \rangle \\
 &= \frac{m_f}{v} \langle f(\mathbf{p}, \sigma); \bar{f}(\mathbf{q}, \zeta) | \bar{f}f H | H(\mathbf{k}) \rangle \\
 &= \frac{m_f}{v} \bar{v}(\mathbf{q}, \zeta) u(\mathbf{p}, \sigma)
 \end{aligned} \tag{4.56}$$

Summing the square of this matrix element over final-state spins gives

$$\begin{aligned}
 \overline{\mathcal{M}}^2_f &\equiv \sum_{\sigma, \zeta} |\mathcal{M}(H \rightarrow f\bar{f})|^2 \\
 &= -\frac{m_f^2}{v^2} \text{tr}[(m_f - i\not{p})(m_f + i\not{q})] \\
 &= \frac{4m_f^2}{v^2} (-p \cdot q - m_f^2)
 \end{aligned} \tag{4.57}$$

The differential decay rate is therefore

$$\begin{aligned} d\Gamma(H \rightarrow f\bar{f}) &= \frac{1}{2k^0} \overline{\mathcal{M}}_f^2 (2\pi)^4 \delta^4(p+q-k) \frac{d^3p d^3q}{4p^0 q^0 (2\pi)^6} \\ &= \frac{2m_f^2}{v^2 k^0} (-p \cdot q - m_f^2) (2\pi)^4 \delta^4(p+q-k) \frac{d^3p d^3q}{4p^0 q^0 (2\pi)^6} \quad (4.58) \end{aligned}$$

In the Higgs rest frame the fermions clearly come out back-to-back and with energies equal to half the Higgs mass. Owing to the rotational symmetry of the problem, the decay probability in the rest frame is also independent of the direction of the outgoing fermion–antifermion pair. Including the potential sum over final-state color, the total Higgs decay rate into a particular flavor of fermion in the Higgs rest frame becomes

$$\begin{aligned} \Gamma(H \rightarrow f\bar{f}) &= \frac{m_H}{8\pi} \left(\frac{m_f}{v} \right)^2 N_c \left(1 - \frac{4m_f^2}{m_H^2} \right)^{3/2} \\ &\approx \frac{m_H}{8\pi} \left(\frac{m_f}{v} \right)^2 N_c \quad \text{for } m_f \ll m_H \quad (4.59) \end{aligned}$$

This decay rate is clearly very sensitive to the final-state fermion mass, and as advertised is largest for the heaviest fermions. An upper bound on the ratio of the Higgs lifetime to the Higgs mass may be obtained by using the above result and including just those fermions that are known to be lighter than is the Higgs. This includes all standard model fermions except the t quark. By far the largest contribution from these fermions comes from the heaviest – the b quark. Neglecting the ratio m_b^2/m_H^2 and using the present (2006) 95% confidence-level lower bound, $m_H > 114$ GeV, on the Higgs mass gives the bound

$$\begin{aligned} \Gamma_{\text{tot}} &> \Gamma(H \rightarrow b\bar{b}) \\ &= (3.5 \times 10^{-5}) m_H \\ &> 4 \text{ MeV} \quad (4.60) \end{aligned}$$

This represents an upper limit for the Higgs lifetime of $\tau(H) < 1.6 \times 10^{-22}$ s. Such a particle could not make a track that is long enough to be visible in a detector and so must be detected through its decay products.

4.3.2 Decays into gauge bosons

We consider the W^+W^- final state before computing the Z^0Z^0 case in order to postpone the complication due to statistics when the final particles are identical.

4.3.2.1 The decay $H \rightarrow W^+W^-$

The matrix element for this process is evaluated as usual by expanding all fields in terms of creation and annihilation operators:

$$\begin{aligned}
 \mathcal{M}(H \rightarrow W^+W^-) &= \langle W^+(\mathbf{p}, \sigma); W^-(\mathbf{q}, \zeta) | \mathcal{H}_g(0) | H(\mathbf{k}) \rangle \\
 &= \frac{2M_W^2}{v} \langle W^+(\mathbf{p}, \sigma); W^-(\mathbf{q}, \zeta) | W_\mu^+ W^{-\mu} H | H(\mathbf{k}) \rangle \\
 &= \frac{2M_W^2}{v} \epsilon_\mu^*(\mathbf{q}, \zeta) \epsilon^{*\mu}(\mathbf{p}, \sigma) \quad (4.61)
 \end{aligned}$$

Squaring and summing over gauge-boson spins using Eq. (1.119) gives

$$\begin{aligned}
 \overline{\mathcal{M}}_g^2 &\equiv \sum_{\sigma, \zeta} |\mathcal{M}(H \rightarrow W^+W^-)|^2 \\
 &= \left(\frac{2M_W^2}{v} \right)^2 \left[2 + \frac{(p \cdot q)^2}{M_W^4} \right] \quad (4.62)
 \end{aligned}$$

The differential decay rate becomes

$$\begin{aligned}
 d\Gamma(H \rightarrow W^+W^-) &= \frac{1}{2k^0} \left(\frac{2M_W^2}{v} \right)^2 \left[2 + \frac{(p \cdot q)^2}{M_W^4} \right] \\
 &\quad \times (2\pi)^4 \delta^4(p + q - k) \frac{d^3p d^3q}{2p^0 2q^0 (2\pi)^6} \quad (4.63)
 \end{aligned}$$

The kinematics of this collision are identical to those for the fermion–antifermion decay: the gauge bosons are back-to-back in the Higgs rest frame with an energy of $m_H/2$. The decay distribution in the rest frame is also completely isotropic. The total Higgs decay rate in the Higgs rest frame for this particular channel becomes

$$\begin{aligned}
 \Gamma(H \rightarrow W^+W^-) &= \frac{m_H}{16\pi} \frac{m_H^2}{v^2} \left[1 - 4 \left(\frac{M_W^2}{m_H^2} \right) + 12 \left(\frac{M_W^2}{m_H^2} \right)^2 \right] \sqrt{1 - \frac{4M_W^2}{m_H^2}} \\
 &\approx \frac{m_H}{16\pi} \frac{m_H^2}{v^2} \quad \text{for } M_W \ll m_H \quad (4.64)
 \end{aligned}$$

An interesting feature about this result is that it is proportional to the square of the Higgs mass, $(m_H/v)^2$, rather than to the square of the mass of the final-state particle as was the case for Higgs decays into fermions, Eq. (4.59). Since the ratio $(m_H/v)^2$ is essentially the Higgs self-coupling, λ , (c.f. Eq. (2.38)), this reflects the fact that the longitudinal component of the massive gauge bosons originate as components of the scalar doublet that are “eaten” by the Higgs mechanism. This allows a simple interpretation for the two terms in the square bracket in the last equality of Eq. (4.63):

the factor of 2 corresponds to the two transverse polarization states of the W meson which couple with a strength proportional to the gauge coupling, $g_2 \approx M_W/v$, and the remaining term represents the momentum-dependent coupling of the longitudinal “Goldstone mode” that couples proportional to the Higgs self-coupling as above.

This decay rate clearly cannot contribute at all unless the Higgs mass is at least twice that of the W boson, $m_H > 2M_W = 161$ GeV.

4.3.2.2 The decay $H \rightarrow Z^0 Z^0$

The only difference between this decay and the decay into W pairs lies in the statistics of the final two-boson state. Since the two final Z^0 particles are identical, there are two changes to be made to the calculation.

- (i) In the evaluation of the matrix element, \mathcal{M} , of Eq. (4.61), there are two effects which cancel one another. First, the numerical coefficient of the interaction Hamiltonian, Eq. (4.55), is half as large for Z^0 s as it is for W^\pm s. This introduces a factor of $\frac{1}{2}$ in the matrix element relative to the W^+W^- result.
- (ii) On the other hand, there are now two ways in which the fields, $Z_\mu Z^\mu$, can create the two particles in the final state. These two ways arise because either field can create any given Z^0 . This introduces a factor of 2 in the matrix element relative to the W^+W^- result. This factor of 2 cancels the previous factor of $\frac{1}{2}$.
- (iii) Next, the range of the final integration over the solid angle of the direction of one of the outgoing particles need only be 2π steradians rather than the usual 4π since it is impossible to distinguish which Z boson heads in which direction. This reduces the decay rate relative to the WW final state by a factor of 2. (Alternatively, if we perform the integral over 4π steradians, we have double-counted each possible final state, and must divide by a factor of 2 to undo this double counting.)

The final result is therefore

$$\Gamma(H \rightarrow Z^0 Z^0) = \frac{1}{2} \Gamma(H \rightarrow W^+ W^-) \Big|_{M_W \rightarrow M_Z} \quad (4.65)$$

that is, the rate for decay into Z^0 s is the same as it is separately to decay into each of the two charge states, W_1 or W_2 , that make up the W^\pm , except for the substitution of M_Z for M_W .

4.4 Problems

[4.1] **W width at finite fermion mass**

Calculate the rate $\Gamma(W^- \rightarrow e^- \bar{\nu}_e)$ without assuming $m_e \ll M_w$.

Use

$$\mathcal{L} = ie_W W_\mu^- \bar{e} \gamma^\mu (g_V + g_A \gamma_5) \nu + \text{h.c.}, \quad g_V = g_A = 1$$

[4.1.1] Show that the matrix element is

$$\langle e(p) \bar{\nu}_e(q) | \mathcal{H} | W(k) \rangle = -ie_W \epsilon_\mu(k) \bar{u}(p) \gamma^\mu (g_V + g_A \gamma_5) v(q)$$

[4.1.2] Show that, if $m_\nu \neq 0$ and $m_e \neq 0$, we would get

$$\sum_{\sigma_1 \sigma_2} |\epsilon_\mu(k) \bar{u}(p) \gamma^\mu (g_V + g_A \gamma_5) v(q)|^2 = 4 \left\{ m_e m_\nu (g_V^2 - g_A^2) + (g_V^2 + g_A^2) [-p \cdot q + 2\epsilon \cdot p \epsilon \cdot q] \right\}$$

[4.1.3] Suppose the initial W^- is linearly polarized in the direction \vec{e} . In the W^- rest frame, show that

$$\frac{d\Gamma}{d\cos\theta} = \frac{e_W^2 M_W}{16\pi} (g_V^2 + g_A^2) \left[1 - \left(1 - \frac{m_e^2}{M_W^2} \right) \cos^2\theta \right] \left(1 - \frac{m_e^2}{M_W^2} \right)^2$$

where θ is the angle between \vec{p} and \vec{e} . Assume $m_\nu = 0$.

[4.1.4] Show that the unpolarized rate is

$$\Gamma(W^- \rightarrow e^- \bar{\nu}_e) = \frac{e_W^2 M_W}{12\pi} (g_V^2 + g_A^2) \left(1 + \frac{m_e^2}{2M_W^2} \right) \left(1 - \frac{m_e^2}{M_W^2} \right)^2$$

[4.2] **Decay of the top quark**

Consider the top quark, with a mass of $m_t \simeq 174$ GeV.

[4.2.1] Identify the only interaction term in the Lagrangian which is linear in the top quark. Can a single insertion of this interaction term cause the top quark to decay? What are the decay products?

[4.2.2] Write an expression for the matrix element for the dominant top-quark decay process.

[4.2.3] Find a compact expression for the square of the matrix element, summing over final-state spin or helicity states and averaging over the initial top-quark helicity state.

[4.2.4] Compute the width of the top quark. Neglect the masses of any other fermions in comparison to the top-quark mass, but treat the masses of W and Z bosons as comparable to the top-quark mass. You should be able to find an analytic expression for the decay rate. Then, substitute in physical values and express the answer in GeV.

[4.3] Gamma-matrix identities

Prove the following useful formulae involving gamma matrices. You should need only the relations, $\gamma^\mu \gamma^\nu = 2\eta^{\mu\nu} - \gamma^\nu \gamma^\mu$, and $\eta_\mu^\mu = 4$.

$$\not{k}\not{k} = k^2$$

$$\not{k}\not{p}\not{k} = 2p \cdot k \not{k} - k^2 \not{p}$$

$$\gamma^\mu \gamma_\mu = 4$$

$$\gamma^\mu \not{k} \gamma_\mu = -2\not{k}$$

$$\gamma^\mu \not{p} \not{k} \gamma_\mu = 4p \cdot k$$

$$\gamma^\mu \not{p} \not{k} \not{q} \gamma_\mu = -2\not{q} \not{k} \not{p}$$

5

Leptonic weak interactions: decays

The next simplest application of the standard model to understanding the properties of the observed elementary particles is to compute the decay lifetimes of the other weakly interacting particles of the model. The only remaining particles that do not participate in the strong interactions are the leptonic fermions. This chapter is devoted to a calculation of their decay properties.

The purpose of this chapter is threefold. Two of these are straightforward. Lepton decays furnish our first example of a “second-order” decay that proceeds *via* a virtual particle, and so provide a good motivation for a full description of the Feynman rules of the theory. This calculation also provides some insight into the observed properties of real leptons and so allows more contact with experimentally accessible quantities. Indeed, the weak decays of the known fundamental particles provide much of our current information concerning the electroweak couplings. The third and final motivation is to provide the first illustration of the utility of the technique of effective Lagrangians for computing the virtual effects of heavy particles.

5.1 Qualitative features

The six flavors of fundamental leptons are $e, \mu, \tau, \nu_e, \nu_\mu$, and ν_τ . Four of these are absolutely stable in the standard model by virtue of exact or extremely good approximate conservation laws of the model. The stable species are the three neutrino types and the electron. They are absolutely stable because they are each the lightest particles that carry nonzero values for a conserved quantum number. They cannot decay because any potential decay product would have to be lighter than the decaying particle, and would have to carry a nonzero value for the quantum number in question. No such particles exist

by the very assumption that the original particle is the lightest one that carries this quantum number.

The neutrinos, being massless, are the lightest particles that carry the appropriate lepton number: L_e , L_μ , and L_τ . The electron is similarly stable since it is the lightest particle that carries electric charge. One might wonder whether the stability of the neutrinos might be suspect because of the anomalies in the conservation of L_e , L_μ , and L_τ discussed in Subsection 2.5.3. This turns out not to be so; there are three anomaly-free quantum numbers in the standard model, $L_e - L_\mu$, $L_\mu - L_\tau$, and $B - L$, which (together with the fact that the neutrinos are all lighter than any particle carrying a nonzero baryon number) are sufficient to ensure the stability of all three neutrinos. Note however that if the standard model is enlarged by relaxing renormalizability to allow dimension-5 operators, as discussed in Chapter 10, then lepton numbers are generically violated and the neutrinos may not be absolutely stable. However, estimates of their lifetimes are so long that the question of their stability is not experimentally interesting.

The decay properties of the remaining two leptons, μ and τ , are computed here. The first step is to determine which interactions are responsible for their decays. In this regard, notice that in the absence of the charged-current fermion interactions, the symmetry group of the leptonic sector of the standard model would be larger than $U_e(1) \times U_\mu(1) \times U_\tau(1)$. They would in particular include a separate symmetry under the rotation of the muon, say, by a phase that is independent of the muon neutrino. This would imply the separate conservation of the number of muon (minus antimuons) and muon neutrinos (minus muon antineutrinos), and so imply the stability of the μ (and similarly of the τ). It follows that any process which results in μ or τ decay must necessarily involve the lepton charged-current interaction at least once. The dominant contribution to the decay will be that one which involves the fewest interactions.

For definiteness consider τ^- decay. In order to involve the minimum number of interactions – one – there must be a potential decay product, $|\beta\rangle$, for which the matrix element

$$\langle\beta|\mathcal{H}_{cc}(0)|\tau^-\rangle = -\frac{ig_2}{2\sqrt{2}}\langle\beta|W_\mu^+ \bar{\nu}_\tau\gamma^\mu(1+\gamma_5)\tau|\tau^-\rangle \neq 0 \quad (5.1)$$

The only state, $|\beta\rangle$, for which this matrix element is not zero is $|\beta\rangle = |W^-; \nu_\tau\rangle$. This cannot be a decay product for a τ^- particle, since the W^- boson is more massive than is the τ^- .

It follows that τ^- - (and μ^- -) decay must arise at at least second order in the perturbative expansion of Eq. (3.26). That is, the dominant contribution

to a decay $\tau^- \rightarrow \beta$ must proceed via the matrix element

$$\mathcal{M}(\tau^- \rightarrow \beta) = \frac{(-i)^2}{2!} \int d^4x \langle \beta | T[\mathcal{H}_I(x)\mathcal{H}_I(0)] | \tau^- \rangle + \dots \quad (5.2)$$

if not at higher order.

From the above considerations, the interaction term which destroys the τ^- particle must be the charged-current Hamiltonian appearing in Eq. (5.1). Besides destroying the τ^- , this interaction also creates W^- and ν_τ particles. The second interaction term must therefore destroy the created W^- particle, in order to produce a final state that involves only particles that are less massive than the initial τ^- lepton. As is demonstrated in some detail in Section 4.2, the only interaction that can cause a transition from a W^- particle to lighter particles is once again the charged-current fermion interaction. These interactions destroy the W^- and produce a fermion and an antifermion, for instance, $e^- \bar{\nu}_e$. The τ -neutrino that is produced when the τ lepton is destroyed must appear in the final state to carry off the nonzero L_τ of the original τ^- .

The dominant decay processes must therefore be three-body decays, of the form $\tau^- \rightarrow f_m \bar{f}_n \nu_\tau$, in which f_m and f_n are any two fermions that are related to one another through the charged-current interactions and which are lighter than the initial τ^- . The rate for this decay is given to first approximation by Eq. (5.2), in which the relevant terms in the interaction Hamiltonian are

$$\mathcal{H}_I \subset -ie_W \left(W_\mu^+ \bar{\nu}_\tau \gamma^\mu (1 + \gamma_5) \tau + \sum_{mn} U_{mn}^* W_\mu^- \bar{f}_n \gamma^\mu (1 + \gamma_5) f_m \right) \quad (5.3)$$

The matrix U_{mn} in this expression is the same as that used in Section 4.2 and represents the unit matrix, δ_{mn} , if f_m and f_n are leptons and the Kobayashi–Maskawa matrix, V_{mn} if they are quarks. As before, $e_W \equiv g_2/2\sqrt{2} = e/(2\sqrt{2} \sin \theta_W)$.

Many of the properties of μ^- and τ^- decays follow from these general observations before any detailed calculations are performed.

5.1.1 μ^- decays

- (i) For μ^- decays there is only one combination of three fermions for which the sum of the masses is smaller than the μ^- mass itself, and is therefore kinematically allowed. The three particles in the final state are completely determined by the conservation laws for the decay. The final state must include the electron, e^- , since this is the only negatively charged particle that is lighter than the μ^- . Conservation

of L_e and L_μ then dictate that the remaining two fermions must be ν_μ to carry off the initial muon number, and $\bar{\nu}_e$ to cancel the electron number of the final electron. The dominant decay must therefore be

$$\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu \quad (5.4)$$

- (ii) Of the three particles in the final state, only the electron is detectable (without heroic efforts) since the neutrinos interact so weakly as to make them easily leave any detector without interacting at all. The observable of most interest is therefore the decay rate as a function of the final electron's quantum numbers. Since Eq. (5.4) is a three-body decay, the electron can emerge with a continuous range of energies, with energy conservation satisfied by having the remainder of the initial muon's energy shared by the remaining neutrinos. One of the goals of the next section is to compute the number of electrons of any given energy that emerge from a sample of decaying muons.
- (iii) Counting the coupling constants and (2π) s associated with the decay rate allows a simple estimate of its size and so of the muon lifetime. This estimate compares reasonably well with the more detailed calculation to follow. The decay involves two insertions of \mathcal{H}_I , which is linear in g_2 , so it follows that $\mathcal{M} \propto g_2^2$.

There is another factor that must be included as well, the suppression associated with the necessity to produce and destroy a virtual W^- boson. As is justified in more detail in what follows, this suppression is given by a factor of $1/M_W^2$ in the amplitude. This factor is the relativistic analog of the familiar energy denominators of non-relativistic quantum mechanical perturbation theory (c.f. Eq. (1.34) for example.)

Including this factor gives the estimate $\mathcal{M} \sim g_2^2/M_W^2$. Since the typical energy available to the final particles in the muon rest frame is m_μ , the integral of the squared matrix element over phase space may be estimated by including the appropriate power of m_μ . Since our estimate for $|\mathcal{M}|^2$ has dimension M^{-4} and the decay rate is dimension M , the power of m_μ required by dimensional analysis is m_μ^5 .

It remains to find the power of (2π) arising from the phase space integration. Each of the three integrals over final particle momenta introduces $(2\pi)^{-3}$, but there is a $(2\pi)^4$ from the energy and momentum conserving delta function. There are two independent solid angle integrations, each contributing $\sim (2\pi)$. The total power of (2π) is therefore $(2\pi)^{-3}$.

The decay rate is therefore of order

$$\begin{aligned}
 \Gamma(\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu) &\sim \frac{|\mathcal{M}|^2 m_\mu^5}{(2\pi)^3} \\
 &\sim \frac{g_2^4}{(2\pi)^3} \frac{m_\mu^5}{M_W^4} \\
 &\sim \frac{2\alpha^2}{\pi \sin^4 \theta_W} \frac{m_\mu^5}{M_W^4} \\
 &\sim 2 \times 10^{-15} m_\mu \\
 &\sim 2 \times 10^{-16} \text{ GeV}
 \end{aligned} \tag{5.5}$$

corresponding to a lifetime of $\tau(\mu) \sim 3 \times 10^{-9}$ s. Unlike in the previous chapter we take here the value $\alpha \approx 1/137$ for the electromagnetic fine structure constant that is appropriate to low energies compared to the weak scale. (The scale dependence of α is discussed in Section 7.4.)

This differs from the measured lifetime of $\tau_{\text{exp}} = 2.2 \times 10^{-6}$ s by some three orders of magnitude, motivating the more careful calculation performed below. Notice that an extremely relativistic particle with a lifetime of a microsecond can travel several hundred meters before decaying. Muons therefore live long enough to escape the region immediately surrounding the interaction point and can enter the surrounding detector for observation.

- (iv) The branching fractions for differing final states in μ^- decay may also be simply estimated. As argued above, the decay into $e\nu\bar{\nu}$ is the only one that may proceed to second order in the interactions of the model. This will therefore have a branching fraction of essentially $\approx 100\%$.

There will be other decay products available, and so deviations from the 100% branching fraction, to the extent that higher-order processes are possible. One such process that arises once three powers of the interaction Hamiltonian are allowed is the decay $\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu \gamma$, in which a photon is emitted by the initial muon or by the final electron. Apart from all of the factors of coupling constants that already arise for the purely three-body decay, the matrix element for emitting an additional photon involves an extra factor of the electromagnetic coupling, e , and the phase-space integration involves an extra $(2\pi)/(2\pi)^3$ (the numerator from an angular integration, the denominator from the momentum integration measure).

The estimate of the branching fraction for decays with an extra photon in the final state is therefore $B(\mu^- \rightarrow e\bar{\nu}\nu\gamma) \sim e^2/(2\pi)^2 \sim$

2×10^{-3} . Having the photon pair produce an electron–positron pair – $\mu^- \rightarrow e^- e^+ e^- \bar{\nu} \nu$ – would bear another factor of $e^2/(2\pi)^2$ of suppression, for a branching fraction $\sim 10^{-5}$. These are in rough agreement with the measured branching fractions

$$\begin{aligned} B_{\text{exp}}(\mu^- \rightarrow e^- \bar{\nu} \nu) &= \sim 100\% \\ B_{\text{exp}}(\mu^- \rightarrow e^- \bar{\nu} \nu \gamma) &= (1.4 \pm 0.4)\% \\ B_{\text{exp}}(\mu^- \rightarrow e^- \bar{\nu} \nu e^+ e^-) &= (3.4 \pm 0.4) \times 10^{-5} \end{aligned} \quad (5.6)$$

The photon branching fraction is larger than our estimate because it turns out to be enhanced by a factor of $\log(m_\mu/m_e)$, for reasons we will discuss in Subsection 6.7.2.

5.1.2 τ^- decays

- (i) The tau lepton differs from the muon only through the size of its mass. The arguments of the preceding section for muons therefore apply equally well to taus to the extent that they do not rely crucially on the value of the initial lepton’s mass.

Because of its larger mass, the τ lepton can decay at second order in \mathcal{H}_{cc} to many more three-fermion final states than could the relatively light muon. It has two purely leptonic decays: $\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$ and $\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$. At the quark level it can also decay into either $\tau^- \rightarrow \bar{u} d \nu_\tau$ or $\tau^- \rightarrow \bar{u} s \nu_\tau$. All of the other quark combinations are ruled out by energy conservation. (The $\bar{c}s$ and $\bar{c}d$ combination superficially appears to be just possible since the charm and strange quark masses sum to a value just below the tau mass. The c quark nevertheless cannot contribute to τ decays because real hadrons are bound states of these quarks and all of these bound states that contain a single charmed quark are too heavy to be produced as a tau decay product.)

- (ii) The decay rate for the τ is easily estimated given the decay rate of the muon. All of the estimates that lead to Eq. (5.5) apply equally well to tau decays and so the same result holds here. In particular the ratio of the tau decay rate to the muon decay rate must scale like the fifth power of the ratio of their masses. Using the experimentally observed muon lifetime therefore gives

$$\begin{aligned} \tau(\tau) &\sim \left(\frac{m_\mu}{m_\tau} \right)^5 \tau(\mu) \\ &\sim 1.6 \times 10^{-12} \text{ s} \end{aligned} \quad (5.7)$$

This estimate is just about right (to within the accuracy of the estimate) since the observed τ lifetime is $(2.906 \pm 0.011) \times 10^{-13}$ s. The factor of about 5 is expected because of the five allowed decay products for the τ (the $\bar{u}d$ counts as 3 because there are three available colors.) A relativistic particle with this lifetime can travel a tenth of a millimeter or more before decaying, which can be a visible displacement with the proper experimental setup.

- (iii) As is noted above, the tau meson has more decay channels open to it than does the muon just by virtue of the fact that it is so much heavier. Predictions for the τ^- branching fractions may be made simply by counting the degrees of freedom available in each channel. These predictions are quite robust since they rely on few (if any) of the details of any potentially poorly-measured parameters of the model. One of these predictions, that follows simply from the observation that the tau decays via a virtual W boson and from the universal nature of the couplings of the W , is that the branching fraction for the two leptonic decays must be equal. This and other predictions are summarized in the following.

$$\begin{aligned}
 B(\tau \rightarrow e\bar{\nu}\nu) &= B(\tau \rightarrow \mu\bar{\nu}\nu) \\
 &\approx \frac{1}{2 + 3(|V_{ud}|^2 + |V_{us}|^2)} \\
 &\approx 20\% \\
 B(\tau \rightarrow \text{strange hadrons}) &\approx B(\tau \rightarrow \bar{u}s\nu) \\
 &\approx \frac{3|V_{us}|^2}{2 + 3(|V_{ud}|^2 + |V_{us}|^2)} \\
 &\approx 2\% \\
 B(\tau \rightarrow \text{non-strange hadrons}) &\approx B(\tau \rightarrow \bar{u}d\nu) \\
 &\approx \frac{3|V_{ud}|^2}{2 + 3(|V_{ud}|^2 + |V_{us}|^2)} \\
 &\approx 58\%
 \end{aligned} \tag{5.8}$$

By way of comparison, the corresponding experimental numbers are

$$\begin{aligned}
 B_{\text{exp}}(\tau \rightarrow e\bar{\nu}\nu) &= (17.84 \pm 0.06)\% \\
 B_{\text{exp}}(\tau \rightarrow \mu\bar{\nu}\nu) &= (17.37 \pm 0.06)\% \\
 B_{\text{exp}}(\tau \rightarrow \text{strange hadrons}) &\approx (2.7 \pm 0.9)\% \\
 B_{\text{exp}}(\tau \rightarrow \text{non-strange hadrons}) &\approx (62 \pm 4)\%
 \end{aligned} \tag{5.9}$$

The agreement is within the accuracy of the estimate. The hadronic entries in the experimental table here are estimated by summing over all of the observed rates in the Particle Data Book. The errors listed in this case are rough estimates based on the combined errors for each mode. Note that our estimates for the leptonic modes are systematically low. A more detailed calculation turns out to show that the rates for the $\bar{u}d$ and $\bar{u}s$ decay modes receive a positive $O(\alpha_3)$ correction; so the difference from the naive branching fraction estimates allows a determination of the size of α_3 . The size of α_3 determined in this way differs from the determination from the width of the Z^0 boson, discussed at the end of Subsection 4.1.6; but this is expected, as we will discuss in Subsection 7.4.1, and the difference turns out to agree with the predictions of the standard model.

5.2 The calculation

Consider, for simplicity, the decay $\tau^- \rightarrow \nu_\tau \bar{f}_m f_n$ in which none of the initial or final polarizations are measured. We must evaluate the matrix element, Eq. (5.2), using the interaction Hamiltonian, Eq. (5.3). The first term in the interaction Hamiltonian is responsible for destroying the initial τ^- meson and creating the ν_τ . The second term creates the final fermion–antifermion pair, $\bar{f}_m f_n$. Since the total amplitude is the product of two powers of the interaction Hamiltonian, there are two types of contributions, corresponding to which interaction Hamiltonian destroys the τ particle and which creates the $\bar{f}_m f_n$ pair. Each of these turns out to contribute equally to the total amplitude, and so we compute here only one of them and multiply the result by two.

It is convenient to write out the action of the interaction Hamiltonian separately for the τ, ν_τ, W -boson, and $\bar{f}_m f_n$ sectors of the Hilbert space. To this end write the initial and final states as

$$\begin{aligned} |\tau^-\rangle &= |\tau^-\rangle_\tau \otimes |0\rangle_W \otimes |0\rangle_f \\ |\nu_\tau; \bar{f}_m; f_n\rangle &= |\nu_\tau\rangle_\tau \otimes |0\rangle_W \otimes |\bar{f}_m; f_n\rangle_f \end{aligned} \quad (5.10)$$

The utility of writing this dependence out explicitly is that the desired matrix element, Eq. (5.2), then factorizes into three parts, which may be dealt with separately:

$$\begin{aligned} -i\mathcal{M} &= \frac{(-i)^2}{2!} \int d^4x \langle \nu_\tau(\mathbf{l}); \bar{f}_m(\mathbf{q}); f_n(\mathbf{p}) | T [\mathcal{H}_{cc}(x) \mathcal{H}_{cc}(0)] | \tau(\mathbf{k}) \rangle \\ &= 2 \frac{(-i)^2}{2!} (-e_W^2 U_{mn}^*) \int d^4x A^\mu(\mathbf{k}, \mathbf{l}; x) G_{\mu\nu}(x) B^\nu(\mathbf{q}, \mathbf{p}) \end{aligned} \quad (5.11)$$

in which the factors A^μ , B^ν , and $G_{\mu\nu}$ are defined by

$$\begin{aligned} A^\mu(\mathbf{k}, \mathbf{l}; x) &= {}_\tau \langle \nu_\tau(\mathbf{l}) | [\bar{\nu}_\tau \gamma^\mu (1 + \gamma_5) \tau](x) | \tau(\mathbf{k}) \rangle_\tau \\ B^\nu(\mathbf{q}, \mathbf{p}) &= {}_f \langle \bar{f}_m(\mathbf{q}); f_n(\mathbf{p}) | [\bar{f}_n \gamma^\nu (1 + \gamma_5) f_m](0) | 0 \rangle_f \\ G_{\mu\nu}(x) &= {}_w \langle 0 | T [W_\mu^+(x) W_\nu^-(0)] | 0 \rangle_w \end{aligned} \quad (5.12)$$

The first factor of 2 in the last line of Eq. (5.11) is the factor of discussed in the opening paragraph of this section, which corresponds to the two ways in which the interaction terms can destroy the τ : the τ can be destroyed by the interaction at spacetime point x , or by the one at 0.

These matrix elements are evaluated by expanding each field in terms of its creation and annihilation operators and then evaluating the resulting matrix elements of these operators. The evaluation of matrix elements A^μ and B^ν only involves initial or final states and so closely parallels the evaluation of those matrix elements performed in previous chapters. They give

$$\begin{aligned} A^\mu(\mathbf{k}, \mathbf{l}; x) &= \bar{u}_\nu(\mathbf{l}) \gamma^\mu (1 + \gamma_5) u_\tau(\mathbf{k}) e^{i(k-l)x} \\ B^\nu(\mathbf{q}, \mathbf{p}) &= \bar{u}_n(\mathbf{p}) \gamma^\nu (1 + \gamma_5) v_m(\mathbf{q}) \end{aligned} \quad (5.13)$$

5.2.1 The W propagator

The matrix element $G_{\mu\nu}(x)$ of the W field operators that arises in Eq. (5.12) is called the W propagator. It is determined completely by the properties of the W bosons that contribute to it as intermediate states. Its evaluation requires a little more care and is the topic of the present aside.

$G_{\mu\nu}(x)$ may be evaluated by inserting a complete set of one-particle W -states. (For notational simplicity we drop the ubiquitous subscript “ W ” on the Hilbert-space state vectors in this section but it is implicit in all formulae.) Recalling the definition, Eq. (3.17), of the time-ordering operation, T , and inserting a complete set of 1-particle W -boson states between the operators, gives

$$\begin{aligned} G_{\mu\nu}(x) &= \langle 0 | T [W_\mu^+(x) W_\nu^-(0)] | 0 \rangle \\ &= \sum_{\lambda=-1}^1 \int \frac{d^3r}{2E_{\mathbf{r}}(2\pi)^3} \left[\langle 0 | W_\mu^+(x) | W^+(\mathbf{r}, \lambda) \rangle \langle W^+(\mathbf{r}, \lambda) | W_\nu^-(0) | 0 \rangle \theta(x^0) \right. \\ &\quad \left. + \langle 0 | W_\nu^-(0) | W^-(\mathbf{r}, \lambda) \rangle \langle W^-(\mathbf{r}, \lambda) | W_\mu^+(x) | 0 \rangle \theta(-x^0) \right] \\ &= \sum_{\lambda=-1}^1 \int \frac{d^3r}{2E_r(2\pi)^3} \left[\epsilon_\mu \epsilon_\nu^*(\mathbf{r}, \lambda) e^{ir'x} \theta(x^0) + \epsilon_\nu \epsilon_\mu^*(\mathbf{r}, \lambda) e^{-ir'x} \theta(-x^0) \right] \end{aligned} \quad (5.14)$$

Here \mathbf{r} is the momentum of the inserted state, and $E_{\mathbf{r}} = \sqrt{\mathbf{r}^2 + M_W^2}$ is its energy. The four-vector, r'_μ , that appears in the phase $e^{\pm i r'_\mu x}$ is defined with timelike component given by $r'^0 \equiv E_{\mathbf{r}}$.

$\theta(x)$ is the step function that is unity when its argument is positive and is zero when its argument is negative. For the present purposes the following integral representation proves convenient:

$$\theta(x) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{e^{ix\omega}}{\omega - i\epsilon'} d\omega \quad (5.15)$$

ϵ' here denotes a positive infinitesimal that is to be taken to zero at the end of the calculation.

The spin sum may be evaluated using Eq. (1.119):

$$\begin{aligned} \Pi_{\mu\nu}(\mathbf{r}, E_{\mathbf{r}}) &\equiv \sum_{\lambda=-1}^1 \epsilon_\mu(\mathbf{r}, \lambda) \epsilon_\nu^*(\mathbf{r}, \lambda) \\ &= \eta_{\mu\nu} + r'_\mu r'_\nu / M_W^2 \end{aligned} \quad (5.16)$$

Putting this spin sum, and Eq. (5.15), into Eq. (5.14) gives the desired expression for the propagator. After performing a change of integration variables in this result in order to put the coefficient of both of the step functions, $\theta(x^0)$ and $\theta(-x^0)$, into a common form, the result becomes

$$\begin{aligned} G_{\mu\nu}(x) &= -i \int \frac{d^3r}{(2\pi)^3} \frac{d\omega}{2\pi} \Pi_{\mu\nu}(\mathbf{r}, E_{\mathbf{r}}) e^{i\mathbf{r} \cdot \mathbf{x} - i\omega x^0} \\ &\quad \times \frac{1}{2E_{\mathbf{r}}} \left[\frac{1}{E_{\mathbf{r}} - \omega - i\epsilon'} + \frac{1}{E_{\mathbf{r}} + \omega - i\epsilon'} \right] \\ &= -i \int \frac{d^4r}{(2\pi)^4} \frac{\Pi_{\mu\nu}(\mathbf{r}, E_{\mathbf{r}})}{r^2 + M_W^2 - i\epsilon} e^{irx} \end{aligned} \quad (5.17)$$

The four-vector r_μ (as opposed to r'_μ of Eq. (5.14)) that appears in the last equality is defined with time component r^0 equal to the integration variable, ω (as opposed to $E_{\mathbf{r}}$). The infinitesimal, ϵ , appearing here has been rescaled from the original infinitesimal, ϵ' , of Eq. (5.15) by $\epsilon \equiv 2E_{\mathbf{r}} \epsilon' > 0$.

5.2.1.1 Lorentz covariance: an aside

This last expression is almost, but not quite, covariant with respect to Lorentz transformations. The qualification comes because the polarization “tensor,” $\Pi_{\mu\nu}(\mathbf{r}, E_{\mathbf{r}})$, depends on the variable $E_{\mathbf{r}}$ rather than the time component of r_μ : $r^0 = \omega$. This is something of an embarrassment since it would seem to imply a loss of Lorentz invariance for the S -matrix element that is being computed! Happily enough this particular failure of Lorentz invariance is just what is required to cancel another source that has been glossed over

up until this point. (See, however, the discussion in Section 4.1.) This other source of Lorentz non-invariance would have arisen in Eq. (5.3) if the interaction Hamiltonian had been properly identified. The result that is implicit in this equation is that the charged-current interaction Hamiltonian, \mathcal{H}_{cc} , is related to the interaction Lagrangian, \mathcal{L}_{cc} , by $\mathcal{H}_{cc} = -\mathcal{L}_{cc}$ as would usually be the case for a non-derivative interaction. This relation does not hold for the couplings of gauge potentials, however, as is perhaps more familiar in quantum electrodynamics where the non-derivative coupling, $\mathcal{L}_I = A_\mu J^\mu$, produces a non-covariant Coulomb contact interaction in the Hamiltonian.

It is beyond the scope of this book to detail how these two sources of Lorentz non-invariance cancel one another out (see, however, Problem 5.4 for an illustration of what is involved). The final result is simple to state, however. The full calculation is equivalent to neglecting the difference between \mathcal{H}_I and $-\mathcal{L}_I$ and replacing the naively time-ordered W propagator of Eq. (5.14) through Eq. (5.17) by the covariant expression obtained by replacing $\Pi_{\mu\nu}(\mathbf{r}, E_{\mathbf{r}})$ in Eq. (5.17) by $\Pi_{\mu\nu}(r) \equiv \Pi_{\mu\nu}(\mathbf{r}, r^0)$:

$$\begin{aligned}\tilde{G}_{\mu\nu}(x) &= \langle 0|T^* [W_\mu^+(x)W_\nu^-(0)]|0\rangle \\ &= -i \int \frac{d^4r}{(2\pi)^4} \Pi_{\mu\nu}(r) \frac{e^{irx}}{r^2 + M_Z^2 - i\epsilon}\end{aligned}\quad (5.18)$$

The upshot of this aside is that $\tilde{G}_{\mu\nu}(x)$ of Eq. (5.18) must be used in the amplitude of Eq. (5.11) rather than $G_{\mu\nu}(x)$.

5.3 The large-mass expansion

The results for A^μ , B^ν , and $\tilde{G}_{\mu\nu}$ accumulated above, in Eq. (5.13) and Eq. (5.18), may now be combined in Eq. (5.11) for the matrix element $\mathcal{M}(\tau^- \rightarrow \nu_\tau \bar{f}_m f_n)$. The x integral may be performed and gives a momentum conserving delta function, $\int d^4x e^{i(k-l+r)x} = (2\pi)^4 \delta^4(k-l+r)$, giving the following result for \mathcal{M} :

$$\begin{aligned}\mathcal{M}(\tau \rightarrow \nu_\tau \bar{f}_m f_n) &= e_W^2 U_{mn}^* [\bar{u}_\nu(\mathbf{l})\gamma^\mu(1+\gamma_5)u_\tau(\mathbf{k})][\bar{u}_n(\mathbf{p})\gamma^\nu(1+\gamma_5)v_m(\mathbf{q})] \\ &\quad \times \left[\frac{\eta_{\mu\nu} + (k-l)_\mu(k-l)_\nu/M_W^2}{(k-l)^2 + M_W^2 - i\epsilon} \right]\end{aligned}\quad (5.19)$$

All of the techniques of the previous sections may be brought to bear on this expression to evaluate the corresponding τ^- decay rate. A great simplification is possible at this point, however, if it is agreed to neglect any contributions that are suppressed relative to the dominant one by powers of the small quantity $m_\tau^2/M_W^2 \approx 5 \times 10^{-4}$. (The approximation is even better

for the muon where this ratio is 300 times smaller.) In this case the entire W -boson propagator, as represented by the last square bracket in Eq. (5.19), may be expanded in inverse powers of M_W^2 :

$$\frac{\eta_{\mu\nu} + (k-l)_\mu(k-l)_\nu/M_W^2}{(k-l)^2 + M_W^2 - i\epsilon} \approx \frac{\eta_{\mu\nu}}{M_W^2} \quad (5.20)$$

since in the rest frame of the tau meson all of the components of the four-momentum, $k-l$, are at most equal to m_τ .

Within this approximation, the matrix element simplifies to

$$\mathcal{M}(\tau \rightarrow \nu_\tau \bar{f}_m f_n) = \frac{e_W^2 U_{mn}^*}{M_W^2} [\bar{u}_\nu(\mathbf{l}) \gamma^\mu (1 + \gamma_5) u_\tau(\mathbf{k})] [\bar{u}_n(\mathbf{p}) \gamma_\mu (1 + \gamma_5) v_m(\mathbf{q})] \quad (5.21)$$

It is conventional to denote the coupling combination e_W^2/M_W^2 , that appears in this expression, by $G_F/\sqrt{2}$. i.e.

$$\frac{G_F}{\sqrt{2}} = \frac{e_W^2}{M_W^2} = \frac{g_2^2}{8M_W^2} = \frac{1}{2v^2} \quad (5.22)$$

The constant $G_F = 1.1664 \times 10^{-5} \text{ (GeV)}^{-2}$ obtained in this way is for historical reasons called the *Fermi coupling constant*. Indeed, it is the measurement of G_F through comparison of the predicted and measured muon lifetimes that fixes the value of the Higgs *v.e.v.*, v , that is quoted in Appendix A.

Returning to the matrix element, Eq. (5.21), averaging over the two initial spin states of the τ and summing over spins of the final fermions gives

$$\begin{aligned} \overline{\mathcal{M}^2} &= \frac{1}{2} \sum_{\text{spins}} |M(\tau \rightarrow \nu_\tau \bar{f}_m f_n)|^2 \\ &= \frac{G_F^2}{4} |U_{mn}|^2 M^{\mu\nu}(k, l) N_{\mu\nu}(p, q) \end{aligned} \quad (5.23)$$

in which the quantities $M^{\mu\nu}$ and $N^{\mu\nu}$ denote traces over Dirac matrices:

$$\begin{aligned} M^{\mu\nu}(k, l) &\equiv \text{tr}[\gamma^\mu (1 + \gamma_5) u_\tau \bar{u}_\tau \gamma^\nu (1 + \gamma_5) u_\nu \bar{u}_\nu] \\ &= \text{tr}[\gamma^\mu (1 + \gamma_5) (m_\tau - i\not{k}) \gamma^\nu (1 + \gamma_5) (-i\not{l})] \\ &= 8 \left[(\eta^{\mu\nu} k \cdot l - k^\mu l^\nu - k^\nu l^\mu) - i\epsilon^{\mu\nu\lambda\rho} k_\lambda l_\rho \right] \end{aligned} \quad (5.24)$$

and

$$\begin{aligned} N_{\mu\nu}(p, q) &\equiv \text{tr}[\gamma_\mu (1 + \gamma_5) v_m \bar{v}_m \gamma_\nu (1 + \gamma_5) u_n \bar{u}_n] \\ &= -\text{tr}[\gamma^\mu (1 + \gamma_5) (m_m + i\not{q}) \gamma^\nu (1 + \gamma_5) (m_n - i\not{p})] \\ &= 8 \left[(\eta^{\mu\nu} p \cdot q - p^\mu q^\nu - p^\nu q^\mu) - i\epsilon^{\mu\nu\lambda\rho} q_\lambda p_\rho \right] \end{aligned} \quad (5.25)$$

Contracting $M^{\mu\nu}$ with $N_{\mu\nu}$ (using the identity $\epsilon^{\mu\nu\alpha\beta}\epsilon_{\mu\nu\lambda\rho} = 2(\delta_\rho^\alpha\delta_\lambda^\beta - \delta_\lambda^\alpha\delta_\rho^\beta)$) finally gives the simple result,

$$M^{\mu\nu}(k, l)N_{\mu\nu}(q, p) = 256(l \cdot p)(k \cdot q) \quad (5.26)$$

Combining all of the above formulae gives the following differential decay rate:

$$d\Gamma(\tau \rightarrow \nu_\tau \bar{f}_m f_n) = \frac{64G_F^2 |U_{mn}|^2}{2k^0} (l \cdot p)(k \cdot q) (2\pi)^4 \delta^4(p + q + l - k) \frac{d^3l d^3p d^3q}{8l^0 p^0 q^0 (2\pi)^9} \quad (5.27)$$

This expression must now be integrated over phase space to produce the desired differential or total decay rate. We will perform the integration assuming that both the ν_τ and one of the other leptons is massless. This treatment is relevant for $\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$ and for $\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu$, and is not too bad for decays of ν_τ into quarks (where g_3^2 effects which we will not compute are anyway more important than the up quark mass).

Consider for concreteness the case of a purely leptonic τ decay: $\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$. We now compute this differential decay rate as a function of the final muon's energy and mass.

Since we do not observe the neutrino momenta l and q , it is convenient to integrate over them first. Moving everything else outside the q, l integrations leaves the integral,

$$I_{\mu\nu}(k, p) \equiv \int l_\mu q_\nu (2\pi)^4 \delta^4(l + p + q - k) \frac{d^3l d^3q}{4l^0 q^0 (2\pi)^6} \quad (5.28)$$

which we must evaluate. $I_{\mu\nu}(k, p)$ as defined is a second-rank tensor that is a function of k^μ and p^μ only through the combination $w^\mu \equiv (k - p)^\mu$. The most general possible form for such a tensor is

$$I_{\mu\nu}(w) = Aw^2 \eta_{\mu\nu} + Bw_\mu w_\nu \quad (5.29)$$

in which A and B can a priori be arbitrary functions of the Lorentz-invariant variable $w^2 = w^\mu w_\mu$. The scalar variables A and B are much more convenient to deal with than is the full tensor $I_{\mu\nu}$, since they are Lorentz-*invariant* (as opposed to being *covariant*), and so may be simply evaluated in the most convenient reference frame.

A and B turn out to be both determined in terms of a single scalar integral

$$I(w) \equiv \int (2\pi)^4 \delta^4(l + q - w) \frac{d^3l d^3q}{4l^0 q^0 (2\pi)^6} \quad (5.30)$$

To see this, note first that

$$\begin{aligned}
 \eta^{\mu\nu} I_{\mu\nu}(w) &= Aw^2 \eta^{\mu\nu} \eta_{\mu\nu} + B \eta^{\mu\nu} w_\mu w_\nu = (4A + B)w^2 \\
 &= \int l \cdot q (2\pi)^4 \delta^4(l + q - w) \frac{d^3 l d^3 q}{4l^0 q^0 (2\pi)^6} \\
 &= \frac{w^2}{2} I(w)
 \end{aligned} \tag{5.31}$$

In the last equality the identity $w^2 = (l + q)^2 = 2l \cdot q$ is used, which relies on four-momentum conservation as well as the masslessness of the neutrinos, $l^2 = q^2 = 0$. Similarly,

$$\begin{aligned}
 w^\mu w^\nu I_{\mu\nu}(w) &= Aw^2 w^\mu w^\nu \eta_{\mu\nu} + B w^\mu w^\nu w_\mu w_\nu = (A + B)(w^2)^2 \\
 &= \int (w \cdot q)(w \cdot l) (2\pi)^4 \delta^4(l + q - w) \frac{d^3 l d^3 q}{4l^0 q^0 (2\pi)^6} \\
 &= \int (l \cdot q)^2 (2\pi)^4 \delta^4(l + q - w) \frac{d^3 l d^3 q}{4l^0 q^0 (2\pi)^6} \\
 &= \frac{(w^2)^2}{4} I(w)
 \end{aligned} \tag{5.32}$$

which uses $w \cdot l = (l + q) \cdot l = l \cdot q = w \cdot q$. These equations may be solved for A and B , giving $B = 2A = I/6$.

It remains to evaluate $I(w)$. It is first convenient to perform the l integration using the following identity, which holds for any Lorentz-invariant integrand:

$$\int \frac{d^3 l}{2l^0 (2\pi)^3} = \int \frac{d^4 l}{(2\pi)^4} 2\pi \delta(l^2 + m^2) \theta(l^0) \tag{5.33}$$

for our case, $m^2 = 0$. Then

$$\begin{aligned}
 I(w) &= \int (2\pi)^4 \delta^4(l + q - w) 2\pi \delta(l^2) \theta(l^0) \frac{d^4 l}{(2\pi)^4} \frac{d^3 q}{2q^0 (2\pi)^3} \\
 &= \int 2\pi \delta[(w - q)^2] \theta[w^0 - q^0] \frac{d^3 q}{2q^0 (2\pi)^3}
 \end{aligned} \tag{5.34}$$

Notice that the four-vector, w^μ , is timelike, since

$$w^2 = (k - p)^2 = (l + q)^2 = 2l \cdot q = -2|\mathbf{q}||\mathbf{l}|(1 - \cos \theta) \leq 0 \tag{5.35}$$

θ here represents the angle between the three-vectors \mathbf{l} and \mathbf{q} . There must therefore be a reference frame in which the three-vector components of w vanish, $\mathbf{w} = \mathbf{0}$. Define $w^0 = E$ in this frame. Then, also in the same frame, $w^2 = -E^2$ and $(w - q)^2 = w^2 - 2w \cdot q = -E^2 + 2Eq^0$. The last integral is

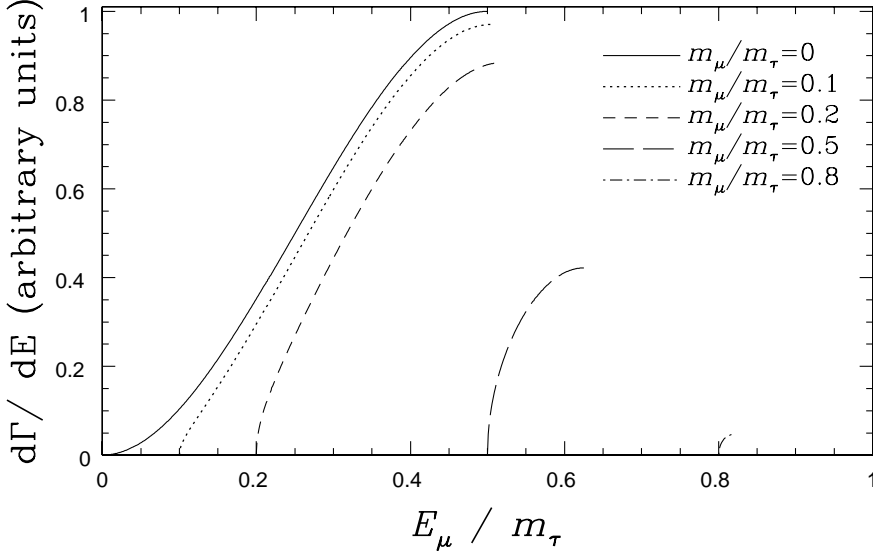


Fig. 5.1 Differential τ, μ decay rate, as function of the charged-lepton energy

most conveniently evaluated in this frame:

$$\begin{aligned} I(w) &= \frac{1}{2\pi} \int \delta[E^2 - 2Eq^0] \theta[E - q^0] q^0 dq^0 \\ &= \frac{1}{8\pi} \theta(-w^2) \end{aligned} \quad (5.36)$$

where the θ function is because the integration only has support provided w is timelike. Clearly, then, $B = 2A = \Theta(-w^2)/(48\pi)$, and

$$I_{\mu\nu}(w) = \frac{1}{96\pi} [\eta_{\mu\nu} w^2 + 2w_\mu w_\nu] \theta(-w^2) \quad (5.37)$$

Inserting this integral into the decay rate, Eq. (5.27), finally gives the differential decay rate as a function of the muon energy and mass (normalized to the tau mass so that $\varepsilon \equiv p^0/m_\tau$ and $\mu \equiv m_\mu/m_\tau$):

$$\begin{aligned} \frac{d\Gamma}{d\varepsilon}(\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau) &= \frac{G_F^2 m_\tau^5}{4\pi^3} \left(\varepsilon - \frac{4\varepsilon^2}{3} + \varepsilon\mu^2 - \frac{2\mu^2}{3} \right) \sqrt{\varepsilon^2 - \mu^2} \\ &\approx \frac{G_F^2 m_\tau^5}{4\pi^3} \varepsilon^2 \left(1 - \frac{4\varepsilon}{3} \right) \quad \text{for } \mu \ll 1 \end{aligned} \quad (5.38)$$

The shape of this curve as a function of ε is plotted in Figure 5.1.

The kinematically allowed range for the muon energy is $0 < p^0 < (m_\tau^2 + m_\mu^2)/(2m_\tau)$. This may be seen from the condition that $k - p$ be timelike as seen in the tau rest frame: $(k - p)^2 = k^2 - 2k \cdot p + p^2 = -m_\tau^2 + 2m_\tau p^0 -$

$m_\mu^2 < 0$. Integrating p^0 over this range then gives the total rate for τ^- decays into this channel (neglecting all fermion masses):

$$\Gamma(\tau^- \rightarrow \nu_\tau \bar{f}_m f_n) = \frac{G_F^2 m_\tau^5}{192\pi^3} |U_{mn}|^2 \quad (5.39)$$

Some final comments about this result.

- (i) Equation (5.39), applied to muon decay, may be compared to the estimate of Section 5.1 to see if the discrepancy of this estimate with the experimental value persists or is instead an artifact of the inaccuracy of the earlier estimate. The full and approximate results are

$$\begin{aligned} \Gamma_{\text{calc}}(\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu) &= \frac{G_F^2 m_\mu^5}{192\pi^3} \\ &= \frac{g_2^4}{3 \cdot 2^9 \cdot (2\pi)^3} \frac{m_\mu^5}{M_W^4} \\ \text{and } \Gamma_{\text{est}}(\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu) &\sim \frac{g_2^4}{(2\pi)^3} \frac{m_\mu^5}{M_W^4} \end{aligned} \quad (5.40)$$

The approximate estimate has missed the numerical factor of $2^{-9} \cdot 3^{-1} = 1/1536$, which provides the missing three orders of magnitude. This illustrates a general *caveat* for the order-of-magnitude estimates: they are useful for judging the rough size of a rate but are not a substitute for a real calculation. The source of this large number in the full calculation is in the integration of the squared amplitude over phase space. The estimate of this integration using simple dimensional analysis and counting of 2π s is the weakest part of the arguments of Section 5.1. Although it furnishes reasonable accuracy for the two-body decays of the previous sections, it can be potentially more of a problem in decays which involve more final-state particles, since the rates for these processes involve a multidimensional integration over phase space.

The full result, Eq. (5.39), gives the following μ^- and τ^- decay rates into leptons:

$$\begin{aligned} \Gamma_{\text{tot}}(\mu^-) &= \Gamma(\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu) \\ &= \frac{G_F^2 m_\mu^5}{192\pi^3} \\ &= 3.009 \times 10^{-19} \text{ GeV} \\ \text{so } \tau(\mu^-) &= 2.187 \times 10^{-6} \text{ s} \end{aligned} \quad (5.41)$$

and

$$\begin{aligned}
 \Gamma_{\text{tot}}(\tau^-) &= \left[2 + 3(|V_{ud}|^2 + |V_{us}|^2) \right] \Gamma(\tau^- \rightarrow e^- \bar{\nu}_e \nu_\mu) \\
 &= 5 \cdot \frac{G_F^2 m_\tau^5}{192\pi^3} \\
 &= 2.025 \times 10^{-12} \text{ GeV} \\
 \text{so } \tau(\tau^-) &= 3.25 \times 10^{-13} \text{ s}
 \end{aligned} \tag{5.42}$$

Corrections to the muon lifetime should be smaller than a percent or so since they are purely electromagnetic. The corrections to the τ lifetime might be somewhat larger since they can include strong-interaction corrections for the hadronic final states. For comparison, the measured lifetimes are

$$\begin{aligned}
 \tau_{\text{exp}}(\mu^-) &= (2.197\,03 \pm 0.000\,04) \times 10^{-6} \text{ s} \\
 \tau_{\text{exp}}(\tau^-) &= (2.906 \pm 0.011) \times 10^{-13} \text{ s}
 \end{aligned} \tag{5.43}$$

The comparison for the μ^- is not really fair, since the value of G_F is determined from this width. However, with G_F so determined, the τ^- width is fair game. Again we emphasize that the τ^- width is not that close to the prediction (10% discrepancy); this is because of strong interaction physics. The partial widths to leptons are in very good agreement with the predictions of our formulae.

- (ii) The shape of the differential decay probability, $d\Gamma/dE$, for decays into leptons as a function of the charged-lepton energy, E , is given in Figure 5.1. It vanishes like E^2 as $E \rightarrow 0$ and rises monotonically to a maximum at the *endpoint*, i.e. the maximum energy that is kinematically available to the charged-lepton (roughly half of the mass of the decaying particle in the present case). The most probable energy for the outgoing charged-lepton is therefore its endpoint value. The E^2 -dependence for small E is also easily understood. It arises partly from the phase space measure for relativistic fermions, $d^3p/p^0 \approx E dE$, and partly from the proportionality to E of the squared matrix element, $\overline{\mathcal{M}}^2$ of Eq. (5.27).

5.4 Feynman rules

The general pattern for the perturbative evaluation of general scattering amplitudes and decay rates is similar to the examples that have been encountered up to this point. In each case the desired matrix element is found by expressing it in terms of the fields of the theory which are then themselves

expressed in terms of the corresponding creation and annihilation operators. The resulting matrix elements of these operators may then be evaluated by applying the rules of Section 1.1.

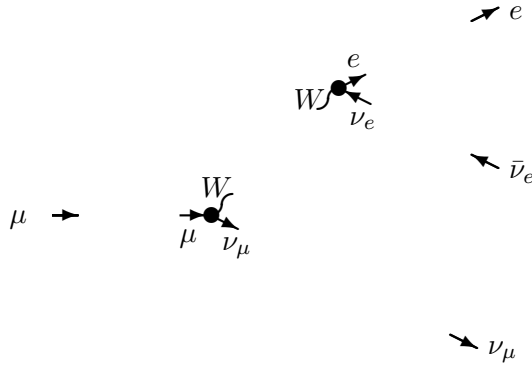
In each case we pick up standard factors corresponding to the polarization vectors or spinors in the expansions of the fields, and to the coupling constants and Dirac matrices of the interaction Lagrangians. These rules may be very graphically summarized in a way that allows the corresponding amplitude to be straightforwardly written down. The procedure is to associate a line to the propagation of every particle in a particular matrix element. These lines end whenever the corresponding particle is created or destroyed by one of the creation or annihilation operators of the matrix element of interest. The lines drawn in this way form a *Feynman graph* (or *Feynman diagram*) that is associated with the given matrix element.

Explicitly, the relation between the procedure we have followed so far, and the drawing of a Feynman graph, is as follows. First, one chooses the initial and final states under consideration. For the case of μ^- decay, this consisted of a μ^- in the initial state, $|\mu^- \rangle$, and $e^-, \bar{\nu}_e, \nu_\mu$ in the final state, $\langle e^- \bar{\nu}_e \nu_\mu |$. These are represented graphically by putting the end of a line on the left-hand side of the graph for each initial state particle and the end of a line on the right-hand side for each final state particle, labeled with the particle type:



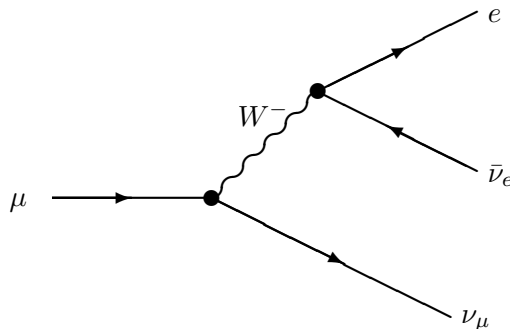
Fermions are given arrows pointing right for particles and left for antiparticles. Some people let time flow from the bottom to the top of the figure rather than from the left to the right.

Next, one must determine what insertions of the interaction Hamiltonian are involved. Each H_I insertion is represented by a dot (vertex) with the stubs of lines coming out, labeled as dictated by the fields appearing in H_I . Fermion fields are given an arrow entering the vertex if the field operator e, μ, ν is involved and an arrow leaving the vertex if the antifield operator $\bar{e}, \bar{\mu}, \bar{\nu}$ is involved. For the case of μ^- decay, the graph is now



The locations on the page of the vertices in the graph are arbitrary, and are usually chosen so that the final graph will look nice.

Finally, each field in an H_I insertion provides either a creation or an annihilation operator. These must be contracted either with initial or final state particles, or with each other. When an annihilation operator destroys an initial-state particle, a line is drawn between the incoming line end and the line stub on that vertex; similarly with creation operators and final state particles. These lines are called *incoming lines* and *outgoing lines*, respectively. Two fields in H_I insertions which combine to form a propagator are represented in the graph by attaching the line stubs on the vertices with an *internal line*. In all cases the line ends connected by a line must be of the same particle type, and for fermions they must have compatible arrow directions. By convention fermions are drawn with solid lines, electroweak gauge bosons are wiggly lines, and gluons are curly lines. We will draw scalars with dashed lines – conventions here are less uniform. The result is that the diagram corresponding to the μ^- decay process we have analyzed is



We emphasize again that the exact location on the page of the vertices and lines is arbitrary and is usually chosen to make the picture easy to draw.

Feynman graphs with a given initial and final state, and sets of H_I insertions and creation- and annihilation-operator pairings which can induce the transition from the initial and final state, are in one-to-one correspondence. Drawing Feynman graphs represents a particularly efficient and visual way of finding the possible ways in which an initial state can become a final state. Therefore, the possible processes contributing to a matrix element \mathcal{M} for a process involving a given initial and final state may be found by drawing all possible Feynman graphs which have external lines appropriate to the initial and final states, and involving the vertices corresponding to the interactions of the theory of interest. Furthermore, \mathcal{M} itself can be found from the graphs; it is the sum over each graph of an expression which can be determined by replacing each element (line and vertex) of the graph with an expression determined by the *Feynman rules* of the theory. Each graph must also be multiplied by a symmetry factor, which is precisely the graph theoretic symmetry factor of the Feynman graph.

We first present a table of the *Feynman rules* of the standard model. That is, we present a list of the proper factors that should be associated with each internal line, external line, and vertex in order for a graph to reproduce the corresponding standard-model amplitude. We then consider two examples for which the operator calculation has been done in earlier sections in order to illustrate how the graphical method correctly reproduces matrix elements.

5.4.1 External lines

This section lists the factors that are associated with the external lines of a Feynman graph – and so the initial or final particles of an amplitude.

5.4.1.1 Incoming lines (initial states)

The following factors give the (momentum-space) Feynman rules appropriate to an incoming spin-zero, spin-half, or spin-one particle. The arrows on the fermion lines indicate the direction of fermion-number flow, the dot indicates where the line attaches to an interaction vertex.

Spin-zero

$$\text{---} \text{---} \text{---} \text{---} \text{---} \bullet \qquad 1 \qquad (5.44)$$

Spin-half fermion

$$\text{---} \longrightarrow \bullet \qquad u_i(\mathbf{p}, \sigma) \qquad (5.45)$$

Spin-half antifermion


 $\bar{v}_i(\mathbf{p}, \sigma) \quad (5.46)$

Spin-one


 $\epsilon_\mu(\mathbf{p}, \sigma) \quad (5.47)$

5.4.1.2 Outgoing lines (final states)

The momentum-space Feynman rules for an outgoing spin-zero, spin-half, or spin-one particle are similarly

Spin-zero


 $1 \quad (5.48)$

Spin-half fermion


 $\bar{u}_i(\mathbf{p}, \sigma) \quad (5.49)$

Spin-half antifermion


 $v_i(\mathbf{p}, \sigma) \quad (5.50)$

Spin-one


 $\epsilon_\mu^*(\mathbf{p}, \sigma) \quad (5.51)$

These are the factors needed to compute the matrix element \mathcal{M} . When integrating \mathcal{M}^2 over final state momenta one must use the phase space measure $d^3p/(2p^0[2\pi]^3)$, and there is a factor of $1/(2p^0)$ from the state normalization of each incoming particle.

5.4.2 Internal lines

The momentum space description for an internal line is given by the *propagator* for the corresponding particle. The propagators for the three lowest-spin particles of interest are listed here.

Notice that the Dirac indices on the spin-half propagator are such that Dirac matrix multiplication orders propagators and vertices *oppositely* to the order they would have if they are ordered consecutively along a fermion line in the direction of fermion-number flow. Here and in the expressions for vertices, when some index (color, for instance) is not explicitly displayed, it is contracted with a δ function between the lines which carry that index.

The spin-one propagator given here depends on whether the spin-one particle has a mass or not. For massive spin-one particles *unitary gauge* is presented as was used for the W and Z bosons in the text. This propagator is less useful for massless particles like the photon or gluons, however, since it has a singular limit as the particle mass tends to zero. For massless particles we use instead the propagator in what is called the renormalizable ξ *gauge*. We are free to choose a gauge that is different from unitary gauge for the photon and gluons because the unitary gauge condition, Eq. (2.30), does *not* fix the electromagnetic or $SU_c(3)$ gauge invariance. Notice that the ξ -gauge propagator tends to the unitary gauge result as ξ tends to infinity. The special cases $\xi = 1$ and $\xi = 0$ carry the special names of *Feynman gauge* and *Landau gauge* respectively.

ξ gauge is also useful for massive spin-one particles in higher-loop calculations because of its better behavior as $p^2 \rightarrow \infty$. In this case there are extra Feynman rules involving “unphysical scalars” and “ghosts” that must also be included. As these are not necessary for the tree-level computations that are encountered in this book, a description of the full ξ -gauge Feynman rules are reserved for Appendix D.

Spin-zero

$$\bullet \text{---} \text{---} \text{---} \text{---} \bullet \quad G(p) = -i \int \frac{d^4 p}{(2\pi)^4} \frac{1}{p^2 + m^2 - i\epsilon} \quad (5.52)$$

Spin-half

$$j \bullet \longrightarrow \bullet i \quad G_{ij}(p) = -i \int \frac{d^4 p}{(2\pi)^4} \left[\frac{-i\not{p} + m}{p^2 + m^2 - i\epsilon} \right]_{ij} \quad (5.53)$$

Spin-one (unitary gauge)

$$\mu \bullet \text{~~~~~} \bullet \nu \quad G_{\mu\nu}(p) = -i \int \frac{d^4 p}{(2\pi)^4} \frac{\eta_{\mu\nu} + \frac{p_\mu p_\nu}{m^2}}{p^2 + m^2 - i\epsilon} \quad (5.54)$$

Spin-one (ξ gauge)

$$\mu \bullet \text{~~~~~} \bullet \nu \quad G_{\mu\nu}(p) = -i \int \frac{d^4 p}{(2\pi)^4} \frac{\eta_{\mu\nu} + (\xi - 1) \frac{p_\mu p_\nu}{p^2 + \xi m^2}}{p^2 + m^2 - i\epsilon} \quad (5.55)$$

5.4.3 Vertices

The Feynman rules that differentiate the standard model from any other theory of interacting spin-one, spin-half, and spin-zero particles are those

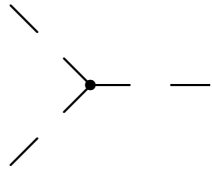
that describe the *vertices* or interactions of the theory. There is a separate vertex for each type of interaction that is given in Section 2.4.

They are all tabulated here for completeness. All momenta are taken as being directed into the vertex. We do not include labels whenever they are connected by a delta function in an obvious way (for instance, color indices in the $Hf\bar{f}$ coupling), or for momentum assignments when the Feynman rule does not depend on them in a complicated way.

Some numerical coefficients in denominators in the following expressions are printed in boldface. We do this for diagrams where there are always multiple ways to attach the external lines to the vertex; in practice, these factors are almost always canceled by the combinatorics of the number of ways to form a graph. Many other references absorb these factors into the computation of the combinatorial factor for the diagram.

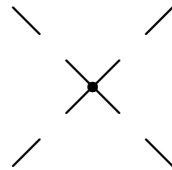
5.4.3.1 Higgs self-couplings

H^3 coupling



$$\left(-3i \frac{m_H^2}{\mathbf{6}v} \right) (2\pi)^4 \delta^4(k + l + p) \quad (5.56)$$

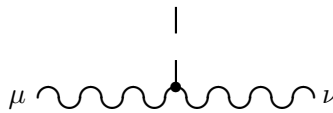
H^4 coupling



$$\left(-3i \frac{m_H^2}{\mathbf{24}v^2} \right) (2\pi)^4 \delta^4(k + l + p + q) \quad (5.57)$$

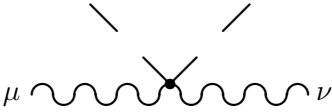
5.4.3.2 Higgs–gauge boson couplings

HW^+W^- coupling



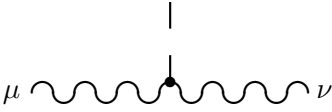
$$\left(-2i \frac{M_W^2}{v} \right) \eta_{\mu\nu} (2\pi)^4 \delta^4(k + l + p) \quad (5.58)$$

$H^2 W^+ W^-$ coupling



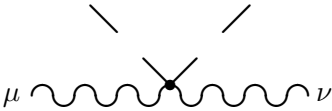
$$\left(-2i \frac{M_W^2}{2v^2} \right) \eta_{\mu\nu} (2\pi)^4 \delta^4(k + l + p + q) \quad (5.59)$$

 $H Z^2$ coupling



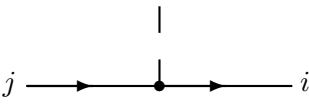
$$\left(-2i \frac{M_Z^2}{2v} \right) \eta_{\mu\nu} (2\pi)^4 \delta^4(k + l + p) \quad (5.60)$$

 $H^2 Z^2$ coupling



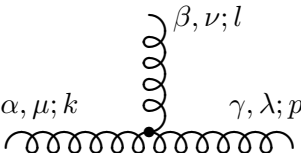
$$\left(-2i \frac{M_Z^2}{4v^2} \right) \eta_{\mu\nu} (2\pi)^4 \delta^4(k + l + p + q) \quad (5.61)$$

5.4.3.3 Higgs-fermion couplings **$H f \bar{f}$ coupling**



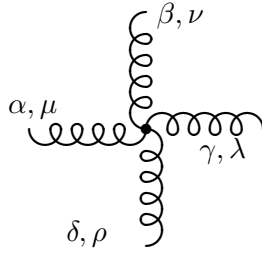
$$\left(-i \frac{m_f}{v} \right) \delta_{ij} (2\pi)^4 \delta^4(k + l + p) \quad (5.62)$$

5.4.3.4 Gluon self-couplings **G^3 coupling**



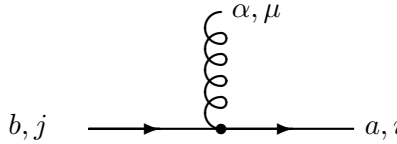
$$+ \frac{g_3}{6} f_{\alpha\beta\gamma} [(k - p)_\nu \eta_{\mu\lambda} + (l - k)_\lambda \eta_{\mu\nu} + (p - l)_\mu \eta_{\nu\lambda}] (2\pi)^4 \delta^4(k + l + p) \quad (5.63)$$

G^4 coupling



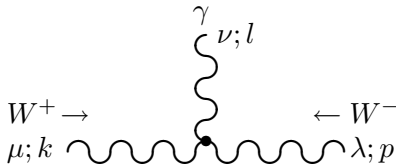
$$\begin{aligned}
 & -i \frac{g_3^2}{24} \left[f_{\xi\alpha\beta} f_{\xi\gamma\delta} (\eta_{\mu\lambda} \eta_{\nu\rho} - \eta_{\mu\rho} \eta_{\nu\lambda}) \right. \\
 & \quad + f_{\xi\alpha\gamma} f_{\xi\beta\delta} (\eta_{\mu\nu} \eta_{\lambda\rho} - \eta_{\mu\rho} \eta_{\nu\lambda}) \\
 & \quad \left. + f_{\xi\alpha\delta} f_{\xi\beta\gamma} (\eta_{\mu\nu} \eta_{\lambda\rho} - \eta_{\nu\rho} \eta_{\mu\lambda}) \right] \\
 & \quad \times (2\pi)^4 \delta^4(k + l + p + q) \quad (5.64)
 \end{aligned}$$

5.4.3.5 Gluon-fermion couplings **$Gf\bar{f}$ coupling**



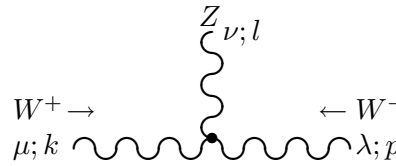
$$-\frac{g_3}{2} (\lambda_\alpha)_{ab} (\gamma^\mu)_{ij} (2\pi)^4 \delta^4(k + l + p) \quad (5.65)$$

5.4.3.6 Electroweak boson self-couplings **$W^+W^-\gamma$ coupling**



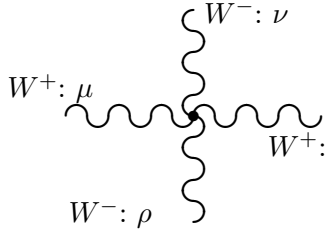
$$\begin{aligned}
 & ie [(k-p)_\nu \eta_{\mu\lambda} + (l-k)_\lambda \eta_{\mu\nu} + (p-l)_\mu \eta_{\nu\lambda}] \\
 & \quad \times (2\pi)^4 \delta^4(k + l + p) \quad (5.66)
 \end{aligned}$$

 W^+W^-Z coupling



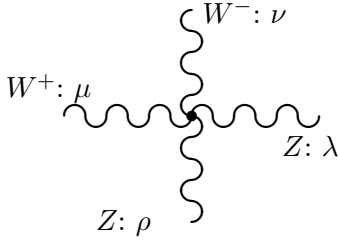
$$\begin{aligned}
 & ie \cot \theta_W [(k-p)_\nu \eta_{\mu\lambda} + (l-k)_\lambda \eta_{\mu\nu} \\
 & \quad + (p-l)_\mu \eta_{\nu\lambda}] (2\pi)^4 \delta^4(k + l + p) \quad (5.67)
 \end{aligned}$$

$W^+W^-W^+W^-$ coupling



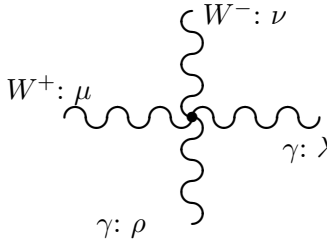
$$\frac{ig_2^2}{4} [2\eta_{\mu\lambda}\eta_{\nu\rho} - \eta_{\mu\rho}\eta_{\nu\lambda} - \eta_{\mu\nu}\eta_{\lambda\rho}] \times (2\pi)^4 \delta^4(k+l+p+q) \quad (5.68)$$

 $W^+W^-Z^2$ coupling



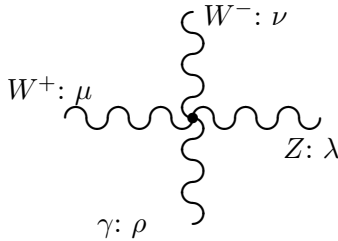
$$-\frac{ie^2 \cot^2 \theta_W}{2} [2\eta_{\mu\nu}\eta_{\lambda\rho} - \eta_{\mu\rho}\eta_{\nu\lambda} - \eta_{\mu\lambda}\eta_{\nu\rho}] \times (2\pi)^4 \delta^4(k+l+p+q) \quad (5.69)$$

 $W^+W^-\gamma^2$ coupling



$$-\frac{ie^2}{2} [2\eta_{\mu\nu}\eta_{\lambda\rho} - \eta_{\mu\rho}\eta_{\nu\lambda} - \eta_{\mu\lambda}\eta_{\nu\rho}] \times (2\pi)^4 \delta^4(k+l+p+q) \quad (5.70)$$

 $W^+W^-Z\gamma$ coupling



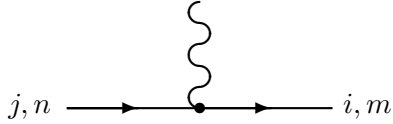
$$-ie^2 \cot \theta_W [2\eta_{\mu\nu}\eta_{\lambda\rho} - \eta_{\mu\rho}\eta_{\nu\lambda} - \eta_{\mu\lambda}\eta_{\nu\rho}] \times (2\pi)^4 \delta^4(k+l+p+q) \quad (5.71)$$

5.4.3.7 Fermion electroweak couplings

The coupling combinations e_W and e_Z used here are the same as those used elsewhere in the text: $e_Z = e/(\sin \theta_W \cos \theta_W) = \sqrt{g_1^2 + g_2^2}$ and $e_W =$

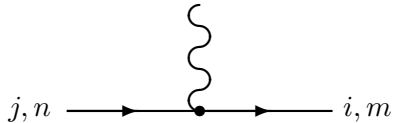
$e/(2\sqrt{2}\sin\theta_W) = g_2/(2\sqrt{2})$. g_V and g_A are the quantum number combinations $g_V = \frac{1}{2}T_3 - Q\sin^2\theta_W$ and $g_A = \frac{1}{2}T_3$, listed explicitly in Table 2.1 and Table 4.1. The matrix U_{mn} is the unit matrix δ_{mn} when “ m ” and “ n ” are leptons and is the Kobayashi–Maskawa matrix, V_{mn} , when they are quarks.

$W^+ f_n \bar{f}_m$ coupling



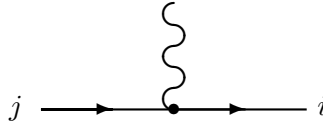
$$- e_W U_{mn} [\gamma^\mu (1 + \gamma_5)]_{ij} (2\pi)^4 \delta^4(k + l + p) \quad (5.72)$$

$W^- f_n \bar{f}_m$ coupling



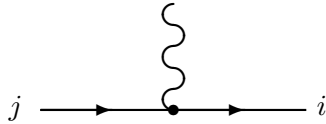
$$- e_W U_{nm}^* [\gamma^\mu (1 + \gamma_5)]_{ij} (2\pi)^4 \delta^4(k + l + p) \quad (5.73)$$

$Z f \bar{f}$ coupling



$$- e_Z [\gamma^\mu (g_V + g_A \gamma_5)]_{ij} (2\pi)^4 \delta^4(k + l + p) \quad (5.74)$$

$\gamma f \bar{f}$ coupling



$$- e Q_f [\gamma^\mu]_{ij} (2\pi)^4 \delta^4(k + l + p) \quad (5.75)$$

5.4.4 The rules

These rules allow a Feynman graph to be converted into an S -matrix element by the following steps.

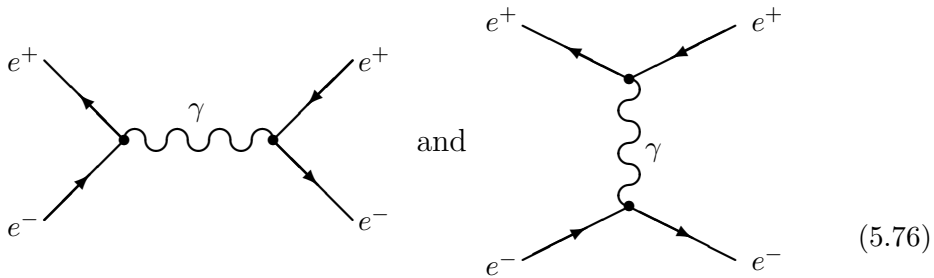
- (i) Draw all graphs that can connect the desired initial and final states using only those vertices that can contribute to the order of perturbation theory that is desired – recall that each vertex is proportional to the coupling constant of the corresponding interaction so each additional vertex costs extra powers of the couplings.
- (ii) For each graph, replace each internal line, external line, and vertex by the expression given above.

- (iii) Integrate the result over the four-momentum flowing through all of the internal lines (corresponding to summing over all virtual intermediate states), and sum over all Dirac and Lorentz indices.
- (iv) If the graph contains n vertices then divide its contribution by $n!$ (c.f. the denominator of Eq. (3.22).) If there are p distinct ways of forming the given graph using the same set of interactions and initial and final states, then multiply the contribution of the graph by p . The product of these factors is called the *symmetry factor* of the graph.
- (v) Multiply the result by a factor of -1 for each closed fermion (or ghost, see Appendix D) loop in the graph.
- (vi) When comparing different matrix elements for the same process, there can be a relative minus sign if the fermion lines connect together differently (see Section 6.6).
- (vii) To convert the resulting S -matrix element to a matrix element \mathcal{M} , multiply by i and remove the overall energy-momentum conserving delta function $(2\pi)^4\delta^4(p_\alpha - p_\beta)$ (with p_α and p_β the sums over all incoming and outgoing momenta, respectively).

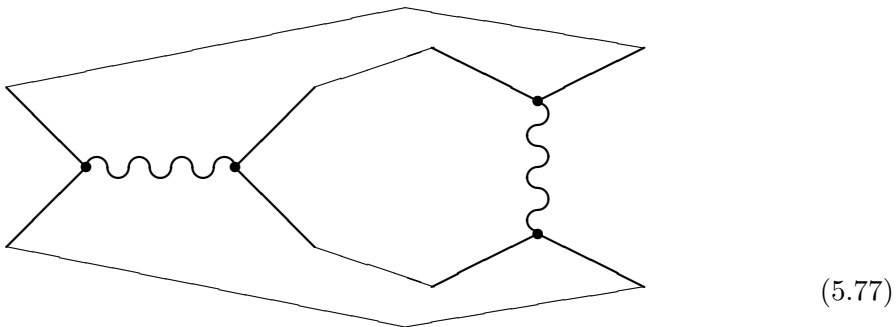
A few of these items demand some clarification. To clarify item (iv): if a graph contains, say, two vertices and they are of distinct types, for instance, a $W^+\bar{\nu}_\mu\mu$ vertex and a $W^-\bar{e}\nu_e$ vertex, then there is a factor of $1/2!$ from the multiple vertices and a factor of 2 from the different choices of which vertex generates the $W^+\bar{\nu}_\mu\mu$ interaction and which is the $W^-\bar{e}\nu_e$ interaction. An alternate way of expressing item (iv) is to say that, for each time the *same type* of vertex appears n times, there is a factor of $1/n!$; and the result is multiplied by the number of ways of constructing the graph out of the required types of vertices. Also, when a vertex has several of the same field operator, there are generally multiple ways that the graph can be formed, corresponding to different choices for which field operator does each job. For a simple (though hardly physically realizable!) example, consider the scattering process $HH \rightarrow HH$. The $HHHH$ vertex mediates this process, but any of the four H operators can annihilate the first H incoming state, any of the remaining three H operators can annihilate the other, and either of the two remaining H operators can create the first H final state, leading to $4 \times 3 \times 2 \times 1 = 24$ ways to build the graph, canceling the $1/24$ denominator in the Feynman rule for the vertex. Therefore, we would get $\mathcal{M} = 3m_H^2/v^2$ (plus the contribution of other diagrams).

To clarify item (vi): a relative minus sign arises whenever the pairing off of the creation and annihilation operators for fermions requires an odd

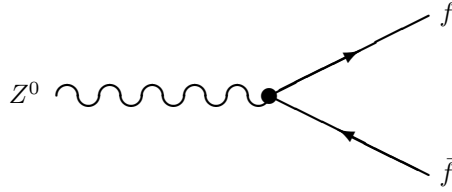
number of anti-commutations of those operators. This is also the origin of the rule, item (v). The sure-fire way to determine whether the matrix element contributions associated with two diagrams have a relative minus sign, due to such fermionic operator anticommutation, is to draw one diagram next to the mirror image of the other, and connect the lines corresponding to the same external particles. Then count the number of fermionic loops in this picture. Next, do the same with either of the original diagrams and itself. If the number of fermion loops differs by an odd number, there is a relative -1 in the original diagrams' contribution to the matrix element. To give an example, consider the following two diagrams for the scattering process, $e^-e^+ \rightarrow e^-e^+$:



Is there a relative sign? To find out, we mirror-image the second element, and connect the lines for the final-state e^+ particles, the final-state e^- particles, the initial-state e^+ particles, and the initial-state e^- particles:



Then we count how many fermionic loops there are. The fermion lines form one big loop. Squaring either of the original diagrams gives two separate loops. Therefore, there is a (-1) relative factor between the diagrams, and when we compute the interference between these diagrams, we will have to include an extra factor of (-1) beyond what the Feynman rules otherwise provide.

Fig. 5.2 The Feynman graph for $Z^0 \rightarrow f\bar{f}$.

5.4.4.1 Example: Z^0 decay

As an illustration of the application of these rules, we wish to recompute the matrix element for the Z^0 to decay into a fermion–antifermion pair. The corresponding graphs start with a single Z^0 boson external line and end with a fermion and antifermion external line.

The simplest set of vertices and internal lines that can connect these initial and final states is a single neutral current vertex, Eq. (5.74). The required graph therefore is as in Figure 5.2.

Using the above rules the S -matrix for this process becomes

$$\begin{aligned} \langle f\bar{f}|S|Z^0\rangle &= \frac{1}{1!} [\epsilon_\mu(\mathbf{k}, \lambda)] [\bar{u}(\mathbf{p}, \sigma)] \\ &\quad \times \left[-e_Z \gamma^\mu (g_V + g_A \gamma_5) (2\pi)^4 \delta^4(p + q - k) \right] [v(\mathbf{q}, \zeta)] \end{aligned} \quad (5.78)$$

The number of independent ways of forming this graph with these vertices is in this case $p = 1$, and the denominator $1!$ corresponds to the graph having only a single vertex.

The matrix element \mathcal{M} is obtained by stripping off the energy-momentum conserving delta function $(2\pi)^4 \delta^4(p + q - k)$ from this expression and multiplying by i . The result obtained in this way is identical to Eq. (4.8) derived by directly evaluating the matrix element.

5.4.4.2 Example: μ^- decay

Perhaps a less trivial example is the amplitude for μ^- decay into $e^- \bar{\nu}_e \nu_\mu$. In this case the relevant graph has a single-fermion initial external line for the μ^- and has two fermion and one antifermion final external lines. Two charged-current vertices are also required. The graph therefore is as in Figure 5.3.

The S -matrix associated with this graph is

$$\langle e\bar{\nu}\nu|S|\mu\rangle = \frac{2}{2!} \int \frac{d^4 r}{(2\pi)^4} \bar{u}(1) \left[-e_W \gamma^\mu (1 + \gamma_5) (2\pi)^4 \delta^4(l + r - k) \right] u(\mathbf{k})$$

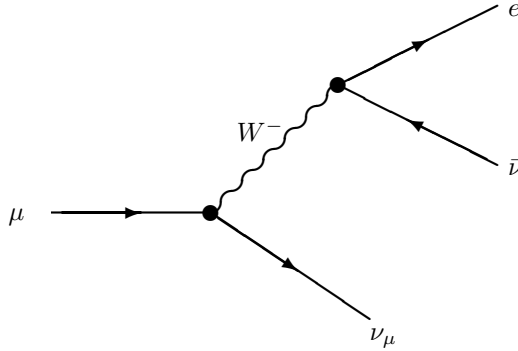


Fig. 5.3 The Feynman graph for the decay $\mu \rightarrow e\bar{\nu}\nu$.

$$\begin{aligned} & \times \bar{u}(\mathbf{p}) \left[-e_W \gamma^\nu (1 + \gamma_5) (2\pi)^4 \delta^4(p + q - r) \right] v(\mathbf{q}) \\ & \times \left[\frac{-i}{r^2 + m^2 - i\epsilon} \left(\eta_{\mu\nu} + \frac{r_\mu r_\nu}{M_W^2} \right) \right] \end{aligned} \quad (5.79)$$

Since this graph contains two vertices, the denominator in the first line of Eq. (5.79) is $2!$. The numerator of the same term is 2, corresponding to the two equal contributions depending on which charged-current interaction vertex destroys the muon and which creates the electron–antineutrino pair.

Grouping terms and identifying the matrix element, $\mathcal{M}(\mu \rightarrow e\bar{\nu}\nu)$, from this S -matrix element then gives the same result as is found in Eq. (5.19) by operator methods.

5.5 Problems

[5.1] Z decay including a Higgs

The first limits on the Higgs boson mass from LEP I came from considering the process $Z \rightarrow H\bar{f}f$. Here m_H must be lighter than M_Z , so that the process is allowed.

[5.1.1] How many interaction vertices are needed for a Z boson to decay to a final state containing a Higgs boson? Assuming the Higgs boson mass is of order 40 GeV, what is the dominant process which leads to the final state $H\bar{f}f$? What propagator is involved in this process?

[5.1.2] Perform a crude parametric estimate of the width of the Z boson to decay to this final state, based on counting powers of couplings and 2π .

[5.1.3] Evaluate the partial width of the Z boson into the $H\bar{f}f$ final state, leaving the final momentum integrations unperformed if you cannot evaluate them. You may treat the fermion masses as small, $m_f \ll m_H, M_Z$, but do not neglect the ratio m_H/M_Z .

[5.1.4] What fraction of the Z boson width does this decay represent? What are the decay products observed in the detector? (Modern vertex detectors have enough resolution to distinguish the path a b quark propagates along before it in turn decays; so you should treat the b quark as a valid final state.)

[5.2] Neutron lifetime

Compute the lifetime of the neutron, in the approximation where the vertex between the W boson, the neutron, and the proton, is the same as the vertex between the W boson, a down quark (in the neutron), and an up quark (in the proton), except that the γ_5 factor is rescaled by a factor g_A .

[5.2.1] Making the approximations $M_W \gg m_p, m_n \gg Q \equiv m_n - m_p \sim m_e$, show that the neutron β -decay rate is given in the neutron rest-frame by

$$\frac{d\Gamma}{dE_e} = \frac{G_F^2 |V_{ud}|^2}{2\pi^3} [\mathcal{F} + \mathcal{G}] E_e \sqrt{E_e^2 - m_e^2} (Q - E_e) \sqrt{(Q - E_e)^2 - m_\nu^2}$$

Here m_ν is a hypothetical neutrino mass and the *Fermi* and *Gamow-Teller* terms \mathcal{F} and \mathcal{G} are defined by

$$\mathcal{F} = |\bar{u}_p(\mathbf{p}) \gamma^0 (g_V + g_A \gamma_5) u_n(\mathbf{p}')|^2$$

and

$$\mathcal{G} = |\bar{u}_p(\mathbf{p}) \vec{\gamma} (g_V + g_A \gamma_5) u_n(\mathbf{p}')|^2$$

Assume that only the outgoing electron energy, E_e , is measured.

[5.2.2] Evaluate \mathcal{F} and \mathcal{G} in the approximation that the nucleon does not recoil; i.e., $\mathbf{p}_p = 0$ (or is $\ll m_n$) in the neutron rest frame.

[5.2.3] Plot the quantity

$$y = \left[\frac{d\Gamma/dE_e}{E_e \sqrt{E_e^2 - m_e^2}} \right]^{1/2}$$

as a function of the electron energy, E_e , for the two cases $m_\nu = 0$ and $m_\nu = 10$ eV. This is known as a Curie plot. How do the two graphs differ?

[5.2.4] Evaluate numerically the total life time of the neutron, neglecting m_ν and using the numerical values, $g_A = 1.269$, $m_p = 938.27200$ MeV,

$m_n = 939.565 \text{ MeV}$, $m_e = 0.511 00 \text{ MeV}$; $(1 \text{ fm})^{-1} = 197.326 96 \text{ MeV}$,
 $1 \text{ s} = 2.9979 \times 10^{23} \text{ fm}$.

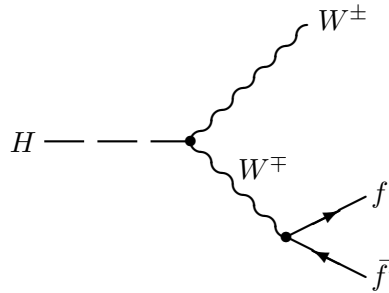
[5.3] **Higgs decay to $Wf\bar{f}$**

In Section 4.3, we found that, for $m_H < 2M_W$, the dominant Higgs boson decay mechanism is $H \rightarrow b\bar{b}$. However, the decay width is very small due to the very small b quark mass.

In the mass regime $M_W < m_H < 2M_W$, another mechanism is

$$H \rightarrow Wf\bar{f}'$$

via the diagram



where f and \bar{f}' are a pair of fermions which could result from the decay of a W^\mp (that is, for W^- they are $e^-\bar{\nu}_e$, $\mu^-\bar{\nu}_\mu$, $\tau^-\bar{\nu}_\tau$, $d\bar{u}$, $s\bar{c}$). For this problem, you should systematically ignore the fermion masses (except m_t , which is so heavy that the top quark does not participate anyway). However, you cannot neglect the W -boson mass M_W . Label the initial momentum p , the final W -boson momentum as q , the momentum on the virtual W -boson propagator as r , and the final fermion and antifermion momenta as k and l (so $r = k + l$).

[5.3.1] **Matrix element**

Argue that exactly half the width, via this process, will be from the case with a W^+ in the final state, and half from the case with a W^- . (Is there a symmetry at play here?) Having made this argument, concentrate on the width when it is a W^+ appearing in the final state. Remember to multiply by 2 at the end of the problem.

Write down the matrix element for this process, before summing on the external state spins and polarizations.

[5.3.2] **Squared matrix element**

Evaluate the final spin and polarization summed, squared matrix element. Carry out all Dirac traces to get an expression which is an algebraic function of the relevant particle four-momenta. It will turn

out to be convenient *not* to contract all the Lorentz indices, however; leave the factors of the form $(\eta^{\mu\nu} + r^\mu r^\nu / M_W^2)$, from the W^- propagator, in this form.

[5.3.3] Integration on fermionic momenta

Write down the width as an integral over final-state momenta, of the squared matrix element.

Holding r fixed, carry out the integration over the final-state fermionic momenta. That is, perform the integrals over k, l . The integration is similar to the one we encountered for $I_{\mu\nu}(r)$ in the text. The resulting expression should be proportional to $(r^2 \eta_{\mu\nu} - r_\mu r_\nu)$. It should now be straightforward to perform the rest of the Lorentz index contractions.

[5.3.4] Total width

Express the total width as a single integral. Re-write your answer by factoring all dimensionful quantities out of the integral, so it depends only on the dimensionless parameter m_H/M_W and the integration variable, which might for instance be p_W^0/M_W . If you cannot do the integral by hand, you will have to find some way of evaluating it numerically. In this case, only attempt the numerical integral for $m_H/M_W < 2$.

Make a plot of $\Gamma(H \rightarrow b\bar{b})$ and $\Gamma(H \rightarrow W f \bar{f})$ as a function of m_H , using $m_b = 3.1$ GeV in evaluating $\Gamma(H \rightarrow b\bar{b})$.

At what mass m_H does the new decay become the dominant one? What happens as m_H approaches $2M_W$? Can you explain why this happens, both in terms of the diagrammatics and physically?

[5.4] **The miracle of Lorentz invariance** Consider the following Lagrangian density for a real scalar field, $\phi(x)$, that is coupled to a classical background current, $J^\mu(x)$:

$$\mathcal{L} = -\frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 - g J^\mu \partial_\mu \phi \quad (5.80)$$

[5.4.1] Construct the canonical Hamiltonian density for this problem and show that it is given by

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\text{int}}$$

with

$$\mathcal{H}_0 = \frac{1}{2} \left[\pi^2 + (\nabla \phi)^2 + m^2 \phi^2 \right] \quad (5.81)$$

$$\mathcal{H}_{\text{int}} = g J^0 \pi + \frac{g^2}{2} J^0 J^0 + g \mathbf{J} \cdot \nabla \phi = +g J^\mu \partial_\mu \phi - \frac{g^2}{2} J^0 J^0 \quad (5.82)$$

Here $\pi = \dot{\phi} - g J^0$ is the canonical momentum. Notice in particular how the interaction Hamiltonian is *not* Lorentz-invariant.

[5.4.2] Find the propagator

$$G_{\mu\nu}(x, x') \equiv \langle 0 | T [\partial_\mu \phi(x) \partial_\nu \phi(x')] | 0 \rangle$$

and show that it can be written in the following way:

$$G_{\mu\nu} = -i \partial_\mu \partial'_\nu \int \frac{d^4 p}{(2\pi)^4} \frac{e^{ip \cdot (x-x')}}{p^2 + \mu^2 - i\epsilon} + \Delta_{\mu\nu}(x, x') \quad (5.83)$$

Explicitly compute the function $\Delta_{\mu\nu}(x, x')$ in this equation, and show that it is *not* Lorentz-covariant.

[5.4.3] Compute the vacuum transition-matrix element, $\langle 0 | S | 0 \rangle$, to second order in the current, $J^\mu(x)$, and show that the above two sources of Lorentz-non-covariance cancel one another. This shows that the final Lorentz-invariant result is equivalent to what would have been obtained if we had simply used the naive expression $\mathcal{H}_{\text{int}} = -\mathcal{L}_{\text{int}}$ and used the naive propagator $\tilde{G}_{\mu\nu} = G_{\mu\nu} - \Delta_{\mu\nu}$. Modified time ordering which produces this propagator is often called the “ T^* -ordering,” and denoted $\tilde{G}_{\mu\nu}(x, x') \equiv \langle 0 | T^* [\partial_\mu \phi(x) \partial_\nu \phi(x')] | 0 \rangle$.

6

Leptonic weak interactions: collisions

The only applications of the standard model discussed up to this point have been calculations of the decay rates for the unstable weakly interacting elementary particles of the model. These are important applications since much of what is known about the fundamental interactions of nature comes from the basic properties of the particles involved, including their decay products and lifetimes. As we have seen, the standard model is able to do a good job of accounting for these properties to within the accuracy of current measurements, at least within the leptonic sector.

There are other applications which the model must also describe, however. Prominent among these are reactions that are observed within particle accelerators. This is, after all, how these unstable particles are produced. This chapter is meant to present some of the standard model predictions for the results of elementary-particle collisions among leptons and electroweak bosons. We focus here on these particles since their collisions are understandable with the fewest complications. Hadronic collisions are the topic of Chapter 9.

e^+e^- -annihilation processes are the lepton collisions that have been of particular interest since these have been studied in great detail near and beyond the Z^0 resonance. The precision of these measurements has been used to test the model with exquisite precision. For this reason the reaction $e^+e^- \rightarrow f\bar{f}$ is examined in some detail.

Neutrino-electron scattering is another purely leptonic process of experimental interest. Beams of electron-type neutrinos (produced for instance in a nuclear reactor or the Sun), or muon-type neutrinos (produced by pion decay downstream of a target area within an accelerator), can be collided with electrons and the resulting collision rates compared with the predictions of the theory. Neutrino collisions with electrons also take place within the large neutrino observatories and must be understood in order to understand

neutrino oscillations in solar, reactor, atmospheric, and beam experiments, and to understand neutrinos from supernovae, such as those observed from Supernova 1987A (and any more that are yet to come).

6.1 The Mandelstam variables

Before launching into a detailed calculation of the collision rates in various accelerators, some notational points must first be made. In any two-body scattering process in which only the momenta and energies of the scattering particles are observed (as opposed to their spins etc.) there are precisely two relativistically invariant variables on which Lorentz-invariant observables like cross sections can depend. There is a conventional choice for these variables that is outlined in this section.

Consider, then, a two-body process of the form $a + b \rightarrow c + d$ in which particles a, b, c , and d have four-momenta p_k^μ and masses $m_k^2 = -p_k^2$, with $k = a, b, c, d$. These four-momenta are arbitrary future-directed timelike (or possibly null) vectors that are subject only to the condition of four-momentum conservation:

$$p_a + p_b = p_c + p_d \quad (6.1)$$

If only momenta and energies are measured in this reaction then the cross section must depend only on the four four-momenta of the problem: p_a through p_d . Being Lorentz-invariant, the cross section $d\sigma(a + b \rightarrow c + d)$ can only depend on the independent Lorentz-invariant combinations that can be constructed from these momenta. Since the square of each of these four-vectors is a constant – being equal to the mass of the corresponding particle – the Lorentz-invariant combinations that contain the kinematic information (such as the directions traveled by each particle) are the six inner products: $p_k \cdot p_l$ with k and l running over particle types a to d with $k \neq l$. Since four-momentum conservation, Eq. (6.1), allows any one of the p_k to be eliminated in terms of the others only three of these inner products need a priori be considered as being distinct. If, for example, four-momentum conservation is chosen to eliminate p_d then the three inner products could be chosen to be $p_a \cdot p_b$, $p_a \cdot p_c$ and $p_b \cdot p_c$.

Instead of directly using these inner products, it is conventional to use the following equivalent three combinations, known as *Mandelstam variables* or *Mandelstam invariants*:

$$\begin{aligned} s &\equiv -(p_a + p_b)^2 \\ &= -2p_a \cdot p_b + m_a^2 + m_b^2 \end{aligned}$$

$$\begin{aligned}
t &\equiv -(p_a - p_c)^2 \\
&= 2p_a \cdot p_c + m_a^2 + m_c^2 \\
u &\equiv -(p_a - p_d)^2 \\
&= 2p_a \cdot p_d + m_a^2 + m_d^2
\end{aligned} \tag{6.2}$$

These invariants may also be re-expressed in terms of the other four-momenta using four-momentum conservation:

$$\begin{aligned}
s &= -(p_c + p_d)^2 \\
&= -2p_c \cdot p_d + m_c^2 + m_d^2 \\
t &= -(p_d - p_b)^2 \\
&= 2p_b \cdot p_d + m_b^2 + m_d^2 \\
u &= -(p_c - p_b)^2 \\
&= 2p_b \cdot p_c + m_b^2 + m_c^2
\end{aligned} \tag{6.3}$$

Now, given the masses of all of the particles involved, a two-body collision should be completely described in terms of two invariant parameters. These could be chosen to be the collision energy and scattering angle as seen in the center-of-mass frame, for example. There must therefore be a relationship amongst the three Mandelstam invariants. This relationship is easily derived if the definitions for s , t , and u in Eq. (6.2) are added to one another:

$$\begin{aligned}
s + t + u &= -2p_a \cdot (p_b - p_c - p_d) + 3m_a^2 + m_b^2 + m_c^2 + m_d^2 \\
&= +2p_a^2 + 3m_a^2 + m_b^2 + m_c^2 + m_d^2 \\
&= m_a^2 + m_b^2 + m_c^2 + m_d^2
\end{aligned} \tag{6.4}$$

These variables can be related to the basic kinematic quantities in any given reference frame, such as the overall energy of the collision and the scattering angles, etc. There are two frames that are of the most practical interest. These are the *center-of-mass frame* – or CM frame for short – defined as the rest frame of the timelike four-vector $p_a + p_b$, and the *lab frame*, defined as the rest frame of particle “a.” The lab frame terminology is appropriate for “fixed target” experiments in which a beam of high-energy particles impinge on a target at rest. Center-of-mass variables are useful both because they are frequently simpler, and because many modern experiments are beam-on-beam experiments where the center-of-mass frame is the same as the frame of the particle detector.

Lab frame: The lab frame is defined as the rest frame of particle “a”:

$$E_a \equiv p_a^0 = m_a; \quad \text{and} \quad \mathbf{p}_a = 0 \tag{6.5}$$

In this frame inner products of four-vectors with p_a have a very simple form: $p_a \cdot p_b = -m_a E_b$. s , t , and u are therefore directly related to the energies of particles “ b ”, “ c ” and “ d ” in this frame:

$$\begin{aligned} s &= +2m_a E_b + m_a^2 + m_b^2, & \text{lab frame} \\ t &= -2m_a E_c + m_a^2 + m_c^2, & \text{lab frame} \\ u &= -2m_a E_d + m_a^2 + m_d^2, & \text{lab frame} \end{aligned} \quad (6.6)$$

Once the energies E , and hence the magnitudes of the three-momenta, $|\mathbf{p}| = \sqrt{E^2 - m^2}$, are determined from these relations, the angular information may next be obtained from these same variables. Denote the angle between the direction of the incoming particle, \mathbf{p}_b , and the directions of the outgoing particles, \mathbf{p}_c and \mathbf{p}_d , by θ_c^* and θ_d^* . Then the angular information is obtained from t and u as expressed in Eq. (6.3). To see this use $p_b \cdot p_c = -E_b E_c + \mathbf{p}_b \cdot \mathbf{p}_c = -E_b E_c + p_b p_c \cos \theta_c^*$. (The notation used here has $p_b \equiv |\mathbf{p}_b|$ with the understanding that the context will keep p_b defined in this way from being confused with the four-vector p_b .) The Mandelstam invariants t and u therefore become

$$\begin{aligned} t &= -2E_b E_d + 2p_b p_d \cos \theta_d^* + m_b^2 + m_d^2, & \text{lab frame} \\ &\approx -2E_b E_d (1 - \cos \theta_d^*) & (\text{ultra-relativistic}) \\ u &= -2E_b E_c + 2p_b p_c \cos \theta_c^* + m_b^2 + m_c^2, & \text{lab frame} \\ &\approx -2E_b E_c (1 - \cos \theta_c^*) & (\text{ultra-relativistic}) \end{aligned} \quad (6.7)$$

CM frame: The CM frame is defined as the frame in which the three-momenta of the initial particles (and so also of the final particles) are equal and opposite:

$$\mathbf{p}_a + \mathbf{p}_b = 0 \quad \text{and so} \quad E_a^2 - m_a^2 = E_b^2 - m_b^2 \quad (6.8)$$

In this frame the invariant s is simply the square of the total energy of the collision:

$$\begin{aligned} s &= (E_a + E_b)^2; & \text{CM frame} \\ &= (E_c + E_d)^2; & \text{CM frame} \end{aligned} \quad (6.9)$$

Clearly knowledge of s therefore completely determines the energies and the magnitudes of the three-momenta of all particles in this frame.

The directional information lies in t and u . Defining the angle θ as the angle between the direction of the initial particle “ a ” and the

direction of the outgoing particle “ c ” in the CM frame, we have

$$\begin{aligned}
 t &= -2E_a E_c + 2p_a p_c \cos \theta + m_a^2 + m_c^2; & \text{CM frame} \\
 &\approx -2E_a E_c (1 - \cos \theta); & (\text{ultra-relativistic}) \\
 u &= -2E_a E_d + 2p_a p_d \cos(\pi - \theta) + m_a^2 + m_d^2 \\
 &= -2E_a E_d - 2p_a p_d \cos \theta + m_a^2 + m_d^2; & \text{CM frame} \\
 &\approx -2E_a E_d (1 + \cos \theta); & (\text{ultra-relativistic})
 \end{aligned} \tag{6.10}$$

These expressions also indicate the range of values over which s , t , and u may run. Inspection of Eq. (6.6) and Eq. (6.9) shows that s , t , and u must lie within the following kinematically allowed ranges:

$$\begin{aligned}
 s &\geq \max[m_a^2 + m_b^2; m_c^2 + m_d^2] \\
 t &\leq \min[m_a^2 + m_c^2; m_b^2 + m_d^2] \\
 u &\leq \min[m_a^2 + m_d^2; m_b^2 + m_c^2]
 \end{aligned} \tag{6.11}$$

6.2 e^+e^- annihilation: calculation

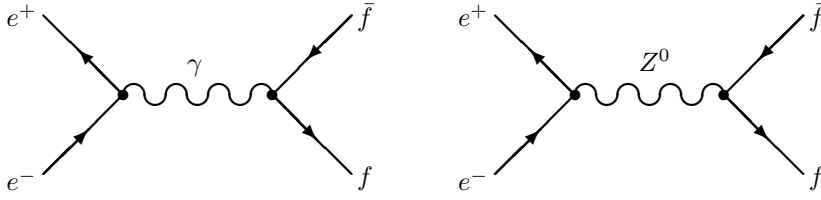
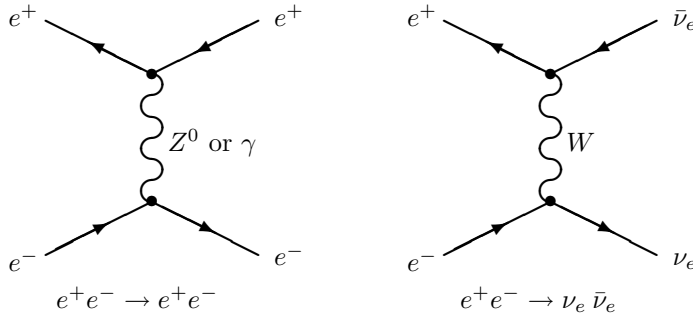
Consider now the collision process $e^+e^- \rightarrow f\bar{f}$. The cross section for this process is computed in this section for unpolarized initial electrons and with the spin of the final-state fermions unmeasured. This calculation is meant to provide an explicit illustration of how such cross sections are determined, as well as to derive formulae for the cross section that have applications in later sections and are of interest in themselves. Since most applications involve energies well in excess of 1 GeV – and the most interesting application is for $s \simeq M_Z \simeq 90$ GeV – the masses of the fermion final states are neglected to good approximation in this section.

Provided that the final state particles are not electrons or electron neutrinos, the standard model scattering amplitude is dominated by two Feynman diagrams, shown in Figure 6.1.

Should the final state particles be e^+e^- or $\nu_e\bar{\nu}_e$ then there are additional graphs such as those in Figure 6.2 that must also be included. Therefore, we postpone treatment of these final states to Section 6.5.

Using the Feynman rules of the previous chapter, the Z^0 -exchange graph of Figure 6.1 has the following matrix element:

$$\begin{aligned}
 \mathcal{M}_{e^+e^- \rightarrow f\bar{f}} &= \frac{2(-e_Z)^2}{2!} [\bar{v}_e(\mathbf{p}') \gamma^\mu \Gamma_{Ze} u_e(\mathbf{p})] [\bar{u}_f(\mathbf{k}) \gamma^\nu \Gamma_{Zf} v_f(\mathbf{k}')] \\
 &\quad \times \left[\frac{1}{(p+p')^2 + M_Z^2 - i\epsilon} \left(\eta_{\mu\nu} + \frac{(p+p')_\mu (p+p')_\nu}{M_Z^2} \right) \right]
 \end{aligned} \tag{6.12}$$

Fig. 6.1 The Feynman graphs for the process $e^+e^- \rightarrow f\bar{f}$.Fig. 6.2 Additional graphs for $e^+e^- \rightarrow e^+e^-$ and $e^+e^- \rightarrow \nu_e\bar{\nu}_e$.

In this equation, Γ_{zf} denotes the Dirac matrix that specifies the Z -boson's neutral-current couplings to fermion type “ f ”:

$$\begin{aligned}\Gamma_{zf} &= g_V + g_A\gamma_5 \\ &= g_L P_L + g_R P_R\end{aligned}\tag{6.13}$$

P_L and P_R are the projection matrices onto left- and right-handed helicity as defined in Eq. (1.76) and Eq. (1.77). The coupling constants g_L and g_R are the more convenient combinations to use if the fermions involved are ultra-relativistic, since in this limit helicity is a conserved quantum number. They are given in terms of g_A and g_V by

$$g_L = g_V + g_A = T_3 - Q \sin^2 \theta_W \tag{6.14}$$

$$g_R = g_V - g_A = -Q \sin^2 \theta_W \tag{6.15}$$

The contribution of the photon-exchange graph is also easily obtained from Eq. (6.12) by making a few substitutions. First, the unitary-gauge Z^0 propagator must be replaced with the ξ -gauge photon propagator appropriate for a massless particle, $M_\gamma = 0$. Next the gauge coupling constant, e_Z , must be replaced by the electromagnetic one, $e_\gamma = e$. Also Γ_{zf} must be replaced by $\Gamma_{\gamma f}$, which has the same form as in Eq. (6.13) with the

numerical constants g_V and g_A of the neutral current replaced by the values $q_V = q_L = q_R = Q$ and $q_A = 0$ relevant for the electromagnetic current.

An immediate simplification of this amplitude is possible if the fermion masses are neglected relative to M_Z , as is the case here. This is because the $(p + p')_\mu(p + p')_\nu$ term in the Z -propagator may be dropped, since it contributes to the S -matrix an amount that is proportional to $g_A^2 m_f^2 / M_Z^2$. The same terms in the photon propagator may also be dropped for any fermion masses, since the axial couplings to the photon vanish, $q_A = 0$, for all fermion types.

The total matrix element, $\mathcal{M}(e^+e^- \rightarrow f\bar{f})$, is the sum of the photon- and Z -exchange contributions and therefore becomes:

$$\begin{aligned} \mathcal{M}(e^+e^- \rightarrow f\bar{f}) = & - \sum_{V=Z,\gamma} e_V^2 [\bar{v}_e(\mathbf{p}') \gamma^\mu \Gamma_{Ve} u_e(\mathbf{p})] [\bar{u}_f(\mathbf{k}) \gamma_\mu \Gamma_{Vf} v_f(\mathbf{k}')] \\ & \times \left[\frac{1}{(p + p')^2 + M_V^2 - i\epsilon} \right] \end{aligned} \quad (6.16)$$

Averaging the square of this matrix element over the four initial spin states (two each for each incoming particle) and summing over the final spins (and, if necessary, colors) gives the following result:

$$\begin{aligned} \overline{\mathcal{M}^2} &= \frac{1}{4} \sum_{\text{spins}} \sum_{\text{colors}} |\mathcal{M}(e^+e^- \rightarrow f\bar{f})|^2 \\ &= N_c \sum_{V=Z,\gamma} \sum_{V'=Z,\gamma} e_V^2 e_{V'}^2 \frac{K^{\mu\nu}(k, k') P_{\mu\nu}(p, p')}{(s - M_V^2)(s - M_{V'}^2)} \end{aligned} \quad (6.17)$$

in which $s = -(p + p')^2$ has been used, the “ $i\epsilon$ ” terms have been dropped, and N_c is as usual $N_c = 1$ if “ f ” is a lepton and $N_c = 3$ if “ f ” is a quark. $K^{\mu\nu}$ and $P_{\mu\nu}$ represent the following Dirac traces:

$$\begin{aligned} P^{\mu\nu} &\equiv \text{tr}[\gamma^\mu \Gamma_{Ve} u_e \bar{u}_e(\mathbf{p}) \gamma^\nu \Gamma_{V'e} v_e \bar{v}_e(\mathbf{p}')] \\ &= -\text{tr}[\gamma^\mu \Gamma_{Ve} \not{p} \gamma^\nu \Gamma_{V'e} \not{p}'] \\ &= -2[(g_{eL} g'_{eL} + g_{eR} g'_{eR})(p^\mu p'^\nu + p^\nu p'^\mu - p \cdot p' \eta^{\mu\nu}) \\ &\quad + i(g_{eL} g'_{eL} - g_{eR} g'_{eR}) \epsilon^{\mu\nu\lambda\rho} p_\lambda p'_\rho] \end{aligned} \quad (6.18)$$

and

$$\begin{aligned} K^{\mu\nu} &\equiv \text{tr}[\gamma^\mu \Gamma_{Vf} v_f \bar{v}_f(\mathbf{k}') \gamma^\nu \Gamma_{V'f} u_f \bar{u}_f(\mathbf{k})] \\ &= -\text{tr}[\gamma^\mu \Gamma_{Vf} \not{k}' \gamma^\nu \Gamma_{V'f} \not{k}] \\ &= -2[(g_{fL} g'_{fL} + g_{fR} g'_{fR})(k^\mu k'^\nu + k^\nu k'^\mu - k \cdot k' \eta^{\mu\nu}) \\ &\quad - i(g_{fL} g'_{fL} - g_{fR} g'_{fR}) \epsilon^{\mu\nu\lambda\rho} k_\lambda k'_\rho] \end{aligned} \quad (6.19)$$

The prime on the coupling constants g_L and g_R indicates it is the coupling appropriate to gauge boson V' .

Because the denominators involve the Mandelstam variable s , this process is conventionally referred to as an s -channel process.

Contracting these last results with one another gives the intermediate result

$$\begin{aligned}
 K^{\mu\nu} P_{\mu\nu} &= 16 \left[(g_{fL} g'_{fL} g_{eL} g'_{eL} + g_{fR} g'_{fR} g_{eR} g'_{eR}) (p \cdot k') (p' \cdot k) \right. \\
 &\quad \left. + (g_{fL} g'_{fL} g_{eR} g'_{eR} + g_{fR} g'_{fR} g_{eL} g'_{eL}) (p \cdot k) (p' \cdot k') \right] \\
 &= 4 \left[(g_{fL} g'_{fL} g_{eL} g'_{eL} + g_{fR} g'_{fR} g_{eR} g'_{eR}) u^2 \right. \\
 &\quad \left. + (g_{fL} g'_{fL} g_{eR} g'_{eR} + g_{fR} g'_{fR} g_{eL} g'_{eL}) t^2 \right] \quad (6.20)
 \end{aligned}$$

The last equality uses the ultra-relativistic approximation to Eq. (6.2) and Eq. (6.3) for the Mandelstam invariants as applied to this reaction: $s = -2p \cdot p' = -2k \cdot k'$, $t = 2p \cdot k = 2p' \cdot k'$, and $u = 2p \cdot k' = 2p' \cdot k$.

Combining these results gives the spin-averaged squared matrix element

$$\begin{aligned}
 \overline{\mathcal{M}}^2 &= N_c \left(\left| \sum_{V=Z,\gamma} e_V^2 \frac{g_{eL} g_{fL}}{s - M_V^2} \right|^2 u^2 + \left| \sum_{V=Z,\gamma} e_V^2 \frac{g_{eR} g_{fR}}{s - M_V^2} \right|^2 u^2 \right. \\
 &\quad \left. + \left| \sum_{V=Z,\gamma} e_V^2 \frac{g_{eL} g_{fR}}{s - M_V^2} \right|^2 t^2 + \left| \sum_{V=Z,\gamma} e_V^2 \frac{g_{eR} g_{fL}}{s - M_V^2} \right|^2 t^2 \right) \quad (6.21)
 \end{aligned}$$

This last formula has a simple physical interpretation that might have been expected for massless – i.e. ultra-relativistic – fermions. Eq. (6.21) gives the rate for the collision process as the sum of the incoherent scattering rates in which the initial and final fermions have definite helicity.

Also, as is easily seen from Eq. (6.10), the limits where $u \rightarrow 0$ or $t \rightarrow 0$ correspond for ultra-relativistic fermions to the cases where the scattering angle, θ , between the directions of the incoming electron, e^- , and the outgoing fermion, f , in the CM frame approach zero ($t \rightarrow 0$) or π ($u \rightarrow 0$). In this case the direction of motion of both the incident and final particles are parallel or antiparallel. An argument identical to that given in Subsection 4.1.5 then implies that the conservation of angular momentum along this common direction of motion is only consistent with conservation of helicity for specific choices for the relative helicities of the initial and final fermions. This is seen in the squared matrix element, Eq. (6.21), by the way that each of the terms for definite helicities vanishes either for $t = 0$ or for $u = 0$.

With Eq. (6.21) in hand, the cross section for $e^+e^- \rightarrow f\bar{f}$ is easily computed. From the definition, Eq. (3.41), of the differential cross section, we have

$$\begin{aligned} d\sigma(e^+e^- \rightarrow f\bar{f}) &= \frac{1}{2p^0 2p'^0 f} \overline{\mathcal{M}}^2 (2\pi)^4 \delta^4(p + p' - k - k') \frac{d^3k d^3k'}{(2\pi)^6 2k^0 2k'^0} \\ \text{with } f &\equiv \frac{-p \cdot p'}{p^0 p'^0} v_{\text{rel}} \\ &\approx \frac{s}{2p^0 p'^0} \quad (\text{ultra-relativistic}) \end{aligned} \quad (6.22)$$

so combining all of the above results gives

$$\begin{aligned} d\sigma(e^+e^- \rightarrow f\bar{f}) &= \frac{8\pi^2 \alpha^2}{s} N_c \left([|A_{\text{LL}}(s)|^2 + |A_{\text{RR}}(s)|^2] u^2 \right. \\ &\quad \left. + [|A_{\text{LR}}(s)|^2 + |A_{\text{RL}}(s)|^2] t^2 \right) d\chi \end{aligned} \quad (6.23)$$

in which the helicity amplitudes $A_{ij}(s)$, with $i, j = \text{L, R}$, are given by

$$A_{ij} = \frac{1}{\sin^2 \theta_w \cos^2 \theta_w} \left(\frac{g_{ei} g_{fj}}{s - M_Z^2} \right) + \frac{Q_e Q_f}{s} \quad (6.24)$$

with g_{fi} the coupling strengths of left- and right-handed particles to the Z^0 , and with $d\chi$ denoting the Lorentz-invariant measure on phase-space,

$$\begin{aligned} d\chi &\equiv (2\pi)^4 \delta^4(p + p' - k - k') \frac{d^3k d^3k'}{(2\pi)^6 2k^0 2k'^0} \\ &= (2\pi)^4 \delta^4(p + p' - k - k') \frac{d^3k}{(2\pi)^3 2k^0} 2\pi \delta(k'^2) \theta(k'^0) \frac{d^4k'}{(2\pi)^4} \\ &= 2\pi \delta[(p + p' - k)^2] \theta(p^0 + p'^0 - k^0) \frac{d^3k}{(2\pi)^3 2k^0} \\ &= -\frac{1}{8\pi s} \delta(s + t + u - m_a^2 - m_b^2 - m_c^2 - m_d^2) du dt \end{aligned} \quad (6.25)$$

This gives the final form for the invariant cross section:

$$\begin{aligned} \frac{d\sigma}{du dt}(e^+e^- \rightarrow f\bar{f}) &= -\frac{\pi \alpha^2}{s^2} N_c \left([|A_{\text{LL}}(s)|^2 + |A_{\text{RR}}(s)|^2] u^2 \right. \\ &\quad \left. + [|A_{\text{LR}}(s)|^2 + |A_{\text{RL}}(s)|^2] t^2 \right) \delta(s + t + u) \end{aligned} \quad (6.26)$$

Evaluating this in the CM frame gives the differential cross section as a function of the angle between e^- and f directions, which is called the scattering angle θ :

$$\begin{aligned} \frac{d\sigma}{\sin \theta d\theta}(e^+e^- \rightarrow f\bar{f}) &= \frac{\pi \alpha^2 s N_c}{8} \left\{ [|A_{\text{LL}}(s)|^2 + |A_{\text{RR}}(s)|^2] (1 + \cos \theta)^2 \right. \\ &\quad \left. + [|A_{\text{LR}}(s)|^2 + |A_{\text{RL}}(s)|^2] (1 - \cos \theta)^2 \right\} \end{aligned} \quad (6.27)$$

This last result uses the relations $s = 4E^2$, $t = -2E^2(1 - \cos \theta)$, and $u = -2E^2(1 + \cos \theta)$ that connect the Mandelstam variables to θ . Integrating θ over its range $0 < \theta < \pi$ using $\int_0^\pi (1 \pm \cos \theta)^2 \sin \theta \, d\theta = \frac{8}{3}$ gives the total rate:

$$\sigma(e^+e^- \rightarrow f\bar{f}) = \frac{\pi\alpha^2 s N_c}{3} \left(|A_{LL}(s)|^2 + |A_{RR}(s)|^2 + |A_{LR}(s)|^2 + |A_{RL}(s)|^2 \right) \quad (6.28)$$

6.3 e^+e^- annihilation: applications

The energy dependence of this cross section for electron–positron annihilation is largely governed by the s dependence of the polarized amplitudes $A_{ij}(s)$. There are naturally three regions to consider depending on the relative size of contributions due to photon- and Z^0 -exchange. We consider each region successively in this section.

6.3.1 Low energies: $e^+e^- \rightarrow \text{hadrons}$

For CM-frame energies that are very small compared to $M_Z = 90$ GeV – yet still large compared to the fermion masses, m_e and m_f – the amplitudes $A_{ij}(s)$ are well approximated by the contribution due to photon exchange:

$$A_{LL} \approx A_{LR} \approx A_{RL} \approx A_{RR} \approx \frac{Q_e Q_f}{s} \quad (6.29)$$

In this limit the electron–positron annihilation rate reduces to the form found in quantum electrodynamics,

$$\left. \frac{d\sigma}{du dt} (e^+e^- \rightarrow f\bar{f}) \right|_{\gamma\text{-exchange}} = -\frac{2\pi\alpha^2}{s^2} Q_e^2 Q_f^2 N_c \left(\frac{u^2 + t^2}{s^2} \right) \delta(s + t + u) \quad (6.30)$$

which becomes, in the CM frame,

$$\left. \frac{d\sigma}{\sin \theta d\theta} (e^+e^- \rightarrow f\bar{f}) \right|_{\gamma\text{-exchange}} = \frac{\pi\alpha^2}{2s} Q_e^2 Q_f^2 N_c (1 + \cos^2 \theta) \quad (6.31)$$

The total rate similarly reduces to the result familiar from QED,

$$\sigma(e^+e^- \rightarrow f\bar{f})|_{\gamma\text{-exchange}} = \frac{4\pi\alpha^2}{3s} Q_e^2 Q_f^2 N_c \quad (6.32)$$

6.3.1.1 $\mu^+\mu^-$ production

To get a feeling for the size of these numbers, consider $\mu^+\mu^-$ production at energies $\sqrt{s} = 1$ GeV. At these energies the ratio s/M_Z^2 is $s/M_Z^2 \approx 10^{-4}$ and

$m_\mu^2/s \approx 10^{-2}$, so Eq. (6.32) provides a perfectly adequate description. In this case, using $\alpha = 1/137$ and $Q_e^2 = Q_\mu^2 = N_c = 1$ gives

$$\begin{aligned}\sigma(e^+e^- \rightarrow \mu^+\mu^-) &= \frac{4\pi\alpha^2}{3s} \\ &= 2.23 \times 10^{-4} (\text{GeV})^{-2} \\ &= 87 \text{ nb}\end{aligned}\tag{6.33}$$

The units of the final line are *nanobarns* with a *barn* defined to be 10^{-24}cm^2 . Now, an accelerator *luminosity* is defined as the rate at which the accelerator can deliver incident particles per unit area of beam. This is useful because the product of the cross section and luminosity gives the number of events which can be expected per unit time. Luminosity is usually quoted in inverse cm^2s ; for instance, the LEP I experiment achieved $2.4 \times 10^{31}/\text{cm}^2\text{s}$, but at a much higher energy than 1 GeV. A machine designed to study the 1 GeV energy range in detail, the VEPP-2000, has a luminosity of $10^{32}/\text{cm}^2\text{s} = 0.1/\text{nb.s}$, enough to produce about 9 $\mu^+\mu^-$ pairs per second.

For the purposes of comparison, a strong interaction cross section is roughly a typical strong interaction scale raised to the power that is dictated by dimensional analysis: $\sigma_{\text{str}} \sim \Lambda_{\text{QCD}}^{-2} \sim 40 (\text{GeV})^{-2} \sim 20 \text{ mbarn}$. We take the strong scale $\Lambda_{\text{QCD}} \approx 150 \text{ MeV}$ in this estimate.

6.3.1.2 Hadron production

There is an immediate application of these low-energy results that takes advantage of the fact that in this energy range the energy dependence of the cross section is the same for all particle types in the final state. To use this fact it is convenient to compute the cross section for producing hadrons in low-energy e^+e^- annihilations, normalized by the muon pair-production rate. The complication is that quarks and gluons interact very strongly with each other at low energies, as we discuss in Part III. In fact, the interactions are so strong that quarks and gluons are not valid external states for a reaction; instead they stick together into bound states called hadrons. However, at suitably high energies, $\sqrt{s} > 1 \text{ GeV}$ or so, the strong coupling is weaker and perturbation theory begins to be useful. At these energies the process of producing quarks and gluons and the process of their combining into hadronic bound states are approximately independent. Rather than summing over a complete set of hadronic final states, one can sum over color-neutral quark and gluon final states and ignore the question of how these project onto the hadronic states. For energies large enough to justify perturbation theory the dominant terms in

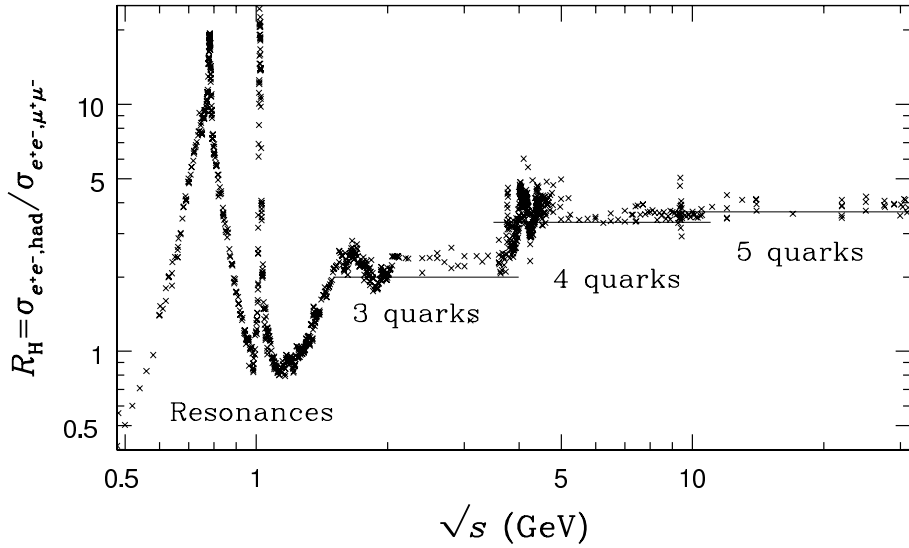


Fig. 6.3 The measured pair-production ratio, R_H .

the final-state sum are then the quark–antiquark pairs. The cross section for this process may therefore be computed by summing Eq. (6.32) over all quark flavors with masses small enough to allow pair production at the given CM energy, \sqrt{s} .

This gives the following expression for the ratio

$$\begin{aligned}
 R_H &\equiv \frac{\sigma(e^+e^- \rightarrow \text{hadrons})}{\sigma(e^+e^- \rightarrow \mu^+\mu^-)} \\
 &\approx \frac{3s}{4\pi\alpha^2} \sum_{q, 2m_q < \sqrt{s}} \sigma(e^+e^- \rightarrow q\bar{q}) \\
 &= 3 \sum_{q, 2m_q < \sqrt{s}} Q_q^2
 \end{aligned} \tag{6.34}$$

The overall factor of 3 here is due to the number of colors available to each quark type. The approximation that is used in the second line is the low-energy expression, Eq. (6.32), for the cross section in which fermion masses are neglected relative to \sqrt{s} . The neglect of fermion masses implies that Eq. (6.34) should not be expected to hold in the immediate vicinity of a threshold, $\sqrt{s} \approx 2m_q$.

Clearly $R_H(s)$ is independent of energy between mass thresholds to the extent that photon exchange dominates the production cross section. Measurement of its value gives an indication of the number of quark degrees of

freedom that are available at the given energy. It gives, in particular, an experimental indication of the number of colors, N_c .

A plot of the experimental value for this ratio is given in Figure 6.3 (from data compiled and made freely available by the Particle Data Group). The solid lines in the figure represent Eq. (6.34) evaluated using u, d, s (three quarks), u, d, s, c (four quarks), and u, d, s, c, b (five quarks). At low energies, the hadronic character of the final state is important and the cross section has a number of peaks and troughs. Above this region, the cross section is reasonably well approximated by Eq. (6.34) (provided $N_c = 3$), including steplike features at $s \simeq 2m_c$ and $s \simeq 2m_b$. (The figure does not show very high but, rather, narrow spikes at these points, which arise because of $c\bar{c}$ and $b\bar{b}$ bound states.)

6.3.2 Intermediate energies: asymmetries

The range of CM interaction energies in the neighborhood of 10 GeV or so is an intermediate range within which the neutral current contribution to the cross section is still small, but is large enough to be detectable. In this energy range we must keep the subdominant term in the expansion of the helicity amplitudes, $A_{ij}(s)$, in powers of s/M_Z^2 :

$$A_{ij}(s) \approx \frac{Q_e Q_f}{s} - \frac{1}{\sin^2 \theta_w \cos^2 \theta_w} \frac{g_{ei} g_{fj}}{M_Z^2} \quad (6.35)$$

Since the small M_Z^{-2} effect of the neutral-current interaction must be picked out of a background of electromagnetic events it helps to focus on some sort of observable to which the electromagnetic interactions do not contribute at all. A natural choice for such an observable would be anything which measures either C or P violation, since this is a symmetry of the electromagnetic, but not of the neutral-current, interaction.

6.3.2.1 Left-right asymmetry

An example of this type of an observable is given by the comparison of cross sections as the helicity of the initial electron-positron pair is varied, since the amplitude only develops a dependence on helicity through the neutral-current couplings. A candidate example might be to take the difference between the cross section measured for left- and right-handed electrons colliding with unpolarized positrons:

$$\mathcal{A}_{\text{LR}} = \frac{\sigma(e_L) - \sigma(e_R)}{\sigma(e_L) + \sigma(e_R)} \quad (6.36)$$

This is known as the *left-right asymmetry* and may be computed using Eq. (6.28) and Eq. (6.35). The leading contribution arises from the interference of the neutral-current and electromagnetic amplitudes:

$$\begin{aligned} \mathcal{A}_{\text{LR}}(e^+e^- \rightarrow f\bar{f}) &= \frac{[|A_{\text{LL}}|^2 + |A_{\text{LR}}|^2] - [|A_{\text{RL}}|^2 + |A_{\text{RR}}|^2]}{[|A_{\text{RL}}|^2 + |A_{\text{RR}}|^2] + [|A_{\text{LL}}|^2 + |A_{\text{LR}}|^2]} \\ &\simeq -\left(\frac{s}{M_Z^2}\right) \left[\frac{(g_{e\text{L}}^2 - g_{e\text{R}}^2)(g_{f\text{L}}^2 + g_{f\text{R}}^2)}{2Q_e Q_f \sin^2 \theta_W \cos^2 \theta_W} \right] \end{aligned} \quad (6.37)$$

For $\sqrt{s} = 25$ GeV the ratio $s/M_Z^2 = 0.08$, so this asymmetry can be in the neighborhood of an 8% effect at these energies.

6.3.2.2 Forward-backward asymmetry

A similar kind of asymmetry that is also sensitive to C-violating interactions is the *forward-backward asymmetry*, \mathcal{A}_{FB} . This is defined in the following way. Suppose the particle detector is imagined to be enclosed within a sphere which is centered at the collision point. If the direction of motion of the initial electron is taken to define the north pole of this sphere, then \mathcal{A}_{FB} is given by the difference in cross sections, call them σ_{\pm} , between those collisions for which the final fermion, f , emerges in the northern $-0 < \theta < \frac{\pi}{2}$ – and southern $-\frac{\pi}{2} < \theta < \pi$ – hemispheres of this sphere, normalized by the total cross section. That is to say, for

$$\sigma_{\pm} \equiv \pm \int_0^{\pm 1} \frac{d\sigma}{d\cos\theta} d\cos\theta \quad (6.38)$$

we have

$$\begin{aligned} \mathcal{A}_{\text{FB}} &= \frac{\sigma_+ - \sigma_-}{\sigma_+ + \sigma_-} \\ &= \frac{3}{4} \left(\frac{[|A_{\text{LL}}(s)|^2 + |A_{\text{RR}}(s)|^2] - [|A_{\text{LR}}(s)|^2 + |A_{\text{RL}}(s)|^2]}{[|A_{\text{LL}}(s)|^2 + |A_{\text{RR}}(s)|^2] + [|A_{\text{LR}}(s)|^2 + |A_{\text{RL}}(s)|^2]} \right) \\ &\approx -\left(\frac{s}{M_Z^2}\right) \left[\frac{3(g_{e\text{L}}^2 - g_{e\text{R}}^2)(g_{f\text{L}}^2 - g_{f\text{R}}^2)}{8Q_e Q_f \sin^2 \theta_W \cos^2 \theta_W} \right] \end{aligned} \quad (6.39)$$

The approximation used in the third line here is the use of the approximate form, Eq. (6.35), for the helicity amplitudes, $A_{ij}(s)$.

Notice that although \mathcal{A}_{LR} and \mathcal{A}_{FB} are proportional at low energies to the squared difference of the left- and right-handed electron couplings, they each sample a different combination of the couplings of the pair-produced fermions. It was measurements of these asymmetries which first gave convincing evidence of the existence of the Z^0 boson, before any accelerator had sufficient energy to create one directly.

6.3.3 High energies: asymptotic forms

The next simplest limit takes $s \gg M_Z^2$, so the helicity amplitudes may be approximated by

$$A_{ij}(s) \approx \left[Q_e Q_f + \frac{g_{ei} g_{fj}}{\sin^2 \theta_w \cos^2 \theta_w} \right] \frac{1}{s} \quad (6.40)$$

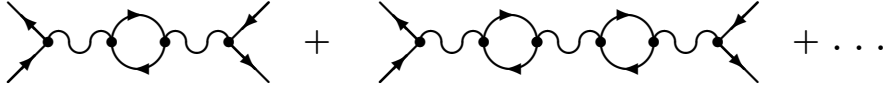
In this limit the energy dependence of the cross section is precisely as it is in the photon-dominated case, but with the electromagnetic coupling constants $Q_e Q_f$ replaced by the combination in the square bracket of Eq. (6.40) above.

At these energies the photon and Z^0 exchange graphs differ only in the strength of their couplings. This is the signature of electroweak unification; at high energies the weak and electromagnetic interactions are indeed very similar in form.

6.4 The Z boson resonance

Even a superficial inspection of Eq. (6.26), Eq. (6.27), or Eq. (6.28) indicates that there is a problem in the regime where the exchange momentum, $r \equiv (p + p')$, approaches the Z boson mass shell, $r^2 = -M_Z^2$, where the intermediate Z boson has the right four-momentum to be a real (as opposed to virtual) particle. As $s = -r^2$ approaches M_Z^2 , the cross section apparently diverges. This indicates a failure of the perturbative expansion for the S -matrix. After all, the S -matrix elements are bounded by the general requirement of unitarity, so their perturbative approximations must also be bounded, or must be bad approximations. For this reason, in this section we will have to make a digression into the topic of higher-order perturbative corrections. We will find that these corrections are essential to resolving this puzzle.

There is a large body of knowledge concerning the higher-order perturbative corrections to the lowest-order expressions described up to this point throughout this book. As has already been seen, these corrections are an important part of the agreement of the standard model with experiment, particularly for the properties of the Z^0 boson, due to the accuracy of the experimental results that are now available. This is doubly true in the resonance regime, $s \simeq M_Z^2$, where the radiative corrections are already essential at leading order, and where exquisitely precise measurements, conducted by the LEP I experiment at CERN and the SLC experiment at SLAC, have thoroughly explored the physics and provided some of the highest-precision experiments, and tightest tests, of the standard model.

Fig. 6.4 Important corrections near $s = M_Z^2$.

6.4.1 Corrections near resonance

From the fact that the perturbative S -matrix near the Z^0 -mass shell diverges, it follows that there must be additional, supposedly higher-order, graphs that nevertheless contribute in an important way, when $s \simeq M_Z^2$. It is the purpose of this subsection to identify these contributions and to find their size, in order to have a good approximation to the lowest-order cross section near the pole.

The graphs that are the source of the difficulty are graphs of the form of Figure 6.4, which can be thought of as modifying the Z^0 -boson propagator. Although these graphs are superficially suppressed relative to Figs. 6.2 by additional powers of the small electroweak coupling constants, they can be of comparable size for Z^0 -boson four-momenta, r_μ , that lie within the immediate neighborhood (i.e. within $O(\alpha M_Z^2)$) of $r^2 = -M_Z^2$.

The reason that corrections are needed here, is because of a cancellation in the denominator of the propagator, $r^2 + M_Z^2 \simeq 0$, which renders the propagator much larger than the usual size $\sim 1/M_Z^2$. Each “loop” of the form shown in Figure 6.4 (the loop is the pair of fermionic propagators going in a circle, including the loop-momentum integration and the vertex factors) gives a contribution which is of the order αM_Z^2 (as can be argued on dimensional grounds). Each loop also leads to one more appearance of the gauge-boson propagator, introducing a factor of $1/(r^2 + M_Z^2)$. When $r^2 + M_Z^2 \sim \alpha M_Z^2$, the addition of a loop is not a suppressed correction. While there are other ways of adding loops, they do not lead to a new factor of $1/(s - M_Z^2)$ per loop, and we can therefore continue to neglect them.

Denote the contribution of one of these loops, after all polarization sums have been performed, as $M_Z^2 \alpha(x + iy)$, with x and y pure numbers of order 1. We know on dimensional grounds that this is the right general form for the correction. It is easy to see that the inclusion of these graphs does remove the divergence of the S -matrix near the Z^0 -mass shell. The whole set of graphs with 0, 1, 2, \dots “loops” inserted can be summed as a geometric series, and

appears as a correction in the denominator of the Z^0 -boson propagator:

$$\begin{aligned}
 (0 \text{ loop}) + (1 \text{ loop}) + \cdots &\propto \frac{1}{r^2 + M_Z^2} + \frac{M_Z^2 \alpha (x + iy)}{(r^2 + M_Z^2)^2} + \frac{[M_Z^2 \alpha (x + iy)]^2}{(r^2 + M_Z^2)^3} + \cdots \\
 &= \frac{1}{r^2 + M_Z^2 (1 - \alpha (x + iy))} \quad (6.41)
 \end{aligned}$$

A radiative (loop) correction which can be understood as a correction of a propagator in this way is called a *self-energy correction*, because it represents a correction to the propagating particle's energy due to its interaction with the vacuum (with its own field).

The correction to the propagator is *complex*, and in particular it moves the singular point of the propagator away from the real point $r^2 + M_Z^2 = 0$ and out to a complex point:

$$r^2 = -M_Z^2 [1 - \alpha (x + iy)] \quad (6.42)$$

However, since the Z^0 boson four-momentum, r_μ , must necessarily be real, it can never satisfy Eq. (6.42), and so the corresponding source of the divergence of the S -matrix does not arise. Since these corrections are only significant for s within $O(\alpha M_Z^2)$ of $s = M_Z^2$, none of the discussions of the previous sections need be modified. From the above considerations it is clearly the imaginary part of the contribution from diagrams like Figure 6.4 that is the most important. A real shift, e.g. αx in Eq. (6.42), can be re-interpreted as a shift in the Z^0 boson mass squared, $M_Z^2(\text{physical}) = M_Z^2(1 - \alpha x)$. In fact, it is the combination $M_Z^2(1 - \alpha x)$ which we “measure” as the true mass of the Z boson – a point we will return to in Subsection 7.4.1. An imaginary shift cannot be similarly absorbed. The next step is to determine how to compute the size of this shift reliably. The main conclusion to be argued is that the imaginary part of the shift in the position of the pole of the propagator is simply related to the mass and total decay width of the Z^0 . Once this is established, the results of Chapter 4 may be used immediately to compute the size of the shift.

For these purposes we take advantage of the small range of momenta for which these corrections are appreciable. It is therefore a good approximation to take the corrections to the Z^0 propagator near the mass shell, due to Figure 6.4, to be independent of momentum. That is, we are neglecting the r^2 dependence of y above. It will be convenient to redefine y as $-iM_Z^2 \alpha y \equiv -iM_Z \Delta$. That is to say

$$\frac{1}{r^2 + M_Z^2 - i\epsilon} \rightarrow \frac{1}{r^2 + M_Z^2 - iM_Z \Delta} \quad (6.43)$$

where Δ is a constant with the dimensions of mass that is much smaller than the Z^0 mass itself: $\Delta \sim O(\alpha M_Z)$. The infinitesimal, ϵ , has been dropped

since its role is to indicate how to avoid the singularity at $r^2 = -M_Z^2$ in the integration over the component r^0 . This is no longer necessary since the additional term, $-iM_Z\Delta$, shifts the singularity off of the (real) integration axis.

The remainder of the argument is to relate the parameter Δ to the properties of the particles “going around in the loop.” The main observation is that the last two terms in the denominator of Eq. (6.43) may be written to lowest order in α as a perfect square:

$$M_Z^2 - iM_Z\Delta = \left(M_Z - \frac{i\Delta}{2}\right)^2 + O(\alpha^2) \quad (6.44)$$

which is completely equivalent to a shift of the Z^0 mass by a small imaginary part. If the arguments used to derive the propagator from the sum over virtual particle states in Subsection 5.2.1 are now run backwards, the shift of the pole of the propagator implies that the intermediate Z^0 states evolve in time with a small negative imaginary part for their mass. Since the time dependence of these particle states, in the Z rest frame, is

$$\begin{aligned} |Z(t)\rangle &= e^{-iM_Z t} |Z(0)\rangle \\ &\rightarrow e^{-iM_Z t - \frac{1}{2}\Delta t} |Z(0)\rangle \end{aligned} \quad (6.45)$$

the probability that the Z^0 particle survives as a function of time therefore becomes

$$\begin{aligned} p(t) &= |\langle Z(t)|Z(0)\rangle|^2 \\ &= e^{-\Delta t} p(0) \end{aligned} \quad (6.46)$$

This implies that Δ should be identified with the full decay width for the Z^0 particle as computed in Section 4.1 (*c.f.* Eq. (4.38) and Eq. (4.39)):

$$\begin{aligned} \Delta &= \Gamma_Z \\ &= \frac{e_Z^2}{12\pi} M_Z \sum_f (g_V^2 + g_A^2) N_c \end{aligned} \quad (6.47)$$

In a nutshell, then, the net effect of all of the higher-order graphs such as those of Figure 6.4 for the results of Section 6.2 is to replace the denominator $(s - M_Z^2)$ by

$$\frac{1}{s - M_Z^2} \rightarrow \frac{1}{s - M_Z^2 - iM_Z\Gamma_Z} \quad (6.48)$$

Two comments concerning this replacement are in order.

- (i) Notice that, as advertised, because Γ_Z is $O(\alpha M_Z)$, the difference between the corrected propagator and the original one is completely negligible *except* when $s - M_Z^2 = O(\alpha M_Z^2)$. This implies that none of

the results of the previous sections are affected by this change (except at an order where there are other corrections anyway).

- (ii) Although the S -matrix elements for $e^+e^- \rightarrow f\bar{f}$ no longer diverge after this substitution, they do become very large. In fact, for s precisely equal to M_Z^2 , the Z^0 -exchange contributions to the helicity amplitudes, $A_{ij}(s)$, are larger than those due to photon decay by a factor of $1/\alpha \sim 100$. This implies that there is an enormous enhancement of the Z^0 -exchange amplitude for s in the immediate vicinity of M_Z^2 , and so photon exchange may be neglected for these energies. The s -dependence of the squared helicity amplitudes then acquires the classic Lorentzian, or *Breit–Wigner* lineshape of a resonance:

$$\begin{aligned} d\sigma(e^+e^- \rightarrow Z^0 \rightarrow f\bar{f}) &\propto \left| \frac{1}{s - M_Z^2 - iM_Z\Gamma_Z} \right|^2 \\ &= \frac{1}{(s - M_Z^2)^2 + M_Z^2\Gamma_Z^2} \end{aligned} \quad (6.49)$$

6.4.2 Application: $e^+e^- \rightarrow f\bar{f}$ near resonance

Using this replacement in Eq. (6.24) and neglecting photon exchange gives the following approximation for the helicity amplitudes, $A_{ij}(s)$, that holds near resonance (“near resonance” here means for $|\sqrt{s} - M_Z| \leq O(\sqrt{\alpha M_Z^2}) \approx 10$ GeV):

$$\begin{aligned} |A_{ij}(s)|^2 &\approx \left| \frac{g_{ei}g_{fj}}{\sin^2\theta_W \cos^2\theta_W} \frac{1}{s - M_Z^2 - iM_Z\Gamma_Z} \right|^2 \\ &= \frac{g_{ei}^2 g_{fj}^2}{\sin^4\theta_W \cos^4\theta_W} \frac{1}{(s - M_Z^2)^2 + M_Z^2\Gamma_Z^2} \end{aligned} \quad (6.50)$$

The resulting expressions for the differential and integrated cross sections in the CM frame are

$$\begin{aligned} \frac{d\sigma}{\sin\theta d\theta}(e^+e^- \rightarrow Z^0 \rightarrow f\bar{f}) \Big|_{\text{res}} &= \frac{\pi\alpha^2}{8\sin^4\theta_W \cos^4\theta_W} N_c \frac{s}{(s - M_Z^2)^2 + M_Z^2\Gamma_Z^2} \\ &\quad \times \left([g_{eL}^2 g_{fL}^2 + g_{eR}^2 g_{fR}^2](1 + \cos\theta)^2 \right. \\ &\quad \left. + [g_{eL}^2 g_{fR}^2 + g_{eR}^2 g_{fL}^2](1 - \cos\theta)^2 \right) \end{aligned} \quad (6.51)$$

and

$$\begin{aligned} \sigma_{\text{res}}(e^+e^- \rightarrow Z^0 \rightarrow f\bar{f}) &= \frac{\pi\alpha^2}{3\sin^4\theta_W \cos^4\theta_W} N_c \frac{s}{(s - M_Z^2)^2 + M_Z^2\Gamma_Z^2} \\ &\quad \times (g_{eL}^2 + g_{eR}^2)(g_{fL}^2 + g_{fR}^2) \end{aligned} \quad (6.52)$$

A particularly clean prediction on resonance is possible for the asymmetries, \mathcal{A}_{LR} and \mathcal{A}_{FB} , respectively defined by Eq. (6.36) and Eq. (6.38), since the common resonant energy dependence drops out of these cross section ratios:

$$\begin{aligned}
 \mathcal{A}_{\text{LR}}|_{\text{res}} &= \frac{g_{e\text{L}}^2 - g_{e\text{R}}^2}{g_{e\text{L}}^2 + g_{e\text{R}}^2} \\
 &= \frac{\frac{1}{4} + \sin^2 \theta_{\text{W}}}{\frac{1}{4} + \sin^2 \theta_{\text{W}} + 2 \sin^4 \theta_{\text{W}}}, \\
 \mathcal{A}_{\text{FB}}|_{\text{res}} &= \frac{(g_{e\text{L}}^2 - g_{e\text{R}}^2)(g_{f\text{L}}^2 - g_{f\text{R}}^2)}{(g_{e\text{L}}^2 + g_{e\text{R}}^2)(g_{f\text{L}}^2 + g_{f\text{R}}^2)} \\
 &= \frac{(g_{f\text{L}}^2 - g_{f\text{R}}^2)}{(g_{f\text{L}}^2 + g_{f\text{R}}^2)} \mathcal{A}_{\text{LR}}|_{\text{res}}
 \end{aligned} \tag{6.53}$$

6.4.2.1 Factorization

Resonant amplitudes and cross sections have a particularly simple form when evaluated right on the central point of the resonance, $s = M_Z^2$. On resonance the Z^0 propagator may be simplified using the spin-sum identity, Eq. (1.119):

$$\frac{1}{r^2 + M_Z^2 - iM_Z\Gamma_Z} \left(\eta_{\mu\nu} + \frac{r_\mu r_\nu}{M_Z^2} \right) \rightarrow \frac{i}{M_Z\Gamma_Z} \sum_{\lambda=-1}^1 \epsilon(\mathbf{r}, \lambda) \epsilon^*(\mathbf{r}, \lambda) \tag{6.54}$$

so the scattering matrix element of Eq. (6.12) may be written as (c.f. Eq. (4.8) for the Z^0 decay matrix element)

$$\begin{aligned}
 \mathcal{M}(e^+e^- \rightarrow f\bar{f})|_{\text{res}} &= -\frac{ie_Z^2}{M_Z\Gamma_Z} \sum_{\lambda=-1}^1 [\bar{v}_e(\mathbf{p}') \gamma^\mu \Gamma_{Ze} u_e(\mathbf{p}) \epsilon^*(\mathbf{p} + \mathbf{p}', \lambda)] \\
 &\quad \times [\bar{u}_f(\mathbf{k}) \gamma^\nu \Gamma_{Zf} v_f(\mathbf{k}') \epsilon(\mathbf{p} + \mathbf{p}', \lambda)] \\
 &= -\frac{i}{M_Z\Gamma_Z} \sum_{\lambda=-1}^1 [ie_Z \bar{v}_e(\mathbf{p}') \gamma^\mu \Gamma_{Ze} u_e(\mathbf{p}) \epsilon^*(\mathbf{p} + \mathbf{p}', \lambda)] \\
 &\quad \times [-ie_Z \bar{u}_f(\mathbf{k}) \gamma^\nu \Gamma_{Zf} v_f(\mathbf{k}') \epsilon(\mathbf{p} + \mathbf{p}', \lambda)] \\
 &= +\frac{i}{M_Z\Gamma_Z} \sum_{\lambda=-1}^1 [\mathcal{M}(Z^0 \rightarrow e^+e^-)]^* [\mathcal{M}(Z^0 \rightarrow f\bar{f})]
 \end{aligned} \tag{6.55}$$

That is, on resonance the cross section for $e^+e^- \rightarrow f\bar{f}$ factorizes into the product of the amplitude for the process $e^+e^- \rightarrow Z^0$ and the decay $Z^0 \rightarrow f\bar{f}$ for a Z^0 particle at rest, $E_Z = M_Z$. This is as would be expected if a real Z^0 is produced and then decays. The Z^0 so produced does not appear with its spin in a pure state but instead is prepared in a state that is described

by a density matrix which has each of the three possible spin states equally weighted. Since this is indeed how real Z^0 s are produced this justifies the choice made in Eq. (4.10) for the spin state of an unpolarized Z^0 sample.

The cross section for this process has a similar factorized form:

$$\begin{aligned}\sigma(e^+e^- \rightarrow Z^0 \rightarrow f\bar{f}) &= \frac{e_Z^4}{48\pi} N_c \frac{1}{\Gamma_Z^2} (g_{eL}^2 + g_{eR}^2)(g_{fL}^2 + g_{fR}^2) \\ &= \frac{12\pi}{M_Z^2} \frac{\Gamma(Z \rightarrow e^+e^-)}{\Gamma_Z} \frac{\Gamma(Z \rightarrow f\bar{f})}{\Gamma_Z}\end{aligned}\quad (6.56)$$

Here we used $g_L^2 + g_R^2 = 2(g_V^2 + g_A^2)$. Note that this result is precisely the Breit-Wigner result for (ultra-relativistic) scattering through a p -wave (spin-one) resonance,

$$\sigma = \frac{16\pi}{s} \frac{(2s_Z + 1)}{(2s_e + 1)(2s_{\bar{e}} + 1)} \frac{\Gamma(Z \rightarrow e^+e^-)}{\Gamma_Z} \frac{\Gamma(Z \rightarrow f\bar{f})}{\Gamma_Z} \quad (6.57)$$

where $(2s_Z + 1) = 3$ is the number of spin states of the Z boson, and $(2s_e + 1) = (2s_{\bar{e}} + 1) = 2$ are the number of spin states of the incoming particles. This latter factor appears because we are computing the spin-averaged cross section; had we computed the cross section for a specific spin state of the e^+ and e^- , no such factor would appear.

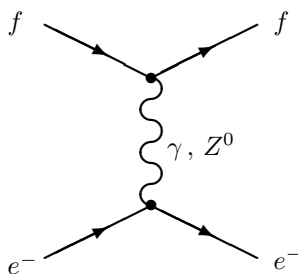
Equation (6.56) has a natural physical interpretation. In order to bring this interpretation out, rewrite Eq. (6.56) in terms of the total cross section for Z^0 production on resonance, defined by summing the above result over all possible fermion–antifermion final states. Then, using $\sum_f \Gamma(Z \rightarrow f\bar{f}) = \Gamma_Z$, we find

$$\sigma(e^+e^- \rightarrow Z^0 \rightarrow f\bar{f}) = \sigma_{\text{tot}} B(Z \rightarrow f\bar{f}) \quad (6.58)$$

Here $B(Z \rightarrow f\bar{f}) = \Gamma(Z \rightarrow f\bar{f})/\Gamma_Z$ is the Z^0 branching fraction into the final fermion–antifermion pairs of flavor “ f .” The total cross section, σ_{tot} , is itself given explicitly by

$$\begin{aligned}\sigma_{\text{tot}}(e^+e^- \rightarrow Z^0) &= \frac{12\pi}{M_Z^2} B(Z \rightarrow e^+e^-) \\ &= 1.52 \times 10^{-4} (\text{GeV})^{-2} \\ &= 59.4 \text{ nb}\end{aligned}\quad (6.59)$$

where we used the experimental values for the width and branching fraction. Now both the total cross section, σ_{tot} , and the cross section $\sigma(e^+e^- \rightarrow Z^0 \rightarrow f\bar{f})$ give the rate per unit incident flux per unit time for the corresponding reactions. Also, the branching fraction, $B(Z^0 \rightarrow f\bar{f})$, gives the probability that any given Z^0 , once produced, decays into an $f\bar{f}$ pair. Equation (6.58)

Fig. 6.5 The Feynman graph for $e^- f \rightarrow e^- f$.

therefore declares that the probability for $f\bar{f}$ production is given on resonance by the product of the probability of creating a Z^0 with the probability of this Z^0 decaying into $f\bar{f}$.

Notice that the total cross section given by Eq. (6.59) translates into roughly one Z^0 produced per second using the luminosity of the LEP-I collider at CERN: $17 \times 10^{30} \text{ cm}^{-2} \text{ s}^{-1} = 1.7 \times 10^{-2} / \text{nb.s}$. We will see in Subsection 6.7.2 that Eq. (6.59) turns out to be missing a rather substantial correction, which brings the actual cross section down by about 28% from the one we have computed.

6.5 *t*-channel processes: crossing symmetry

Next, consider the process $e^- f \rightarrow e^- f$, with f any fermion other than e^+ , e^- , $\bar{\nu}_e$, or ν_e . This type of scattering is dominantly mediated by the exchange of a virtual Z boson or by photon-exchange as in Figure 6.5.

There are a great many practical situations for which this cross-section is of interest. Some of these are: (i) elastic scattering of electrons and muons: $e^- \mu^- \rightarrow e^- \mu^-$; (ii) elastic muon-neutrino collisions with electrons, $e^- \nu_\mu \rightarrow e^- \nu_\mu$; or (iii) various hadronic processes considered in more detail in Chapter 9. Inelastic processes such as $e^- \nu_\mu \rightarrow \mu^- \nu_e$ may also be described by the result derived below provided that all fermion masses may be neglected.

Consider therefore the evaluation of Figure 6.5 for the contribution of a vector boson V to the reaction $e^-(\mathbf{p})f(\mathbf{p}') \rightarrow e^-(\mathbf{k})f(\mathbf{k}')$. Inspection of the Feynman rules of the previous chapter gives a matrix element for this process of

$$\begin{aligned} \mathcal{M}(e^- f \rightarrow e^- f) &= - \sum_{V=Z,\gamma} e_V^2 [\bar{u}_e(\mathbf{k}) \gamma^\mu \Gamma_{Ve} u_e(\mathbf{p})] [\bar{u}_f(\mathbf{k}') \gamma_\mu \Gamma_{Vf} u_f(\mathbf{p}')] \\ &\quad \times \left[\frac{1}{(p-k)^2 + M_V^2 - i\epsilon} \right] \end{aligned} \quad (6.60)$$

This expression is very similar to Eq. (6.16). In fact, Eq. (6.60) may be obtained from Eq. (6.16) by the simple substitution of $\bar{u}_e(\mathbf{k})$ for $\bar{v}_e(\mathbf{p}')$ and $u_f(\mathbf{p}')$ for $v_f(\mathbf{k}')$, as well as with the replacements $p'_\mu \rightarrow -k_\mu$, $k \rightarrow k'$, and $k'_\mu \rightarrow -p'_\mu$. The signs of the four-momenta are reversed whenever an incoming particle becomes an outgoing particle or vice versa.

This correspondence allows the results of Section 6.2 to be used directly to give the spin-summed, squared matrix element for the present process. If we take the Mandelstam variables for elastic e^-f scattering (in the ultra-relativistic limit) as

$$\begin{aligned} s &= -(p + p')^2 \approx -2p \cdot p' \\ t &= -(p - k)^2 \approx +2p \cdot k \\ u &= -(p - k')^2 \approx +2p \cdot k' \end{aligned} \quad (6.61)$$

then the desired result is obtained from Eq. (6.21) with the following substitution:

$$\begin{aligned} s &= -2p \cdot p' \rightarrow +2p \cdot k = t \\ t &= +2p \cdot k \rightarrow +2p \cdot k' = u \\ u &= +2p \cdot k' \rightarrow -2p \cdot p' = s \end{aligned} \quad (6.62)$$

The result therefore becomes

$$\begin{aligned} \overline{\mathcal{M}}^2 &= \left(\left| \sum_{V=Z,\gamma} e_V^2 \frac{g_{eL} g_{fL}}{t - M_V^2} \right|^2 s^2 + \left| \sum_{V=Z,\gamma} e_V^2 \frac{g_{eR} g_{fR}}{t - M_V^2} \right|^2 s^2 \right. \\ &\quad \left. + \left| \sum_{V=Z,\gamma} e_V^2 \frac{g_{eL} g_{fR}}{t - M_V^2} \right|^2 u^2 + \left| \sum_{V=Z,\gamma} e_V^2 \frac{g_{eR} g_{fL}}{t - M_V^2} \right|^2 u^2 \right) \end{aligned} \quad (6.63)$$

This expression admits the same simple interpretation in terms of polarization amplitudes as did Eq. (6.21).

Also in analogy with the earlier treatment, the differential cross section becomes

$$\begin{aligned} \frac{d\sigma}{du dt}(e^-f \rightarrow e^-f) &= -\pi\alpha^2 \left([|A_{LL}(t)|^2 + |A_{RR}(t)|^2] \right. \\ &\quad \left. + [|A_{LR}(t)|^2 + |A_{RL}(t)|^2] \frac{u^2}{s^2} \right) \delta(s + t + u) \end{aligned} \quad (6.64)$$

Here the helicity amplitudes are

$$A_{ij}(t) = \sum_V \frac{e_V^2}{e^2} \left(\frac{g_{ei} g_{fj}}{t - M_V^2} \right) \quad (6.65)$$

Because the Mandelstam variable t appears in the denominator, the process considered here is generally referred to as a *t*-channel process, as opposed to the *s*-channel process of the previous sections.

This cross section does not encounter problems when $t - M_V^2$ goes to zero, because the kinematically allowed range for t is $-s < t \equiv -Q^2 \leq 0$ for ultra-relativistic fermions. On the other hand, the cross-section does diverge as $t \rightarrow 0$. This is the familiar Coulomb divergence of the cross section, which again expresses a breakdown of an approximation we have made in deriving $d\sigma/dt$. Very small t corresponds to very small scattering angles, which classically would occur for large impact parameters. In principle, the long range of the Coulomb interaction ensures that multiple scatterings must be included to obtain an accurate determination of very-small-angle scattering amplitudes. These higher-order complications are typically not important when discussing the differential cross section as a function of angle, since it is typically true that the experiment of interest cannot distinguish scattering at sufficiently small angles from no scattering occurring at all.

In deriving the matrix element for this process, we were able to learn almost everything by recycling the results of the *s*-channel calculation of $e^+e^- \rightarrow f\bar{f}$. This recycling was possible because of a symmetry, called *crossing symmetry*, between processes with the same species, but where species move between initial and final states.

Suppose that we have computed the spin sum of $|\mathcal{M}|^2$ for some process. Form another process by making a series of exchanges, where an incoming particle/antiparticle is replaced with an outgoing antiparticle/particle or vice versa. Each external (incoming or outgoing) state in the original process is assigned to the corresponding external state in the new process. The new process is called a “crossing” of the old process, and its matrix element squared $|\mathcal{M}|^2$ is related to the old process’s value by making the following substitutions:

- (i) replace the momentum of each particle in the first process with the corresponding momentum of the analog particle in the second, with a minus sign if a particle has switched between incoming and outgoing;
- (ii) multiply by (-1) for each fermion which switches between incoming and outgoing.

The origin of the minus sign is the following. When a fermion in the initial state is replaced with an anti-fermion in the final state, the squared matrix element goes from containing $\bar{u}(p)u(p) = -i\not{p} + m$ to containing $\bar{v}(k)v(k) = -i\not{k} + m$. This is not the same as we would get by the substitution $p \rightarrow -k$, it has in addition an overall minus sign. In the case we just considered, there were two such minus signs, so this rule did not matter.

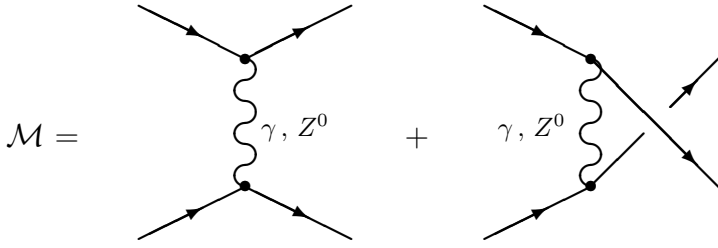


Fig. 6.6 The “uncrossed” and “crossed” graphs for e^-e^- scattering.

We can quickly see that these rules give us the relation we already found between the matrix elements for $e^+(p)e^-(p') \rightarrow f(k)\bar{f}(k')$ and $e^-(p)f(p') \rightarrow e^-(k)f(k')$. The rule says we must multiply by $(-1)^2 = 1$ and make the substitutions, $p \rightarrow p$, $p' \rightarrow -k$, $k \rightarrow k'$, $k' \rightarrow -p'$ in the expression for \mathcal{M}^2 for the former process. Similarly, we can quickly find the matrix element for the process $e^-(p)\bar{f}(p') \rightarrow e^-(k)\bar{f}(k')$; we must multiply by $(-1)^2 = 1$ and make the substitutions, $p \rightarrow p$, $p' \rightarrow -k$, $k \rightarrow -p'$, and $k' \rightarrow k'$. In the ultra-relativistic limit, this changes the Mandelstam variables via

$$\begin{aligned} s &= -2p \cdot p' \rightarrow 2p \cdot k = t \\ t &= 2p \cdot k \rightarrow -2p \cdot p' = s \\ u &= 2p \cdot k' \rightarrow 2p \cdot k' = u \end{aligned} \quad (6.66)$$

Applying these substitutions to Eq. (6.21) gives the differential cross section for $e^- \bar{f} \rightarrow e^- \bar{f}$.

6.6 Interference: Møller scattering

Crossing symmetry makes it possible to recycle a small number of calculations into a complete list of desired matrix elements. However, not all matrix elements are as simple as $e^-e^+ \rightarrow f\bar{f}$. Two additional complications are possible; interference effects and bosons in external states. We will handle them in turn.

Consider the process $e^-e^- \rightarrow e^-e^-$, called Møller scattering. The new complication is that, because the initial and final particles are identical, there are two ways that the electron fields can create the final electrons, which must be summed over. This is equivalent to including both Feynman diagrams of Figure 6.6 – the *uncrossed* and *crossed* graphs – in the amplitude.

There are also two mutually compensating factors of two that arise. One is a factor of $\frac{1}{2}$ in the amplitude relative to the e^-f scattering result that arises because there are no longer two equivalent ways of evaluating the graph depending on whether the f -type vertex appears in the interaction at the spacetime point “ x ” or at the point $x = 0$. The other factor is a factor of 2 in the amplitude relative to the e^-f result that corresponds to the two different interaction operators that can now destroy each of the initial electrons.

The result is to simply add the contribution of the crossed graph to Eq. (6.60). The matrix element associated with the crossed graph is obtained from the matrix element found earlier for e^-f scattering by multiplying by an overall factor of -1 – due to the antisymmetry of fermi statistics for electrons – and then interchanging the four-momenta of the final particles: $k \leftrightarrow k'$. The easiest way to understand the factor of (-1) is to remember that a final state with two fermions in it is antisymmetric under exchange of the fermion labels, $|\mathbf{k}, \mathbf{k}'\rangle = -|\mathbf{k}', \mathbf{k}\rangle$, see Eq. (1.4).

This prescription may be equivalently formulated in terms of the helicity amplitudes, A_{ij} , that appear in the cross section of Eq. (6.65). For scattering between particles of identical helicity, the required substitution is given by $A_{LL}(t) \rightarrow A_{LL}(t) + A_{LL}(u)$ and $A_{RR}(t) \rightarrow A_{RR}(t) + A_{RR}(u)$. Since the amplitudes with mixed helicities, $A_{LR} = A_{RL}$, cannot interfere with one another, the correct replacement for them is

$$\begin{aligned} u^2 \left[|A_{LR}(t)|^2 + |A_{RL}(t)|^2 \right] &= 2u^2 |A_{LR}(t)|^2 \\ &\rightarrow 2 \left[|uA_{LR}(t)|^2 + |tA_{LR}(u)|^2 \right] \end{aligned} \quad (6.67)$$

The differential cross section therefore is

$$\begin{aligned} \frac{d\sigma}{du dt}(e^-e^- \rightarrow e^-e^-) &= -\frac{\pi\alpha^2}{s^2} \left(|A_{LL}(t) + A_{LL}(u)|^2 s^2 + |A_{RR}(t) + A_{RR}(u)|^2 s^2 \right. \\ &\quad \left. + |uA_{LR}(t)|^2 + |tA_{LR}(u)|^2 \right) \delta(s+t+u) \end{aligned} \quad (6.68)$$

In the low-energy limit this expression simplifies to the QED result for ultra-relativistic *Møller scattering*:

$$\begin{aligned} \frac{d\sigma}{du dt}(e^-e^- \rightarrow e^-e^-) &= -\frac{2\pi\alpha^2}{s^2} \left(\left| \frac{s}{t} + \frac{s}{u} \right|^2 + \left| \frac{u}{t} \right|^2 + \left| \frac{t}{u} \right|^2 \right) \delta(s+t+u) \\ &= -\frac{2\pi\alpha^2}{s^2} \left(\frac{s^2+u^2}{t^2} + \frac{s^2+t^2}{u^2} + \frac{2s^2}{ut} \right) \delta(s+t+u) \end{aligned} \quad (6.69)$$

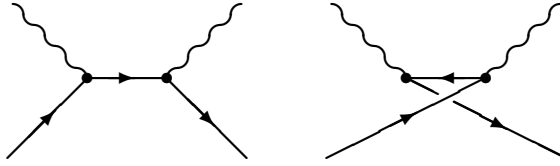


Fig. 6.7 Feynman graphs for Compton scattering.

which becomes, in the CM frame,

$$\frac{d\sigma}{\sin\theta d\theta}(e^-e^- \rightarrow e^-e^-) = \frac{\pi\alpha^2}{s} \left[\frac{1 + \cos^4\left(\frac{\theta}{2}\right)}{\sin^4\left(\frac{\theta}{2}\right)} + \frac{1 + \sin^4\left(\frac{\theta}{2}\right)}{\cos^4\left(\frac{\theta}{2}\right)} + \frac{2}{\sin^2\left(\frac{\theta}{2}\right)\cos^2\left(\frac{\theta}{2}\right)} \right] \quad (6.70)$$

Note however that, to determine the *total* cross section, one should either integrate over only half of available outgoing angles, or integrate over all angles and divide by two, to eliminate a double counting – the final state when an electron emerges (in the CM frame) with angle θ also has an electron emerging with angle $\pi - \theta$, and is therefore identical to the final state where the electron emerges with angle $\pi - \theta$.

We can easily use this result, together with crossing, to find the cross section for *Bhabha scattering*, $e^-e^+ \rightarrow e^-e^+$, see Problem 6.1.

6.7 Processes involving photons

So far we have only considered scattering processes in which all external lines are fermions. To be complete, we will briefly discuss what happens when an external line is a massless gauge boson. The reason we concentrate on a massless gauge boson is that, for the massive case, the polarization sum, Eq. (1.119), $\sum_\lambda \epsilon_\mu(p)\epsilon_\nu^*(p) = \eta_{\mu\nu} + p_\mu p_\nu/m^2$, is obviously Lorentz-covariant; but the corresponding massless sum, Eq. (1.132), is not obviously so. However, it turns out that there is a simple, Lorentz-invariant way to determine polarization averaged cross sections, thanks to gauge invariance.

6.7.1 Compton scattering, $e^-\gamma \rightarrow e^-\gamma$

Consider then the process $e^-\gamma \rightarrow e^-\gamma$. There are two Feynman graphs for this process, shown in Figure 6.7. Labeling the incoming electron momentum

\mathbf{p}_1 , the incoming photon momentum \mathbf{p}_2 , the outgoing electron momentum \mathbf{p}_3 , and the outgoing photon momentum \mathbf{p}_4 , the corresponding matrix element is

$$\begin{aligned} \mathcal{M} = & e^2 \epsilon_\mu(\mathbf{p}_2, \lambda) \epsilon_\nu(\mathbf{p}_4, \lambda') \\ & \times \bar{u}(\mathbf{p}_3, \sigma') \left(\gamma^\nu \frac{-i(\not{p}_1 + \not{p}_2) + m_e}{(p_1 + p_2)^2 + m_e^2} \gamma^\mu + \gamma^\mu \frac{-i(\not{p}_3 - \not{p}_2) + m_e}{(p_3 - p_2)^2 + m_e^2} \gamma^\nu \right) u(\mathbf{p}_1, \sigma) \end{aligned} \quad (6.71)$$

We will eventually sum over polarizations for each photon, using Eq. (1.132),

$$\sum_\lambda \epsilon_\mu(\mathbf{p}_2, \lambda) \epsilon_\alpha^*(\mathbf{p}_2, \lambda) = \eta_{\mu\alpha} + p_\mu \bar{p}_\alpha + \bar{p}_\mu p_\alpha \quad (6.72)$$

At first sight this is worrying, since \bar{p}^μ is not uniquely defined; it is not obvious that the polarization-summed cross section will be Lorentz-invariant.

In fact, Lorentz invariance follows from the fact that, if we substitute $\epsilon^\mu(p_2) \rightarrow p_2^\mu$ in the matrix element, we get zero:

$$\begin{aligned} & \bar{u}(\mathbf{p}_3, \sigma') \left(\not{\epsilon}' \frac{-i(\not{p}_1 + \not{p}_2) + m_e}{(p_1 + p_2)^2 + m_e^2} \not{p}_2 + \not{p}_2 \frac{-i(\not{p}_3 - \not{p}_2) + m_e}{(p_3 - p_2)^2 + m_e^2} \not{\epsilon}' \right) u(\mathbf{p}_1, \sigma) \\ = & \bar{u}(\mathbf{p}_3, \sigma') \left(\not{\epsilon}' \frac{-i(\not{p}_1 + \not{p}_2) + m_e}{(p_1 + p_2)^2 + m_e^2} [(\not{p}_1 + \not{p}_2 - im_e) - (\not{p}_1 - im_e)] \right) u(\mathbf{p}_1, \sigma) \\ & + \bar{u}(\mathbf{p}_3, \sigma') \left([(\not{p}_2 - \not{p}_3 + im_e) + (\not{p}_3 - im_e)] \frac{-i(\not{p}_3 - \not{p}_2) + m_e}{(p_3 - p_2)^2 + m_e^2} \not{\epsilon}' \right) u(\mathbf{p}_1, \sigma) \\ = & \bar{u}(\mathbf{p}_3, \sigma') \left(\not{\epsilon}' \frac{-i(p_1 + p_2)^2 - im_e^2}{(p_1 + p_2)^2 + m_e^2} + \frac{i(p_3 - p_2)^2 + im_e^2}{(p_3 - p_2)^2 + m_e^2} \not{\epsilon}' \right) u(\mathbf{p}_1, \sigma) \\ = & 0 \end{aligned} \quad (6.73)$$

Here $\epsilon' \equiv \epsilon(p_4, \lambda')$. In passing from the second to third lines, we have used the Dirac equation, $(\not{p}_1 - im_e)u(\mathbf{p}_1) = 0$ and $\bar{u}(\mathbf{p}_3)(\not{p}_3 - im_e) = 0$.

Because of this relation, when we carry out the summation over spin states, the $\bar{p}_\mu p_\alpha$ terms in Eq. (6.72) will not contribute; we are therefore free to make the substitution

$$\sum_\lambda \epsilon_\mu(p, \lambda) \epsilon_\alpha^*(p, \lambda) \rightarrow \eta_{\mu\alpha} \quad (6.74)$$

The same holds for the final state photon, as we can quickly verify by substituting $\epsilon'_\mu \rightarrow (p_4)_\mu$ in Eq. (6.71). This not only provides a rather substantial simplification, but it also makes manifest the Lorentz invariance of the spin summed cross section.

The cancellation in Eq. (6.73), while at first sight rather remarkable, turns out to be an absolutely general property of external photon lines; when summing over all possible diagrams contributing to the matrix element,

replacing $\epsilon_\mu(p) \rightarrow p_\mu$ always gives zero, so Eq. (6.74) may always be used. The physical origin of this property is gauge invariance. To ensure gauge invariance, it was necessary to ensure that the electromagnetic gauge field A_μ always couples to a conserved current; $\mathcal{L}_{\text{int}} = \int d^4x A_\mu J^\mu$ with $\partial_\mu J^\mu = 0$ (see Subsection 1.5.2). In Fourier space, current conservation becomes $p_\mu J^\mu(p) = 0$, precisely the relation we need, because the current J^μ is what the polarization tensor ϵ_μ is contracted against.

Now we proceed with computing the spin averaged matrix element. To simplify expressions we will take the ultra-relativistic $m_e \rightarrow 0$ limit. Squaring Eq. (6.71), performing the spin sums, and using Eq. (6.74) for the polarization sums, gives

$$\begin{aligned} \overline{\mathcal{M}}^2 &= \frac{e^4}{4} \text{tr} \not{p}_1 \left(\gamma^\mu \frac{\not{p}_1 + \not{p}_2}{(p_1 + p_2)^2} \gamma^\nu + \gamma^\nu \frac{\not{p}_3 - \not{p}_2}{(p_3 - p_2)^2} \gamma^\mu \right) \times \\ &\quad \times \not{p}_3 \left(\gamma_\nu \frac{\not{p}_1 + \not{p}_2}{(p_1 + p_2)^2} \gamma_\mu + \gamma_\mu \frac{\not{p}_3 - \not{p}_2}{(p_3 - p_2)^2} \gamma_\nu \right) \\ &= \frac{e^4}{4} \text{tr} \not{p}_1 \left(\gamma^\mu \frac{\not{p}_1 + \not{p}_2}{s} \gamma^\nu + \gamma^\nu \frac{\not{p}_3 - \not{p}_2}{u} \gamma^\mu \right) \not{p}_3 \left(\gamma_\nu \frac{\not{p}_1 + \not{p}_2}{s} \gamma_\mu + \gamma_\mu \frac{\not{p}_3 - \not{p}_2}{u} \gamma_\nu \right) \end{aligned} \quad (6.75)$$

Expanding gives four terms. The first term is

$$\begin{aligned} \frac{e^4}{4} \text{tr} \not{p}_1 \gamma^\mu \frac{\not{p}_1 + \not{p}_2}{s} \gamma^\nu \not{p}_3 \gamma_\nu \frac{\not{p}_1 + \not{p}_2}{s} \gamma_\mu &= \frac{e^4}{s^2} \text{tr} \not{p}_1 (\not{p}_1 + \not{p}_2) \not{p}_3 (\not{p}_1 + \not{p}_2), \\ &= \frac{e^4}{s^2} \text{tr} \not{p}_1 \not{p}_2 \not{p}_3 \not{p}_2 \\ &= \frac{8e^4}{s^2} (p_1 \cdot p_2 p_2 \cdot p_3) \\ &= -\frac{2e^4}{s^2} (su) = -2e^4 \frac{u}{s} \end{aligned} \quad (6.76)$$

where we use repeatedly the gamma matrix identities of Problem 4.3, and the approximation $p_1^2 = 0$. The second term is

$$\begin{aligned} \text{tr} \left[\not{p}_1 \gamma^\mu \frac{\not{p}_1 + \not{p}_2}{s} \gamma^\nu \not{p}_3 \gamma_\mu \frac{\not{p}_3 - \not{p}_2}{u} \gamma_\nu \right] &= -2 \text{tr} \left[\not{p}_1 \not{p}_3 \gamma^\nu \frac{\not{p}_1 + \not{p}_2}{s} \frac{\not{p}_3 - \not{p}_2}{u} \gamma_\nu \right] \\ &= 8 \frac{(p_1 + p_2) \cdot (p_3 - p_2)}{su} \text{tr} \not{p}_1 \not{p}_3 \\ &= 0, \end{aligned} \quad (6.77)$$

because $(p_1 + p_2) \cdot (p_3 - p_2) = (t/2 + s/2 + u/2 + 0) = 0$. The third term is the Hermitian conjugate of the second, and also vanishes; the fourth term's

evaluation is similar to the first. The result is

$$\overline{\mathcal{M}}^2(e^- \gamma \rightarrow e^- \gamma) = -2e^4 \left(\frac{u}{s} + \frac{s}{u} + \dots \right) \quad (6.78)$$

This is positive, because $u < 0$ and $s > 0$. In this expression the ellipses indicate terms which vanish as $m^2 \rightarrow 0$ with s, t , and u fixed, which are required in order to properly capture the entire Compton-scattering cross section even in the ultra-relativistic limit (see Problem 6.6 for details). These additional terms are required because there are also ms hidden within the definitions of s, t , and u , and these conspire to ensure that the terms neglected in Eq. (6.78) compete with those that are included. This expression is more useful once it is used to obtain the cross section for e^+e^- annihilation, using crossing symmetry.

It is elementary to apply crossing symmetry to determine the annihilation rate $e^-e^+ \rightarrow \gamma\gamma$. Labeling the incoming momenta \mathbf{p}_1 and \mathbf{p}_2 and the final momenta \mathbf{p}_3 and \mathbf{p}_4 , the momenta are reassigned via $p_3 \rightarrow -p_2$, $p_2 \rightarrow -p_3$. The Mandelstam variables are changed according to

$$\begin{aligned} s &= -2p_1 \cdot p_2 \rightarrow +2p_1 \cdot p_3 = t \\ t &= +2p_1 \cdot p_3 \rightarrow -2p_1 \cdot p_2 = s \\ u &= +2p_1 \cdot p_4 \rightarrow +2p_1 \cdot p_4 = u \end{aligned} \quad (6.79)$$

Furthermore, there is a factor of $(-1)^1$ because one fermion is reassigned from the final to the initial state. Therefore, the spin averaged matrix element squared is

$$\overline{\mathcal{M}}^2(e^-e^+ \rightarrow \gamma\gamma) = +2e^4 \left(\frac{t}{u} + \frac{u}{t} \right) \quad (6.80)$$

Since both t and u are negative, the (-1) ensures that the result is positive. This result manifestly has the right behavior if we interchange the two photon states, corresponding to $t \leftrightarrow u$. Using it to obtain the cross section leads to the correct ultra-relativistic limit

$$\frac{d\sigma}{dt du} = -\frac{2\pi\alpha^2}{s^2} \left(\frac{t}{u} + \frac{u}{t} \right) \delta(s + t + u) \quad (6.81)$$

and so

$$\frac{d\sigma}{\sin\theta d\theta} = \frac{2\pi\alpha^2}{s} \left(\frac{1 + \cos^2\theta}{\sin^2\theta} \right) \quad (6.82)$$

Since the photons are identical particles, to find the total cross section one must integrate over only half of the space of final-state angles.

The interaction of quarks with gluons is almost the same as the electromagnetic processes we have considered here; the main added complication

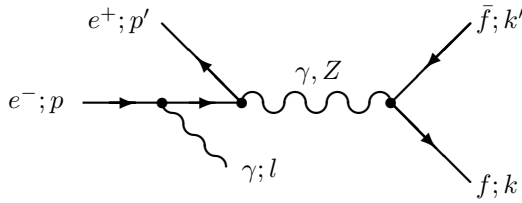


Fig. 6.8 Diagram for scattering with a photon emission.

is the appearance of color factors. However, when the mutual interactions of gluons via the three or four gluon vertices, Eq. (5.63) and Eq. (5.64), are involved, then the polarization summation issues are more complicated and it is not permitted to make the substitution, Eq. (6.74). We postpone further discussion on this point to Chapter 9.

6.7.2 Radiated photons

Another application where photons appear in the final state is when one or more photons are radiated from a participating particle (initial or final) in a scattering process, for instance via the Feynman graph depicted in Figure 6.8.

Naively, such a process is suppressed by $\sim e^2/(2\pi)^2 = \alpha/\pi \ll 1$ with respect to the diagram without a photon emission, and should therefore be negligibly small. However, this is not entirely correct. The probability to emit a photon in a scattering process is enhanced by *large logarithms*, which can make it quite important. In particular, at the end of this subsection, we will see that the on-resonance cross section for $e^+e^- \rightarrow Z \rightarrow f\bar{f}$ is reduced by about 1/4 due to these processes.

Consider the process shown in Figure 6.8. This is one of four diagrams which must be summed, and then squared, to find the photon emission rate. For general angles, when the center-of-mass-frame angle between the photon and any other particle is large, the naive estimate that the photon emission rate is suppressed by α/π , is correct. However this turns out not to be true in the special kinematic regime in which the photon is approximately collinear with another particle, for instance, when the angle between \mathbf{p} and \mathbf{l} in the CM frame is small.

In this regime, the four possible diagrams (with the photon attached to any of the four external states) all contribute, but they do not contribute equally. When the photon emerges from the e^- , the denominator of the

intermediate electron propagator is $1/((p-l)^2 + m_e^2) = 1/(-2p \cdot l) \simeq 1/(2p^0 l^0(1 - \cos \theta_{\mathbf{pl}}))$. Since $(1 - \cos \theta_{\mathbf{pl}}) \simeq 0$, this amplitude is enhanced, relative to the others. It turns out in this case that it is simplest to work in terms of the two transverse polarization states, in which case the amplitude in question is $\sim \sin(\theta_{\mathbf{pl}})/(1 - \cos \theta_{\mathbf{pl}}) \sim 1/\theta_{\mathbf{pl}}$, while the other amplitudes are not enhanced at small $\theta_{\mathbf{pl}}$. Therefore, it is permissible to drop the other amplitudes to determine the leading behavior in this small angle region.

Let us evaluate just the part of the square of the matrix element involving the photon emission, to compare with the case without photon emission. Where, without photon emission, we have the quantity $\sum_{\sigma} u(p, \sigma) \bar{u}(p, \sigma) = (-i\not{p} + m)$, we now have the quantity

$$\begin{aligned}
 & -e^2 \sum_{\lambda} \frac{-i\not{p} + i\not{l} + m}{(p-l)^2 + m^2} \not{\epsilon}^*(\lambda) (-i\not{p} + m) \not{\epsilon}(\lambda) \frac{-i\not{p} + i\not{l} + m}{(p-l)^2 + m^2} \\
 &= \frac{-e^2}{(2p \cdot l)^2} \sum_{\lambda} (-i\not{p} + i\not{l} + m) \left\{ 2(i\not{p} + m) - 2i\epsilon \cdot p \not{\epsilon}^* \right\} (-i\not{p} + i\not{l} + m) \\
 &= \frac{-e^2}{(2p \cdot l)^2} \sum_{\lambda} [-4il \cdot p \not{l} - 2i\epsilon \cdot p (-i\not{p} + i\not{l} + m) \not{\epsilon}^* (-i\not{p} + i\not{l} + m)] \\
 &\simeq \frac{ie^2 \not{l}}{p \cdot l} + \sum_{\lambda} \frac{e^2 |\epsilon \cdot p|^2}{(p \cdot l)^2} (-i\not{p} + i\not{l} + m)
 \end{aligned} \tag{6.83}$$

where in the second step we used $l^2 = 0$ and $p^2 + m^2 = 0$, and in the last step we dropped a term $\propto (\epsilon \cdot p)(p \cdot l)/(p \cdot l)^2$, which is subleading in $\theta_{\mathbf{pl}}$. Since $\sum_{\lambda} |\epsilon \cdot p|^2 \simeq (p^0)^2 \theta_{\mathbf{pl}}^2$ and $l \cdot p \simeq -l^0 p^0 \theta_{\mathbf{pl}}^2/2$, the remaining two terms are of the same order. Since l is almost collinear with p , for the rest of the calculation it is adequate to substitute $\not{l} \rightarrow (l^0/p^0)\not{p}$. Defining $l^0/p^0 \equiv x$, so x is the fraction of the electron's energy carried off by the photon, we find that the matrix element in the collinear limit is approximately given by

$$(-i\not{p}) \rightarrow (-i\not{p}) \frac{2e^2(1 + (1-x)^2)}{x^2(p^0)^2 \theta_{\mathbf{pl}}^2} \tag{6.84}$$

Being slightly more careful,

$$-l \cdot p \simeq p^0 l^0 \left(1 - \cos \theta_{\mathbf{pl}} + \frac{m_e^2}{2(p^0)^2} \right) \simeq \frac{(p^0)^2 x}{2} \left(\theta_{\mathbf{pl}}^2 + \frac{m_e^2}{(p^0)^2} \right) \tag{6.85}$$

This correction at very small angles exists in both terms, and is important in cutting off the otherwise divergent angular integral. We can now do most of the integral over the photon momentum;

$$\int \frac{d^3 l}{(2\pi)^3 2l^0} = \frac{1}{8\pi^2} \int_{-1}^1 d \cos \theta_{\mathbf{pl}} l^0 dl^0 \tag{6.86}$$

The integral over the angle is

$$\int_{-1}^1 d \cos \theta_{\mathbf{p}\mathbf{l}} \frac{1}{1 + \frac{m_e^2}{2p^2} - \cos \theta_{\mathbf{p}\mathbf{l}}} = \log \left(\frac{4p^2}{m_e^2} \right) + O(1) = \log \left(\frac{s}{m_e^2} \right) + O(1) \quad (6.87)$$

There is an unknown constant in this expression, because the approximations we have made break down at large angles. This constant could be found by making a more careful treatment, which included the interference between emissions from different lines at large angles. Nevertheless, our simplified treatment is sufficient to show that photon emission is logarithmically dominated by small opening angles, with the log cut off by the mass of the emitting particle; and it is sufficient to find the coefficient of that log.

The integration over l^0 is more delicate, because the emission of the photon potentially changes the kinematics of the rest of the diagram. One must recompute the remainder of the diagram, but changing the momentum carried by the incoming electron line from p to $(1-x)p$. For the case of emission from a final state line, f or \bar{f} , this does not matter. For the case of emission from an incoming line, e^+ or e^- , this *can* matter. In general, it matters if $x \sim 1$, so the kinematics is substantially disturbed. In the case of scattering through a resonance, such as the Z boson, an energy loss of $l^0 \sim \Gamma_Z$, or $x \sim \Gamma_Z/M_Z$, is important, as we discuss below. Nevertheless, for the present purposes we will ignore this complication. The likelihood to emit a photon with energy larger than ω is

$$\frac{e^2}{8\pi^2} \log \frac{s}{m_e^2} \int_{\omega/p}^1 \frac{x dx}{x^2} [1 + (1-x)^2] = \frac{\alpha}{2\pi} \log \left(\frac{s}{m_e^2} \right) \left[\log \left(\frac{s}{4\omega^2} \right) - \frac{3}{2} \right] \quad (6.88)$$

This likelihood is not necessarily small, because it is amplified by the product of two logarithms which may each individually be large. We should also add the probabilities to emit from each of the other three legs, giving a total likelihood of

$$\frac{\sigma(\omega)}{\sigma} \simeq \frac{\alpha}{\pi} \left[\log \left(\frac{s}{m_e^2} \right) + Q_f^2 \log \left(\frac{s}{m_f^2} \right) + O(1) \right] \left[\log \left(\frac{s}{4\omega^2} \right) - \frac{3}{2} \right] \quad (6.89)$$

The likelihood to emit a photon during a scattering can in fact be quite large. For instance, on the Z pole, for $\omega = m_e$, and considering only emissions from the incoming electrons, this evaluates to approximately 1.2.

Once again, we find a higher-order effect (more powers of α) which is just as large as the original effect. In fact, the result presented appears to be sick; if we ask for the rate to scatter with the emission of an arbitrarily soft photon, that rate is logarithmically divergent. It also appears that the

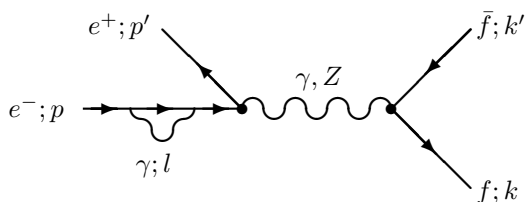


Fig. 6.9 A radiative correction to e^+e^- annihilation.

total rate for the scattering to occur, after we add the possibility of this photon emission, is much larger than we had previously computed. The situation becomes still worse as we consider multiple-photon emission. This suggests that, as in Section 6.4, there may be some additional, formally higher order, diagrams which can compete with the lowest-order one and must be somehow resummed. This proves to be the case.

Consider the diagram of Figure 6.9. When it interferes with the diagram of Figure 6.1, this modifies the rate of scattering *without* photon emission. The interference of those diagrams is structurally very similar to the square of the diagram of Figure 6.8. In particular, the photon in the loop, in Figure 6.9, contributes the same collinear and soft divergences as the real, emitted photon of Figure 6.8. There is one major difference, however; the sign turns out to be opposite. Therefore, the diagram of Figure 6.9 reduces the probability for the scattering process, without the photon emission.

Because the soft and collinear singularities are identical, the cross section to emit an extra photon, and the *reduction* of the cross section to scatter without a photon emission, are approximately the same size. Therefore, the *total* cross section, with or without photon emission, is unchanged up to $O(\alpha)$ correction, except for the correction to the kinematics because of the energy carried away by the photons.

The correct way to handle the computation of the total cross section, and the likelihood to emit any number of photons above some energy threshold, has been solved by Bloch and Nordsieck. One must introduce a cut-off frequency ω , below which a photon cannot be detected. (Realistically, all experiments have such a cut-off.) Then, the likelihood *not* to emit a collinear photon is given, approximately, by

$$\frac{\sigma(e^+e^- \rightarrow f\bar{f} + 0\gamma)}{\sum_n \sigma(e^+e^- \rightarrow f\bar{f} + n\gamma)} \simeq \exp(-\chi) \quad (6.90)$$

with

$$\chi = \frac{\alpha}{\pi} \left[\log \left(\frac{s}{m_e^2} \right) + Q_f^2 \log \left(\frac{s}{m_f^2} \right) + O(1) \right] \left[\log \left(\frac{s}{4\omega^2} \right) - \frac{3}{2} \right] \quad (6.91)$$

This is the expression for s channel exchange. In t channel exchange, it is t , not s , which controls the size of the collinear logarithm.

This phenomenon of photon emission from the initial state is referred to as *initial-state radiation*, or ISR. Let us see its effects on scattering processes. In a typical s channel scattering process, the kinematics is significantly changed if one of the initial particles radiates a photon carrying $x \sim 1$ of the energy. In this case, the second bracketed quantity should be replaced with $\simeq 1$. The change to the cross section can be several percent – not huge, but important when precision is required.

On the other hand, when the e^+e^- particles' energies are tuned to lie on the Z pole, the radiation of a photon with energy $\Gamma_Z/2$, the half-width of the resonance, is enough to move the scattering off resonance and reduce the cross section substantially. Therefore, the on-resonance cross section for $e^+e^- \rightarrow f\bar{f}$ is reduced by a factor of approximately

$$\begin{aligned} \frac{\sigma_{\text{tot}}(e^+e^- \rightarrow Z^0)}{\text{Eq. (6.59)}} &\simeq \exp \left(\frac{-\alpha}{\pi} \log \left(\frac{M_Z^2}{m_e^2} \right) \left[\log \left(\frac{M_Z^2}{\Gamma_Z^2} \right) - \frac{3}{2} \right] \right) \\ &\simeq 0.72 \end{aligned} \quad (6.92)$$

where in the numerical evaluation we used $\alpha \simeq 1/133$, a compromise between its value at the scale M_Z , where $\alpha \simeq 1/128$, and at m_e , where $\alpha \simeq 1/137$. Combining this with the fact that 30% of Z^0 decays are to leptons, the *hadronic* cross section on the Z^0 pole is expected to be only $\simeq 30$ nb. The actual hadronic cross section on resonance is 30.5 nb, in agreement with a more careful calculation.

We illustrate the impact of initial state radiation on the cross section near the Z resonance in Figure 6.10. The initial-state-radiation corrected cross section is an integral over the energy lost to photons, of the probability for that energy loss due to initial-state radiation times the cross section at the reduced energy. The correction is very important in the agreement between theory and data, as shown by the inclusion of the cross section data from the four LEP experiments. Note that at high energies, the cross section actually exceeds the uncorrected value. This is because initial state radiation can lower the e^+e^- pair's energy to lie on resonance, enhancing the cross section.

Initial-state radiation also plays a prominent role in the physics of high-energy hadronic collisions, which we will discuss in Section 9.2, especially Subsection 9.2.3.

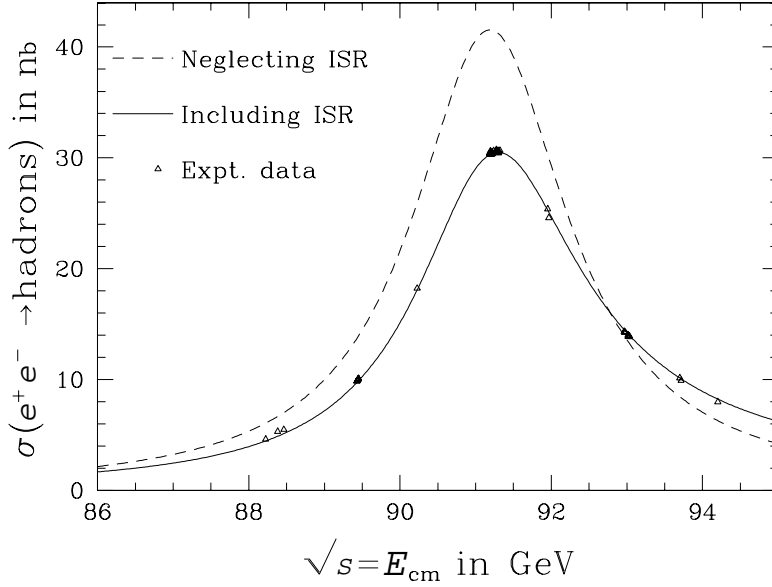


Fig. 6.10 e^+e^- hadronic cross section near the Z^0 resonance.

6.8 Problems

[6.1] Crossing symmetry

Use crossing symmetry to derive the cross section for Bhabha scattering in the ultra-relativistic limit, but still at $s \ll M_Z^2$:

$$\frac{d\sigma}{du dt}(e^-e^+ \rightarrow e^-e^+) = \frac{-2\pi\alpha^2}{s^2} \left(\left| \frac{u}{s} + \frac{u}{t} \right|^2 + \frac{t^2}{s^2} + \frac{s^2}{t^2} \right) \delta(s+t+u)$$

[6.2] Electron–neutrino scattering

[6.2.1] The process $e^-\bar{\nu}_e \rightarrow f_m\bar{f}_n$ (f neither e^- nor ν_e) proceeds via a single diagram. Draw that diagram and show that, taking all external states to be massless, it yields a cross section of

$$\frac{d\sigma}{du dt} = -\frac{G_F^2}{\pi} N_c |U_{nm}|^2 \frac{u^2}{s^2} \left(\frac{M_W^2}{s - M_W^2} \right)^2 \delta(s+t+u)$$

Use crossing symmetry to find $d\sigma/du dt$ for $e^-\nu_\mu \rightarrow \mu^-\nu_e$ from this result.

[6.2.2] Compute the matrix element squared for $e^-\nu_e \rightarrow e^-\nu_e$. Careful: there are two diagrams, one involving Z exchange and one involving W exchange. Use crossing symmetry to find the $e^-\bar{\nu}_e \rightarrow e^-\bar{\nu}_e$ and $e^-e^+ \rightarrow \nu_e\bar{\nu}_e$ matrix elements. Beware; there are two diagrams, and a

relative minus sign between them due to the different ways the initial- and final-state fermions connect to each other.

[6.3] Supernova neutrinos

Neutrinos and antineutrinos were observed in 1987 from the supernova in the nearby Large Magellanic Cloud. They were detected by observing their interactions with electrons and nuclei in a large tank of water. The neutrino energies were ~ 10 MeV and in this energy range the relevant processes are elastic scattering with electrons and “quasi-elastic” (i.e. low energy) inverse beta decay: $\bar{\nu}_e + p \rightarrow n + e^+$ with the hydrogen nuclei.

[6.3.1] Why not consider the interactions with the oxygen nucleus, O^{16} ?

$$\begin{aligned} \text{e.g. :} \quad & \bar{\nu}_e + {}_8^{16}\text{O} \rightarrow {}_7^{16}\text{N} + e^+ \\ \text{or :} \quad & \nu_e + {}_8^{16}\text{O} \rightarrow {}_9^{16}\text{F} + e^- \end{aligned}$$

(Hint: *Nucl. Phys.* **166** (1971) page 60.)

[6.3.2] In these reactions the interactions are detected by observing the Cherenkov radiation of the final charged particles. Given that the index of refraction of water is $n = 1.33$, what is the minimum energy that a particle of mass m must have to radiate Cherenkov light? Why not consider also elastic scattering of neutrinos by the hydrogen and oxygen nuclei?

[6.3.3] In the standard model, draw all of the tree level (i.e. no loops) Feynman graphs that contribute to the following processes:

- (i) $\nu_\mu + e^- \rightarrow \nu_\mu + e^-$
- (ii) $\bar{\nu}_\mu + e^- \rightarrow \bar{\nu}_\mu + e^-$
- (iii) $\nu_e + e^- \rightarrow \nu_e + e^-$
- (iv) $\bar{\nu}_e + e^- \rightarrow \bar{\nu}_e + e^-$

[6.3.4] *Kinematics*

- (i) Show that, for a neutrino of energy ω , the angle between the direction of the incident neutrino momentum and the scattered electron momentum in the lab frame, θ , is related to the same angle in the center-of-mass frame, φ , by

$$\tan \theta = \frac{\sin \varphi}{1 + \cos \varphi} \frac{\sqrt{m^2 + 2m\omega}}{m + \omega}$$

This implies that the electron–neutrino interactions produce electrons that tend to move away from the supernova (with an angular spread of $\Delta\theta \sim \sqrt{m/\omega} \sim 1/\sqrt{20}$ for 10 MeV neutrinos)

regardless of the scattering probability $d\sigma/d\cos\varphi$ in the center-of-mass frame.

- (ii) Show that the energy of the final electron in the lab frame, E , is given in terms of the incident neutrino energy and scattering angle by

$$E = m \left[\frac{(m + \omega)^2 + \omega^2 \cos^2 \theta}{(m + \omega)^2 - \omega^2 \cos^2 \theta} \right]$$

For fixed ω , what is the difference between the energy at $\theta = 0$ and that when $\theta = \sqrt{(m/\omega)} \ll 1$? This represents the energy range of the scattered electrons from this process.

[6.3.5] In the effective Fermi theory of weak interactions (basically, the limit $s, |t|, |u| \ll M_W^2$ so they can be dropped from the W and Z boson propagators), the relevant interaction Hamiltonian density is

$$\mathcal{H}_I = \frac{G_F}{\sqrt{2}} \left\{ [\bar{\nu}_e \gamma^\mu (1 + \gamma_5) e] [\bar{e} \gamma_\mu (1 + \gamma_5) \nu_e] - \frac{1}{2} \sum_{m=e, \mu, \tau} [\bar{\nu}_m \gamma^\mu (1 + \gamma_5) \nu_m] [\bar{e} \gamma_\mu (\rho + \gamma_5) e] \right\}$$

where $\rho = 1 - 4 \sin^2 \theta_W$.

The first term can be rewritten, using the Fiertz identities of Problem 1.6, as

$$[\bar{\nu} \gamma^\mu (1 + \gamma_5) e] [\bar{e} \gamma_\mu (1 + \gamma_5) \nu] = [\bar{\nu}_e \gamma^\mu (1 + \gamma_5) \nu_e] [\bar{e} \gamma_\mu (1 + \gamma_5) e]$$

so

$$\mathcal{H}_I = \frac{G_F}{\sqrt{2}} \sum_{m=e, \mu, \tau} [\bar{\nu}_m \gamma^\mu (1 + \gamma_5) \nu_m] [\bar{e} \gamma_\mu (h_V^{(m)} + \gamma_5 h_A^{(m)}) e]$$

$$\text{where : } h_A^{(e)} = 1 - \frac{1}{2} = \frac{1}{2}$$

$$h_A^{(\mu, \tau)} = -\frac{1}{2}$$

$$\text{and : } h_V^{(e)} = 1 - \frac{1}{2} \rho = \frac{1}{2} + 2 \sin^2 \theta_W$$

$$h_V^{(\mu, \tau)} = -\frac{1}{2} \rho = -\frac{1}{2} + 2 \sin^2 \theta_W$$

Use this result to show that the matrix element for neutrino-electron scattering is

$$\mathcal{M}(\nu e \rightarrow \nu e) = \frac{G_F}{\sqrt{2}} [\bar{u}(q') \gamma^\mu (1 + \gamma_5) u(q)] [\bar{u}(p') \gamma_\mu (h_V^{(m)} + h_A^{(m)} \gamma_5) u(p)]$$

and

$$\mathcal{M}(\bar{\nu}e \rightarrow \bar{\nu}e) = \frac{G_F}{\sqrt{2}} [\bar{v}(q)\gamma^\mu(1 + \gamma_5)v(q')] [\bar{u}(p')\gamma_\mu(h_V^{(m)} + h_A^{(m)}\gamma_5)u(p)]$$

[6.3.6] Averaging over the initial electron spin and summing over all final spins, show that

$$\begin{aligned} \overline{\mathcal{M}}^2 = 16G_F^2 \Big\{ & (h_V \pm h_A)^2(p \cdot q)(p' \cdot q') \\ & + (h_V \mp h_A)^2(p \cdot q')(p' \cdot q) + m^2(h_V^2 - h_A^2)(q \cdot q') \Big\} \end{aligned}$$

in which the upper sign corresponds to $\nu e \rightarrow \nu e$ and the lower sign to $\bar{\nu}e \rightarrow \bar{\nu}e$.

[6.3.7] Using

$$d\sigma = \frac{1}{-2p \cdot q v_{\text{rel}}} \overline{\mathcal{M}}^2 (2\pi)^4 \delta^4(p + q - p' - q') \frac{d^3\mathbf{p}' d^3\mathbf{q}'}{(2\pi)^6 2p'^0 2q'^0}$$

(i) Show that the differential cross-section in the center-of-mass frame is (neglecting the electron mass)

$$\frac{d\sigma}{d(\cos \varphi)} = \frac{G_F^2}{2\pi} \omega_{\text{cm}}^2 \left\{ (h_V \pm h_A)^2 + \frac{1}{4}(h_V \mp h_A)^2(1 - \cos \varphi)^2 \right\}$$

where ω_{cm} is the incident neutrino energy, and φ is the angle between the incident neutrino momentum and the final electron momentum, as before. What is the most probable center-of-mass scattering angle?

(ii) Show that the total cross section is (in terms of the lab energy of the neutrino)

$$\sigma = \frac{G_F^2 \omega m}{2\pi} \left[(h_V \pm h_A)^2 + \frac{1}{3}(h_V \mp h_A)^2 \right] (1 + O(m/\omega))$$

Using the values of the parameters $h_V^{(m)}, h_A^{(m)}$ given earlier and $\sin^2 \theta_W \approx 1/4$, calculate the ratios

$$\begin{aligned} \sigma(\nu_e e \rightarrow \nu_e e) : \sigma(\bar{\nu}_e e \rightarrow \bar{\nu}_e e) : \sigma(\nu_\mu e \rightarrow \nu_\mu e) : \\ \sigma(\bar{\nu}_\mu e \rightarrow \bar{\nu}_\mu e) : \sigma(\nu_\tau e \rightarrow \nu_\tau e) : \sigma(\bar{\nu}_\tau e \rightarrow \bar{\nu}_\tau e) \end{aligned}$$

[6.3.8] For nucleon-neutrino scattering ($\bar{\nu}_e + p^+ \rightarrow n + e^+$) at low energies, the weak current has matrix elements:

$$\langle n | J_{\text{had}}^\mu | p \rangle = \bar{u}_n \gamma^\mu (g_V + g_A \gamma_5) u_p$$

with $g_V = 1$ and $g_A \simeq 1.269$. Using

$$\mathcal{H}_{\text{weak}} = \frac{G_F}{\sqrt{2}} V_{ud} [\bar{\nu}_e \gamma_\mu (1 + \gamma_5) e] J_{\text{had}}^\mu$$

show that

$$\begin{aligned} \mathcal{M}(p^+ \bar{\nu}_e \rightarrow n e^+) &= \frac{G_F}{\sqrt{2}} V_{ud} \bar{v}_{(\nu)}(\mathbf{q}) \gamma^\mu (1 + \gamma_5) v_{(e)}(\mathbf{q}') \\ &\quad \times \bar{u}_n(\mathbf{p}') \gamma_\mu (g_V + g_A \gamma_5) u_p(\mathbf{p}) \end{aligned}$$

where V_{ud} is the relevant Kobayashi–Maskawa matrix element.

[6.3.9] Treat the nucleon mass $m_p \gg \omega$ the neutrino energy.

- (i) Neglecting $m_n - m_p$ and the electron mass, what is the lab energy of the final electron as a function of scattering angle and incident neutrino energy? (This is almost a trick question.)
- (ii) Show that the differential scattering cross section as a function of the lab-frame scattering angle, θ , between the incident neutrino direction and the final positron direction is

$$\frac{d\sigma}{d(\cos \theta)} \approx \frac{G_F^2 \omega^2}{2\pi} \left\{ (g_V^2 + 3g_A^2) + (g_V^2 - g_A^2) \cos \theta \right\}$$

(Neglect ω/m_N , m_e/ω , and $(m_n - m_p)/\omega$.) Notice the angular distribution of final electrons is different than that in the case of electron–neutrino scattering, and is effectively independent of θ (recall $(g_V^2 + 3g_A^2)/(g_V^2 - g_A^2) \approx -10$).

- (iii) Calculate the total cross section as a function of the lab frame neutrino energy, ω .

[6.3.10] Using the given values for $h_A^{(m)}$, $h_V^{(m)}$, and g_V, g_A , and $\omega = 10$ MeV, calculate what percentage of observed events are expected to be due to $\bar{\nu}_e p$, $\nu_e e$, $\bar{\nu}_e e$, $\nu_\mu e$, $\bar{\nu}_\mu e$, $\nu_\tau e$, and $\bar{\nu}_\tau e$ scattering, assuming the supernova emits equal numbers of all types of neutrinos.

Don't forget that only two of the protons in a water molecule are in hydrogen! That is, there are two protons but ten electrons in a water molecule.

[6.3.11] *ν_μ -Matter interactions* In neutrino–matter scattering experiments at accelerators, $\nu_\mu + e$ and $\nu_\mu +$ nucleon interactions are observed. In this case the neutrinos come from pion decay and so are much more energetic than from the supernova. Their energies are governed by the beam energy and are generally much greater than the nucleon mass in the nucleon rest frame. Consider the following four reactions:

$$\begin{aligned} \nu_\mu + e^- &\rightarrow \nu_\mu + e^- \\ \nu_\mu + n &\rightarrow p + \mu^- \\ \nu_\mu + n &\rightarrow \nu_\mu + n \\ \nu_\mu + p &\rightarrow \nu_\mu + p \end{aligned}$$

in the limit of small momentum transfer. In the approximation that the nucleon (and lepton) masses can be neglected, what is the total cross-section for each of these processes in the center-of-mass frame?

For the last two reactions use the neutral current interaction

$$\mathcal{H}_{\text{nc}} = -\frac{G_F}{\sqrt{2}}[\bar{\nu}\gamma_\mu(1+\gamma_5)\nu]J_{\text{nc}}^\mu$$

$$\text{and } \langle N|J_{\text{nc}}^\mu|N\rangle = \frac{1}{(2\pi)^3}\bar{u}\gamma^\mu(k_V + k_A\gamma_5)u$$

with $k_V = \frac{1}{2} - 2\sin^2\theta_W$ if N is a proton and $k_V = -\frac{1}{2}$ if N is a neutron.
 [6.3.12] What is the ratio of their cross sections as a function of the neutrino energy in the lab frame? (Your answer should behave as $\sigma_N/\sigma_e \sim (m_N/m_e) \sim 10^3$ which shows why nucleon–neutrino scattering is so much easier to observe than electron–neutrino scattering.)

[6.4] Higgsstrahlung

At LEP II, the dominant mode used to search for the Higgs boson was “Higgsstrahlung,” $e^+e^- \rightarrow Z^* \rightarrow HZ$, where Z^* just means a Z intermediate state with an energy significantly different from the resonant energy M_Z .

Compute the (integrated) cross section for this process as a function of the center-of-mass energy s , the Z -boson mass m_Z , and the Higgs-boson mass m_H . Treat $m_e = 0$, but do not treat either m_Z or m_H as small compared to \sqrt{s} . Compare it to the cross section to go into $q\bar{q}$ final states.

List the most common final states (the Z and H bosons both subsequently decay). What features of the decay, if any, clearly distinguish it from a scattering $e^+e^- \rightarrow q\bar{q}$ with q any quark type?

[6.5] Resonances

The $\Upsilon(4s)$ is a narrow resonance caused by a $b\bar{b}$ bound state. Its mass and width are $m_{\Upsilon(4s)} = 10.580$ GeV, $\Gamma_{\Upsilon(4s)} = 14$ MeV, with a branching fraction to electrons of $B(\Upsilon(4s) \rightarrow e^+e^-) = 2.8 \times 10^{-5}$. It is experimentally useful because the $\Upsilon(4s)$ decays with almost 100% probability via $\Upsilon(4s) \rightarrow B\bar{B}$, with B a meson containing a \bar{b} quark and \bar{B} a meson containing a b quark. This gives a convenient way to produce B meson pairs approximately at rest, which has been exploited by the B-factories, BaBar and Belle.

What is the cross section for $e^+e^- \rightarrow B\bar{B}$ on the $\Upsilon(4s)$ resonance? Hint: the spin of the $\Upsilon(4s)$ is 1. Use Eq. (6.57).

What, approximately, is the correction to this cross section formula due to the radiation of soft photons from the e^+ and e^- ?

[6.6] Compton scattering

[6.6.1] Repeat the calculation for unpolarized photon–electron scattering without neglecting the electron mass, to show that the spin-summed and averaged matrix element generalizes to

$$\overline{\mathcal{M}}^2 = 2e^4 \left[\frac{p \cdot k'}{p \cdot k} + \frac{p \cdot k}{p \cdot k'} - 2m^2 \left(\frac{1}{p \cdot k} - \frac{1}{p \cdot k'} \right) + m^4 \left(\frac{1}{p \cdot k} - \frac{1}{p \cdot k'} \right)^2 \right]$$

where electron (photon) four-momenta are denoted by $p(k)$ and final-state quantities carry a prime. Show that even though this naively approaches the result quoted in the main text, $-2e^4[(s/u) + (u/s)]$, if $m^2 \rightarrow 0$ with s, t , and u fixed, the kinematical relation between the initial and final photon energies and the final photon scattering angle θ (in the initial electron's rest frame) $k^0/k^{0'} = 1 + (k^0/m)(1 - \cos \theta)$, implies the nominally sub-dominant combination

$$m^2 \left[\frac{1}{p \cdot k} - \frac{1}{p \cdot k'} \right] = 1 - \cos \theta$$

contributes even when $k^0 \gg m$. Use this to show that the differential cross section in the electron rest frame is given by

$$\frac{d\sigma}{\sin \theta d\theta} = \frac{\pi \alpha^2}{m^2} \left(\frac{E'}{E} \right)^2 \left[\frac{E'}{E} + \frac{E}{E'} - \sin^2 \theta \right]$$

where $E = k^0$ and $E' = k^{0'}$.

[6.6.2] Repeat the calculation for photon–electron scattering, but this time without averaging (summing) over the initial (final) photon polarization. Denoting by ε_i and ε_f the polarization vector of the initial and final photon, show that the polarized cross section is given in the rest frame of the initial electron by the Klein–Nishina formula:

$$\frac{d\sigma}{\sin \theta d\theta} = \frac{\pi \alpha^2}{2m^2} \left(\frac{E'}{E} \right)^2 \left[\frac{E'}{E} + \frac{E}{E'} + 4(\varepsilon_f \cdot \varepsilon_i)^2 - 2 \right]$$

Show that the sum over photon polarizations gives $\sum_{\text{if}} (\varepsilon_f \cdot \varepsilon_i)^2 = 1 + \cos^2 \theta$, and so reproduces the above result for the unpolarized cross section. In this form this result can be adapted to describe the important process of bremsstrahlung – the radiation of a photon by a charged

fermion as it moves in the Coulomb field of a nucleus. The cross section for bremsstrahlung can be computed by replacing the initial photon of the above calculation with the appropriate Fourier component of the initial Coulomb field.

[6.6.3] Show that *regardless of the photon energy*, in the limit of small scattering angles ($\theta \rightarrow 0$) the polarized differential cross section reduces to the Thompson formula

$$\frac{d\sigma}{\sin\theta d\theta} = \frac{2\pi\alpha^2}{m^2} (\varepsilon_f \cdot \varepsilon_i)^2$$

This is also the result for all angles in the limit $E \ll m$. The fact that this varies inversely with m^2 resolves a puzzle as to why ultra-relativistic muons and electrons behave so differently within detectors. After all, since electrons and muons have exactly the same gauge interactions within the standard model, any difference between their properties in a detector can only be due to their different mass and one might naively expect that this mass difference should become unimportant for ultra-relativistic particles. The above result shows that their small-angle Compton scattering (and bremsstrahlung) cross-sections differ dramatically even in the ultra-relativistic limit. Since it is the bremsstrahlung due to numerous small-angle scattering events which dominates the energy loss of an ultra-relativistic charged particle passing through matter, this proportionality of this cross section to $1/m^2$ ensures that muons lose their energy much less efficiently than electrons, and so are much more penetrating when they pass through a detector.

7

Effective Lagrangians

A great deal of particle phenomenology, including both the properties of the observed particles and their reactions in many accelerators, deals with energy scales that are very small in comparison with the mass of the weak vector bosons, M_W or M_Z . A technique that has been used to good effect at various points in the previous chapters is the expansion of low-energy scattering amplitudes in inverse powers of the W - or Z -boson masses. This expansion greatly simplified the corresponding calculations and was justified in each case by the fact that the typical energies involved in the amplitudes under consideration were much smaller than M_W and M_Z .

Concrete examples where this type of expansion is justified are given by the weak decays of a light meson such as the muon, as was computed in Chapter 5, since the energy scales involved are much smaller than the mass of the virtual W boson that mediates these decays. A similar simplification is justified in Chapter 6 in the amplitudes for electron-positron annihilation at energies that are low compared to the Z -boson mass.

All of these examples furnish special cases of the general technique of low-energy expansions. This technique appears ubiquitously throughout physics because many physical systems have the property that they involve two (or more) degrees of freedom that each have very different masses. As is clear from the examples from Chapters 5 and 6, it is often of interest in these systems to understand the effects of the heavy degrees of freedom for phenomena that take place at much lower energies. The *effective-Lagrangian* formalism is a particularly efficient tool for organizing such calculations.

The purpose of this chapter is to systematize these low-energy approximation schemes through the introduction and use of effective Lagrangian techniques. We do so using the low-energy *Fermi theory* of the weak interactions as a vehicle for this discussion. The Fermi theory is of interest in its own right both for use in performing practical low-energy calculations and

for its historical role, since an early version was written down during the 1930s as a theory of the beta decay of nucleons, then modified in the 1950s to include parity-violation effects.

7.1 Physics below M_W : the spectrum

The first step in properly constructing the low-energy effective Lagrangian is to identify *all* of the light degrees of freedom that are to be included. For the standard model this would include all of the particles that are much lighter than M_W . The complete list of these particles and the fields that represent them is as follows.

- (i) The photon as represented by its gauge potential, A_μ , and abelian field strength, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.
- (ii) The eight gluons, labeled by the color index $\alpha = 1, \dots, 8$, and represented by the gauge potentials, G_μ^α , and non-abelian field strengths, $G_{\mu\nu}^\alpha = \partial_\mu G_\nu^\alpha - \partial_\nu G_\mu^\alpha + g_3 f_{\beta\gamma}^\alpha G_\mu^\beta G_\nu^\gamma$.
- (iii) All of the charged leptons, e, μ , and τ , and neutrinos, ν_e, ν_μ , and ν_τ . The charged leptons are each represented by a single Dirac spinor field while the neutrinos may be represented by Majorana fields.
- (iv) Finally, all but one – the top – of the quarks, u, c, d, s, b . Each of these is represented by a Dirac spinor field.

The most important part of the low-energy effective Lagrangian that governs the mutual interactions of these particles is given (at tree level) by simply discarding all of those terms in the standard model Lagrangian that involve any heavy fields. The resulting Lagrangian may be denoted \mathcal{L}_0 and has the form

$$\mathcal{L}_0 = \mathcal{L}_{\text{kin}} + \mathcal{L}_{\text{str}} + \mathcal{L}_{\text{em}} \quad (7.1)$$

Here \mathcal{L}_{kin} represents the kinetic and mass terms for all of the fermion fields:

$$\mathcal{L}_{\text{kin}} = - \sum_f \bar{f}(\not{\partial} + m_f)f \quad (7.2)$$

\mathcal{L}_{str} contains both the kinetic terms and self-interactions of the eight species of gluons, as well as the terms which describe the strong interactions of the five light-quark flavors with these gluons:

$$\mathcal{L}_{\text{str}} = -\frac{1}{4}G_{\mu\nu}^\alpha G_\alpha^{\mu\nu} - \frac{g_3^2 \Theta_3}{64\pi^2} \epsilon^{\mu\nu\lambda\rho} G_{\mu\nu}^\alpha G_{\lambda\rho}^\alpha + \frac{ig_3}{2} \sum_q G_\mu^\alpha \bar{q} \gamma^\mu \lambda_\alpha q \quad (7.3)$$

\mathcal{L}_{em} similarly describes the kinetic terms for the photon together with the electromagnetic couplings of all of the light charged particles:

$$\mathcal{L}_{\text{em}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + ie \sum_f Q_f A_\mu \bar{f} \gamma^\mu f \quad (7.4)$$

These are the only interactions within the standard model that do not involve at least one heavy particle such as a W or Z boson, a Higgs boson, or a top quark. These are also the most general renormalizable couplings that can be written down for this particle content. Following the logic of Subsection 1.2.2, they may therefore be expected to be the interactions that dominate in processes that take place exclusively at energies that are much smaller than the masses of these heavy particles.

Of course, just because a particle is heavy compared to the energy scale under consideration does not mean that its effects may be completely ignored. Although energy conservation forbids the direct production in low-energy reactions of very heavy particles, they nonetheless contribute through their virtual effects. In fact, as was seen in Chapter 5, many fundamental low-energy processes, such as the decays of light particles such as the μ^- or τ^- , can only be understood through physics at a much heavier scale, since they proceed by the virtual emission and subsequent decay of a virtual W boson.

Within the effective-Lagrangian framework all such virtual effects are described by various non-renormalizable effective interactions. The following sections are devoted to their construction.

7.2 The Fermi theory

The goal is to construct the effective interactions in the low-energy Lagrangian that reproduce the virtual effects of heavy particles such as the W , Z , or Higgs bosons or the top quark. We will start with the low-energy effects due to the W couplings to light particles. In order to do so it is convenient to reconsider the expressions of Section 5.3 that describe the standard-model amplitude for tau or muon decay. We do so in order to reconstruct those features that are of most interest at energies that are low compared to the W mass.

Recall, then, Eq. (5.19) for the τ decay amplitude:

$$\begin{aligned} \mathcal{M}(\tau \rightarrow \nu_\tau \bar{f}_m f_n) &= e_W^2 U_{mn}^* [\bar{u}_\nu(\mathbf{l}) \gamma^\mu (1 + \gamma_5) u_\tau(\mathbf{k})] [\bar{u}_n(\mathbf{p}) \gamma^\nu (1 + \gamma_5) v_m(\mathbf{q})] \\ &\quad \times \left[\frac{\eta_{\mu\nu} + (k-l)_\mu (k-l)_\nu / M_W^2}{(k-l)^2 + M_W^2 - i\epsilon} \right] \end{aligned} \quad (7.5)$$

Section 5.3 makes the basic observation that four-momentum conservation implies that all of the components of the four-momenta in the problem are, in the rest frame of the decaying meson, less than or of the order of this decaying meson's mass. This fact, together with the small size of the mass ratios, $m_\tau^2/M_W^2 \approx 5 \times 10^{-4}$ and $m_\mu^2/M_W^2 \approx 2 \times 10^{-6}$, is then used to simplify the scattering amplitude by expanding the W -boson propagator in inverse powers of M_W^2 :

$$\frac{\eta_{\mu\nu} + (k-l)_\mu(k-l)_\nu/M_W^2}{(k-l)^2 + M_W^2 - i\epsilon} \approx \frac{\eta_{\mu\nu}}{M_W^2} \left[1 - \frac{(k-l)^2}{M_W^2} + \dots \right] + \frac{(k-l)_\mu(k-l)_\nu}{M_W^4} [1 + \dots] \quad (7.6)$$

in which the ellipsis represents terms that are higher order in $(k-l)^2/M_W^2$ than the terms displayed. Keeping only the lowest order term in this expansion gives the following result for Eq. (7.5):

$$\mathcal{M}(\tau \rightarrow \nu_\tau \bar{f}_m f_n) = \frac{G_F}{\sqrt{2}} U_{mn}^* [\bar{u}_\nu(\mathbf{l}) \gamma^\mu (1 + \gamma_5) u_\tau(\mathbf{k})] [\bar{u}_n(\mathbf{p}) \gamma_\mu (1 + \gamma_5) v_m(\mathbf{q})] \quad (7.7)$$

As in previous chapters the Fermi coupling constant, G_F , is defined by the coupling combination $G_F/\sqrt{2} = e_W^2/M_W^2 = 1/(2v^2)$.

The key point is that this matrix element, Eq. (7.7), is what would have been produced at lowest order in perturbation theory,

$$\mathcal{M}(\tau \rightarrow \nu_\tau \bar{f}_m f_n) = \langle \nu_\tau; \bar{f}_m; f_n | \mathcal{H} | \tau \rangle \quad (7.8)$$

by the following *effective interaction*:

$$\mathcal{H} = -\mathcal{L} = + \frac{G_F}{2\sqrt{2}} U_{mn}^* [\bar{\nu}_\tau \gamma^\mu (1 + \gamma_5) \tau] [\bar{f}_n \gamma_\mu (1 + \gamma_5) f_m] + \text{c.c.} \quad (7.9)$$

It is worth remarking that this type of interaction is *not* renormalizable (c.f. Section 1.2 for the significance of and criteria for renormalizability). This simply reflects the fact that non-renormalizable interactions are characterized by coupling constants which have dimensions of a negative power of mass (when $\hbar = c = 1$). The coupling G_F in Eq. (7.9) has this property precisely because this interaction must vanish as the mass of the virtual W boson becomes arbitrarily large. This is a feature that is typical of an effective interaction that represents the low-energy effects of a heavy virtual particle.

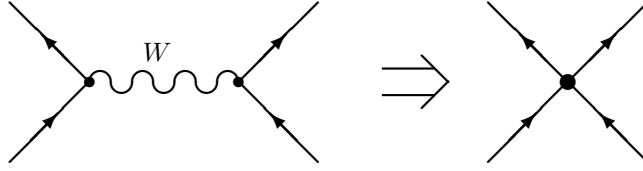


Fig. 7.1 The tree graph that generates the Fermi Lagrangian.

7.2.1 Effective charged-current interactions

This same approximation, in which inverse powers of M_W^2 are neglected, may be applied in the same way to *all* of the other decay and scattering amplitudes that are mediated by the exchange of a W boson between light initial and final states. The result for all of these possible reactions is efficiently summarized by directly constructing the relevant effective Lagrangian once and for all, and then only afterwards computing the particular matrix element that is of interest for a particular application.

Following the same steps as in the previous example, it is easy to see the form of the effective Lagrangian that arises in this way to lowest order in perturbation theory. Recall that the W boson couples to light particles (in this case the fermions) through an interaction term of the form of Eq. (2.88):

$$\mathcal{L}_{cc} = e_W W_\mu C^\mu + \text{h.c.} \quad (7.10)$$

in which the *charged current*, C^μ , is defined by

$$C^\mu = \sum_{m=1}^3 \left[i \bar{e}_m \gamma^\mu (1 + \gamma_5) \nu_m + \sum_{n=1}^3 i U_{nm}^* \bar{d}_m \gamma^\mu (1 + \gamma_5) u_n \right] \quad (7.11)$$

The corresponding effective charged-current interaction is obtained to lowest order in M_W^{-2} by evaluating the general tree-level graph of Figure 7.1, and approximating the internal (heavy) W -boson line by $\eta_{\mu\nu}/M_W^2$. The external-line factors corresponding to the (light) initial and final particles are *not* included but are left as operators in the effective interaction that is being constructed for the low-energy theory.

The effective interaction that is generated in this way by the virtual exchange of a W particle is the *Fermi Lagrangian*:

$$\mathcal{L}'_{cc} = \frac{G_F}{\sqrt{2}} C^\mu C_\mu^* \quad (7.12)$$

with the charged currents given by Eq. (7.11).

It is easily verified that the evaluation of the rate for μ^- or τ^- decay, or the cross section for $e^- \bar{\nu}_e$ scattering using the Fermi Lagrangian, Eq. (7.12), agrees with the leading low-energy approximation to these processes as calculated in Chapters 5 and 6.

7.2.2 Effective neutral-current interactions

The other heavy bosons in the standard model whose masses are at least as heavy as M_W may be handled in precisely the same way, by evaluating the graphs that correspond to Figure 7.1 with the intermediate W boson replaced by a Z or Higgs boson. In these cases the couplings of these bosons to light fermions have the form of Eq. (2.70) or Eq. (2.100). That is,

$$\mathcal{L}_{\text{nc}} = e_Z Z_\mu N^\mu \quad (7.13)$$

with the *neutral current*, N^μ , given by

$$\begin{aligned} N^\mu &= \sum_{\text{fermions}} i \bar{f} \gamma^\mu (T_3 P_L - Q \sin^2 \theta_W) f, \\ &= \frac{1}{2} \left[J_3^\mu - 2 \sin^2 \theta_W J_{\text{em}}^\mu \right] \end{aligned} \quad (7.14)$$

In this last expression, $J_{\text{em}}^\mu = i \bar{f} \gamma^\mu Q f$ is the electromagnetic current and $J_3^\mu = i \bar{f} \gamma^\mu T_3 (1 + \gamma_5) f$ is the conserved current corresponding to the third component of weak isospin, T_3 .

The effects of the virtual exchange of a Z boson due to these couplings to light fermions can, for energies small compared to M_Z , be reproduced by the following effective interactions:

$$\begin{aligned} \mathcal{L}'_{\text{nc}} &= + \frac{e_Z^2}{2M_Z^2} N^\mu N_\mu \\ &= + \rho \frac{G_F}{\sqrt{2}} \left[J_3^\mu - 2 \sin^2 \theta_W J_{\text{em}}^\mu \right] \left[J_{3\mu} - 2 \sin^2 \theta_W J_{\text{em}\mu} \right] \end{aligned} \quad (7.15)$$

where the “ ρ -parameter” is defined as

$$\rho \equiv \frac{M_W^2}{M_Z^2 \cos^2 \theta_W} \quad (7.16)$$

It can be seen here to give a measure of the relative strength of the charged-current and neutral-current weak interactions at low energies. Within the standard model, and at leading order in α , $\rho = 1$ (see Eq. (2.52)).

7.2.3 Virtual Higgs exchange

The Higgs–Yukawa couplings are given by

$$\mathcal{L}_{H-f} = HS_H \quad (7.17)$$

with

$$S_H = - \sum_{\text{fermions}} \frac{m_f}{v} \bar{f} f \quad (7.18)$$

Virtual Higgs exchange is therefore reproduced at low energies by the following effective four-fermion interaction:

$$\begin{aligned} \mathcal{L}'_{\text{Higgs}} &= + \frac{1}{2m_H^2} S_H S_H \\ &= \sum_f \sum_{f'} \frac{m_f m_{f'}}{2m_H^2 v^2} \bar{f} f \bar{f}' f' \\ &= \frac{G_H}{\sqrt{2}m_H^2} \sum_f \sum_{f'} m_f m_{f'} \bar{f} f \bar{f}' f' \end{aligned} \quad (7.19)$$

These effective interactions due to the exchange of a virtual Higgs boson are generally of much less interest than are the effective charged-current and neutral-current interactions of the previous sections. This is because they are generally so small as to be impossible to detect in low-energy processes.

One reason that they are so small is that they involve two factors of the extremely small Higgs–Yukawa coupling, m_f/v . This suppression is on top of the overall factor of m_H^{-2} that arises due to the virtual heavy Higgs particle, and leads to an effective suppression by *four* powers of a heavy mass rather than by two as would be naively expected on dimensional grounds for a four-fermion operator.

This type of interaction would be difficult to detect, however, even if it were not so heavily suppressed by large masses. That is because – as we see in more detail in the following section – the effective interaction, Eq. (7.19), shares the selection rules of the lowest-order Lagrangian of Eq. (7.1) through Eq. (7.4). Its contribution to any particular reaction will therefore tend to be swamped by competing processes that are not suppressed at all by powers of heavy masses.

This is in contrast to the charged-current and neutral-current effective interactions, which do not directly compete with the lowest-order strong and electromagnetic interactions. They do not compete because these effective interactions violate flavor conserving and P- and C- selection rules of the lowest-order interactions. This is a reflection of the fact that only the W and Z couplings break these symmetries in the standard model. It follows that flavor-changing observables – such as the meson decays of Chapter 5 –

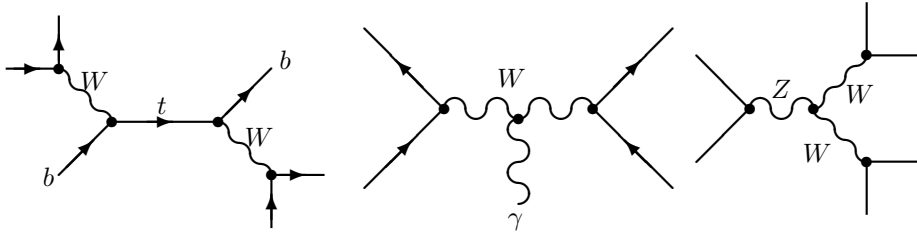


Fig. 7.2 Higher-order tree graphs that contribute to the effective Lagrangian.

or parity or charge-conjugation violating reactions – such as the left-right and forward-backward asymmetries of Chapter 6 – are guaranteed to receive their dominant contributions only from the effective charged- and neutral-current weak interactions.

It is clearly of some interest to understand the accidental and approximate selection rules and conservation laws of the leading terms, Eq. (7.1) through Eq. (7.4), of the low-energy effective action. We therefore turn to these in Section 7.3. Before doing so, a few more odds and ends must be settled concerning other potential tree-level effects for the low-energy effective action, including any potential effects of virtual top quarks. Loop effects are considered in more detail in Section 7.4.

7.2.4 Other tree-level contributions

Equation (7.12), Eq. (7.18), and Eq. (7.19) derived above do not exhaust the effective interactions of the low-energy theory. Indeed, more and more complicated interactions can arise from more and more complicated graphs involving internal lines that represent heavy particles. Figure 7.2 indicates several such graphs that arise at “tree level” (no loops in the graph), all of which generate effective interactions that involve six external light-fermion fields or four light fermionic and one bosonic field.

As may be seen from the figure, it is at this level that a virtual heavy fermion, such as a top quark, first makes its tree-level appearance. The important observation is that, starting just from light initial particles, the standard-model interactions only permit a heavy top quark to be virtually produced together with either a top antiquark or a W boson. In either case, at least two heavy-particle propagators must therefore appear in the amplitude. Besides being suppressed by additional powers of the small-gauge and Yukawa couplings associated with each vertex, the additional heavy-

particle propagators contribute an additional suppression by extra powers of inverse heavy masses like $(M_W)^{-2}$ or $(m_t)^{-2}$.

Because they involve additional suppressing powers of $1/M_{\text{heavy}}^2$, such interactions can generally be ignored in comparison with the virtual effects we have already considered. This will *not necessarily* remain true when we consider diagrams containing loops; we address this point in Section 7.4.

7.3 Physics below M_W : qualitative features

To summarize the results of the previous two sections, interactions at energies that are small compared to M_W are described up to leading order in M_W^{-2} (and M_Z^{-2} etc.) by the effective Lagrangian,

$$\begin{aligned}\mathcal{L}_{\text{eff}} &= \mathcal{L}_0 + \mathcal{L}_{\text{wk}} \\ &= \mathcal{L}_{\text{kin}} + \mathcal{L}_{\text{str}} + \mathcal{L}_{\text{em}} + \mathcal{L}_{\text{wk}}\end{aligned}\tag{7.20}$$

in which \mathcal{L}_{kin} , \mathcal{L}_{str} , and \mathcal{L}_{em} are given by Eq. (7.2) through Eq. (7.4), and the *weak-interaction Lagrangian*, \mathcal{L}_{wk} , represents the sum of the charged-current and neutral-current interactions,

$$\mathcal{L}_{\text{wk}} = \mathcal{L}'_{\text{cc}} + \mathcal{L}'_{\text{nc}}\tag{7.21}$$

of Eq. (7.12) and Eq. (7.15). The Higgs interactions are ignored here for reasons that were touched on in the previous section and are about to be made more clear in what follows. This low-energy Lagrangian has many important qualitative features, some of which we list.

- (i) First and foremost, the particle content of the model includes precisely those particles that are observed at low energies. This conclusion is subject to the caveat that the observed hadrons are to be interpreted as bound states of the fundamental quarks and gluons of the strongly-interacting sector – a point we return to in Chapter 8. Moreover, the most important interactions at low energies, described by the Lagrangian \mathcal{L}_0 , may be characterized as being the most general renormalizable interactions that are consistent with this observed particle content and gauge symmetries. This is significant inasmuch as it automatically incorporates precisely the experimentally observed conservation laws and selection rules as accidental symmetries that are not put in by hand. This point is elaborated in subsequent items. This gives us confidence that we understand the origin of these symmetries as natural consequences of general principles together with the given particle content.

- (ii) The Lagrangian of Eq. (7.20) has the general form of a collection of fundamental fermions, the quarks and leptons, interacting under the influence of three forces each of fundamentally different strength: the strong force \mathcal{L}_{str} , the electromagnetic force \mathcal{L}_{em} , and the charged- and neutral-current “weak” force, \mathcal{L}_{wk} . This agrees with the basic features of low-energy physics that had been distilled from experiment over the years before the advent of the standard model.
- (iii) Beyond grouping the three basic interactions, this effective interaction allows a fundamental understanding of their relative strengths. In particular it gives a natural explanation of why the weak interactions are so weak. Unlike the strong and electromagnetic interactions, whose strength is governed by the size of the corresponding dimensionless gauge-coupling constants, g_3 and e , the weak interactions are controlled by a dimensionful coefficient, G_{F} . The weak interactions are therefore seen to be the weakest of the three forces, not because the underlying coupling is small, but because they are suppressed by a large mass scale, namely $G_{\text{F}}^{-\frac{1}{2}}$. That is, the weak interactions were found, historically, to be weak just because they had only been probed at energies much smaller than their characteristic energy scale, M_W or M_Z .

This understanding was a prerequisite for the partial unification of the electromagnetic and weak interactions that occurs in the standard model since, as we have seen in Chapter 6, these interactions have basically the same strength at energies high compared to M_W and M_Z .

- (iv) The low-energy action, Eq. (7.20), inherits the success of the standard model in understanding the selection rules of the various interactions with respect to the discrete symmetries of parity, time-reversal, or charge-conjugation. In particular, the Lagrangian \mathcal{L}_0 enjoys many more of the “accidental” symmetries that are discussed in Section 2.5 than does the entire standard model Lagrangian, since many of these accidental symmetries are broken only by the charged W -boson couplings.

The only term in \mathcal{L}_0 that violates C, P or, T is the P- and T-violating gluon self-coupling term

$$\mathcal{L}_{\Theta} = -\frac{g_3^2 \Theta_3}{64\pi^2} \epsilon^{\mu\nu\lambda\rho} G_{\mu\nu}^{\alpha} G_{\lambda\rho}^{\alpha} \quad (7.22)$$

This term is something of an embarrassment for the model, since there is absolutely no experimental evidence for P or T breaking by

the electromagnetic or strong interactions; in fact, Θ_3 is bounded by the non-observation of the neutron electric-dipole moment at the level of $|\Theta_3| < 10^{-9}$. It is possible that this term simply happens to be this small. However, as we discuss in Subsection 11.4.2, it appears that it would require a rather remarkable coincidence for this to happen. This is one of the so-called “naturalness” problems with the standard model and it has received considerable attention. In particular there are some proposals for extensions of the standard model which would make the smallness of Θ_3 automatic. We discuss one such “solution” to the “problem” of the smallness of Θ_3 in Subsection 11.4.2.

The analog of Eq. (7.22) for the electromagnetic interactions need not be considered in \mathcal{L}_0 , however. This is because Eq. (7.22) is a total derivative – i.e. it has the form $\partial_\mu V^\mu$ for some four-vector field V^μ , which vanishes in vacuum – and so does not contribute at all to any physical process. (Subsection 11.4.2 explains why the same is not true of the Θ_3 term.)

Given the small size for Θ_3 , the model successfully predicts that the electromagnetic and strong interactions at low energies preserve each of C, P, and T. The charged- and neutral-current weak interactions do not respect these symmetries, although only the charged currents violate C, P, or T and only do so in the very specific way described by the complex phase of the KM matrix.

- (v) A similar story holds for the accidental continuous symmetries of the standard model. Since many of these symmetries are broken only by the W couplings they are inherited as accidental symmetries by the low-energy terms, \mathcal{L}_0 , of the model.

Recall that the exact accidental symmetries of the standard model are $U_e(1) \times U_\mu(1) \times U_\tau(1) \times U_B(1)$, corresponding to the exact conservation (up to electroweak and mixed gravitational anomalies) of electron, muon, tau, and baryon numbers. These symmetries are automatically inherited as accidental symmetries of the low-energy Lagrangian, Eq. (7.20).

The dominant part of the theory at low energies, \mathcal{L}_0 , involving just the strong and electromagnetic interactions, enjoys more symmetries than does the entire standard model. All of the candidate accidental symmetries of the standard model that only failed because they did not preserve the charged-current W couplings are legitimate invariances of \mathcal{L}_0 . From Section 2.5 these symmetries consist of separate phase rotations of each species of Dirac fermion in the low-energy model. That is, in the absence of the charged-current weak

interactions, the accidental symmetry group of the low-energy theory is enlarged from $[U(1)]^4$ to $[U(1)]^{11}$, corresponding to the separate conservation of each type of fermion flavor: i.e. “electron number,” “electron–neutrino number,” “up-ness,” “down-ness,” and all of their analogs for the successive generations. To the extent that these symmetries are exact, the lightest particle carrying each quantum number is stable. Therefore, such particles can only decay *via* the charged-current weak interactions. The decays are therefore suppressed by $1/M_W^4$ in every case.

- (vi) Of course, not all of the features of the standard model are obvious just given the low-energy limit represented by Eq. (7.20). In particular, the underlying $SU_L(2) \times U_Y(1)$ -invariance is not manifest in the low-energy theory because we have kept only parts of various $SU_L(2)$ -multiplets in the low-energy theory. For example, the bottom quark and photon are included, and yet the top quark and W boson are not. Consequences of the pattern of $SU_L(2) \times U_Y(1)$ symmetry breaking such as the existence of the custodial $SU(2)$ of the Higgs sector are also not easy to understand from the low-energy point of view. In particular, there is no understanding purely within the low-energy theory of why the ρ -parameter of the neutral-current Lagrangian, Eq. (7.15), should happen to be one (or very close to it), as it is in the standard model.

The particle content of the low-energy theory, together with general principles, are not sufficient to determine the value of the ρ parameter, or certain other parameters. This makes these parameters particularly useful to probe the underlying structure of the standard model.

7.4 Running couplings

Another place where heavy particles can affect low-energy phenomena is by appearing in *loops*, when calculations are carried out beyond the leading order. So far we have avoided discussing calculations beyond leading order, partly because we have only aspired to low-order accuracy, but partly because their consideration involves substantial new technical complications. In particular, the momenta circulating in loops can be arbitrarily high even if the phenomenon under consideration involves only low-energy scales; and in some cases this only leads to a suppression of the correction by powers of the coupling $(g/4\pi)^2$. These same cases also involve ultraviolet divergences, which took physicists some time to understand and control. We will not give

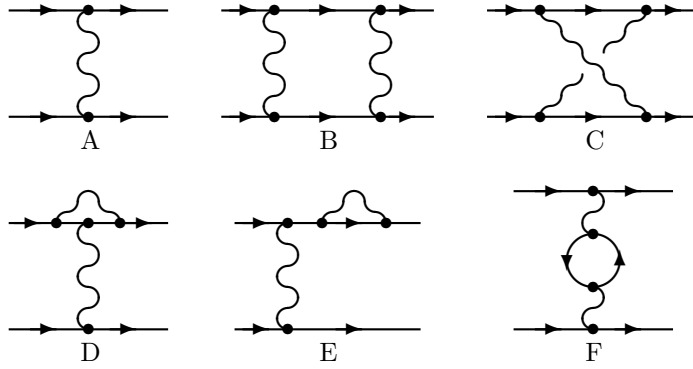


Fig. 7.3 One-loop scattering processes.

a complete exposition of this subject, but it is too important to leave out entirely, so we will outline the relevant physics here.

7.4.1 An example problem

Consider as an example problem, the electromagnetic scattering of two fermion species f_1, f_2 (for instance, $e^- \mu^-$ scattering). The leading-order diagram, which is $O(e^2)$, is shown in Figure 7.3 A. The possible $O(e^4)$ corrections to the matrix element are shown as B–F in the same figure. (In diagram D it should be understood that the correction can occur at either vertex; in diagram E the correction can occur on any of the four external lines. Diagram F could also have a charged boson in the loop.) The square of the leading-order matrix element \mathcal{M}_A is $|\mathcal{M}_A|^2 \sim O(e^4)$, while the interference of these higher-order matrix elements with the leading one, $|\mathcal{M}_A \mathcal{M}_{B-F}| \sim O(e^6)$. In addition there is an $O(e^3)$ process with an extra emitted photon, whose square is also $O(e^6)$. We have discussed this process in Subsection 6.7.2: it is infrared divergent but that divergence is canceled by an IR divergence in diagrams D and E.

Some of these corrections are ultraviolet (UV) finite. This is true of diagram B, which evaluates to

$$\begin{aligned}
 \mathcal{M}_B = ie^4 Q_{f_1}^2 Q_{f_2}^2 \int \frac{d^4 l}{(2\pi)^4} \frac{-i\eta_{\mu\nu}}{l^2 - i\epsilon} \frac{-i\eta_{\alpha\beta}}{(l + p - k)^2 - i\epsilon} \\
 \times \bar{u}(k) \gamma^\alpha \frac{-i(-i\not{p} - i\not{l} + m_1)}{(p + l)^2 + m_1^2 - i\epsilon} \gamma^\mu u(p) \\
 \times \bar{u}(k') \gamma^\beta \frac{-i(-i\not{p}' + i\not{l} + m_2)}{(p' - l)^2 + m_2^2 - i\epsilon} \gamma^\mu u(p') \quad (7.23)
 \end{aligned}$$

To see that this is well behaved at large l , consider the behavior when $l \gg p, p', k, k'$ the external momenta. At very large l , its l dependence is summarized by

$$\int d^4l \frac{\not{l} \cdots \not{l} \cdots \not{l} \cdots}{l^2 l^2 l^2} \quad (7.24)$$

with the \not{l} from the two fermion propagators and the $1/l^2$ from the four propagators making up the loop. This has the behavior d^4l/l^6 , which is well behaved at large l . Therefore there is no problem in principle with performing the integral and evaluating this correction. A version of this diagram, with the photons replaced by W bosons, is relevant to flavor-changing processes discussed in Subsection 9.4.2.

The situation is different for diagram D. This diagram acts as a modification of diagram A, in which the vertex γ^μ in diagram A is replaced with

$$e^2 Q_{f_1}^2 \int \frac{d^4l}{(2\pi)^4} \frac{-i\eta_{\alpha\beta}}{l^2 - i\epsilon} \gamma^\alpha \frac{-i(-i\not{p} - i\not{l} + m)}{(p+l)^2 + m^2 - i\epsilon} \gamma^\mu \frac{-i(-i\not{p}' - i\not{l} + m)}{(p'+l)^2 + m^2 - i\epsilon} \gamma^\beta \quad (7.25)$$

The large l behavior is now $\sim \int d^4l \not{l} \not{l} / l^6$, or $\int d^4l / l^4$, which is logarithmically divergent. In other words, it appears that the correction from this process will be infinite.

7.4.2 Regularization, power counting

The appearance of such apparent divergences does not necessarily mean that the theory is nonsense. What it does mean is that the theory is not well defined unless some procedure is taken which removes or “cuts off” the ultraviolet region of such integrations. This is called *regularization* of the theory. There are strong results in quantum field theory which show that the nature of the regularization is unimportant, provided the scale of the cut-off is sufficiently large and the way the cut-off is imposed does not violate any symmetries of the theory. The regularized theory is finite. However, if a diagram was divergent before imposing a cut-off, its contribution remains large after the cut-off is imposed. For instance, if a diagram is log divergent in the absence of a cut-off, then with a cut-off its value is of order $\log(\Lambda/m)$, with Λ the cut-off scale and m some physical scale. Such large contributions need to be understood and included in the way we do calculations. We must also make sure that no results for physical measurements depend on Λ .

Therefore we need to catalog the diagrams in which UV divergences can occur. To do this one follows the procedure outlined previously for diagrams B and D, and estimates the convergence of the momentum integrals as if they were only one-dimensional integrals, a procedure called *power counting*. According to this line of argument, a graph diverges like Λ^P , as $\Lambda \rightarrow \alpha$ where P is an integer called the graph's *superficial degree of divergence*. P may be regarded as the sum of:

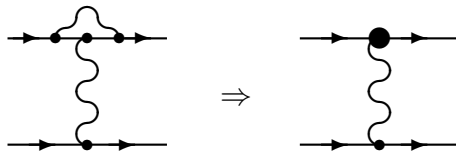
- +4 for each $\int d^4l$ (ie each loop) in the graph;
- +1 for each explicit power of loop momentum, l^μ , in the Feynman rules of the vertices in the graph;
- -1 for each fermionic propagator, $\sim 1/l$, which depends on a loop momentum;
- -2 for each bosonic propagator, $\sim 1/l^2$, which depends on a loop momentum;
- -1 for each power of non-loop momentum for which the loop integral is known to be proportional.

If $p > 0$ then the graph diverges like a power of Λ ; if $p = 0$ the graph diverges logarithmically; in $p < 0$ the graph is “superficially finite” (but may actually diverge if it contains subgraphs for which $p \geq 0$).

For example, for graph B above, there is one loop, all internal lines depend on loop momenta, and there are no powers of momenta in any of the vertices. Consequently, $P_B = 4 + 0 - 2 - 4 + 0 = -2$, and the graph converges. For diagram D, there is one loop, but one of the boson propagators does not participate, so $P_D = 4 + 0 - 2 - 2 + 0 = 0$, making the graph logarithmically divergent. For graph F we have only fermions in the single loop so $P_F^{\text{naive}} = 4 + 0 - 2 + 0 + 0 = 2$. However, gauge invariance implies the loop integral is proportional to two powers of the internal boson (non-loop) momentum, so $P_F = 4 + 0 - 2 + 0 - 2 = 0$, implying logarithmic divergence.

Large l in momentum space corresponds to small separation in position space (by the uncertainty principle), and this allows the UV-divergent part of a loop graph to be described in terms of a local interaction. Pictorially, the divergent part of a loop can be obtained by shrinking the loop in the graph and replacing it by an effective pointlike vertex, as in the figure, with a Λ -dependent coefficient.

For instance, since the vertex correction loop had $p = 0$ and has two fermions and a gauge boson emerging from it, it can be replaced for Λ -counting purposes by a vertex of the form $\log \Lambda (\bar{f} r^\mu f') V_\mu$, where f, f' and V_μ represent the fields entering the loop.



This gives us another way to estimate P for a given loop. Write down the effective operator obtained by shrinking the loop to a point, and count the mass dimension of all of the fields (and derivatives) which appear. If the result is a dimension- D operator, then the coefficient has dimension $P = 4 - D$, where P is the loop's superficial degree of divergence.

This last estimate shows that superficially divergent graphs, $P \geq 0$, correspond to effective interactions for which $D \leq 4$, which are precisely the interactions which already appear in a renormalizable lagrangian. For this reason divergent Λ -dependence can be absorbed into the couplings of a renormalizable theory.

7.4.3 Renormalization

We will refer to the coupling constants which appear in the Lagrangian as the *bare couplings* of the theory, and subscript them with a 0, so the electromagnetic coupling in the Lagrangian will be written e_0 . Experiments measure the sum of the bare coupling and the contribution of loops. For instance, the electromagnetic coupling α is determined by measuring low-energy interactions of electrons – essentially, the process we considered in Figure 7.3. The measured α is given by the sum of the diagrams shown there, including the large UV contributions of D, E, and F. To avoid large and cut-off-dependent corrections, we need a procedure for absorbing such contributions automatically by a re-definition of the couplings and field normalizations, a procedure called *renormalization*. The easiest way to explain renormalization is to illustrate it with an explicit example. Therefore we will carry out part of the calculation of the electromagnetic scattering problem.

Our power counting rules show that diagrams D, E, and F are potentially divergent. The degree of divergence of D is 0, so it can be – and is – logarithmically divergent. The degree of divergence of E is +1, and of F is +2. However, the result for E turns out to be linear in the momentum entering the loop, and the result of F is quadratic in the momentum entering the loop, so the degrees of divergence are only logarithmic. Let us briefly see why.

As we argued, the divergent part of a graph is local, and it must respect the symmetries of the theory. In particular, it must satisfy Lorentz and gauge symmetry. The potentially linearly divergent part of E contains an odd power of gamma matrices, and so must the result. But it must

behave like a Lorentz-invariant local operator quadratic in fermion fields, and the lowest dimension operator with an odd power of γ is $\bar{\psi}\not{\partial}\psi$, which contains a derivative—that is, a power of the external momentum. So E is linear in the external momentum, which reduces its superficial degree of divergence by 1. Similarly for diagram F : the result must be gauge-invariant, and the lowest-order Lorentz-invariant gauge-invariant operator, quadratic in the gauge field, is the field strength squared, $A_\mu(q^2\eta^{\mu\nu} - q^\mu q^\nu)A_\nu$, which is quadratic in the momentum. Gauge invariance implies that the divergent piece must be quadratic in the external momentum, and is therefore only log divergent.

Furthermore, it turns out that the log divergences arising from D and E cancel, due to an identity called the *Ward identity*, which follows from gauge invariance. (We will not give the details.) Physically, this cancellation occurs because the corrections look respectively like $\bar{\psi}\not{A}\psi$ and $\bar{\psi}\not{\partial}\psi$. Gauge invariance requires that these appear in the combination $\bar{\psi}\not{D}\psi$; the size of a divergent correction to $\bar{\psi}\not{A}\psi$ must be Qe times the correction on the propagator, and this ensures that their contributions to the scattering rate cancel.

Now we must focus on diagram F . The contribution of this diagram is the same as of diagram A , except with the propagator replaced with

$$e^2 \left[\frac{-i\eta_{\mu\nu}}{q^2 - i\epsilon} \right] \Rightarrow e^2 \left[\frac{-i\eta_{\mu\alpha}}{q^2 - i\epsilon} \right] (i\Pi^{\alpha\beta}(q)) \left[\frac{-i\eta_{\beta\nu}}{q^2 - i\epsilon} \right]$$

$$\text{where } \Pi^{\alpha\beta}(q) \equiv ie^2 \int \frac{d^4l}{(2\pi)^4} \text{tr} \left[\frac{-i(-i\not{l} + m)}{l^2 + m^2 - i\epsilon} \gamma^\alpha \frac{-i(-i\not{l} - i\not{q} + m)}{(l+q)^2 + m^2 - i\epsilon} \gamma^\beta \right]$$
(7.26)

$\Pi^{\alpha\beta}(q)$ is referred to as the photon self-energy. Before going any further in evaluating it, it pays to figure out everything we can about it on symmetry grounds alone. Lorentz symmetry tells us that it must be of the form $\eta^{\mu\nu}A(q^2) + B(q^2)q^\mu q^\nu$, since these are the only rank-two tensors that can be formed which depend only on q . Gauge invariance requires that a gauge boson only couple to a conserved current, which means $q_\alpha \Pi^{\alpha\beta}(q) = 0$ must be satisfied. This imposes $A = -q^2 B$, so

$$\Pi^{\alpha\beta}(q) = (q^2 \eta^{\alpha\beta} - q^\alpha q^\beta) \Pi(q^2) \quad (7.27)$$

with $\Pi(q^2)$ a scalar function we need to determine. This confirms the claim above that $\Pi^{\alpha\beta}(q)$ must be quadratic in q .

The evaluation of $\Pi(q^2)$ is rather technical. We will take it up at the end of this section; here we just quote the result:

$$\Pi(q^2) = -\frac{e^2}{12\pi^2} \log \left(\frac{\Lambda^2}{M^2} \right)$$

$$\begin{aligned}
M^2 &\equiv \exp \left(6 \int_0^1 dx \, x(1-x) \log [m^2 + x(1-x)q^2] \right) \\
M^2 &\rightarrow \begin{cases} m^2 & m^2 \gg |q^2| \\ q^2 e^{-5/3} & |q^2| \gg m^2 \end{cases} \quad (7.28)
\end{aligned}$$

Here Λ is the UV cut-off scale, that is, roughly, the scale beyond which the effects of UV modes are discarded.

Measurements of the electromagnetic coupling at macroscopic scales find $e^2 = 4\pi/137.036$. However, this is not e_0^2 , the coupling appearing in the Lagrangian. Instead, it is the result from summing all graphs which contribute to scattering in the $q^2 \rightarrow 0$ limit. The leading order interaction is of the form

$$e_0^2 \left(\frac{-i\eta_{\mu\nu}}{q^2} \right) J^\mu J^\nu \quad (7.29)$$

and the contribution from diagram F is

$$\begin{aligned}
&e_0^4 \left(\frac{-i\eta_{\mu\alpha}}{q^2} \right) \frac{-i(q^2\eta^{\alpha\beta} - q^\alpha q^\beta) \log(\Lambda^2/m_e^2)}{12\pi^2} \left(\frac{-i\eta_{\beta\nu}}{q^2} \right) J^\mu J^\nu \\
&= \left(-\frac{e_0^4 \log(\Lambda^2/m_e^2)}{12\pi^2} \right) \left(\frac{-i}{q^2} \right) \left(\eta_{\mu\nu} - \frac{q_\mu q_\nu}{q^2} \right) J^\mu J^\nu \quad (7.30)
\end{aligned}$$

The appearance of $q_\mu q_\nu$ here is harmless because the currents are conserved, and so $q_\mu J^\mu = 0$. Therefore, we can sum these two contributions to find

$$e_{\text{meas}}^2 = e_0^2 - \frac{e_0^4}{12\pi^2} \log \left(\frac{\Lambda^2}{m_e^2} \right) + O(e^6) \quad (7.31)$$

Now for the crux of renormalization theory. Every electromagnetic interaction involves one photon propagator for every power of e_0^2 . For each diagram in which a propagator appears, there is another diagram which is identical except that a self-energy insertion $\Pi^{\alpha\beta}$ appears on that propagator. Therefore, e_0^2 always appears with a $-(e_0^4/12\pi^2) \log(\Lambda^2/m_e^2)$ correction, with m_e^2 some physical scale. When we carry out any arbitrary calculation of an electromagnetic process in terms of e_0^2 , if we invert Eq. (7.31) and substitute e_0^2 in terms of e_{meas}^2 , all reference to Λ will always cancel. Therefore all physical electromagnetic processes can be computed in terms of e_{meas}^2 and will be Λ independent.

To see how this works explicitly, consider high energy electromagnetic scattering with large momentum transfer, $q^2 \gg m_e^2$. Evaluating the same diagrams as above gives a matrix element of the form

$$\mathcal{M} \propto e_0^2 - \frac{e_0^4}{12\pi^2} \log \left(\frac{\Lambda^2}{q^2 e^{-5/3}} \right) + O(e_0^6) \quad (7.32)$$

Eliminating e_0 in favor of e_{meas} by substituting Eq. (7.31), this becomes,

$$\mathcal{M} \propto e_{\text{meas}}^2 - \frac{e_{\text{meas}}^4}{12\pi^2} \log \left(\frac{m_e^2}{q^2 e^{-5/3}} \right) \quad (7.33)$$

We replaced e_0^4 by e_{meas}^4 in the second term. This is correct, but because the difference is $O(e^6)$, we would have to work to one higher loop order to prove it.

The renormalization procedure just outlined eliminates all dependence on the cut-off. But it does not necessarily optimize the convergence of the perturbative expansion (the expansion in powers of e_{meas}). This is because the coefficients in this series can contain large logarithms, such as the $\log(m_e^2/q^2)$ in Eq. (7.33), which is around 22 for $q^2 = M_Z^2$, for instance. Many of these logarithms can be absorbed into the definition of the renormalized coupling itself. For instance, all dependence on the cut-off Λ would be equally well removed if we were instead to define a “running” renormalized coupling by

$$e_{\text{R}}^2(\mu) = e_0^2 - \frac{e_0^4}{12\pi^2} \log \left(\frac{\Lambda^2}{\mu^2} \right) \quad (7.34)$$

where μ is an arbitrary scale we choose for later convenience. Using this in the above scattering amplitude gives

$$\mathcal{M} \propto e_{\text{R}}^2(\mu) - \frac{e_{\text{R}}^4(\mu)}{12\pi^2} \log \left(\frac{\mu^2}{q^2 e^{-5/3}} \right) \quad (7.35)$$

If our interest is in scattering with momentum transfer near $q = q_0$, then it might be convenient to choose $\mu^2 = q_0^2 e^{-5/3}$, in which case the amplitude becomes

$$\mathcal{M} \propto e_{\text{R}}^2(q_0) - \frac{e_{\text{R}}^4(q_0)}{12\pi^2} \log \left(\frac{q_0^2}{q^2} \right) \quad (7.36)$$

With this choice for μ there are no explicit large logarithms in the series for $\mathcal{M}(q^2)$ (for q^2 near q_0^2) once it is expressed in powers of $e_{\text{R}}(q_0)$. The same would be true for other processes involving momentum transfers $q^2 \sim q_0^2$ once these are expressed in terms of the renormalized coupling $e^2(q_0)$. The utility of this approach is that all of the large logarithms are then contained within the μ -dependence of e_{R} itself, and these can be determined quite generally, once and for all, as we shall now see.

To make this prescription useful, we need to know how e_{R} varies with μ . This can be found by eliminating e_0 in terms of a physical quantity like

e_{meas} , using Eq. (7.31):

$$e_{\text{R}}^2(\mu^2) = e_{\text{meas}}^2 - \frac{e_{\text{meas}}^4}{12\pi^2} \log \left(\frac{m_e^2}{\mu^2} \right) + O(e_{\text{meas}}^6) \quad (7.37)$$

A slightly more general way of expressing this same result is obtained by differentiating Eq. (7.37) with respect to μ and using the fact that e_{meas}^2 – the result of an experiment – must be μ independent:

$$\mu^2 \frac{de_{\text{R}}^2(\mu)}{d\mu^2} \equiv \beta(e_{\text{R}}^2) = \frac{e_{\text{R}}^4}{12\pi^2} + O(e_{\text{R}}^6) \quad (7.38)$$

The function $\beta(e^2)$ is called the *beta function*, and once it is known integration of Eq. (7.38) gives $e_{\text{R}}(\mu^2)$.

Superficially it seems silly to differentiate Eq. (7.37) to obtain Eq. (7.38) if we are simply going to integrate it again to get $e_{\text{R}}(\mu)$. But a more careful analysis shows that it really is $e_{\text{R}}(\mu)$ which appears on the right-hand side in Eq. (7.38). Therefore, Eq. (7.38) has a broader domain of validity than does Eq. (7.37). This is because Eq. (7.38) relies only on the quantity $e_{\text{R}}(\mu)$ being small for the range of μ of interest, while Eq. (7.37) requires both e_{meas} and $e_{\text{meas}} \log(\mu^2/m_e^2)$ to be small. In fact this difference allows us to trust the general solution to Eq. (7.38), which is

$$\frac{1}{e_{\text{R}}^2(\mu)} = \frac{1}{e_{\text{R}}^2(\mu_0)} + \frac{1}{12\pi^2} \log \left(\frac{\mu_0^2}{\mu^2} \right) \quad (7.39)$$

and to keep this result for $e_{\text{R}}^2(\mu)$ to all orders in $e_{\text{R}}^2(\mu_0) \log(\mu^2/\mu_0^2)$, and *not* simply drop all terms which are higher order than $e_{\text{R}}^4(\mu_0)$.

The above result allows the determination of the beta function for the electromagnetic coupling, including the contribution from loops involving all charge carriers. Dropping the subscript “R” one finds

$$\beta(e^2) = b_{\text{em}} \frac{e^4}{4\pi}, \quad b_{\text{em}} = \frac{1}{3\pi} \left[\frac{1}{9} n_{-1/3} + \frac{4}{9} n_{2/3} + n_{-1} \right] \quad (7.40)$$

with $n_{-1/3}$ the number of charge- $-1/3$ Dirac fermions, $n_{2/3}$ the number of charge- $2/3$ Dirac fermions, and n_{-1} the number of charge- -1 Dirac fermions, remembering to count each color separately. A species contributes to the beta function when its mass is less than the scale- μ . The solution of Eq. (7.38) for $\alpha = e^2/4\pi$ is

$$\frac{1}{\alpha(\mu_1)} = \frac{1}{\alpha(\mu_2)} + b_{\text{em}} \log \left(\frac{\mu_2^2}{\mu_1^2} \right) \quad (7.41)$$

The other coupling of relevance for low-energy phenomena is the QCD coupling, for which a similar calculation gives

$$\frac{1}{\alpha_3(\mu_1)} = \frac{1}{\alpha_3(\mu_2)} + b_3 \log \left(\frac{\mu_2^2}{\mu_1^2} \right), \quad b_3 = \frac{1}{12\pi} (2n_q - 11N), \quad (7.42)$$

where n_q is the number of quark species and $N = 3$ is the number of colors. Notice that for $n_q \leq 16$ the sign of b_3 is opposite of the sign of b_{em} . Therefore, while $\alpha(\mu)$ grows with increasing scale μ , α_3 shrinks with scale. This latter property, called *asymptotic freedom*, is central to QCD, and is discussed in much more detail in Chapter 8.

The other couplings of the standard model also renormalize: the complete set of one loop beta functions appears in Subsection 11.4.1. The beta function in a general theory is now known through two loops, and $\beta(e^2)$ and $\beta(g_3^2)$ are known to four loops.

7.4.4 An application: the W -boson mass

To understand the utility of all this, it is useful to see a concrete example in which it improves the reliability of a calculation in the standard model. Consider the standard model relationship for the mass of the W boson:

$$\begin{aligned} M_W &= \frac{g_2 v}{2} = \frac{e}{2^{5/4} \sin \theta_W \sqrt{G_F}} \\ &= \left(\frac{\pi \alpha}{\sqrt{2} G_F \sin^2 \theta_W} \right)^{\frac{1}{2}} \end{aligned} \quad (7.43)$$

All three of the quantities on the right-hand side of Eq. (7.43) are measurable in low-energy experiments that may be performed below the threshold for producing real W bosons. This equation may then be used to predict the mass for the W boson, a result that may be compared with the measured mass in order to test the model.

The experimental values for these quantities are given in Table 7.1. The difference between the prediction and the measurement is around 3 GeV out of 80. Whether this agreement/disagreement is adequate/inadequate depends on just how accurate the prediction is. In the absence of large logarithms the corrections to Eq. (7.43) should be roughly electroweak in size: $\alpha/(4\pi \sin^2 \theta_W) \approx 3 \times 10^{-3}$ – say, a percent. This turns out to be a fair estimate of the error and the prediction and measurement of the W mass would be unacceptably different if the first calculation given in Table 7.1 were the whole story.

Table 7.1 *Computed and measured W^\pm mass*

Variable	Measurement	Scale	Value
α	Josephson effect	≈ 500 keV	$1/137.035\,999\,76$ (50)
α	(after running to M_W)	≈ 80 GeV	$1/128.1$
G_F	Muon lifetime	≈ 100 MeV	$1.166\,39(1) \times 10^{-5}$ GeV $^{-2}$
$\sin^2 \theta_W$	Electroweak fit	≈ 90 GeV	$0.231\,13(15)$
$M_{W\text{ calc}}$	(without logs)	77.5 GeV	
$M_{W\text{ calc}}$	(with logs)	80.20 GeV	
$M_{W\text{ exp}}$		80.423(39) GeV	

The main point is that the electromagnetic coupling constant, as determined by the Josephson effect, is the renormalized coupling at the scale $\mu = m_e$. To avoid large logarithms, this must be run up from m_e to the W mass. With the renormalized couplings defined appropriately this may be done by simply applying Eq. (7.41), in which the integers, n_Q , only count the number of Dirac fermions that have electric charge, Q , and have a mass that is smaller than μ . As μ is raised above a charged-particle threshold the coefficient b_{em} is modified to include this new degree of freedom. This prescription gives the following result:

$$\frac{1}{\alpha(M^2)} = \frac{1}{\alpha(m_e^2)} - \frac{1}{3\pi} \sum_f Q_f^2 \log \left(\frac{M^2}{m_f^2} \right) \quad (7.44)$$

Applying this formula to scales $M < 1$ GeV is inappropriate because it computes the running of α using quarks but at such low energies the relevant degrees of freedom are the electrically charged hadrons into which quarks bind, which we shall meet in more detail in the next chapter. It is nevertheless possible to infer how the coupling runs at these energies by using a powerful relation (not explained here) between the running of the coupling α induced by hadrons and the ratio R , encountered in the last chapter, of hadronic to muonic cross sections in e^+e^- scattering (see Eq. (6.34) and Figure 6.3). Using this relation the experimental measurement of R allows an inference as to how α runs.

At the W mass we find that $\alpha(\mu^2 = M_W^2) = 1/128.1$. The one-loop correction to G_F turns out not to involve large logs and so to the present approximation G_F need not be similarly run from m_μ to M_W . (This is because the coupling G_F arises due to a heavy particle, the W boson. Its radiative corrections therefore always involve the heavy scale M_W , even if the

momenta in question are small.) This gives the recomputed values in Table 7.1, which include the large logs. The large logarithm in this case makes the difference between agreement and disagreement of the standard model with experiment! (A more complete calculation together with a number of standard model measurements to determine the parameters of the model yield a value for the W mass which agrees with experiment at the level of the experimental error, that is, at the 0.05% level as of the date of writing (2006).)

7.4.5 Evaluation of the self-energy*

We next return to see how the quantity $\Pi(q^2)$ is evaluated. Since this evaluation is quite technical, and the technical details are not used in subsequent chapters, this subsection may be skipped in a first reading. We include it to give the interested reader an introduction to the techniques which must be used to conduct calculations beyond the leading order.

To begin, since we know that $\Pi^{\alpha\beta}(q) = (\eta^{\alpha\beta}q^2 - q^\alpha q^\beta)\Pi(q)$, we can eliminate the tensorial structure by contracting against $\eta_{\alpha\beta}$:

$$\eta_{\alpha\beta}\Pi^{\alpha\beta}(q) = (D-1)q^2\Pi(q) \quad (7.45)$$

Here $D = 4$ is the dimensionality of spacetime, which we introduce because we will need to keep track of it later. The contraction simplifies Eq. (7.26) to

$$(D-1)q^2\Pi = ie^2 \int \frac{d^4l}{(2\pi)^4} \text{tr} \left[\frac{-i(-i\not{l} + m)}{l^2 + m^2 - i\epsilon} \gamma^\alpha \frac{-i(-i\not{l} - i\not{q} + m)}{(l+q)^2 + m^2 - i\epsilon} \gamma_\alpha \right] \quad (7.46)$$

The γ matrices simplify using $\gamma^\alpha\gamma^\mu\gamma_\alpha = (2-D)\gamma^\mu$ and $\gamma^\alpha\gamma_\alpha = D$, giving

$$(D-1)q^2\Pi(q) = 4ie^2 \int \frac{d^4l}{(2\pi)^4} \frac{(2-D)l \cdot (q+l) - Dm^2}{[(l+q)^2 + m^2 - i\epsilon][l^2 + m^2 - i\epsilon]} \quad (7.47)$$

We will need two tricks and one deep method to evaluate this. The first trick is called the *Feynman parameter* trick. As the integrand stands, $q \cdot l$ appears in the denominator, which makes doing the l integral difficult. We can fix this as follows. Note that

$$\frac{1}{ab} = \int_0^1 dx \frac{1}{[ax + b(1-x)]^2} \quad (7.48)$$

and use this on the denominator:

$$(D-1)q^2\Pi(q) = 4ie^2 \int \frac{d^4l}{(2\pi)^4} \int_0^1 dx \frac{(2-D)l \cdot (q+l) - Dm^2}{[xl^2 + (1-x)(l+q)^2 + m^2 - i\epsilon]^2} \quad (7.49)$$

Now rewrite $xl^2 + (1-x)(q+l)^2 = [l + (1-x)q]^2 + x(1-x)q^2$. Then it is natural to move the x integration outside the l integral and shift l by $(1-x)q$, leaving

$$4ie^2 \int_0^1 dx \int \frac{d^4 l}{(2\pi)^4} \frac{(2-D)l^2 - (2-D)x(1-x)q^2 - Dm^2}{[l^2 + x(1-x)q^2 + m^2 - i\epsilon]^2} + O(l \cdot q) \quad (7.50)$$

We have not written out the term proportional to $l \cdot q$ because it is odd both under $l \rightarrow -l$ and $x \rightarrow 1-x$, and so integrates to zero.

The next trick is called *Wick rotation*, and consists of rotating the contour of the l^0 integration. The l^0 integration, writing non-covariantly, is of the form

$$-i \int \frac{dl^0}{2\pi} \frac{1 \text{ or } -(l^0)^2}{[l^2 + x(1-x)q^2 + m^2 - i\epsilon - (l^0)^2]^2} \quad (7.51)$$

Consider l^0 as a complex parameter. The poles in the complex l^0 plane are just below the axis at $\sqrt{l^2 + m^2 + x(1-x)q^2}$ and just above the axis at $-\sqrt{l^2 + m^2 + x(1-x)q^2}$. Therefore, we can rotate the contour to run along the imaginary axis from $-i\infty$ to $i\infty$. Defining $l_E^0 = -il^0$, the integral becomes

$$(D-1)q^2 \Pi(q) = -4e^2 \int_0^1 dx \int \frac{d^4 l_E}{(2\pi)^4} \frac{(2-D)l_E^2 - (2-D)x(1-x)q^2 - Dm^2}{[l_E^2 + x(1-x)q^2 + m^2]^2}, \quad (7.52)$$

with $l_E^2 = (l_E^0)^2 + l^2$. The four-vector (l_E^0, \mathbf{l}) behaves as a positive metric (Euclidean) variable. Henceforth we will drop the subscript.

The rotation we just performed implicitly assumed that $q^2 > -4m^2$. When this is not so, the poles in the l^0 plane are purely imaginary for some values of x and $|\mathbf{l}|$, which will lead to an imaginary part in the self-energy – in fact, exactly the imaginary part we encountered in Section 6.4. The value of the imaginary part can be determined from the positive q^2 expression via analytic continuation.

The integral we obtain diverges at large L . At this point we need to regulate, that is, to remove the contributions of very large l , in some way. However, this must be done in a way that respects all symmetries of the theory, particularly gauge invariance. The easiest procedure, simply cutting off the integral at some l_E value, breaks gauge invariance. This is because a gauge transformation changes p^μ to $p^\mu + A^\mu$, that is, shifts momenta. We must maintain invariance under momentum shifts, which a simple cut-off does not.

The deep technique we need is called *dimensional regularization*. The idea is as follows. The integral we need to conduct has an integrand which gets smaller at large l as $1/l^4$. But the volume of momentum space grows as $l^3 dl$, which is just fast enough to compensate. However, if there were $D = 4 - 2\epsilon$

dimensions, rather than 4, then the volume of momentum space would be smaller in the ultraviolet, and the integral would converge. We can at least formally define the integral in $4 - 2\epsilon$ dimensions, evaluate it, identify the physical scale where the cut-off is occurring, and continue towards $\epsilon \rightarrow 0$. This is the reason we have kept factors of D in the calculation so far.

To compare different dimensionalities, we must introduce some scale μ where the momentum space volume is equivalent between the 4-dimensional and $4 - 2\epsilon$ -dimensional theories. That is, we have to replace $\int d^4l$ with $\mu^{2\epsilon} \int d^{4-2\epsilon}l$ to get the dimensions right. The scale Λ where the size of momentum space becomes significantly smaller than in 4 dimensions is where $(\mu/\Lambda)^{2\epsilon}$ becomes significantly smaller than 1, roughly $\Lambda^2 = \mu^2 e^{1/\epsilon}$ or $\log(\Lambda^2/\mu^2) = 1/\epsilon$. Therefore we will interpret factors of $1/\epsilon$ in the result as logarithms of the UV cut-off, $\log(\Lambda^2/\mu^2)$ (up to a constant).

If this procedure looks mysterious to you, you are not alone. One should not try to make any physical interpretation of changing the dimensionality of spacetime. It is just a trick for getting integrals to converge in the UV, which has the advantage that it is gauge-invariant. This is important, because it looks like the integral we are faced with, Eq. (7.52), will not give something proportional to q^2 , as we believe it must. In particular, there is one contribution proportional to m^2 rather than q^2 , and another which behaves as $l^2 d^4l/l^4$ which looks to be quadratically divergent, which we also argued cannot happen. However, in any gauge-invariant regularization, these problems will take care of themselves.

Now let us carry out the integrations. Since the integral in Eq. (7.52) depends only on l^2 , we may write

$$\mu^{2\epsilon} \int \frac{d^{4-2\epsilon}l}{(2\pi)^{4-2\epsilon}} = \mu^{2\epsilon} (2\pi)^{-4+2\epsilon} \int d\Omega_D \int l^{3-2\epsilon} dl \quad (7.53)$$

with $\int d\Omega_D$ the integral over angular coordinates in $4 - 2\epsilon$ dimensions. The first integration gives the volume of the sphere in D dimensions, which turns out to be

$$\int d\Omega_D = \frac{2\pi^{D/2}}{\Gamma(D/2)} \quad (7.54)$$

with Γ the usual gamma function, $\Gamma(n+1) = n!$. (One can check using $\Gamma(1/2) = \sqrt{\pi}$, $\Gamma(n+1) = n\Gamma(n)$ that this expression agrees with the well-known results in 1, 2, 3, and 4 dimensions, in which the area of the sphere is 2, 2π , 4π , and $2\pi^2$ respectively.) The integral we need to know, in order to evaluate the radial piece, is

$$\int \frac{l^{2x-1} dl}{(l^2 + \alpha)^y} = \alpha^{x-y} \frac{\Gamma(x)\Gamma(y-x)}{2\Gamma(y)} \quad (7.55)$$

This result strictly only holds for $x > 0$ and $y - x > 0$; otherwise we define the answer by analytic continuation.

Using Eq. (7.54) and Eq. (7.55) to evaluate Eq. (7.52) gives

$$\begin{aligned}
& -4e^2 \int_0^1 dx \mu^{2\epsilon} \int \frac{d^D l}{(2\pi)^D} \frac{(2-D)l^2 - (2-D)x(1-x)q^2 - Dm^2}{(l^2 + x(1-x)q^2 + m^2)^2} \\
&= \frac{4e^2}{(2\pi)^4} \left((2\pi)^2 \mu^2 \right)^\epsilon \int_0^1 dx \int d\Omega_D \\
& \quad \times \left((D-2) \int \frac{l^{2(3-\epsilon)-1} dl}{(l^2 + x(1-x)q^2 + m^2)^2} + Dm^2 \int \frac{l^{2(2-\epsilon)-1} dl}{(l^2 + x(1-x)q^2 + m^2)^2} \right. \\
& \quad \left. + (2-D)x(1-x)q^2 \int \frac{l^{2(2-\epsilon)-1} dl}{(l^2 + x(1-x)q^2 + m^2)^2} \right) \\
&= -\frac{4e^2}{(4\pi)^2} \left(4\pi\mu^2 \right)^\epsilon \int_0^1 dx \\
& \quad \times \left([m^2 + x(1-x)q^2]^{1-\epsilon} \frac{2(-1+\epsilon)\Gamma(3-\epsilon)\Gamma(-1+\epsilon)}{\Gamma(2)\Gamma(2-\epsilon)} \right. \\
& \quad - m^2 [x(1-x)q^2 + m^2]^{-\epsilon} \frac{2(2-\epsilon)\Gamma(2-\epsilon)\Gamma(\epsilon)}{\Gamma(2)\Gamma(2-\epsilon)} \\
& \quad \left. + x(1-x)q^2 [x(1-x)q^2 + m^2]^{-\epsilon} \frac{2(1-\epsilon)\Gamma(2-\epsilon)\Gamma(\epsilon)}{\Gamma(2-\epsilon)\Gamma(2)} \right) \quad (7.56)
\end{aligned}$$

Now using $n\Gamma(n) = \Gamma(n+1)$ wherever available, the m^2 terms cancel between the first two expressions, and the whole simplifies to

$$(3-2\epsilon)q^2\Pi(q) = -q^2 \frac{e^2}{2\pi^2} \int_0^1 dx x(1-x) \left(\frac{4\pi\mu^2}{m^2 + x(1-x)q^2} \right)^\epsilon (3-2\epsilon)\Gamma(\epsilon) \quad (7.57)$$

The result is proportional to q^2 , as required by gauge invariance.

Since we want the answer in four spacetime dimensions, we must take the small ϵ limit of this expression. The small ϵ behavior of $\Gamma(\epsilon)$ is

$$\Gamma(\epsilon) = \frac{1}{\epsilon} - \gamma_E + O(\epsilon) \quad (7.58)$$

The $1/\epsilon$ behavior is easily understood: $\Gamma(1+\epsilon)$ is almost $0! = 1$, and $\Gamma(\epsilon) = \Gamma(1+\epsilon)/\epsilon$. In fact the function $\Gamma(x)$ has a $1/\epsilon$ -type pole at all negative integers. The presence of the $1/\epsilon$ factor means that we need other quantities to $O(\epsilon)$ accuracy; in particular, note that

$$x^\epsilon = e^{\epsilon \ln(x)} \simeq 1 + \epsilon \ln(x) + O(\epsilon^2) \quad (7.59)$$

which allows us to handle the bracketed quantity to obtain

$$\Pi(q^2) = -\frac{e^2}{2\pi^2} \left(\frac{1}{\epsilon} + \gamma_E + \ln 4\pi + \ln \frac{\mu^2}{m^2} \right) \int_0^1 dx x(1-x) \ln \frac{m^2}{m^2 + x(1-x)q^2} \quad (7.60)$$

As discussed above, we absorb the Λ dependence (or $1/\epsilon$ dependence) into the definition of the coupling, e_0^2 . There is always an ambiguity when doing so, as to how much of the finite part of Π also to absorb into e_0^2 . The “modified minimal subtraction” or $\overline{\text{MS}}$ scheme is defined by absorbing the $[1/\epsilon + \gamma_E + \ln 4\pi]$ into e_0^2 , leaving a (subtracted) self-energy of

$$\Pi(q^2) = \frac{-e^2}{12\pi^2} \ln \frac{\mu^2}{M^2} \quad (7.61)$$

with M^2 defined as in Eq. (7.28).

7.5 Higgsless effective theory

Having seen how physics below M_W is efficiently described by an effective field theory, we can ask, what happens if the Higgs boson is very heavy compared to the W boson? Presumably, it should be possible to construct an effective theory valid for energy scales above M_W and M_Z , but below the Higgs mass. For scales below m_t , one can also remove the top quark from the low-energy fields as well. What would such an effective theory look like?

There is another reason for considering this kind of effective theory, independent of the wish to identify the main ways through which a heavy Higgs boson enters into observables within the standard model. This reason asks, what if the Higgs is absent from the theory altogether? After all, there is no direct experimental evidence for a Higgs particle to date (2006). What should we expect if we keep all of the standard model fields *except* the Higgs? This question is not purely academic, since there are explicit models for which the electroweak gauge group is spontaneously broken by a composite field which is bound together by a new type of strong interaction (often called “technicolor”).

Such an effective theory contains some but not all of the fields in certain $SU_L(2) \times U_Y(1)$ representations. That is, at scales below m_t , the theory contains the left-handed bottom quark but not the left-handed top, even though they fit in the same $SU_L(2)$ doublet, Q . At scales below m_H , the theory contains three of the four components of the Higgs doublet, which have become the longitudinal components of the W^\pm and Z bosons, but not the fourth component, the Higgs boson. Therefore the effective theory

must explicitly break $SU_L(2) \times U_Y(1)$. More properly, $SU_L(2) \times U_Y(1)$ is *non-linearly realized*. The effective theories we will discuss can therefore contain renormalizable operators which would be forbidden in the standard model Lagrangian. The result, as we shall see, is that the effective theory contains a host of possible effective operators and is less predictive than the full model. Unlike the standard model, it *must* also break down at energies not too far above $M_W/\sqrt{\alpha}$, since at energies higher than this certain cross sections become impossibly large owing to the absence of the Higgs particle in the effective theory, indicating that the effective description *must* break down below this energy.

7.5.1 Realizations of broken gauge symmetries*

A key assumption in writing down an effective field theory is the low-energy particle content, since this is required in order to identify exhaustively what kinds of interactions are possible having any given dimension. This assumption is also crucial in another way, since it governs how symmetries can restrict the form of the low-energy theory.

The main casualty following from omitting the Higgs is $SU_L(2) \times U_Y(1)$ invariance, at least in the linear realization we use throughout this book. This symmetry must change because without the physical Higgs particle there is no $SU_L(2)$ multiplet which can contain the three longitudinal modes of the W and Z bosons, corresponding to the three would-be Goldstone bosons of the Higgs doublet. (A Goldstone boson is a scalar field of a type described in more detail in Subsection 8.3.5.) What happens to the symmetry is most easily seen by taking the limit where the Higgs gets heavy within the standard model, and integrating it out explicitly, leaving an effective theory containing all of the other standard model particles. This can be done without running into problems with a non-perturbative Higgs self-coupling if the Higgs mass is not too large – say $m_H < (400\text{--}500)$ GeV.

Such studies show that there are two equivalent ways to think about the three spontaneously-broken generators of $SU_L(2) \times U_Y(1)$ in the resulting low-energy theory.

- *Non-linear realization* include the three would-be Goldstone boson scalars in the low-energy theory, and require the Lagrangian to be invariant under *non-linearly realized* gauge transformations; or
- *No gauge symmetry* do not include the would-be Goldstone boson scalars at all in the low-energy theory and completely ignore the spontaneously broken gauge symmetries.

The second case can be regarded as the first, evaluated in an appropriate unitary gauge.

Each of these formulations is most useful for different kinds of applications. The non-linearly-realized form is most useful when computing loop graphs in the low-energy theory, which arise once one goes beyond leading order in the low-energy expansion. The second formulation is more convenient for identifying the kinds of independent effective interactions which arise within the low-energy Lagrangian, and when exploring their physical consequences in tree graphs.

It is simplest to see how this works in a toy example of an abelian gauge field, although the same arguments also generalize to the non-abelian case. Consider for these purposes a minor modification of the scalar electrodynamics model introduced in Subsection 1.6.2, consisting of two complex charged scalar fields, ϕ and χ , coupled to electromagnetism with the following Lagrangian,

$$\begin{aligned}\mathcal{L} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - D_\mu\phi^*D^\mu\phi - D_\mu\chi^*D^\mu\chi - V(\phi,\chi) \\ V(\phi,\chi) &= \lambda^2(\phi^*\phi - v^2)^2 + m^2\chi^*\chi + \cdots\end{aligned}\tag{7.62}$$

where the ellipses represent various quartic interactions involving χ (whose form does not concern us), and the covariant derivatives are $D_\mu\phi = \partial_\mu\phi - ieA_\mu\phi$ and $D_\mu\chi = \partial_\mu\chi - ieqA_\mu\chi$. The $U(1)$ gauge symmetry of this model acts on the fields as $\phi \rightarrow e^{ie\omega}\phi$, $\chi \rightarrow e^{ieq\omega}\chi$ and $A_\mu \rightarrow A_\mu + \partial_\mu\omega$.

The potential V is minimized when the fields take on values $\phi = v$ and $\chi = 0$, which spontaneously breaks the gauge symmetry, and leads to a mass $m_\chi = m$ for the real and imaginary parts of the χ field, and mass $2\lambda v$ for the real part of ϕ . Im ϕ is the would-be Goldstone boson, which through the Higgs mechanism is “eaten” by the gauge boson to give it its mass, $m_A = ev$. We choose parameters $e \ll \lambda$ and $m \ll \lambda v$ to ensure $m_\chi, m_A \ll m_H$, and integrate out the heavy scalar H .

Our goal is to identify the low-energy Lagrangian and how it realizes the gauge symmetry, and this is most easily done explicitly by writing $\phi = H e^{i\xi}$, so that the gauge transformation properties are $H \rightarrow H$ and $\xi \rightarrow \xi + e\omega$. The low-energy theory obtained after removing H then consists of all possible Lorentz-invariant interactions

$$\mathcal{L}_{\text{eff}} = \sum_I c_I \mathcal{O}_I(\tilde{\chi}, \tilde{A}_\mu)\tag{7.63}$$

built from the low-energy fields $\tilde{\chi} = e^{-iq\xi}\chi$ and $\tilde{A}_\mu = A_\mu - \partial_\mu\xi/e$.

Now comes the main point: the form of the low-energy interactions, \mathcal{O}_I , are completely *unconstrained* by the broken gauge interactions! It is unconstrained because both of the fields $\tilde{\chi}$ and \tilde{A}_μ do not transform at all, since their ξ -dependence is chosen to cancel the transformation properties of the original fields χ and A_μ . Furthermore, the fields $\tilde{\chi}$ and \tilde{A}_μ reduce to the original fields χ and A_μ in unitary gauge, which in these variables is defined by the gauge-fixing condition $\xi = 0$.

We see from this example the two equivalent ways described above to think about spontaneously-broken gauge symmetries in a low-energy effective theory. Non-linear realization corresponds to keeping the would-be Goldstone bosons (ξ) in the low-energy theory in addition to the low-energy fields (χ and A_μ). In this picture the couplings of ξ to the other fields are restricted by the non-linearly-realized gauge symmetry. Conversely, in the alternative formulation the low-energy theory does not depend on ξ at all, and depends only on the low-energy fields ($\tilde{\chi}$ and \tilde{A}_μ). (In this picture the would-be Goldstone particles are the longitudinal spin states of the explicitly massive spin-one fields.) The Lagrangian is completely unconstrained by any spontaneously-broken gauge symmetries because these have been by setting $\xi = 0$.

The same conclusions apply to more complex non-abelian gauge groups, such as for $SU_L(2) \times U_Y(1)$, for which the non-linearly-realized transformations can also be characterized quite generally. The result for the three standard model would-be Goldstone bosons is very similar to the non-linear transformation of the pseudoscalar meson multiplet, \mathcal{M} , under the spontaneously broken chiral symmetries of QCD in the effective theory defined in Subsection 8.3.6 for use below the QCD scale. The quarks and leptons similarly transform under the broken gauge symmetries in a complicated way involving the would-be Goldstone boson fields, in much the same way as do the baryon fields under the broken chiral symmetries.

7.5.2 Application to the standard model

We now summarize the kinds of effective interactions which arise in the low-energy effective theory without a Higgs particle. To do so we use the formulation wherein we ignore both the would-be gauge bosons and the spontaneously-broken gauge symmetries. The field content of the theory consists of the three neutrinos, ν_n , plus Dirac spinors representing the charged leptons, ℓ_m , and quarks, u_m and d_m . There is no Higgs particle, and the massless gauge bosons are the gluons and photon: G_μ^α and A_μ for the gauge symmetry $SU_c(3) \times U_{\text{em}}(1)$. Finally, there are massive vector fields, W and

Z . If the t -quark is included then such a theory would apply at energies above the t -quark mass, and if it is not then it could apply above the Z boson mass.

Lowest-dimension interactions

Since the standard model works so well, it is useful to write the effective theory in the form $\mathcal{L} = \mathcal{L}_{\text{SM}} + \delta\mathcal{L}$, where \mathcal{L}_{SM} denotes the standard model Lagrangian density with the heavy fields H , t set to zero. δ represents the effects of virtual H and t particles, and we can take it to be small. We further imagine working in the physical mass basis of fields, for which the mass and kinetic terms appearing in \mathcal{L}_{SM} are already diagonalized. The lowest-dimension terms which are possible in $\delta\mathcal{L}$ given the above field content and gauge symmetries have dimension two:

$$\delta\mathcal{L}_{\text{dim}2} = -w\tilde{m}_W^2 W_\mu^* W^\mu - \frac{z}{2} \tilde{m}_Z^2 Z_\mu Z^\mu \quad (7.64)$$

corresponding to small (but otherwise arbitrary) corrections to the W and Z boson masses. Here we scale out for convenience factors of the standard model mass parameters, \tilde{m}_W^2 and \tilde{m}_Z^2 , so that $|w|, |z| \ll 1$ denote small dimensionless parameters. We write tildes over all standard-model parameters because, as we shall see, the presence of the new terms in $\delta\mathcal{L}$ changes the connection between these parameters and observable quantities like physical particle masses. For instance, this means that $\tilde{m}_W = \tilde{m}_Z \tilde{c}_w$, where $\tilde{c}_w = \cos \tilde{\theta}_w$, even though the physical masses and mixing angles, M_W , M_Z and θ_w , need not satisfy this relation (even at tree level) when $w, z \neq 0$.

Since the W and Z mass terms arise from a single term in \mathcal{L}_{SM} – i.e. the kinetic term of the Higgs doublet – we expect that one combination of w and z can be reabsorbed into \mathcal{L}_{SM} by suitably redefining fields and parameters, leaving only one combination of w and z to have physically-observable implications. This expectation is borne out by explicit calculation below.

The next-lowest-dimension effective interactions arise at dimension three, corresponding to arbitrary corrections to the fermion mass matrices. Unfortunately, since there is already a separate parameter in the standard model for each term in these mass matrices, the presence of these new dimension-3 terms are difficult to detect. The main exception to this statement is the presence of Majorana neutrino masses,

$$\delta\mathcal{L}_{\text{dim}3} = -\frac{1}{2} (m_\nu)_{ij} \bar{\nu}_i P_L \nu_j + \text{h.c.} \quad (7.65)$$

which are no longer forbidden by $SU_L(2) \times U_Y(1)$ invariance. The diagonalization of such masses introduce new CKM-type matrices into the leptonic

charged-current interactions, about which we have much more to say in Chapter 10.

Still more interactions arise at dimension four, including

$$\begin{aligned} \delta\mathcal{L}_{\text{dim } 4} = & -\frac{a}{4}\hat{F}_{\mu\nu}\hat{F}^{\mu\nu} - \frac{b}{2}\hat{W}_{\mu\nu}^*\hat{W}^{\mu\nu} - \frac{c}{4}\hat{Z}_{\mu\nu}\hat{Z}^{\mu\nu} + \frac{k}{2}\hat{F}_{\mu\nu}\hat{Z}^{\mu\nu} \\ & - \frac{\tilde{e}}{\tilde{s}_w\tilde{c}_w}\hat{Z}_\mu\bar{f}_i\gamma^\mu(P_L\delta\tilde{g}_{ij}^L + P_R\delta\tilde{g}_{ij}^R)f_j \\ & - \frac{\tilde{e}}{\sqrt{2}\tilde{s}_w}\hat{W}_\mu^*\bar{f}_i\gamma^\mu(P_L\delta\tilde{h}_{ij}^L + P_R\delta\tilde{h}_{ij}^R)f'_j + \dots \end{aligned} \quad (7.66)$$

where (as before) tildes denote the parameters which appear in the standard model part of the Lagrangian and the ellipses include other dimension-4 terms, like kinetic terms for fermions and self-interactions amongst the spin-one particles. We drop these terms here for simplicity, since they are less well-constrained than are the ones which are explicitly written. We anticipate our later comparison with experiments and imagine all of the dimensionless effective parameters appearing here, a, b, \dots to be small.

We write ‘hats’ over the boson fields in the above expressions since the kinetic and mass terms for these must be re-diagonalized owing to the presence of $\delta\mathcal{L}$. Working to linear order in the small effective couplings we see that this is accomplished by the following redefinitions:

$$\begin{aligned} \hat{A}_\mu &= \left(1 - \frac{a}{2}\right) A_\mu + k Z_\mu \\ \hat{Z}_\mu &= \left(1 - \frac{c}{2}\right) Z_\mu \\ \hat{W}_\mu &= \left(1 - \frac{b}{2}\right) W_\mu \end{aligned} \quad (7.67)$$

The W and Z mass terms then become $-M_W^2 W_\mu^* W^\mu - \frac{1}{2}M_Z^2 Z_\mu Z^\mu$, with the physical masses now given in terms of the Lagrangian parameters by

$$M_W^2 = \tilde{m}_W^2(1 + w - b) \quad M_Z^2 = \tilde{m}_Z^2(1 + z - c) \quad (7.68)$$

Observable implications

Having made this redefinition, it is convenient to identify the combination of parameters which appears in observables. We must first identify the three parameters which define the standard model part of the electroweak physics, and we follow standard practice by choosing these to be the best-measured quantities: the physical Z mass, M_Z , the electromagnetic coupling, $\alpha = e^2/4\pi$, and Fermi’s constant as measured in muon decays, G_F .

To this end we may use Eq. (7.68) to eliminate everywhere \tilde{m}_Z in favor of M_Z . We similarly write the fermion electromagnetic couplings in terms of the newly-diagonalized field A_μ , to find

$$\mathcal{L}_{\text{em}} = -\tilde{e} \left(1 - \frac{a}{2}\right) A_\mu i \bar{f} \gamma^\mu Q f \quad (7.69)$$

and from this we see that the physical electric charge is given by

$$e = \tilde{e} \left(1 - \frac{a}{2}\right) \quad (7.70)$$

which we use from here on to eliminate \tilde{e} in favor of e . Finally, calculating the Fermi constant which would be inferred from measurements of muon decay we find (to linear order)

$$\frac{G_F}{\sqrt{2}} = \frac{\tilde{e}^2}{8\tilde{s}_w^2 \tilde{c}_w^2 \tilde{m}_Z^2} (1 - w + \Delta_e + \Delta_\mu) \quad (7.71)$$

where

$$\Delta_\ell = \left[\sum_i \left| \delta_{\nu_i \ell} + \delta h_{\nu_i \ell}^L \right|^2 \right]^{1/2} - 1 \approx 2 \text{Re} \delta h_{\nu_\ell}^L \quad (7.72)$$

Here the sum is over all three neutrino flavors and the last approximate equality holds if we linearize in the “new” left-handed lepton couplings, $\delta h_{\nu_i \ell}^L$. Finally, if we define the physical weak mixing angle, $c_w = \cos \theta_w$ and $s_w = \sin \theta_w$, in terms of the measured quantities e , M_Z , and G_F , by

$$\frac{G_F}{\sqrt{2}} \equiv \frac{e^2}{8s_w^2 c_w^2 M_Z^2} \quad (7.73)$$

then we may use these expressions to eliminate \tilde{s}_w and \tilde{c}_w in terms of s_w and c_w to get

$$\tilde{s}_w^2 = s_w^2 \left[1 + \frac{c_w^2}{c_w^2 - s_w^2} (a - c - w + z + \Delta_e + \Delta_\mu) \right] \quad (7.74)$$

These redefinitions express the freedom to absorb the new effective couplings into the definitions of standard model fields and couplings, and so once they have been done the remaining parameters express physically-measurable deviations between the predictions of $\mathcal{L}_{\text{SM}} + \delta\mathcal{L}$ and the standard model. For instance, only three of the six quantities a, b, c, k, w , and z appear in physical observables after the above redefinitions, because three of them can be absorbed into the three standard-model operators which describe the gauge-boson masses and kinetic terms. A conventional parameterization of

the three physical combinations of parameters is given by

$$\begin{aligned}
\alpha S &= 4s_w^2 c_w^2 \left[a - c - k \left(\frac{c_w^2 - s_w^2}{c_w s_w} \right) \right] \\
\alpha T &= w - z \\
\alpha U &= 4s_w^2 \left[a - \frac{b}{s_w^2} + c \left(\frac{c_w^2}{s_w^2} \right) - 2k \left(\frac{c_w}{s_w} \right) \right]
\end{aligned} \tag{7.75}$$

where α is the physical electromagnetic coupling.

It is now straightforward to find how observables depend on the physical effective couplings, since this can be obtained by perturbing the relevant standard-model expression to linear order in the deviations parameterized by $\delta\mathcal{L}$. For instance, the physical W mass is given by Eq. (7.68) using $\tilde{m}_W = \tilde{m}_Z \tilde{c}_w$, leading to

$$M_W^2 = (M_W^2)_{\text{SM}} \left[1 - \frac{\alpha S}{2(c_w^2 - s_w^2)} + \frac{c_w^2 \alpha T}{c_w^2 - s_w^2} + \frac{\alpha U}{4s_w^2} + \frac{s_w^2(\Delta_e + \Delta_\mu)}{c_w^2 - s_w^2} \right] \tag{7.76}$$

once expressed in terms of the physically-relevant parameters. In this expression $(M_W^2)_{\text{SM}} = M_Z^2 c_w^2 + (\delta M_W^2)_{\text{r.c.}}$ is the standard-model value, including any relevant radiative corrections. (Because these radiative corrections are performed as if the standard model were true, they paradoxically involve an implicit choice of fiducial values for the t -quark and Higgs boson masses, even though these do not appear within the effective theory.)

Almost all other precisely-measured observables come down to measurements of the couplings of the W and Z bosons to fermions, which in the physical basis are

$$\begin{aligned}
\delta\mathcal{L}_{\text{cc}} &= -\frac{e}{\sqrt{2}s_w} W_\mu^* \bar{f}_i \gamma^\mu (P_L \delta h_{ij}^L + P_R \delta h_{ij}^R) f_j' + \text{h.c.} \\
\delta\mathcal{L}_{\text{nc}} &= -\frac{e}{s_w c_w} Z_\mu \bar{f}_i \gamma^\mu (P_L \delta g_{ij}^L + P_R \delta g_{ij}^R) f_j
\end{aligned} \tag{7.77}$$

with the physical charged-current couplings given by $\delta h_{ij}^R = \tilde{\delta h}_{ij}^R$ and

$$\delta h_{ij}^L = \tilde{\delta h}_{ij}^L + V_{ij} \left[-\frac{\alpha S}{4(c_w^2 - s_w^2)} + \frac{c_w^2 \alpha T}{2(c_w^2 - s_w^2)} + \frac{\alpha U}{8s_w^2} - \frac{c_w^2(\Delta_e + \Delta_\mu)}{2(c_w^2 - s_w^2)} \right] \tag{7.78}$$

where V_{ij} is the relevant CKM-type mixing angle for the fermions of interest.

The physical neutral-current couplings are similarly

$$\begin{aligned} \delta g_{ij}^{\text{L(R)}} &= \delta \tilde{g}_{ij}^{\text{L(R)}} + \frac{1}{2} \delta_{ij} g_i^{\text{L(R)}} [\alpha T - \Delta_e - \Delta_\mu] \\ &\quad - \delta_{ij} Q_i \left[\frac{\alpha S}{4(c_w^2 - s_w^2)} - \frac{c_w^2 s_w^2 \alpha T}{c_w^2 - s_w^2} + \frac{c_w^2 s_w^2 (\Delta_e + \Delta_\mu)}{c_w^2 - s_w^2} \right] \end{aligned} \quad (7.79)$$

Here $g_i^{\text{L}} = T_{3i} - Q_i s_w^2$ and $g_i^{\text{R}} = -Q_i s_w^2$ denote the usual standard model Z couplings, with T_{3i} and Q_i denoting the relevant fermion electric charge and diagonal weak isospin.

The implications for observables are quite simple to calculate, because they are obtained by computing the result to lowest order as a function of M_W , $g_{ij}^{\text{L(R)}} = (g_{ij}^{\text{L(R)}})_{\text{SM}} + \delta g_{ij}^{\text{L(R)}}$, etc. and then using the above expressions giving M_W , $\delta_{ij}^{\text{L(R)}}$ in terms of the effective couplings αS , αT , αU , Δ_e , and so on. It typically suffices to work to linear order in these couplings, and this is justified a posteriori by the comparison with precision electroweak observations, which shows that deviations from the standard-model values must be small.

For instance, the ρ -parameter, defined as the relative strength of the low-energy neutral-current and charged-current lepton couplings, receives contributions from δh , δg , and M_W , and is given by

$$\rho = 1 + \alpha T \quad (7.80)$$

Similarly, the quantities Δ_ℓ measure violations of lepton universality, with

$$\begin{aligned} R_\pi &\equiv \frac{\Gamma(\pi \rightarrow e\nu)}{\Gamma(\pi \rightarrow \mu\nu)} = R_\pi^{\text{SM}} [1 + 2\Delta_e - 2\Delta_\mu] \\ R_\tau &\equiv \frac{\Gamma(\tau \rightarrow e\nu\bar{\nu})}{\Gamma(\mu \rightarrow e\nu\bar{\nu})} = R_\tau^{\text{SM}} [1 + 2\Delta_\tau - 2\Delta_\mu] \\ R_{\mu\tau} &\equiv \frac{\Gamma(\tau \rightarrow \mu\nu\bar{\nu})}{\Gamma(\mu \rightarrow e\nu\bar{\nu})} = R_{\mu\tau}^{\text{SM}} [1 + 2\Delta_\tau - 2\Delta_e] \end{aligned} \quad (7.81)$$

Finally, observables in e^+e^- scattering near the Z pole are obtained by using the above coupling in the expression $\Gamma_f = [\alpha M_Z / (6s_w^2 c_w^2)] [|g_{ff}^{\text{L}}|^2 + |g_{ff}^{\text{R}}|^2]$, leading to

$$\begin{aligned} \Gamma_f &= (\Gamma_f)^{\text{SM}} \left[1 + \alpha T - \Delta_e - \Delta_\mu + \frac{2g_f^{\text{L}} \delta \tilde{g}_{ff}^{\text{L}} + 2g_f^{\text{R}} \delta \tilde{g}_{ff}^{\text{R}}}{(g_f^{\text{L}})^2 + (g_f^{\text{R}})^2} \right. \\ &\quad \left. - \left(\frac{2g_f^{\text{L}} + 2g_f^{\text{R}}}{(g_f^{\text{L}})^2 + (g_f^{\text{R}})^2} \right) Q_f \left(\frac{\alpha S}{4(c_w^2 - s_w^2)} - \frac{c_w^2 s_w^2 (\alpha T - \Delta_e - \Delta_\mu)}{c_w^2 - s_w^2} \right) \right] \end{aligned} \quad (7.82)$$

Using these formulae, the success of the comparison of standard-model predictions with observations can be parlayed into quantitative constraints on the size of the effective interactions in $\delta\mathcal{L}$ which are permitted by the data. Although a complete discussion goes beyond the scope of this book, fits to the many well-measured electroweak observables restrict the “oblique” parameters S , T , and U to be in the range (taking $m_H = 117\text{GeV}$, see below)

$$S = -0.13 \pm 0.10 \quad T = -0.17 \pm 0.12 \quad U = 0.22 \pm 0.13 \quad (7.83)$$

where $1\text{-}\sigma$ errors are quoted. The quantities Δ_e and Δ_μ are similarly constrained to be smaller than $O(10^{-3})$. The bounds on the direct couplings, $\delta\tilde{h}_{ij}$ and $\delta\tilde{g}_{ij}$, depend on the fermion species involved, but are comparably strong.

Sensitivity to a heavy top quark and Higgs boson

As an application of the above formalism it is instructive to compute the dominant contribution of virtual top quarks and Higgs bosons to precision electroweak observables. Assuming only that both are heavy compared with the Z boson (as we know to be true for the top quark), this can be done by first computing the effective interactions which describe their virtual effects at low energy, and then asking how the effective couplings in $\delta\mathcal{L}$ depend on their masses, m_t and m_H .

There are two advantages of performing the calculation in this way. The first is that it allows us to be systematic in identifying those places which depend strongly on these masses, since on dimensional grounds the most dominant dependence arises for the coefficients of the lowest-dimension interactions. The second advantage is that this type of calculation is comparatively easy to do. After all, once we find how the effective couplings like S or T depend on m_t and m_H , we may simply use the above expressions for observables in terms of these couplings, and need not laboriously rederive the dependence of each observable on top quarks and Higgs bosons.

Since the strongest dependence on heavy masses is likely to come from the lowest-dimension interactions, we focus first on the dimension-2 effective couplings, w and z , which we expect can depend quadratically on the relevant heavy mass. We have seen that only the combination $\alpha T = w - z$ enters observables, so only this quantity need be computed. Although in principle the Higgs boson could contribute to this quantity through a loop graph in which a W or Z temporarily splits into a Higgs loop, we know that all of the m_H^2 contributions cancel in αT . We know this because αT contributes to the low-energy ρ parameter, but such a contribution is forbidden

for the Higgs within the standard model by the Higgs sector's approximate custodial $SU(2)$ symmetry, discussed in Subsection 2.3.2.

The same is not true for the top quark, however, because the custodial symmetry is broken by the mass difference between the t - and b -quarks. Its size can be determined by computing the top and bottom quarks' contributions to the one loop self-energy of the W and Z bosons. If all fermion masses are neglected except for m_t , then the contribution obtained for T is

$$(\alpha T)_{\text{quad}} \approx \frac{3\alpha}{16\pi \sin^2 \theta_W} \left(\frac{m_t^2}{M_W^2} \right) \simeq 0.0095 \quad (7.84)$$

We have seen that αT contributes to many electroweak observables, and in particular to M_W . If the tree-level mass relation between the W - and Z -boson masses were exact, then the value of $\sin^2 \theta_W$ quoted earlier, together with $M_Z = 91.187$ GeV, would predict $M_W = 79.955$ GeV, which is substantially off from the experimental value, $M_W = 80.423(39)$ GeV. However, including the above correction from the top loop gives $M_W = 80.334$ GeV, in much better agreement.

The dominant m_H dependence can be identified in a very similar way. Since $O(m_H^2)$ contributions to dimension-2 interactions in $\delta\mathcal{L}$ are precluded by the standard model's custodial symmetry, the next-largest contributions (for large m_H) should be $O(\log m_H)$ contributions to dimension-4 interactions. Such contributions (as well as $\log m_t$ terms) are indeed generated at one loop within the standard model by the Higgs contribution to the W and Z self-energies, and show up in $\delta\mathcal{L}$ as contributions to S of the form

$$\begin{aligned} (\alpha S)_{\log} &= \frac{\alpha}{12\pi} \left[\log \left(\frac{m_H^2}{\hat{m}_H^2} \right) - 2 \log \left(\frac{m_t^2}{\hat{m}_t^2} \right) \right] \\ (\alpha T)_{\log} &= \frac{3\alpha}{16\pi} \left[\log \left(\frac{m_t^2}{\hat{m}_t^2} \right) - \log \left(\frac{m_H^2}{\hat{m}_H^2} \right) \right] \end{aligned} \quad (7.85)$$

$$(\alpha U)_{\log} = -\frac{\alpha}{2\pi} \log \left(\frac{m_t^2}{\hat{m}_t^2} \right) \quad (7.86)$$

Here \hat{m}_t and \hat{m}_H represent the fiducial values of the top and Higgs mass which are used to define the standard model predictions which are the reference against which $\delta\mathcal{L}$ is compared.

It is through contributions like this that precision fits of standard model predictions to electroweak measurements acquire a dependence on the Higgs mass. At the time of writing (2006) the success of these fits gives a preference for a comparatively light Higgs, $53 \text{ GeV} < m_H < 213 \text{ GeV}$ at 90% C.L. (as

opposed to the direct-search limit $m_H > 114.4$ GeV) providing at present the only experimental indication we have that the Higgs exists.

Contributions to parameters like T and S can also be generated by new heavy particles in theories beyond the standard model, particularly those for which there is no Higgs boson or the Higgs boson is very heavy. Although contributions to T can be suppressed in these theories if the W and Z couplings are not protected by custodial symmetries, the same is not typically true of S , which tends to strongly constrain such theories because of the good observational bounds on how big S can be. One of the main motivations for taking a doublet Higgs field seriously is the existence of the custodial symmetry and its success in describing the precision electroweak measurements.

Unitarity constraints

Because gauge invariance does not relate the various W and Z couplings in $\delta\mathcal{L}$, severe problems arise if one tries to push too high the mass scale where the Higgsless effective description is to apply. (At some scale new particles must emerge, much as the W boson was necessary to complete the Fermi theory.) In particular the tree-level cross sections for processes like $e^+e^- \rightarrow W^+W^-$ and $W^+W^- \rightarrow W^+W^-$ computed using the generic couplings of this Lagrangian grow like $d\sigma/d\cos\theta \sim \alpha^2 E_{\text{cm}}^2/M_W^4$ at high center-of-mass energies, E_{cm} (see Problem 7.3). But growth this fast with energy is inconsistent with general upper bounds based on the unitarity of the underlying S -matrix. Such an absurd result can only happen if an approximation in the calculation is breaking down, and the fatal approximation in this instance is the low-energy expansion which is always implicit when using an effective field theory.

This tells us that the effective field theory having generic W and Z couplings must break down at energies for which $\alpha E_{\text{cm}}^2/M_W^2$ is of order unity, or E_{cm} is of order 1 TeV. If there is no Higgs boson, then some other new degrees of freedom must intervene at energies at or below this energy in order to cure this problem.

We note in passing that the standard model itself does not have a similar restriction because within the standard model there are cancellations between the graphs contributing to the problematic processes. For instance, for the process $e^+e^- \rightarrow W^+W^-$, cancellations between graphs involving the exchange of a virtual photon, Z boson and neutrino occur when one imposes the special relations amongst the W and Z couplings which gauge invariance enforces. Similarly, as seen in Problem 7.3, Higgs exchange cancels the would-be divergence in the $WW \rightarrow WW$ cross section. It is only

for gauge-invariant couplings that the effective interactions of the massive spin-one particles can make sense up to energies $E_{\text{cm}}^2 \gg M_W^2/\alpha$.

7.6 Problems

[7.1] Running b quark mass

The mass of the b quark, listed in Appendix A, is 4.24 GeV. In Chapter 5, Problem 5.2, you were instructed to use 3.1 GeV, which is substantially smaller. In this problem we see why.

At leading order in α_3 , the scale dependence of the strong coupling $g_3(\mu)$ and of $h_{33}(\mu)$ (the Yukawa coupling responsible for the bottom mass) are, in the regime $m_b < \mu < m_t$,

$$\mu \frac{\partial g_3}{\partial \mu}(\mu) = -\frac{23}{3} \frac{g_3^3(\mu)}{16\pi^2},$$

and

$$\mu \frac{\partial h_{33}}{\partial \mu}(\mu) = -8 \frac{h_{33}(\mu) g_3^2(\mu)}{16\pi^2}$$

The former expression neglects corrections suppressed by a further $g_3^2/4\pi^2$, and the latter neglects corrections smaller by a factor of g_2^2/g_3^2 .

The value of $g_3(\mu)$ is such that $\alpha_3(\mu = M_Z) = 0.118$. The value of h_{33} is determined by the tree-level expression, $m_b = h_{33}v/\sqrt{2}$, if you evaluate $h_{33}(\mu)$ using $\mu = m_b$, that is, $h_{33}(\mu = m_b) = m_b\sqrt{2}/v$ (since this is the energy scale relevant to the physical mass of the b quark).

The Higgs decay process involves the $Hb\bar{b}$ vertex at a large energy scale $\simeq m_H$. Therefore, it is $h_{33}(\mu = m_H)$ which is relevant in evaluating the Higgs decay width.

Evaluate $h_{33}(\mu = m_H)$, assuming $m_H = 120$ GeV. You should be able to solve for $h_{33}(\mu = m_H)$ *explicitly* (that is, analytically without resorting to numerical methods). However, if you get frustrated, solve the differential equations by numerical means.

Using the tree level relation $m_b = h_{33}v/\sqrt{2}$, evaluate what b mass you should use, to evaluate the Higgs decay width.

[7.2] Hypercharge beta function

Find the expression for b_1 the coefficient describing the scale dependence of the hypercharge coupling, in analogy with Eq. (7.41). Argue from its sign that Eq. (7.41) for hypercharge predicts that at some ultraviolet scale Λ , the hypercharge coupling becomes infinite. Using the value for $g_1^2(\mu = M_Z)$

from Appendix A to determine what this scale Λ is. Do you expect that the energy scale Λ you found will ever be experimentally probed?

[7.3] **WW scattering and unitarity**

This problem shows why there *must* be a Higgs boson, or *some* additional fields beyond those which have been detected experimentally to date.

Consider the process $W^+W^- \rightarrow W^+W^-$. For simplicity, work in the limit $g_1 = 0$, which is the same as setting $e = 0$ and $e \cot \theta_w = g_2$. Furthermore, for simplicity, consider only the scattering of longitudinally polarized W bosons into longitudinally polarized W bosons (in the CM frame, which you should use throughout). Write the W energy as E and the W momentum as $p = \sqrt{E^2 - M_W^2}$.

[7.3.1] Define the angle between \mathbf{p}_1 and \mathbf{k}_1 , in the rest frame of the system, to be θ . For a particle with momentum \mathbf{p} , take two of the polarization vectors, the transverse ones, to be purely spatial and orthogonal to \mathbf{p} , that is, $\epsilon(\mathbf{p}, \lambda_1, \lambda_2) \cdot \mathbf{p} = 0$ and $\epsilon^0(\mathbf{p}, \lambda_1, \lambda_2) = 0$. The third, longitudinal, polarization, is defined so that $\epsilon(\mathbf{p}, \lambda_3) \cdot \mathbf{p} = 0$, $\epsilon(\mathbf{p}, \lambda_3) \cdot \epsilon(\mathbf{p}, \lambda_1, \lambda_2) = 0$, $\epsilon^2(\mathbf{p}, \lambda_3) = 1$, and $\epsilon^0(\mathbf{p}, \lambda_3) > 0$.

Working in spherical coordinates with \vec{p}_1 along the z axis, find explicit expressions, in terms of p and θ , for the longitudinal polarization vectors for each particle, that is, for the momenta p_1 , p_2 , k_1 , and k_2 .

[7.3.2] Forget for a moment that there is a Higgs boson.

Find the three tree-level Feynman diagrams which contribute to the matrix element for $W^+W^- \rightarrow W^+W^-$. Draw them. (The fourth diagram, photon exchange, is absent because we took $e = 0$.)

Evaluate each diagram's contribution to the matrix element for longitudinally polarized W bosons to scatter into longitudinally polarized W bosons, at general angle θ . Be *very careful* to find correctly the relative signs of the different matrix elements.

[7.3.3] Show that, in the limit $s \rightarrow \infty$ (the large energy limit), individual matrix elements diverge as E^4/M_W^4 . However, when you add the three matrix elements together, the divergence is ameliorated to “only” E^2/M_W^2 . Find an expression for the sum of the matrix elements.

Unitarity requires that, if \mathcal{M} is decomposed into spherical harmonics,

$$\mathcal{M}(\theta) = \sum_{\ell} \mathcal{M}_{\ell}(2\ell + 1)P_{\ell}(\cos \theta)$$

with P_{ℓ} the ℓ th Legendre polynomial, then

$$\mathcal{M}_{\ell} \leq 16\pi$$

is the maximum value allowed in a partial wave analysis. Show that the result violates this bound for $\ell = 0$ above about $s \simeq (2500 \text{ GeV})^2$.

[7.3.4] Now, find the two additional Higgs-boson mediated graphs. Evaluate their matrix elements, and show that, when summed with the other matrix elements, the combination is well behaved in the large E limit, that is, the matrix element does not grow with a power of E for generic θ . What limit may be placed on m_H to ensure that this asymptotic value satisfies the unitarity bound?

Part III

Applications: hadrons

8

Hadrons and QCD

Up until this point quarks and gluons have been treated in basically the same manner as have leptons, with little acknowledgement of the plain fact that quarks and gluons are never directly seen in experiments in the same way as are electrons, protons, and pions. This voluntary blindness has been possible to the extent that we have always considered processes that have the following two important properties:

- (i) they never involve any strongly interacting particles in the initial state, and
- (ii) they never involve any specific combinations of strongly-interacting particles in the final state – so-called *exclusive* processes.

The only reactions involving hadrons that have been contemplated are *inclusive* ones, i.e. those for which we have summed over all possible combinations of hadrons that could be produced, sometimes subject to some general flavor-conservation rules.

This is obviously a fairly serious handicap, since the vast bulk of the reactions that are seen in experiments involve strongly-interacting particles of one sort or another. In order to be considered a success, the standard model must provide at least a qualitative, but preferably also a quantitative, picture of these processes. The model does indeed provide such a framework. This chapter and Chapter 9 are devoted to outlining to what extent predictions and post-dictions can be made and to what extent they are successful. The goal is to focus here on those calculations that can be made with the minimum of modeling of the unknown dynamics of strongly-coupled physics. The present chapter is devoted to collecting together the general qualitative features of strong-interaction physics as it is understood using quantum chromodynamics (QCD).

8.1 Qualitative features of the strong interactions

Since all of the strongly-interacting particles whose properties are meant to be understood, except for the top quark, are light compared to the W and Z bosons, it is convenient to work to leading order in $1/M_W$ and so formulate the analysis in terms of the low-energy effective Lagrangian that was developed in Chapter 7. This Lagrangian is constructed in just such a way as to efficiently include the dominant contributions of the standard model at energies that are low compared to M_W .

The leading terms of this low-energy Lagrangian are given by \mathcal{L}_0 of Chapter 7:

$$\mathcal{L}_0 = \mathcal{L}_{\text{kin}} + \mathcal{L}_{\text{str}} + \mathcal{L}_{\text{em}} \quad (8.1)$$

The three terms in this expression are given by Eq. (7.2), Eq. (7.3), and Eq. (7.4). This low-energy Lagrangian splits very naturally into a theory of a number of fermion flavors that interact through two kinds of interactions: strong and electromagnetic. The third low-energy interaction – the weak interaction – has coupling strength G_F and so is higher order in $1/M_W^2$ than those given in \mathcal{L}_0 .

The terms in \mathcal{L}_0 that involve just the strongly-interacting particles of the low-energy theory – the quarks and gluons – are given by \mathcal{L}_{str} together with the quark kinetic- and mass-terms. These, taken together, make up the *QCD Lagrangian*

$$\begin{aligned} \mathcal{L}_{\text{QCD}} = \sum_q & \left[-\bar{q}(\not{\partial} + m_q)q + \frac{ig_3}{2} G_\mu^\alpha \bar{q} \gamma^\mu \lambda_\alpha q \right] \\ & - \frac{1}{4} G_{\mu\nu}^\alpha G_\alpha^{\mu\nu} - \frac{g_3^2 \Theta_3}{64\pi^2} \epsilon^{\mu\nu\lambda\rho} G_{\mu\nu}^\alpha G_{\lambda\rho}^\alpha \end{aligned} \quad (8.2)$$

Since experiment requires that the parameter Θ_3 must be extremely small (see Subsection 11.4.2) we take it to be zero in what follows.

For comparison, the terms that involve just the electromagnetically interacting particles – the *QED Lagrangian* – are

$$\mathcal{L}_{\text{em}} = \sum_f \left[-\bar{f}(\not{\partial} + m_f)f + ie A_\mu \bar{f} \gamma^\mu Q f \right] - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \quad (8.3)$$

The forms of these two Lagrangians are clearly very similar. The fermion couplings in each differ only in the current to which each type of gauge boson couples. Their principal difference is that the gluons couple directly to one another while the photon does not. The strength of these two interactions are respectively described in terms of dimensionless gauge couplings, g_3 and e .

8.1.1 Asymptotic freedom and the strong-coupling regime

As measured at energies of the order of the mass of the Z boson, the strong and electromagnetic coupling strengths are

$$\begin{aligned}\alpha_3(\mu^2 = M_Z^2) &\equiv \frac{g_3^2}{4\pi}(\mu^2 = M_Z^2) = 0.12 \\ \alpha(\mu^2 = M_Z^2) &\equiv \frac{e^2}{4\pi}(\mu^2 = M_Z^2) = \frac{1}{128} = 7.8 \times 10^{-3}\end{aligned}\quad (8.4)$$

The strong coupling is stronger (hence the name), although both are small enough at this scale to have some confidence in perturbative methods.

At lower energies, these couplings run differently and so take different values depending on what scale is of interest. For $m_b < \mu < M_Z$ their running is given by Eq. (7.42) to Eq. (7.41), which are reproduced for convenience here:

$$\begin{aligned}\frac{1}{\alpha_i(\mu)} &= \frac{1}{\alpha_i(M_Z)} + b_i \log\left(\frac{M_Z^2}{\mu^2}\right) \\ \text{with : } b_3 &= \frac{1}{12\pi}[2n_q - 33] \\ \text{and : } b_{\text{em}} &= \frac{1}{3\pi}\left[\left(\frac{1}{9}\right)n_{-1/3} + \left(\frac{4}{9}\right)n_{2/3} + n_{-1}\right]\end{aligned}\quad (8.5)$$

In these expressions n_q represents the number of Dirac fermions whose mass is smaller than μ that transform as triplets of $SU_c(3)$. n_q with $Q = -1, 2/3$ and $-1/3$ similarly counts the number of light Dirac fermions in the theory having each of these electric charges. In practice, for μ greater than the b -quark mass ($m_b \approx 5$ GeV), $n_q = 5$, $n_{2/3} = 6$, $n_{-1/3} = 9$, and $n_{-1} = 3$. With these numbers we have $b_3 = -0.61$ and $b_{\text{em}} = +0.71$.

Since b_{em} is positive, the electromagnetic coupling decreases with decreasing μ . Although the running is only logarithmic and so is very slow, in principle perturbative methods become more and more accurate for electromagnetic processes at lower energies.

By contrast b_3 is negative, so the strong coupling gets increasingly strong for lower scales μ . As was mentioned in Chapter 7, a coupling such as g_3 that decreases with increasing μ is said to be *asymptotically free*. Eventually, for μ sufficiently small, the strong coupling must become large enough to invalidate perturbation in g_3 . Putting in the numbers, and remembering to remove the contribution of each fermion species as μ falls below its mass threshold, we find

$$\begin{aligned}\alpha_3(M_Z) &= 0.12 \\ \alpha_3(m_b) &= 0.20\end{aligned}$$

$$\begin{aligned}
\alpha_3(m_c) &= 0.30 \\
\alpha_3(m_s) &= 1.7
\end{aligned}
\tag{8.6}$$

Since the combination $\alpha_3/(4\pi)$ counts loops, we may estimate that, although it is a good approximation to work perturbatively in α_3 for $\mu = M_Z \approx 90$ GeV or $m_b \approx 5$ GeV, and marginally acceptable at $m_c \approx 1.4$ GeV, the perturbative expansion is expected to fail at several hundred MeV, $\mu \sim m_s \approx 200$ MeV. At these scales we should instead expect non-perturbative phenomena to play the dominant role. At these scales, the one-loop approximation to the scale-dependence of the coupling constant, Eq. (8.5), should itself be suspect.

We see that there is a natural scale associated with an asymptotically free interaction such as the QCD coupling, g_3 . This is the scale at which the coupling constant becomes “strong.” Precisely what this scale is will depend on just what is meant by “strong.” For QCD it has become conventional to define the *QCD scale*, Λ_{QCD} , as the point at which the perturbative running takes α_3 to infinity.[†] That is, given the value for the strong coupling at a particular scale – such as in Eq. (8.4) – and ignoring thresholds, Λ_{QCD} is determined by solving Eq. (8.5) (or its higher-loop generalization) for μ once $1/\alpha_3(\mu)$ is set to zero. The result naturally depends on the value chosen for n_q , and is given at one loop for $n_q = 3, 4$, or 5 by

$$\begin{aligned}
\Lambda_{\text{QCD}}^{(n_q)} &= M_Z \exp\left(\frac{1}{2b_3\alpha_3(M_Z)}\right) \\
\text{so : } \Lambda_{\text{QCD}}^{(3)} &= 220 \text{ MeV} \\
\Lambda_{\text{QCD}}^{(4)} &= 140 \text{ MeV} \\
\text{and : } \Lambda_{\text{QCD}}^{(5)} &= 80 \text{ MeV}
\end{aligned}
\tag{8.7}$$

It is, of course, not to be believed that the strong coupling really does diverge at Λ_{QCD} as given in Eq. (8.7). Instead, Λ_{QCD} simply provides a fairly idiosyncratic way of parameterizing the strength of the strong coupling, since knowledge of Λ_{QCD} is completely equivalent, by Eq. (8.7), to knowledge of α_3 at some particular scale. The main point is that there really is a physical scale, in the ballpark of several hundred MeV, at which α_3 does become strong. At this scale, our ability to make firm quantitative predictions for the behavior of QCD becomes quite limited.

For certain quantities, numerical techniques called *lattice QCD* are available, and allow predictions at the $\simeq 5\%$ level. In particular, the masses and

[†] Strictly speaking this is Λ_{QCD} in the *minimal* or *modified minimal subtraction* renormalization scheme.

spins of the QCD bound states, to be discussed in the remainder of this chapter, can be computed from the QCD Lagrangian by lattice methods, and are in good agreement with observations. To some extent, symmetry arguments allow for fairly precise statements. For instance, as we will discuss in Section 9.3, the low-energy interactions of the lightest hadrons can be described fairly accurately by an effective theory, chiral perturbation theory, which is based only on symmetry arguments and a few experimental inputs. However, to a good extent, detailed quantitative understanding is simply lacking. Nevertheless, we must push ahead and see what we can say qualitatively or semi-quantitatively.

8.1.2 Bound states

We know that electromagnetic phenomena are well described by quantum electrodynamics, which is equivalent to the electromagnetic part, \mathcal{L}_{em} , of the low-energy standard model Lagrangian, \mathcal{L}_0 . Even such a weakly-coupled theory has an extremely rich spectrum that includes many bound states such as ordinary atoms, molecules, or states involving more exotic particles such as positronium or muonium. The same should be expected to be true in spades for a more strongly coupled theory such as QCD.

Furthermore, since QED and QCD are so very similar in the form of the couplings between the elementary fermions and the gauge bosons, they can be expected to have very similar predictions for the energy levels and decay widths for states that are heavy enough to lie within the perturbative regime for QCD. In this section we flesh out this similarity to argue that QCD tends to bind the quarks and gluons into color-neutral bound states.

To this end, consider the interaction energy, E_{int} , due to the strong interactions of a quark–antiquark state, $|q_r, \bar{q}_s\rangle$. For simplicity of presentation we suppress here all quark indices except for the color labels r and s . E_{int} represents the difference between the energy of the two-quark state (in the center-of-mass frame) and the energy, $E_0 = E(q) + E(\bar{q})$, of the non-interacting quark and antiquark. In perturbation theory this interaction energy arises for the first time at second order in the quark–gluon interaction Hamiltonian

$$E_{\text{int}} = \int dp \frac{\langle q_r, \bar{q}_s | H_{\text{int}} | p \rangle \langle p | H_{\text{int}} | q_t, \bar{q}_u \rangle}{E(q) + E(\bar{q}) - E(p)}$$

with

$$H_{\text{int}} = -\frac{ig_3}{2} \int d^3x (\bar{q} \gamma^\mu \lambda_\alpha q) G_\mu^\alpha \quad (8.8)$$

Introducing the notation, \mathcal{E} , for the reduced matrix element once all of the

coupling constants and color indices have been made explicit, the result becomes

$$E_{\text{int}} = -\frac{g_3^2}{4} \mathcal{E} \sum_{\alpha=1}^8 (\lambda_\alpha)_{rt} (\lambda_\alpha)_{us} \quad (8.9)$$

where the eight Gell-Mann matrices, λ_α , are given explicitly in Eq. (1.186).

For the purposes of comparison, the electromagnetic interaction energy of the same two-quark state is, to leading order in perturbation theory,

$$E_{\text{int(em)}} = -e^2 Q_1 Q_2 \delta_{rt} \delta_{us} \mathcal{E} \quad (8.10)$$

for the *same* matrix element \mathcal{E} . Q_i is the electric charge of quark type i . The main observation here is that because both matrix elements involve only the initial quarks themselves as well as intermediate massless spin-one particles, the two interaction energies only differ at leading order in their explicit coupling-constant and color factors.

Now, since the electric charge of the antiquark is $-Q$ and since opposite electric charges attract and like-sign charges repel, we know that the sign of the matrix element is positive: $\mathcal{E} > 0$. This ensures that, say, an up quark and an up antiquark would lower their energy through the electromagnetic interaction and so attract one another. We can use this information to learn how the sign of the interaction energy depends on the color combination taken by the two quarks.

Since each quark comes in three colors, there are a total of nine color combinations that a two-quark state could take. These nine combinations transform under $SU_c(3)$ as an octet or a singlet: $\mathbf{3} \otimes \mathbf{\bar{3}} = \mathbf{8} \oplus \mathbf{1}$ (see Appendix B). These combinations may be found by taking the following linear combinations of the quark–antiquark state:

$$\begin{aligned} |q, \bar{q}; \mathbf{1}\rangle &= \sqrt{\frac{1}{3}} \sum_{r=1}^3 |q_r, \bar{q}_r\rangle \\ &= \sqrt{\frac{1}{3}} \text{tr}[|q, \bar{q}\rangle] \\ \text{and} \quad |q, \bar{q}; \mathbf{8}_\alpha\rangle &= \sqrt{\frac{1}{2}} \sum_{r,s=1}^3 (\lambda_\alpha)_{sr} |q_r, \bar{q}_s\rangle \\ &= \sqrt{\frac{1}{2}} \text{tr}[\lambda_\alpha |q, \bar{q}\rangle], \quad \alpha = 1, \dots, 8 \end{aligned} \quad (8.11)$$

Using the following properties of the color matrices, $\text{tr}[\lambda_\alpha \lambda_\beta] = 2\delta_{\alpha\beta}$ and $\sum_\alpha \text{tr}[\lambda_\alpha \lambda_\beta \lambda_\alpha \lambda_\beta] = -4/3$ (see Appendix B), we see that the perturbative energies of interaction within QCD of the octet and singlet quark–antiquark

combinations are

$$\begin{aligned}
 E_{\text{int}}(\mathbf{1}) &= -\frac{g_3^2}{12}\mathcal{E} \sum_{\alpha=1}^8 \text{tr}[\lambda_\alpha \lambda_\alpha] \\
 &= -\frac{4}{3}g_3^2\mathcal{E} \\
 \text{and : } E_{\text{int}}(\mathbf{8}_\beta) &= -\frac{g_3^2}{8}\mathcal{E} \sum_{\alpha=1}^8 \text{tr}[\lambda_\alpha \lambda_\beta \lambda_\alpha \lambda_\beta] \\
 &= +\frac{1}{6}g_3^2\mathcal{E}
 \end{aligned} \tag{8.12}$$

Notice that, given that \mathcal{E} is positive, the color-singlet combination has an attractive interaction energy and the color-octet combination is repulsive.

Just as for electromagnetism, we expect bound states to form in the cases where the interaction energy is negative. In the same way that it is energetically favorable to form *electrically neutral* bound states within QED, we see that it is favorable within perturbative QCD to form bound states that are *colorless*.

This argument can be repeated for other combinations of quarks and gluons in order to decide which combinations can be expected to bind together into bound states. In all cases it is the color-neutral combination which most prefers to bind. The quantum numbers of strongly-interacting bound states should therefore be equivalent to those of the color-singlet combinations of quarks and gluons.

There are two particularly simple color-singlet combinations of quarks and antiquarks. These are the singlet quark-antiquark pair discussed above, together with the completely antisymmetric combination of three quarks (or, equivalently, antiquarks):

$$|q, q, q; \mathbf{1}\rangle = \epsilon^{rst}|q_r, q_s, q_t\rangle \tag{8.13}$$

ϵ^{rst} here is the completely antisymmetric symbol, which is an invariant tensor of $SU_c(3)$. As we shall see, these combinations of quarks have the correct quantum numbers to account for all of the known long-lived hadrons – with no extras left over, and with perhaps one exception. The quark-antiquark combinations describe the strongly interacting mesons such as the pions, kaons and rhos etc. The three-quark configurations correspond to the baryons: the protons, neutrons, deltas, and such. The three antiquark combinations are the antibaryons.

Moving beyond the simplest combinations of quarks, there are a few other ways strongly interacting particles might bind together into hadrons:

glueballs, exotics, and multi-quark objects. We return to this possibility briefly at the end of Subsection 8.3.4 but just note here that if such combinations exist, they produce highly unstable particles, with perhaps a single exception: the pentaquark Θ^+ with strangeness +1.

8.1.3 The confinement hypothesis

The precise one-to-one correspondence between the color-singlet combinations of quarks and antiquarks and the observed spectrum of strongly-interacting hadrons is strong circumstantial evidence in favor of the interpretation of these particles as bound states of quarks and antiquarks. This evidence is further strengthened by the equally perfect agreement between the predicted and observed quantum numbers of these states under the conserved and approximately-conserved quantities of the strong interactions. (More about this in subsequent sections.) Even better, the *pièce de résistance* is furnished by the extremely successful, quantitative understanding of high-energy inelastic hadron scattering that is possible once hadrons are recognized as composites of quarks and gluons. As discussed in Section 9.2, data in these experiments are very well described in terms of hard, incoherent scattering among the constituents. This agreement includes features that are sensitive to the properties of the constituents, and which bear out the properties expected from the quark-composite picture.

There is an experimental fly in all of this theoretical ointment, however. This problem is the complete failure to directly experimentally observe any of the constituent quarks or gluons. That is, *free quarks* have never been observed. After all, even though bound states arise in QED, the constituent electrons and nuclei are themselves frequently observed, as are the processes such as ionization in which constituents are released from their bondage by a sufficiently large transfer of energy. The analogs of these processes have never been seen for strongly interacting particles, even though the putative bound states have been submitted to collision energies that are many times the rest energy of the bound particles themselves.

What appears to be the resolution to this conundrum may be formulated as the following conjecture about the spectrum of QCD.

- *The confinement hypothesis*: the only energy eigenstates of the QCD Hamiltonian which have finite energy are color neutral.

In principle it should be possible to derive this hypothesis from the QCD Lagrangian. At this time, no rigorous derivation exists, so it is not absolutely clear that the confinement hypothesis is a *bone fide* prediction of QCD.

However, a good deal of circumstantial evidence in its favor has accumulated over the years, not least by direct attempts to compute the spectrum numerically, via lattice QCD. A non-rigorous but plausible physical picture for how confinement works has also been developed, and checked to some extent by lattice QCD. Since the confinement hypothesis seems quite bizarre at first sight, it is useful to present this physical picture in at least a little detail.

Consider a bound state, which for the sake of simplicity we take to be a bound state of two heavy quarks. Imagine that we could “grab hold” of the quarks and pull them apart from each other. At short distances, we would encounter a perturbative potential $V \sim \alpha_3(r)/r$. As r increased to the scale $1/\Lambda_{\text{QCD}}$, however, α_3 would become large, and the potential would rise more steeply than as $1/r$. In fact, it is believed that, in the absence of light quarks, the potential would rise linearly with separation, $V \sim \sigma r$. Such behavior is called *linear confinement*. Lattice gauge theory has presented some evidence for this behavior, and even allows an estimate of the *string tension* σ , namely $\sigma \sim 800 \text{ MeV/fermi}$. Therefore, as we attempted to get the quarks further and further apart, a large investment of energy would become necessary. Eventually, this energy would be enough to create a quark–antiquark pair out of the vacuum. The quark would be attracted to the heavy anti-quark and would form a bound state with it; the anti-quark would act similarly with the heavy quark. We could then pull the heavy-quark pair apart from each other as far as we liked, but we would not be creating isolated colored particles. Instead, we would be separating color-neutral bound states, formed from the original particles and extra particles created out of the vacuum.

This thought experiment tells us of an important feature that should be expected of confining bound states in asymptotically free theories. Colored constituents should have apparently contradictory properties. They should behave for small separations as if they are only very weakly interacting, since the coupling at these small distances is weak. They are nonetheless effectively confined within a definite *confinement radius*, r_c , whose size is set by the condition that the strong coupling, $\alpha(\Lambda_c)$, becomes large at the scale $\Lambda_c = 1/r_c$. From what we know about the running of the strong coupling, we expect that Λ_c should be in the neighborhood of the only other scale in this problem: Λ_{QCD} . This would put it in the range of several hundred MeV. The confinement radius is then in the neighborhood of a fermi, 10^{-13} cm .

Notice also that, by the uncertainty principle, particles that are confined to a region with these linear dimensions should have momenta of order $p \approx \Lambda_c$.

For light quarks, this is large compared to the rest mass of the constituent, and the energy is determined by this momentum, $E = p \approx \Lambda_c$. The upshot is that the energy scale which characterizes the mass of such a bound state is set by the confinement scale, Λ_c , provided only that the constituents are quite light in comparison.

8.1.4 The hadronic zoo

We are led to expect that all bound states involving only light constituents – i.e. gluons, up quarks, down quarks, and (possibly) strange quarks – should all cluster in energy independent of the masses of the constituent quarks that are involved. The energy of these bound states, and the spacing between energy levels, should be around Λ_{QCD} in order of magnitude. This is in the vicinity of hundreds of MeV. Conversely, bound states involving heavier quarks, such as c or b , should have masses that are more dominated by the intrinsic mass of the quarks they contain.

Consultation of the *Particle Data Book* shows that, with a few important exceptions such as the pions, which are discussed further below, this expectation is borne out by what is experimentally observed. There are very many strongly-interacting resonances whose energies and energy-spacings range from half to several GeV. Although they tend to be somewhat heavier than Λ_c , they have the expected energy to within the accuracy of the very rough estimate.

It is natural, then, to divide the discussion of the general features of the QCD spectrum into two qualitatively different cases, according to whether the masses of the constituent particles are larger or smaller than the confinement scale. The next two sections deal with each of these cases in turn.

8.2 Heavy quarks

If the bound particles happen to be much heavier than the confinement scale, $m_q > \Lambda_c$, then they can be expected to be non-relativistic. In this case, much of the formalism that has arisen over the years for describing non-relativistic bound states can be expected to apply.

8.2.1 Quarkonia

Consider for definiteness a quark–antiquark bound state in which both the quark and the antiquark are heavy. Concrete examples of this type would be the $b\bar{b}$ or $c\bar{c}$ bound states which are respectively believed to describe the

Table 8.1 *Zero-point momenta for the heavier $q\bar{q}$ systems*

q	m_q	$\alpha_3(m_q)$	$r_b^{-1} = \alpha_3(m_q)m_q$	$\alpha_3(r_b)$	$\alpha_3(r_b)m_q$
b	5000 MeV	0.20	1000 MeV	0.35	1750 MeV
c	1400 MeV	0.30	420 MeV	0.62	870 MeV
s	200 MeV	1.7	340 MeV	0.77	150 MeV

experimentally observed Υ and J/ψ resonances. The main thing to establish is just how widely separated the constituents are within these bound states, since this sets the size of the relevant strong coupling. For heavy quarks such as these, the existence of confinement should be largely irrelevant, at least for the lowest-lying states, since the quark and antiquark pair can be easily contained inside a region of radius r_c without acquiring an appreciable zero-point energy.

To the extent that the size of the bound state is indeed small compared to r_c , it should be very similar to the corresponding electromagnetic particle–antiparticle system in which an e^+e^- pair bind into a *positronium* “atom.” (The quark analogs of this system have borrowed the name and are similarly called *quarkonium* states.) From experience with the electromagnetic problem, we know that the size of this type of bound state is set by the analog of the Bohr radius: $r_b \approx 1/(\alpha_3 m_q)$, in which m_q is the mass of the bound quark. The corresponding zero point momentum is then $p \approx \alpha_3 m_q$, giving a binding energy of order $E_{\text{bind}} = p^2/(2m_q) \approx \alpha_3^2 m_q$.

Clearly, provided that the size of such a bound state is much smaller than the confinement scale, $r_b \ll r_c$ (or equivalently $\alpha_3 m_q \gg \Lambda_c$), this gives a consistent picture of a positronium-like bound state of quarks. As may be seen from the last column in Table 8.1, if we take Λ_c to be in the neighborhood of several hundred MeV, this condition is well satisfied by the $b\bar{b}$ state (*bottomonium*), is acceptable for the $c\bar{c}$ system (*charmonium*), but is at best marginal for $s\bar{s}$.

The $c\bar{c}$ and $b\bar{b}$ states may be labeled with their angular momentum quantum numbers, ℓ and m , their principal quantum number, n , and by the spin quantum numbers, s and s_z , of the quarks. The same spectroscopic nomenclature is commonly used that is used for atoms. That is S, P, D, F, ... represent the $\ell = 0, 1, 2, 3, \dots$ eigenstates, etc.

8.2.2 Quantum numbers

Among the most basic properties of these bound states that can be predicted are their quantum numbers, masses, and lifetimes. The predictions for

quantum numbers are robust in the sense that they do not depend on the details of how the inter-quark potential is modeled. They are also important, both from the point of view of classifying the observed spectrum and for understanding the selection rules that allow a qualitative understanding of the expected rates for various transitions among the various energy levels. For this reason, we defer a discussion of masses and lifetimes to the literature and focus here on the quantum numbers of the expected bound states.

It should be emphasized that these quantum number assignments that are derived in this section are not restricted to bound states that are made up exclusively of heavy quarks. This is because nothing need be assumed about the form for the wavefunctions of the states, or about how complicated the overlap with various many particle states might be. As a result, the conclusions drawn within the present section can be applied quite generally for *any* color-singlet quark combination.

The exact symmetries of the strong interaction Lagrangian, \mathcal{L}_{QCD} , are discussed in some detail in Section 7.3. Since Θ_3 is known from experiment to be extremely small, we take it here to be zero. In this case, the strong interactions conserve separately each of the discrete symmetries, C, P, and T. There is also a $U(1)$ flavor symmetry which ensures the conservation of each species, i.e. up, down, strange, charmed etc., of fermion. Baryon number, B , may be defined in terms of these charges by taking their sum.

Since gluons are neutral under the flavor symmetries, the quantum numbers of the quark bound states may be simply read from the quantum numbers of the *valence quarks* that are being bound. To see this in more detail, taking the quarkonium systems as our example once more, we write the bound-state energy eigenfunction, $|\Psi_{\ell m s s_z}\rangle$, in the following way:

$$|\Psi_{\ell m s s_z}\rangle = \sum_{\sigma, \sigma'} \int_{p, p'} \psi_{\ell m s s_z}(\mathbf{p}, \sigma; \mathbf{p}', \sigma') a_{\mathbf{p}\sigma}^* \bar{a}_{\mathbf{p}'\sigma'}^* |\mathcal{G}\rangle + \cdots \quad (8.14)$$

where $\int_{p, p'}$ means $\int d^3p d^3p' / [(2\pi)^6 2p^0 2p'^0]$. $|\mathcal{G}\rangle$ here represents the gluonic piece of the two quark component of the state, and $\psi_{\ell m s s_z}(\mathbf{p}, \sigma; \mathbf{p}', \sigma')$ represents the amplitude for finding valence quarks having the given linear momentum and spin. The dots represent other terms which involve additional quark-antiquark states.

The main observation is that, since the conserved charges commute with the QCD Hamiltonian, each term in this sum must always involve states which share the same flavor quantum numbers. For the purposes of identifying the quantum numbers of the full state, Ψ , it is therefore sufficient

to identify the quantum numbers of the term that is explicitly displayed in Eq. (8.14).

To take a simple example, given that each quark carries baryon number $1/3$ and each antiquark has baryon number $-1/3$, we have

$$\begin{aligned} [B, a_{\mathbf{p}\sigma}^*] &= \frac{1}{3} a_{\mathbf{p}\sigma}^* \\ \text{and : } [B, \bar{a}_{\mathbf{p}\sigma}^*] &= -\frac{1}{3} \bar{a}_{\mathbf{p}\sigma}^* \end{aligned} \quad (8.15)$$

Since the baryon number of the glue state $|\mathcal{G}\rangle$ is zero, this implies that the baryon number of $|\Psi\rangle$ is the sum of that of the valence quark and antiquark, and so is zero. Schematically,

$$\begin{aligned} B|\Psi\rangle &= \sum \int \psi B a^* \bar{a}^* |\mathcal{G}\rangle \\ &= \sum \int \psi \left[\frac{1}{3} a^* + a^* B \right] \bar{a}^* |\mathcal{G}\rangle \\ &= \sum \int \psi a^* \bar{a}^* \left[\frac{1}{3} - \frac{1}{3} + B \right] |\mathcal{G}\rangle \\ &= 0 \end{aligned} \quad (8.16)$$

The identical argument gives the eigenvalues of the other flavor quantum numbers. Clearly, the three-quark states, $|qqq\rangle$, carry baryon number $B = 1$, and the three antiquark states, $|\bar{q}\bar{q}\bar{q}\rangle$, have $B = -1$ etc.

A less trivial example is given by the quantum numbers of the meson states under C and P. Mesons can be classified according to their parity eigenvalues. If the quark and antiquark in a meson should have the same flavor then they may also be labeled by their eigenvalue under charge-conjugation. Quark-antiquark states that are eigenstates of these symmetries must, on general grounds, have eigenvalues that are determined by their other quantum numbers. This is because the transformation properties of the quark and antiquark are necessarily related to one another. The transformation rules for particle states are given by Eq. (2.104) in Section 2.5:

$$\begin{aligned} \mathcal{P} a_{\mathbf{p},\sigma}^* \mathcal{P}^* &= \eta_p a_{-\mathbf{p},\sigma}^* \\ \mathcal{C} a_{\mathbf{p},\sigma}^* \mathcal{C}^* &= \eta_c \bar{a}_{\mathbf{p},\sigma}^* \end{aligned} \quad (8.17)$$

In these expressions j denotes the spin of the particle involved and σ represents the eigenvalue of the z -component of angular momentum. The overbar on the creation operator indicates that it creates an antiquark. Although we are ultimately interested in the case $j = 1/2$ we will keep the quark spin, j , arbitrary here in order to outline how the final result would change for

different quark spins. The undetermined phases, η , may differ from particle type to particle type.

The crucial point for what follows is that general principles require that the antiparticles must transform with phases that are the *complex conjugates* of those of Eq. (8.17):

$$\begin{aligned}\mathcal{P}\bar{a}_{\mathbf{p},\sigma}^*\mathcal{P}^* &= (-)^{2j}\eta_p^*\bar{a}_{-\mathbf{p},\sigma}^* \\ \mathcal{C}\bar{a}_{\mathbf{p},\sigma}^*\mathcal{C}^* &= \eta_c^*a_{\mathbf{p},\sigma}^*\end{aligned}\quad (8.18)$$

Notice also the additional sign that arises in the parity transformations for half-integer spin particles. This relation between the C and P phases for particles and antiparticles can be seen from the requirement that the fields transform properly under C and P. As was discussed in Section 1.2, causality requires that creation and annihilation operators only appear together in the schematic combination $\psi(x) \sim (a + \bar{a}^*)$. This is only consistent if a^* and \bar{a}^* transform as indicated under C and P.

Eq. (8.17) and Eq. (8.18) imply the following properties for the meson eigenfunctions, $|\Psi_{\ell m s s_z}\rangle$:

$$\begin{aligned}\mathcal{C}|\Psi_{\ell m s s_z}\rangle &= \sum_{\sigma,\sigma'} \int_{\mathbf{p},\mathbf{p}'} \psi_{\ell m s s_z}(\mathbf{p},\sigma;\mathbf{p}',\sigma') \mathcal{C}a_{\mathbf{p}\sigma}^*\bar{a}_{\mathbf{p}'\sigma'}^*|\mathcal{G}\rangle + \cdots \\ &= (\eta_c\eta_c^*) \sum_{\sigma,\sigma'} \int_{\mathbf{p},\mathbf{p}'} \psi_{\ell m s s_z}(\mathbf{p},\sigma;\mathbf{p}',\sigma') \bar{a}_{\mathbf{p}\sigma}^*a_{\mathbf{p}'\sigma'}^*\mathcal{C}|\mathcal{G}\rangle + \cdots\end{aligned}\quad (8.19)$$

Notice that the unknown phase, η_c , only enters through the combination $\eta_c\eta_c^*$, and so drops out of the final result.

The goal is to manipulate the right-hand side of this last equation in order to put it into the form of Eq. (8.14). To this end, we relabel $\mathbf{p} \leftrightarrow \mathbf{p}'$ and $\sigma \leftrightarrow \sigma'$ and interchange the order of $a_{\mathbf{p}\sigma}^*$ and $\bar{a}_{\mathbf{p}'\sigma'}^*$. This last interchange introduces an overall factor of $(-)^{2j}$ owing to the statistics of the quarks.

After these steps, the right-hand side would have the same form as in Eq. (8.14) – assuming the quark and antiquark are the same flavor – except that $\psi_{\ell m s s_z}$ appears with (\mathbf{p},σ) interchanged with (\mathbf{p}',σ') . Now, this wavefunction is either even or odd under the interchange of these arguments depending on the values of ℓ and s . In particular, states with orbital angular momentum, ℓ , acquire a sign $(-)^{\ell}$ when the three-momenta of the quark and antiquark are swapped. The singlet spin state, $s = 0$, is always antisymmetric when the quark spin is half-integer, $j = n + 1/2$, and symmetric when it is integer, $j = n$. The triplet state, $s = 1$, is similarly even under interchange of half-integer quark and antiquark spins, or is odd under interchange of integer-spins. Exchanging quark and antiquark spins within the spin wavefunction therefore introduces a further sign $(-)^{2j+s}$.

Table 8.2 C and P eigenvalues for quark–antiquark bound states

ℓ	s	J	C	$\mathcal{P}(j = \frac{1}{2})$	$\mathcal{P}(j = 1)$
0	0	0	+	–	+
0	1	1	–	–	+
1	0	1	–	+	–
1	1	0, 1, 2	+	+	–

Combining all of these signs, and using $\mathcal{C}|\mathcal{G}\rangle = |\mathcal{G}\rangle$, gives the following result:

$$\begin{aligned}
 \mathcal{C}|\Psi_{\ell m s s_z}\rangle &= (-)^{2j}(-)^\ell(-)^{2j+s} \sum_{\sigma, \sigma'} \int_{p, p'} \psi_{\ell m s s_z}(\mathbf{p}, \sigma; \mathbf{p}', \sigma') a_{\mathbf{p}\sigma}^* \bar{a}_{\mathbf{p}'\sigma'}^* |\mathcal{G}\rangle + \cdots \\
 &= (-)^{\ell+s} |\Psi_{\ell m s s_z}\rangle
 \end{aligned} \tag{8.20}$$

An identical argument may be made for parity:

$$\begin{aligned}
 \mathcal{P}|\Psi_{\ell m s s_z}\rangle &= \sum_{\sigma, \sigma'} \int_{p, p'} \psi_{\ell m s s_z}(\mathbf{p}, \sigma; \mathbf{p}', \sigma') \mathcal{P} a_{\mathbf{p}\sigma}^* \bar{a}_{\mathbf{p}'\sigma'}^* |\mathcal{G}\rangle + \cdots \\
 &= (-)^{2j} \eta_p \eta_p^* \sum_{\sigma, \sigma'} \int_{p, p'} \psi_{\ell m s s_z}(\mathbf{p}, \sigma; \mathbf{p}', \sigma') a_{-\mathbf{p}, \sigma}^* \bar{a}_{-\mathbf{p}', \sigma'}^* \mathcal{P} |\mathcal{G}\rangle + \cdots \\
 &= (-)^{2j} \sum_{\sigma, \sigma'} \int_{p, p'} \psi_{\ell m s s_z}(-\mathbf{p}, \sigma; -\mathbf{p}', \sigma') a_{\mathbf{p}, \sigma}^* \bar{a}_{\mathbf{p}', \sigma'}^* \mathcal{P} |\mathcal{G}\rangle + \cdots \\
 &= (-)^{2j+\ell} \sum_{\sigma, \sigma'} \int_{p, p'} \psi_{\ell m s s_z}(\mathbf{p}, \sigma; \mathbf{p}', \sigma') a_{\mathbf{p}\sigma}^* \bar{a}_{\mathbf{p}'\sigma'}^* |\mathcal{G}\rangle + \cdots \\
 &= (-)^{2j+\ell} |\Psi_{\ell m s s_z}\rangle
 \end{aligned} \tag{8.21}$$

Notice that the parity eigenvalue here depends on the spin, j , of the constituent quarks, and is completely independent of the unknown parity phase η_p .

These results have been summarized for some simple quantum numbers in Table 8.2. In this table J represents the total angular momentum of the bound state and j is the spin of the underlying quarks. The physical case $j = 1/2$ is contrasted with what would be predicted for hypothetical bound states of spin-one quarks.

Since the $\ell = 0$ states correspond to the ground states for the quark–antiquark pair, these should represent the observed stable mesons. (“Stable,” in this context, means stable in the absence of electromagnetic and weak interactions.) Excited states should be expected to be able to decay *via* the strong interactions (unless they are forbidden from doing so by a selection

rule) and so should be unstable resonances. We may therefore predict the properties of the stable mesons from the first two rows of Table 8.2.

This prediction agrees with the observed meson spectrum provided that the quarks carry spin half. There are two features of the spectrum that point to $j = 1/2$. The first is that stable mesons are either spinless or carry spin one, i.e. $J = 0$ or $J = 1$. This is precisely what would be expected for two spin-half constituents. By contrast, spinless quarks could not give $J = 1$ with $\ell = 0$, and spin-one quarks would also predict $\ell = 0$, $J = 2$ states. Spin-half quarks also precisely reproduce the correct P and C quantum numbers. The ground-state spinless mesons are parity-odd and C-even, and the stable spin-one mesons are similarly parity-odd but are also C-odd.

8.3 Light quarks

If the bound particles are very light compared to the confinement scale, $m_q \ll \Lambda_c$, then the bound state is not positronium-like at all. In this case the mass of the bound state is set by zero-point energy, $E \approx p \approx \Lambda_c$, that arises owing to the confinement of the constituents within the confinement radius, r_c . Since $m_q \ll E$, the constituent quarks within such states are highly relativistic, and so may be well approximated as being massless. This condition is appropriate for bound states that contain the up and down quarks. Strange quarks lie in the middle ground between this ultra-relativistic case and the non-relativistic system considered previously.

Weak coupling methods are not expected to give a good approximation within this regime, and so it is much more difficult to extract definite predictions. Much of what can be said in a model-independent way about the spectrum and the interactions for light-quark bound states therefore relies purely on symmetry arguments. The symmetries that are used, and their implications, are the topic of the present section.

8.3.1 Chiral symmetry

We are interested in the symmetry properties of the QCD Lagrangian, Eq. (8.2). Since only the lightest quarks – u , d , and possibly s – are of interest for the present purposes, we work within the low-energy effective theory defined at scales below the charmed-quark mass threshold, $\mu < 2m_c$. In this energy regime, the most important strong interactions are the usual ones, but only involving the lightest three quark types.

There is one main observation upon which all of the rest of the analysis is based. It is that the quarks involved are very light in comparison to the only other scale of the problem, Λ_c or Λ_{QCD} . This suggests the massless limit

as a fruitful first approximation. The effects of nonzero quark masses can then be incorporated perturbatively as small corrections to the massless – or *chiral* – limit.

We therefore split the QCD Lagrangian into a zeroth-order Lagrangian, \mathcal{L}_{ch} , and a mass term, \mathcal{L}_m :

$$\begin{aligned}\mathcal{L}_{\text{QCD}} &= \mathcal{L}_{\text{ch}} + \mathcal{L}_m \\ \text{with } \mathcal{L}_{\text{ch}} &= -\frac{1}{4}G_{\mu\nu}^\alpha G_\alpha^{\mu\nu} - \begin{pmatrix} \bar{u} & \bar{d} & \bar{s} \end{pmatrix} \begin{pmatrix} \not{D} & & \\ & \not{D} & \\ & & \not{D} \end{pmatrix} \begin{pmatrix} u \\ d \\ s \end{pmatrix} \\ \text{and } \mathcal{L}_m &= - \begin{pmatrix} \bar{u} & \bar{d} & \bar{s} \end{pmatrix} \begin{pmatrix} m_u & & \\ & m_d & \\ & & m_s \end{pmatrix} \begin{pmatrix} u \\ d \\ s \end{pmatrix} \quad (8.22)\end{aligned}$$

This is just the three-quark QCD Lagrangian with the quarks written out as an explicit column vector. As usual, a slash indicates contraction with the gamma matrices and D_μ is just the QCD covariant derivative, $D_\mu = \partial_\mu - ig_3 T_\alpha G_\mu^\alpha$, in which T_α are the $SU_c(3)$ generators that are given explicitly in Section 2.1. It is important to keep in mind that these 3×3 matrices act on the color quantum numbers of the quarks and do *not* act in the space spanned by the columns, u , d , and s , that are written out explicitly in Eq. (8.22).

To first approximation the quark mass matrix, M_q ,

$$M_q = \begin{pmatrix} m_u & & \\ & m_d & \\ & & m_s \end{pmatrix} \quad (8.23)$$

may be neglected. Within this approximation the bound states whose properties are of interest should reflect the symmetries of \mathcal{L}_{ch} . Any symmetries of \mathcal{L}_{ch} that are broken by the mass term, \mathcal{L}_m , will then be only approximate symmetries whose implications for the spectrum are corrected order-by-order in the quark masses.

8.3.1.1 Chiral $SU(3)$

The symmetries of \mathcal{L}_{ch} are easily determined following the results of Section 2.5. The kinetic term for n_f flavors of Dirac fermions is invariant under the symmetry group $U(2n_f)$. The group is $U(2n_f)$ rather than $U(n_f)$ because each Dirac fermion is made up of two Weyl (or Majorana) fermions.

In order to preserve the gauge couplings of Eq. (8.23), the candidate symmetry must also commute with the $SU_c(3)$ generators that appear within the covariant derivative. This implies that it cannot act at all on the color indices of the quarks and must act purely in flavor space. It also forbids transformations between triplets and anti-triplets of $SU_c(3)$.

The symmetry subgroup of $U(2n_f)$ that satisfies all of these conditions is the $U_L(n_f) \times U_R(n_f)$ which separately shuffles the left- and right-handed parts of the n_f quark flavors. For the $n_f = 3$ quarks of concrete interest, the action of these symmetry transformations on quark fields may be explicitly written out:

$$\delta \begin{pmatrix} u \\ d \\ s \end{pmatrix} = \left[\frac{i}{2} \omega_L^a \lambda_a P_L + \frac{i}{2} \omega_R^a \lambda_a P_R \right] \begin{pmatrix} u \\ d \\ s \end{pmatrix} \quad (8.24)$$

$P_L = \frac{1}{2}(1 + \gamma_5)$ projects onto the left-handed components of each of the quark fields, and $P_R = \frac{1}{2}(1 - \gamma_5)$ similarly projects onto the right-handed part. For a between 1 and 8, λ_a represents the eight traceless Gell-Mann matrices of Eq. (1.186). Here they are taken to act on the explicitly-written three-by-three column vectors in flavor space (rather than on color space). For $a = 0$, $\lambda_0 = \sqrt{2/3}I$, in which I is the 3×3 unit matrix. The normalization ensures that these matrices satisfy the property $\text{tr}(\lambda_a \lambda_b) = 2\delta_{ab}$. ω_L and ω_R are the nine spacetime independent group parameters for each of the factors of $U_L(3) \times U_R(3)$.

8.3.1.2 Anomalies

Of course, symmetries of the classical Lagrangian need not survive quantization. Before extracting the implications of the approximate symmetry group, $U_L(3) \times U_R(3)$, it is also necessary to check that it is anomaly-free. The criteria for anomaly freedom are spelled out in Subsection 2.5.3. Since we are neglecting all interactions here save for the strong interactions, it suffices to check only those anomalies that involve both $U_L(3) \times U_R(3)$ and $SU_c(3)$ generators. Anomalies owing to the other interactions are dealt with in subsequent sections.

In order to perform the anomaly cancellation analysis, it is convenient to use a different basis of generators of $U_L(3) \times U_R(3)$ than those that are given in Eq. (8.24). It is preferable in this case to define $\omega_L^a = \omega_V^a + \omega_A^a$ and

$\omega_R^a = \omega_V^a - \omega_A^a$. In this case, Eq. (8.24) becomes

$$\delta \begin{pmatrix} u \\ d \\ s \end{pmatrix} = \left[\frac{i}{2} \omega_V^a \lambda_a + \frac{i}{2} \omega_A^a \lambda_a \gamma_5 \right] \begin{pmatrix} u \\ d \\ s \end{pmatrix} \quad (8.25)$$

The vector combination, parameterized by ω_V^a , is independent of γ_5 , and so rotates left- and right-handed quarks in the same way. The axial generators, corresponding to ω_A^a , are explicitly proportional to γ_5 , and so act with opposite sign on left- and right-handed quarks.

Only the axial transformations are potentially anomalous. This is because the vector generators treat left- and right-handed fermions equally, and are therefore automatically anomaly-free by the general arguments of Subsection 2.5.3. Even for the axial generators, there is only one non-trivial condition to be satisfied. That is, since all of the generators of $SU_c(3)$ have zero trace, the only potential anomaly is the mixed anomaly which involves two $SU_c(3)$ generators and one $U_L(3) \times U_R(3)$ generator. This anomaly is proportional to the trace of the corresponding $U_L(3) \times U_R(3)$ generator, so the only symmetries with anomalies are those that are both axial and which have a non-vanishing trace.

The Gell-Mann matrices are explicitly traceless. This is because they generate an $SU(3)$ subgroup of $U(3)$ and so their exponential must have unit determinant. The only generator of $U(3)$ with a nonzero trace is $\lambda_0 = \sqrt{2/3}I$. Being proportional to the unit matrix, this commutes with all of the Gell-Mann matrices, and so generates a $U(1)$ subgroup of $U(3)$. Since λ_0 , together with the Gell-Mann matrices, form a complete basis of 3×3 Hermitian matrices, they span the Lie algebra of $U(3)$. Therefore, $U(3)$ is locally equivalent to $SU(3) \times U(1)$.

The upshot is that there is only a single generator of $U_L(3) \times U_R(3)$ that is anomalous and so does not survive as a symmetry of the spectrum of the quantum theory. This is the axial λ_0 generator: $U_A(1)$. Only the $SU_A(3)$ transformations in $U_A(3)$ are *bona fide* symmetries. Combining this with the vector part of the group, $U_V(3) = SU_V(3) \times U_V(1)$, the full anomaly-free symmetry group of \mathcal{L}_{ch} is therefore $G = SU_L(3) \times SU_R(3) \times U_V(1)$.

We will now explore the implications of this symmetry group for hadrons.

8.3.2 Exact symmetries: baryon number

As we shall see in the next section, although the group $G = SU_L(3) \times SU_R(3) \times U_V(1)$ is a symmetry of the strong interactions amongst massless quarks most of it is broken by the other interactions within the standard

model. It can therefore only be regarded as providing an *approximate* symmetry, to the extent that these other interactions only make small contributions to the physics of interest and so can be regarded as being perturbations to the dominant strong interactions. Historically, this pattern of approximate symmetries had been identified on phenomenological grounds, and it counts to the standard model's credit that it provides an explanation for these symmetries and the pattern of their breaking.

A special role is played by exact symmetries, and the only part of G which remains a symmetry when extended to the rest of the standard model is the $U_V(1)$ subgroup. Since this rotates all quarks by a common phase, it is equivalent to the $U_B(1)$ symmetry which is responsible for baryon-number conservation. For this reason in what follows we use the notations $U_V(1)$ and $U_B(1)$ interchangeably.

8.3.3 Chiral symmetry breaking effects

The symmetry group G is only an approximate symmetry of the full theory. Before applying its symmetries to the hadron spectrum, we pause to consider the size of the effects which can break these symmetries. This is necessary in order to estimate the accuracy of any given prediction.

There are two ways in which the symmetries of \mathcal{L}_{ch} can fail to be symmetries of the entire Lagrangian.

- (i) They can be explicitly broken by other terms in the Lagrangian. In decreasing order of importance, the most important symmetry-breaking terms are the mass terms, \mathcal{L}_m , the electromagnetic interactions, \mathcal{L}_{em} , and the weak interactions, \mathcal{L}_{wk} .
- (ii) They can have mixed anomalies with the electromagnetic or weak gauge interactions.

The accuracy of any particular symmetry prediction depends therefore on the strength of the terms in the Lagrangian which do not respect it, or on the strength of the interaction which is responsible for the anomaly. Since there are several interactions with varying strengths which can spoil any given symmetry, we expect to see a hierarchy of symmetries, each of which is more and more accurate. In order to explore the nature of this hierarchy, we next examine the transformation properties of the various interactions.

8.3.3.1 Mass terms

There are no nonzero fermion-mass terms which are invariant under axial transformations. The quark-mass terms therefore completely break the

$SU_A(3)$ elements of $U_L(3) \times U_R(3)$. It is noteworthy, though, that the up- and down-quark masses are both much smaller than the strange-quark mass. We therefore expect that the $SU_A(2)$ transformations that involve only the u and d quarks will be much less strongly broken than are the rest of the axial transformations.

The mass term also breaks most of the vector transformations. After performing a vector transformation, the change in the quark mass matrix is

$$\delta M_q = \frac{i}{2} \omega_V^a [M_q, \lambda_a] \quad (8.26)$$

Given that none of the quark masses are equal, the only $U_V(3)$ transformations for which this vanishes are those generated by λ_3 , λ_8 , and λ_0 . These correspond to three linear combinations of the generators of up number, down number, and strange number. These three symmetries are preserved by mass effects, but rotations of one quark type into another are not.

It is useful to identify which symmetries would survive in the approximation in which just the strange quark is massive. In order to do so, it is convenient to rewrite the mass matrix as a linear combination of the $U_V(3)$ generators, λ_a :

$$\begin{aligned} M_q &= \begin{pmatrix} m_u & & \\ & m_d & \\ & & m_s \end{pmatrix} \\ &= m_0 \sqrt{\frac{2}{3}} \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} + \frac{m_8}{\sqrt{3}} \begin{pmatrix} 1 & & \\ & 1 & \\ & & -2 \end{pmatrix} + m_3 \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix} \\ &= m_0 \lambda_0 + m_8 \lambda_8 + m_3 \lambda_3 \end{aligned} \quad (8.27)$$

The three parameters, m_0 , m_8 , and m_3 represent the following combinations of quark masses:

$$\begin{aligned} m_0 &= \frac{1}{\sqrt{6}}(m_u + m_d + m_s) \\ m_8 &= \frac{1}{2\sqrt{3}}(m_u + m_d - 2m_s) \\ m_3 &= \frac{1}{2}(m_u - m_d) \end{aligned} \quad (8.28)$$

The utility of this split is that each successive term in Eq. (8.27) preserves fewer and fewer of the symmetries. The term proportional to λ_0 does not break any of the vector symmetries at all, since it commutes with each of them. The λ_8 term commutes with an $SU_I(2) \times U_S(1)$ subgroup. The $SU_I(2)$

factor here is the *isospin* group, which shuffles only the u and d quarks. $U_s(1)$ is the symmetry under which only the strange quark gets rotated by a phase. The generator of this group is the *strangeness* charge, whose normalization is

$$S = \begin{pmatrix} 0 & & \\ & 0 & \\ & & -1 \end{pmatrix} \quad (8.29)$$

For historical reasons it is conventional to assign strangeness $S = -1$ to the strange quark, s . Finally, the last term breaks even this subgroup down to $U_u(1) \times U_d(1) \times U_s(1)$. The only $SU_V(3)$ symmetry-breaking effect of a nonzero strange-quark mass on the vector symmetries is therefore through the term proportional to λ_8 .

An estimate of the size of the corrections to the symmetry predictions that arise from the mass matrix would be m/Λ_c , in which m is the relevant quark mass combination from Eq. (8.28). Λ_c is an estimate of the matrix elements of the unperturbed Hamiltonian within the bound state of interest, which we take of the order of the bound state's mass. Since most of the bound states involving only light quarks have masses in the vicinity of 1 GeV, we expect that $SU_L(3) \times SU_R(3)$ predictions should be accurate to within roughly $m_s/\Lambda_c \approx 20\%$. The $SU_I(2) \times U_s(1)$ subgroup that survives even in the presence of a strange quark mass should be much more accurate, with corrections of order $m_d/\Lambda_c \approx 1\%$.

8.3.3.2 Electromagnetic interactions

Electromagnetic interactions can break flavor symmetries in either of two ways. They may be explicitly broken by the electromagnetic couplings in the Lagrangian, and there may be nonzero mixed electromagnetic anomalies. In either case, one might expect symmetry-breaking effects of the order $\alpha/(4\pi) \approx 0.1\%$.

The vector symmetries are broken only in the first of these ways. The subgroup that remains unbroken by electromagnetic effects is therefore that which commutes with the electric charge operator

$$Q = \begin{pmatrix} \frac{2}{3} & & \\ & -\frac{1}{3} & \\ & & -\frac{1}{3} \end{pmatrix} \quad (8.30)$$

The unbroken symmetry therefore corresponds to a phase rotation of the up quark, u , together with an $SU(2)$ transformation among the d and s quarks.

The axial symmetries, on the other hand, are not explicitly broken by the electromagnetic interactions. This is because the electromagnetic current, $\bar{\psi}\gamma^\mu\psi$, contains a γ^μ , and the γ_5 of the chiral transformation anticommutes across it. However, there may be mixed electromagnetic anomalies. There are potentially two such anomalies depending on whether one or two factors of the electromagnetic current are involved. The two anomaly coefficients are

$$\begin{aligned} A(T, Q, Q) &= \text{tr}(TQ^2) \\ \text{and} \quad A(T, T', Q) &= \frac{1}{2} \text{tr}(\{T, T'\}Q) \end{aligned} \quad (8.31)$$

in which T and T' represent two generators of $SU_A(3)$ and Q is the electromagnetic charge matrix of Eq. (8.30).

For example, in Subsection 9.3.4 we will consider a symmetry which, if unbroken, would unacceptably suppress π^0 decay; namely, the symmetry generated by $T = \frac{1}{2}\lambda_3$. For this generator these two anomaly coefficients become

$$\begin{aligned} A(T, Q, Q) &= 2 \text{tr} \begin{pmatrix} \frac{2}{9} & & \\ & -\frac{1}{18} & \\ & & 0 \end{pmatrix} = \frac{1}{3} \neq 0 \\ \text{and} \quad A(T, T, Q) &= 2 \text{tr} \begin{pmatrix} \frac{1}{6} & & \\ & -\frac{1}{12} & \\ & & 0 \end{pmatrix} = \frac{1}{6} \neq 0 \end{aligned} \quad (8.32)$$

We shall see that this non-vanishing value turns out to be essential to understanding the very short lifetime of the π^0 meson.

We now turn to the applications of these symmetries to the hadron spectrum.

8.3.4 Flavor $SU_f(3)$: the “eightfold way”

One of the most basic implications of a symmetry for the spectrum of any theory is the existence of degeneracies. The prediction is that any two particle states must have the same mass if they are related by a symmetry that leaves the ground state invariant – i.e. one that is not spontaneously broken. As a result, the particle spectrum should organize itself into multiplets of particles with equal masses which form representations of the symmetry group.

This section explores how the low energy hadrons are grouped into representations of the vector subgroup, $SU_V(3)$, of the chiral symmetry group G .

This subgroup is often called *flavor* $SU_f(3)$ (rather than $SU_V(3)$) in much the same way as the strong gauge group is known as the color $SU_c(3)$.

If the symmetry is only approximate rather than exact, as is the case here, then the particles within any such multiplet will have mass splittings that are governed by the strength of the non-invariant interactions. From the estimates of the previous section we see that we therefore expect to find particles in $SU_f(3)$ representations whose masses are split by roughly 20%. Each of these $SU_f(3)$ multiplets may themselves be broken up into $SU_I(2)$ representations whose elements should be degenerate to within roughly 1%.

The experimental fact that the hadron spectrum does indeed seem to be so organized gives us evidence that the QCD ground state does not spontaneously break $SU_f(3)$.

8.3.4.1 Mesons

Consider first the $q\bar{q}$ bound states. From Table 8.2, we know that with spin-half quarks the stable mesons should have the quantum numbers $J^{PC} = 0^{-+}$ and $J^{PC} = 1^{--}$. Each of these should come in any of nine flavors since each quark type can be either u , d , or s .

For the purposes of displaying the $SU_f(3)$ transformation properties, it is convenient to write these nine flavor states as a 3×3 Hermitian matrix:

$$\begin{aligned} \mathcal{M} &= \begin{pmatrix} u \\ d \\ s \end{pmatrix} \begin{pmatrix} \bar{u} & \bar{d} & \bar{s} \end{pmatrix} \\ &= \begin{pmatrix} u\bar{u} & u\bar{d} & u\bar{s} \\ d\bar{u} & d\bar{d} & d\bar{s} \\ s\bar{u} & s\bar{d} & s\bar{s} \end{pmatrix} \end{aligned} \quad (8.33)$$

We would like to determine how these quark pairs transform under $SU_f(3)$. From the quark transformation rules, Eq. (8.25), the matrix \mathcal{M} inherits the following transformation property: $\delta\mathcal{M} = \frac{i}{2}\omega_V^a[\lambda_a, \mathcal{M}]$. The combinations which form irreducible representations of the flavor group may be identified by expanding \mathcal{M} in terms of the generators λ_a , $a = 0, \dots, 8$: $\mathcal{M} = \sum_a \mu_a \lambda_a$. Clearly, since λ_0 is proportional to the unit matrix, it is a flavor-invariant state, or singlet. Also, since the eight remaining Gell-Mann matrices satisfy the commutation relations of $SU_f(3)$, the eight combinations, μ_a , $a = 1, \dots, 8$, form an eight-dimensional representation – a flavor octet. This decomposition of \mathcal{M} into irreducible representations may be summarized by the symbolic equation: $\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{8} \oplus \mathbf{1}$.

The nine pseudoscalar ($J^{PC} = 0^{-+}$) quark–antiquark states having definite $SU_f(3)$ transformation properties, μ_a , carry the following names:

$$\begin{aligned}
 \mathcal{M} &= \sum_{a=1}^9 \mu_a \lambda_a \\
 &= \begin{pmatrix} \mu_3 + \frac{1}{\sqrt{3}}\mu_8 + \sqrt{\frac{2}{3}}\mu_0 & \mu_1 - i\mu_2 & \mu_4 - i\mu_5 \\ \mu_1 + i\mu_2 & -\mu_3 + \frac{1}{\sqrt{3}}\mu_8 + \sqrt{\frac{2}{3}}\mu_0 & \mu_6 - i\mu_7 \\ \mu_4 + i\mu_5 & \mu_6 + i\mu_7 & -\frac{2}{\sqrt{3}}\mu_8 + \sqrt{\frac{2}{3}}\mu_0 \end{pmatrix} \\
 &= \begin{pmatrix} \pi^0 + \frac{1}{\sqrt{3}}\eta_8 + \sqrt{\frac{2}{3}}\eta_0 & \sqrt{2}\pi^+ & \sqrt{2}K^+ \\ \sqrt{2}\pi^- & -\pi^0 + \frac{1}{\sqrt{3}}\eta_8 + \sqrt{\frac{2}{3}}\eta_0 & \sqrt{2}K^0 \\ \sqrt{2}K^- & \sqrt{2}\bar{K}^0 & -\frac{2}{\sqrt{3}}\eta_8 + \sqrt{\frac{2}{3}}\eta_0 \end{pmatrix}
 \end{aligned} \tag{8.34}$$

Comparing these flavor eigenstates with the quark combinations shown in Eq. (8.33) allows us to solve for the valence-quark content of each particle type:

$$\begin{aligned}
 \pi^0 &= \mu_3 \sim \frac{1}{2}(u\bar{u} - d\bar{d}) \\
 \pi^+ &= \frac{1}{\sqrt{2}}(\mu_1 - i\mu_2) \sim u\bar{d} \\
 \pi^- &= \frac{1}{\sqrt{2}}(\mu_1 + i\mu_2) \sim d\bar{u} \\
 K^+ &= \frac{1}{\sqrt{2}}(\mu_4 - i\mu_5) \sim u\bar{s} \\
 K^- &= \frac{1}{\sqrt{2}}(\mu_4 + i\mu_5) \sim s\bar{u} \\
 K^0 &= \frac{1}{\sqrt{2}}(\mu_6 - i\mu_7) \sim d\bar{s} \\
 \bar{K}^0 &= \frac{1}{\sqrt{2}}(\mu_6 + i\mu_7) \sim s\bar{d} \\
 \eta_8 &= \mu_8 = \frac{1}{2\sqrt{3}}(u\bar{u} + d\bar{d} - 2s\bar{s}) \\
 \eta_0 &= \mu_0 = \frac{1}{\sqrt{6}}(u\bar{u} + d\bar{d} + s\bar{s})
 \end{aligned} \tag{8.35}$$

As may be seen from their quark content, the superscripts 0, + and – represent the electric charges of the corresponding states.

The pseudovector states, with $J^{PC} = 1^{--}$, are similarly labeled:

$$\tilde{\mathcal{M}} = \begin{pmatrix} \rho^0 + \frac{1}{\sqrt{3}}\omega_8 + \sqrt{\frac{2}{3}}\omega_0 & \sqrt{2}\rho^+ & \sqrt{2}K^{*+} \\ \sqrt{2}\rho^- & -\rho^0 + \frac{1}{\sqrt{3}}\omega_8 + \sqrt{\frac{2}{3}}\omega_0 & \sqrt{2}K^{0*} \\ \sqrt{2}K^{-*} & \sqrt{2}\bar{K}^{0*} & -\frac{2}{\sqrt{3}}\omega_8 + \sqrt{\frac{2}{3}}\omega_0 \end{pmatrix} \quad (8.36)$$

The valence-quark content of each of these states is identical to that of the corresponding pseudoscalar meson.

The pseudoscalar η_0 and the pseudovector ω_0 are the $SU_f(3)$ singlet states. The pseudoscalar octet consists of $\pi^0, \pi^\pm, K^0, \bar{K}^0, K^\pm$, and η_8 . The 1^{--} octet is $\rho^0, \rho^\pm, K^{0*}, \bar{K}^{0*}, K^{\pm*}$, and ω_8 . In the limit that $SU_f(3)$ is unbroken, all of the mesons within an octet would be degenerate in mass.

Since $(m_d - m_u) \ll (m_s - m_d)$, invariance under $SU_I(2)$ is a much better approximation than $SU_f(3)$. It is therefore useful to know how the $SU_f(3)$ octets break up into isospin representations. Since the isospin generators are given by λ_a , with $a = 1, 2, 3$, the combinations (π^0, π^\pm) and (ρ^0, ρ^\pm) form triplets of $SU_I(2)$. Since the corresponding generators also commute with the strangeness generator, S , these two triplet representations each have $S = 0$. Each octet also contains two isodoublets: (K^+, K^0) and (K^{*+}, K^{0*}) with $S = +1$, and (K^-, \bar{K}^0) and (K^{*-}, \bar{K}^{0*}) with $S = -1$. Finally, both η_8 and ω_8 are isospin singlets having $S = 0$. Each of these isospin multiplets should be degenerate to within much greater accuracy than the octets as a whole.

We return to the size of the expected mass splittings within a flavor multiplet in a subsequent section.

8.3.4.2 Baryons

The transformation properties of the baryons may be found in a similar way. Since the baryons contain three valence quarks, the counting of the states proceeds slightly differently. The typical valence contribution to the baryon state is

$$|\Psi\rangle = \sum_{\sigma_1, \sigma_2, \sigma_3} \sum_{\lambda_1, \lambda_2, \lambda_3} \int \prod_i \frac{d^3 p_i}{2p_i^0 (2\pi)^3} \psi(\mathbf{p}_1, \sigma_1; \dots, \mathbf{p}_3, \sigma_3) a_1^* a_2^* a_3^* |\mathcal{G}\rangle + \dots \quad (8.37)$$

λ_i , σ_i , and \mathbf{p}_i here represent the flavor, spin, and momentum quantum numbers of quark type “ i .”

Since the ground state is expected to have vanishing relative angular momentum for all of the quarks, the spin of the stable baryons should be completely determined by the spin of the quarks they contain. Since each

quark carries spin-half, the potential baryon spin states are determined by the rules for the addition of angular momenta:

$$\frac{1}{2} \otimes \frac{1}{2} \otimes \frac{1}{2} = \frac{1}{2} \oplus \frac{1}{2} \oplus \frac{3}{2} \quad (8.38)$$

There are also potentially $3^3 = 27$ flavor combinations that may be taken by the three quarks in the baryon. Together with the eight possible spin states, this would superficially allow $8 \times 27 = 216$ possible combinations of quantum numbers. Not all of these combinations are actually realized in the spectrum, however, because statistics requires that the entire wavefunction must be completely antisymmetric under the interchange of any two quarks. (No such restriction exists for the mesons because a quark and an antiquark are distinguishable.) Since the color part of the wavefunction is proportional to the completely antisymmetric invariant tensor, ϵ^{rst} , of $SU_c(3)$ – c.f. Eq. (8.13) – it is by itself already completely antisymmetric. This means that the rest of the wavefunction must be completely symmetric. Given no relative angular momentum between quarks in the ground state, the wavefunction is automatically symmetric under the interchange of any two of the quark three-momenta, \mathbf{p}_i . This implies that the flavor and spin parts of the wavefunction must also be completely symmetric.

There are a total of 56 states that satisfy these symmetry conditions. The completely symmetric combination of all three of the quark spins forms a $j = 3/2$ spin state. This must transform as a **10** of $SU_f(3)$ since this is the completely symmetric combination of three triplets. The only other allowed combination is a spin-half $SU_f(3)$ -octet of baryons which has mixed symmetry under the interchange of only spins or only flavors. We would therefore predict a decuplet of spin-3/2 baryons together with an octet of spin-half baryons, which agrees with the observed multiplets of stable baryons. (This adds up to 56 states because there are two allowed spin states for each spin-half baryon and four allowed spin states for each spin-3/2 baryon.)

Notice that this agreement between predictions for the quantum numbers and the observed spectrum relies crucially on the color degree of freedom. It also relies on the requirement of the color neutrality of the bound state, since this is what implies that the state is completely antisymmetric with respect to the interchange of any of its color indices. Historically it was the need to reconcile the observed quantum numbers with the required antisymmetry of the wave function which first led to the hypothesis that quarks carry color.

The baryon octet can be written in the same matrix form as were the meson octets, \mathcal{M} and $\tilde{\mathcal{M}}$. Besides exhibiting the $SU_f(3)$ transformation laws explicitly, this allows the determination of the valence-quark content

of any particular octet baryon. The principal observation that is necessary is that an antisymmetric combination of two quarks must transform just as does an antiquark under $SU_f(3)$. This is because the two may be related using the completely antisymmetric invariant tensor of $SU(3)$: $\bar{q}^i = \epsilon^{ijk} q_j q_k$. For example, $[u, d] = ud - du$ transforms under $SU_f(3)$ in precisely the same way as does \bar{s} .

This quark content may be found by comparing the octet quark matrix,

$$\begin{pmatrix} u[d, s] & u[s, u] & u[u, d] \\ d[d, s] & d[s, u] & d[u, d] \\ s[d, s] & s[s, u] & s[u, d] \end{pmatrix} \quad (8.39)$$

with the λ_a basis of baryons

$$\mathcal{B} = \begin{pmatrix} \Sigma^0 + \frac{1}{\sqrt{3}}\Lambda_8 & \Sigma^+ & p \\ \Sigma^- & -\Sigma^0 + \frac{1}{\sqrt{3}}\Lambda_8 & n \\ \Xi^- & \Xi^0 & -\frac{2}{\sqrt{3}}\Lambda_8 \end{pmatrix} \quad (8.40)$$

The $SU_f(3)$ transformation rules for these states are $\delta\mathcal{B} = \frac{i}{2}\omega_V^a[\lambda_a, \mathcal{B}]$. The antibaryons similarly lie in a *distinct* octet, $\bar{\mathcal{B}}$.

The valence-quark content of these octet baryons therefore is

$$\Sigma^0, \Lambda_8 \sim dus, \quad \Sigma^+ \sim uus, \quad \Sigma^- \sim dds \quad (8.41)$$

$$\Xi^0 \sim uss, \quad \Xi^- \sim dss \quad (8.42)$$

$$\text{and: } p \sim uud, \quad n \sim udd \quad (8.43)$$

The $SU_I(2) \times U_S(1)$ quantum numbers of these states may be seen by comparing with the meson octet and/or by examining the valence quark content. We see that there is an $S = -1$ isotriplet of baryons, (Σ^0 and Σ^\pm); an $S = 0$ isodoublet of nucleons – i.e. the proton and neutron – (n and p); an $S = -2$ isodoublet of *cascade* particles, (Ξ^0 and Ξ^-); and an $S = -1$ isoscalar baryon Λ_8 .

The quantum numbers of the decuplet of spin-3/2 baryons is similarly given by an isospin-3/2, $S = 0$ multiplet,

$$\Delta^- \sim ddd, \quad \Delta^0 \sim ddu, \quad \Delta^+ \sim duu, \quad \Delta^{++} \sim uuu \quad (8.44)$$

an isotriplet $S = -1$ multiplet,

$$\Sigma^{*-} \sim dds, \quad \Sigma^{0*} \sim dus, \quad \Sigma^{+*} \sim uus \quad (8.45)$$

an $S = -2$ isodoublet,

$$\Xi^{*-} \sim dss, \quad \Xi^{0*} \sim uss \quad (8.46)$$

and an $S = -3$ isosinglet,

$$\Omega^- \sim sss \quad (8.47)$$

These ten baryon flavors would be degenerate in the limit of $SU_f(3)$ invariance, but have this degeneracy lifted by correction terms which we explore later. Inspection of the *Particle Data Book* shows that the isospin multiplets that lie within the baryon octet and decuplet are indeed split from one another in mass by roughly 25%, as our order-of-magnitude estimate would indicate. The flavor $SU(3)$ was discovered phenomenologically, carrying the name the *eightfold way*, well before the unraveling of the theory of the strong interactions. Splitting within the isomultiplets themselves are much smaller, and are of the order of the 1% effects that we expect.

8.3.4.3 Glueballs, exotics

So far we have only considered hadrons which can be considered to be composed of a $q\bar{q}$ pair or three quarks. There are other ways of combining quarks and gluons into colorless combinations, which we now briefly address.

Two gluons can combine together in a singlet color combination, forming an object rather prosaically named a *glueball*. Similarly, a gluon can bind a quark–antiquark pair in a color-octet combination, to form an exotic meson (or three quarks and a gluon can form an exotic baryon). However, the masses of such objects are expected to be quite large – for instance, lattice simulations attribute a mass to the lightest glueball, the $J^{PC} = 0^{++}$ glueball, of around 1.6 GeV. Since there are combinations of mesons with the same quantum numbers and much smaller combined masses than any glueballs or exotics, these bound states are expected to be very short lived, which makes their clean experimental identification very difficult. Although strongly interacting states have been observed which might be interpreted as being glueballs, at this time of writing (2006) experimental difficulties have precluded a definitive glueball interpretation for any of them.

Similarly, it is possible to construct colorless combinations of four quarks and an antiquark, called a *pentaquark*, or of six quarks, called a *hexaquark*. There are some theoretical arguments that a $SU_V(3)$ singlet, $J^{PC} = 0^{++}$ hexaquark (containing two up quarks, two down quarks, and two strange quarks) should be strongly bound, but there has never been compelling evidence for such a combination. Similarly, pentaquarks are expected to be heavy and very short-lived, with one possible exception.

A pentaquark would consist of four quarks and an antiquark. The most attractive color channel between two quarks is for their colors to be antisymmetrized, $\epsilon^{rst}q_sq_t$, forming a $\bar{\mathbf{3}}$. If the four quarks combine in this way,

they form two diquark $\bar{\mathbf{3}}_s$, which together with the $\bar{\mathbf{3}}$ of the antiquark can form a color-singlet combination. If the diquark flavors also combine into $\bar{\mathbf{3}}_s$ of $SU_V(3)$, then the problem of combining these is somewhat similar to the case of baryons. The lightest combination is theoretically expected to be the $\bar{\mathbf{10}}$ flavor combination. The isospin singlet state is $S = +1$, consisting of $uudd\bar{s}$, and is called the Θ^+ . Opposite to the case of the baryon decuplet, the other states have a higher average s content and are heavier. Because of its positive total strangeness, the Θ^+ cannot decay into a $\pi^+\Lambda$, but must decay to K^0p or K^+n . The Θ^+ may be only slightly heavier than this combination, which would phase-space restrict the decay, leading to a narrow state.

At this time (2006), the experimental situation is unclear, with some evidence for a state of mass 1540 MeV and a width of at most a few MeV, decaying through this channel, but with other experiments failing to find such a state. If this pentaquark exists, it is the sole narrow QCD bound state not understood as a meson or a baryon. We will not address this state further.

8.3.5 Spontaneous chiral symmetry breaking

Up until this point we have only discussed the experimental implications of the vector subgroup, $SU_f(3)$, of the full chiral symmetry group, $SU_L(3) \times SU_R(3)$. The axial symmetries have consequences that are just as successful as are those of the vector symmetries. These are explored in the present section.

We have already discussed, both in Section 1.4 and in the treatment of the vector symmetries, that the transformation rules of particle states depend crucially on the transformation properties of the ground state. If the ground state is invariant under the symmetry, then the spectrum will fall into multiplets of degenerate (or approximately degenerate) particles that form representations of the symmetry group. This was what we found in the preceding pages with the vector symmetry, $SU_f(3)$, and what we found even earlier for Lorentz transformations themselves. If, on the other hand, the ground state is not a singlet, so the symmetry is spontaneously broken, then particles that are related by symmetry transformations need not have equal masses.

In principle, the transformation properties of the ground state may be determined by directly constructing the lowest-energy eigenstate of the Hamiltonian and applying a symmetry transformation to it. Indeed, although this cannot yet be done for QCD, there are strong theoretical arguments that

the QCD ground state must spontaneously break the chiral symmetry group, $SU_L(3) \times SU_R(3) \times U_B(1)$, down to the vector subgroup, $SU_f(3) \times U_B(1)$.

In the absence of a direct demonstration of this symmetry-breaking pattern, we must in practice infer the ground-state symmetries by comparing the observed spectrum with the symmetry predictions. Since the QCD spectrum is consistent with the approximately degenerate multiplets of the vector subgroup, $SU_f(3) \times U_B(1)$, we infer that if QCD is the correct underlying theory then this vector subgroup cannot be spontaneously broken.

Using the same logic, it is clear that the axial symmetry transformations, $SU_A(3)$, have to be spontaneously broken for consistency with the observed spectrum. This is because, were these symmetry transformations not broken, particles would have to fall into multiplets under both $SU_L(3)$ and $SU_R(3)$ transformations. This is inconsistent with the fact that the number and quantum numbers of the observed stable mesons precisely agree with the particle content that is predicted by the $SU_f(3)$ multiplets that are outlined above.

To see this inconsistency in more detail, notice that quarks and antiquarks respectively transform under $SU_L(3) \times SU_R(3)$ as a $(\mathbf{3}, \mathbf{1}) \oplus (\mathbf{1}, \mathbf{3})$ and a $(\bar{\mathbf{3}}, \mathbf{1}) \oplus (\mathbf{1}, \bar{\mathbf{3}})$ representation. In this notation, the pair of numbers gives the representation of each of the group factors, $SU_L(3)$ and $SU_R(3)$, and each term in the sum represents the right- or left-handed part of the fermion field. If $SU_L(3) \times SU_R(3)$ were not spontaneously broken, then the mesons would therefore have to transform as the product of these representations:

$$[(\mathbf{3}, \mathbf{1}) \oplus (\mathbf{1}, \mathbf{3})] \otimes [(\bar{\mathbf{3}}, \mathbf{1}) \oplus (\mathbf{1}, \bar{\mathbf{3}})] = (\mathbf{8}, \mathbf{1}) \oplus (\mathbf{1}, \mathbf{8}) \oplus (\bar{\mathbf{3}}, \mathbf{3}) \oplus (\mathbf{3}, \bar{\mathbf{3}}) \oplus 2(\mathbf{1}, \mathbf{1}).$$

This represents many more flavor states than the nine that are observed.

A spontaneously-broken (global) symmetry does, however, carry implications for the spectrum of the theory. The most important of these is the existence of a set of massless particles, called *Goldstone particles*. If the symmetry is only an approximate symmetry then these particles are not exactly massless, since the effects that explicitly break the symmetry can generate masses for them. The size of the mass grows with the strength of the explicit symmetry breaking. A not-quite-massless would-be Goldstone particle for an approximate symmetry is often called a *pseudo-Goldstone particle*. In order to see what the predictions would be in the particular case of QCD, it is necessary to make an aside about the properties of these Goldstone particles.

8.3.5.1 Goldstone's theorem

Suppose we have a Lagrangian which is invariant under a global, continuous symmetry whose conserved current is j_μ . The corresponding conserved charge is $Q = \int d^3x j^0$. If we perform the symmetry transformation with a spacetime dependent, infinitesimal parameter $\omega(x)$, the states in the Hilbert space are transformed by the operator $R[\omega] = i \int d^3x \omega(x) j^0(x)$. That is, the change in a state, $|\psi\rangle$, is $\delta|\psi\rangle = R[\omega]|\psi\rangle$, and the change in a field operator is $\delta F(x) = [R(\omega), F(x)]$. Notice that R reduces to $i\omega_0 Q$ in the special case of a constant parameter $\omega = \omega_0$. We introduce here the artifice of a space-dependent symmetry parameter even though the transformation generated by R is only really a symmetry for constant ω in order to sidestep some technical issues, and intend a limit towards constant ω at the end of the analysis.

By definition a symmetry is spontaneously broken if the ground state, $|\Omega\rangle$, is not invariant: $\delta|\Omega\rangle = R[\omega]|\Omega\rangle \neq 0$; but in practice this is not so useful a criterion when working with field theories. For this reason it is usually more convenient to work with an *order parameter*, $\delta F(x)$, which is a field having the following two properties: (i) it is assumed to be the variation of some quantity, $\delta F = [R, F]$, under the symmetry transformation; and (ii) its expectation value is not zero in the ground state: $\langle\Omega|\delta F(x)|\Omega\rangle \neq 0$. The existence of such an order parameter is useful because it is a sufficient condition for the spontaneous breaking of a symmetry, as may be seen because $R[\omega]|\Omega\rangle = 0$ implies $\langle\Omega|\delta F|\Omega\rangle = 0$.

The existence of such an order parameter also has other important implications; most notably it implies the existence of a special ‘Goldstone’ state, $|G\rangle$, whose properties we now explore. To this end suppose that a complete set of four-momentum eigenstates is inserted into the matrix element $\langle\Omega|[R, F]|\Omega\rangle \neq 0$. Since we know the result is nonzero we see that there must exist an energy eigenstate, $|G(p, n)\rangle$, for which the matrix element $\langle G(p, n)|R[\omega]|\Omega\rangle \neq 0$ for ω sufficiently slowly varying. The argument, p , here denotes the four-momentum eigenvalue of this state, $P^\mu|G(p, n)\rangle = p^\mu|G(p, n)\rangle$. The other index, “ n ,” represents all of the other labels of this state.

An equivalent way to say the same thing is that there must be an energy eigenstate, $|G(p, n)\rangle$, for which the matrix element of the conserved current.

$$\begin{aligned} \mathcal{A}^\mu(p; x) &\equiv \langle G(p, n)|j^\mu(x)|\Omega\rangle \\ &= \langle G(p, n)|e^{iPx}j^\mu(0)e^{-iPx}|\Omega\rangle \end{aligned}$$

$$\begin{aligned}
&= e^{ipx} \langle G(p, n) | j^\mu(0) | \Omega \rangle \\
&= f_n p^\mu e^{ipx} \\
&\neq 0
\end{aligned} \tag{8.48}$$

We assume here that the symmetry generator, R , commutes with Poincaré transformations, as is true for any internal symmetry. This, together with the Lorentz invariance of the ground state, implies the state $|G\rangle$ is spinless. The spinlessness of $|G\rangle$ and Poincaré invariance is also used in the last equality to make the dependence of the matrix elements on p and x explicit. Lorentz invariance further implies that the unknown scalar quantity, $f_n \neq 0$, is itself invariant.

The state $|G(p, n)\rangle$ defined in this way is the Goldstone state, and all of its properties follow from this non-vanishing matrix element. We list the most useful ones here.

- (i) Current-conservation, $\partial_\mu j^\mu = 0$, implies that the state, $|G(p, n)\rangle$, is massless, $p^2 = p^\mu p_\mu = 0$, because

$$\begin{aligned}
p^2 f_n e^{ipx} &= -i \partial_\mu \mathcal{A}^\mu(p; x) \\
&= \langle G(p, n) | \partial_\mu j^\mu(x) | \Omega \rangle \\
&= 0
\end{aligned} \tag{8.49}$$

- (ii) As has already been mentioned, given that $\langle G(p, n) | j^0(x) | \Omega \rangle \neq 0$ and j^0 is a rotational scalar, it is possible to show that $|G(p, n)\rangle$ must have spin zero.
- (iii) In a parity-invariant theory $|G(p, n)\rangle$ must be a pseudoscalar if j^μ is an axial vector, as would be required for the axial symmetries of QCD. To see this we use the following transformation properties: $\mathcal{P}|G(p, n)\rangle = \eta|G(p_P, n)\rangle$, $\mathcal{P}|\Omega\rangle = |\Omega\rangle$ and $\mathcal{P}j^\mu(x)\mathcal{P}^* = -P_\nu^\mu j^\nu(x_P)$. The subscript “ P ” denotes the parity-transformed vector: e.g. $x_P^\mu = P_\nu^\mu x^\nu$. The matrix P_ν^μ is as defined in Section (2.5), and reflects the spatial components of any vector on which it acts. The following argument shows that the parity phase, η , must be $\eta = -1$:

$$\begin{aligned}
f_n p^\mu e^{ipx} &= \langle G(p, n) | j^\mu(x) | \Omega \rangle \\
&= \langle G(p, n) | \mathcal{P}^* \mathcal{P} j^\mu(x) \mathcal{P}^* \mathcal{P} | \Omega \rangle \\
&= -\eta^* P_\nu^\mu \langle G(p_P, n) | j^\nu(x_P) | \Omega \rangle
\end{aligned}$$

$$\begin{aligned}
&= -\eta^* f P_\nu^\mu p_P^\nu e^{ip_P x_P} \\
&= -\eta^* f p^\mu e^{ipx}
\end{aligned} \tag{8.50}$$

These Goldstone boson properties hold as exact statements if the spontaneously broken symmetry in question is an exact symmetry of the system. However we know that the symmetry of present interest, $SU_A(3)$, is only an approximate symmetry in QCD, which is broken by terms like quark masses and by electromagnetic interactions in the standard-model action. We should therefore expect that if the $SU_A(3)$ symmetry is spontaneously broken then there are *pseudo-Goldstone* bosons for which the above Goldstone properties are only approximately true. In particular, although we do not expect these states to be precisely massless we do expect them to be systematically lighter than the system's generic particle states.

8.3.5.2 Pions as pseudo-Goldstone bosons

If consistency with the QCD spectrum is only possible if the approximate $SU_A(3)$ transformations are spontaneously broken, there must be a set of eight pseudo-Goldstone bosons in the spectrum. Using arguments like those that were just given, it is possible to infer the quantum numbers of these Goldstone modes from the properties of the corresponding spontaneously broken conserved current.

The current for the $SU_A(3)$ transformations may be constructed from the Lagrangian, \mathcal{L}_{ch} , of Eq. (8.23), and the transformation law, Eq. (8.25), using the formalism of Subsection 1.4.2. Introducing the notation $\mathcal{Q} = (u \ d \ s)^T$, the result is

$$(j_A)^\mu_a = \frac{i}{2} \bar{\mathcal{Q}} \gamma^\mu \gamma_5 \lambda_a \mathcal{Q} \tag{8.51}$$

It is clear from the previous section that the properties of this current imply that the Goldstone bosons must be pseudoscalar particles. Identical arguments also imply that the eight Goldstone modes must transform as an octet under $SU_f(3)$, and those that are charge-conjugation eigenstates must have $C = +1$.

The prediction is that there should be an $SU_f(3)$ octet of $J^{PC} = 0^{-+}$ particles which should be significantly lighter than are the typical hadrons that just involve light quarks. This should be particularly true for the $SU_I(2)$ triplet of mesons within this octet, which are the pseudo-Goldstone bosons for the much less strongly broken axial symmetry, $SU_A(2)$, which just involves the u and d quarks.

Such a multiplet certainly is seen in the spectrum. It is the pseudoscalar octet of Eq. (8.34). Indeed, the very small mass of the lightest members of

this multiplet – the pions π^0 and π^\pm – would otherwise have been a real puzzle from the general qualitative estimates of Section 8.1. We develop a more quantitative picture of their masses, and of the other hadrons, within the next section.

8.3.6 Mass relations

The next step is to quantify the predictions for the mass splittings within the various hadron multiplets in order to see more precisely whether they agree with what is observed. We do so by perturbing in the terms of the Lagrangian that explicitly break $SU_L(3) \times SU_R(3)$. The simplest way to do so is to work within the effective low-energy Lagrangian which is directly expressed in terms of the low-energy degrees of freedom – the mesons and baryons. This effective Lagrangian could be derived, in principle, by computing the properties of these mesons and baryons from the underlying dynamics of QCD. Since this is at present too difficult a task, we must content ourselves with writing down the most general low-energy Lagrangian which is consistent with the symmetries of the underlying theory. Any parameters in the low-energy theory that cannot be determined purely by symmetry arguments must be left as unknown parameters whose values are to be determined by experiment.

8.3.6.1 The meson octet

Since the lightest hadrons are the would-be Goldstone bosons of the pseudoscalar meson octet, we start with the effective action at energies that are sufficiently low as to be below the threshold for producing any other particles. In the chiral limit, that is, for vanishing quark masses and electromagnetic interactions, the resulting Lagrangian must be a function of only the meson fields, \mathcal{M} , of Eq. (8.34) and must be invariant under the chiral $SU_L(3) \times SU_R(3)$ group.

The action of the symmetry group, $SU_L(3) \times SU_R(3)$, on \mathcal{M} is completely determined by the symmetry-breaking pattern $SU_L(3) \times SU_R(3) \rightarrow SU_f(3)$. It is most easily formulated in terms of the variable ξ , which is defined as the exponential of \mathcal{M} :

$$\xi = \exp \left[\frac{i\mathcal{M}}{f} \right] \quad (8.52)$$

f here is a constant with the dimensions of mass which can be expected to be of order the QCD scale, Λ_{QCD} , in size. If L and R are 3×3 unitary matrices which respectively live in the groups $SU_L(3)$ and $SU_R(3)$, then it

can be shown that ξ transforms under $SU_L(3) \times SU_R(3)$ in the same way as does the following combination of quark fields: $Q_L \bar{Q}_R$, or

$$\xi \rightarrow L \xi R^\dagger \quad (8.53)$$

For the unbroken diagonal (i.e. vector) subgroup of $SU_L(3) \times SU_R(3)$ L and R are related by $L = R \equiv U$. In this case, Eq. (8.53) reduces to a similarity transformation,

$$\xi \rightarrow U \xi U^\dagger \quad (8.54)$$

which, when restricted to infinitesimal transformations, becomes

$$\delta \mathcal{M} = \frac{i}{2} \omega_V^a [\lambda_a, \mathcal{M}] \quad (8.55)$$

in agreement with the rule given in the paragraph immediately following Eq. (8.33).

The transformation law, Eq. (8.53), is more complicated for axial transformations, for which $L = R^\dagger \equiv V$. For infinitesimal transformations the axial transformation rule becomes

$$\delta \xi = \frac{i}{2} \omega_A^a \{\lambda_a, \xi\} \quad (8.56)$$

Unlike the transformation law, Eq. (8.55), this last result is neither linear nor homogeneous in the field \mathcal{M} . Indeed, expanding in powers of \mathcal{M} gives

$$\delta \mathcal{M} = f \omega_A^a \lambda_a + \frac{i}{2} \omega_A^a \{\lambda_a, \mathcal{M}\} + O(\mathcal{M}^2) \quad (8.57)$$

It is significant that the first term in the variation of \mathcal{M} should be independent of \mathcal{M} , and so acts to shift \mathcal{M} rather than multiply it by a matrix. This is the classic transformation rule for a Goldstone boson since it shows that there is no preferred value of the field \mathcal{M} corresponding to the vacuum – instead all such values are related by a symmetry transformation (precisely as would be expected for a spontaneously broken symmetry, for which the vacuum is not invariant). As we see in more detail in the next chapter, this transformation implies that \mathcal{M} can only couple through derivative couplings that vanish in the limit when the four-momentum of the boson goes to zero. In particular, this also implies that all of the mesons in \mathcal{M} are massless in the $SU_L(3) \times SU_R(3)$ -invariant limit.

The most general $SU_L(3) \times SU_R(3)$ -invariant Lagrangian that involves just \mathcal{M} , and involves the smallest possible number – two – of derivatives, is

$$\mathcal{L}_{gb} = \frac{f^2}{4} \text{tr} [\xi^\dagger \partial_\mu \xi \xi^\dagger \partial^\mu \xi]$$

$$\begin{aligned}
&= -\frac{1}{4} \text{tr}[\partial_\mu \mathcal{M} \partial^\mu \mathcal{M}] + \frac{1}{24f^2} \text{tr}[\mathcal{M}^2 \partial_\mu \mathcal{M} \partial^\mu \mathcal{M} - \mathcal{M} \partial_\mu \mathcal{M} \mathcal{M} \partial^\mu \mathcal{M}] \\
&\quad + O(\mathcal{M}^6)
\end{aligned} \tag{8.58}$$

In the chiral-symmetry limit, the Lagrangian is seen to be completely determined in terms of the constant f , which is ultimately to be determined by comparison with experiment. This is all the more remarkable since, besides the kinetic terms, Eq. (8.58) contains all of the self-interactions of the meson octet in the chiral limit that involve the fewest numbers of derivatives. Since they are all determined in terms of a single constant, it becomes possible to derive a system of *soft-pion theorems* that relate the low-energy scattering of the light mesons. These consequences for low-energy meson scattering are explored in more detail in Section 9.3.

The next step is to incorporate the symmetry-breaking effects of the various quark masses. We wish to derive the corrections to \mathcal{L}_{gb} that arise to lowest – i.e. linear – order in these masses. In order to determine the form these corrections take, we must determine the $SU_L(3) \times SU_R(3)$ transformation properties of the quark mass terms of Eq. (8.23) and Eq. (8.23).

The main observation is that the quark mass term in the QCD Lagrangian, \mathcal{L}_m , may be written as

$$\begin{aligned}
\mathcal{L}_m &= -\bar{\mathcal{Q}} M_q \mathcal{Q} \\
&= -\text{tr}[M_q(P_L \bar{\mathcal{Q}} \mathcal{Q} P_L + P_R \bar{\mathcal{Q}} \mathcal{Q} P_R)]
\end{aligned} \tag{8.59}$$

As before, \mathcal{Q} denotes the column vector containing the three light quark species, u , d , and s . Under an $SU_L(3) \times SU_R(3)$ transformation, the quark operators that appear in this mass term transform as

$$\begin{aligned}
P_L \bar{\mathcal{Q}} \mathcal{Q} P_L &\rightarrow L(P_L \bar{\mathcal{Q}} \mathcal{Q} P_L) R^\dagger, \\
\text{and : } P_R \bar{\mathcal{Q}} \mathcal{Q} P_R &\rightarrow R(P_R \bar{\mathcal{Q}} \mathcal{Q} P_R) L^\dagger
\end{aligned} \tag{8.60}$$

We must construct a term, \mathcal{L}_M , for the effective Lagrangian involving the meson field, \mathcal{M} , that satisfies the following properties:

- (i) it must have the form $\text{Tr}[M(\mathcal{O} + \mathcal{O}^\dagger)]$, in which the operator $\mathcal{O} = \mathcal{O}(\mathcal{M})$ transforms under $SU_L(3) \times SU_R(3)$ as $\mathcal{O} \rightarrow L \mathcal{O} R^\dagger$, and
- (ii) it must involve the fewest number of derivatives.

The unique solution to these conditions is to choose \mathcal{O} proportional to ξ itself, giving

$$\mathcal{L}_M = \frac{cf^2}{4} \text{tr}[M_q(\xi + \xi^\dagger)]$$

$$= \frac{cf^2}{2} \text{tr} M_q - \frac{c}{4} \text{tr}[M_q \mathcal{M}^2] + \frac{c}{48f^2} \text{tr}[M_q \mathcal{M}^4] + O(\mathcal{M}^6) \quad (8.61)$$

Here, c represents an unknown constant having dimensions of mass. It might reasonably be expected to be of the order of the strong-interaction scale on dimensional grounds. The first term of Eq. (8.61) describes a vacuum energy and may be dropped. The meson interactions in \mathcal{L}_M are of no further interest here but are discussed in more detail in Section 9.3. This leaves the mass term

$$\begin{aligned} \mathcal{L}_M &= -\frac{c}{4} \text{tr}[M_q \mathcal{M}^2] \\ &= -\frac{c}{2} \begin{pmatrix} \pi^0 & \eta_8 \end{pmatrix} \begin{pmatrix} \frac{1}{2}(m_u + m_d) & \frac{1}{2\sqrt{3}}(m_u - m_d) \\ \frac{1}{2\sqrt{3}}(m_u - m_d) & \frac{1}{6}(m_u + m_d + 4m_s) \end{pmatrix} \begin{pmatrix} \pi^0 \\ \eta_8 \end{pmatrix} \\ &\quad - \frac{c}{2}(m_u + m_d)\pi^+\pi^- - \frac{c}{2}(m_u + m_s)K^+K^- - \frac{c}{2}(m_s + m_d)\bar{K}^0 K^0 \end{aligned} \quad (8.62)$$

Some of the meson masses may be read off immediately:

$$\begin{aligned} m_{\pi^\pm}^2 &= (m_u + m_d) \frac{c}{2} \\ m_{K^\pm}^2 &= (m_u + m_s) \frac{c}{2} \\ \text{and : } m_{K^0}^2 &= (m_s + m_d) \frac{c}{2} \end{aligned} \quad (8.63)$$

The mass matrix for the self-conjugate mesons must be diagonalized, since the isospin eigenstates, π^0 and η_8 , mix with one another in the presence of an isospin-breaking mass difference $m_u - m_d$. The eigenvectors of this 2×2 matrix are

$$\begin{pmatrix} \pi^{0'} \\ \eta \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \pi^0 \\ \eta_8 \end{pmatrix} \quad (8.64)$$

where the mass eigenvalues are

$$\begin{aligned} m_{\pi^0}^2 &= \frac{c}{3}[(m_u + m_d + m_s) - R] \\ m_\eta^2 &= \frac{c}{3}[(m_u + m_d + m_s) + R] \\ \text{with } R &= \sqrt{m_u^2 + m_d^2 + m_s^2 - m_u m_d - m_u m_s - m_d m_s} \end{aligned} \quad (8.65)$$

Comparing these meson masses with those that are measured allows a determination of the values of the light quark masses, m_u , m_d , and m_s . This

is one source of the numerical values for the meson masses that appear in the *Particle Data Book*, and are used throughout this book.

The main $SU_1(2)$ -breaking effects are the isospin mass splittings, and the $\pi^0 - \eta_8$ mixing angle, θ :

$$\begin{aligned}
 m_{K^0}^2 - m_{K^\pm}^2 &= (m_d - m_u) \frac{c}{2} \\
 m_{\pi^\pm}^2 - m_{\pi^0}^2 &= \frac{c}{6} [(m_u + m_d - 2m_s) - R] \\
 &\approx \frac{(m_d - m_u)^2}{8m_s} c \\
 \text{and} \quad \tan 2\theta &= \frac{\sqrt{3}(m_d - m_u)}{2m_s - m_u - m_d} \approx \frac{\sqrt{3}}{2} \left(\frac{m_d - m_u}{m_s} \right) \quad (8.66)
 \end{aligned}$$

Since the ratio $(m_d - m_u)/m_s \approx 1/40$, we see that predictions that are based on isospin invariance should be quite accurate. Within this approximation there is no mixing of neutral mesons and the above formulae for the masses of the three isomultiplets within the octet simplify dramatically. Taking $m_d \approx m_u \equiv m$ we find

$$\begin{aligned}
 m_\pi^2 &\approx mc \\
 m_K^2 &\approx \frac{1}{2}(m_s + m)c \\
 m_\eta^2 &\approx \frac{1}{3}(2m_s + m)c \quad (8.67)
 \end{aligned}$$

Notice in particular that, since the unknown constant c only enters as an overall factor, the three independent ratios of these masses are fixed in terms of two parameters, m and m_s . By eliminating these two parameters from the three mass ratios we derive the following *mass formula*:

$$4m_K^2 = 3m_\eta^2 + m_\pi^2 \quad (8.68)$$

This relation is indeed extremely well satisfied by the measured masses. Using $m_\pi = 137$ MeV, $m_K = 496$ MeV, and $m_\eta = 549$ MeV, we find $4m_K^2 = 0.98 \text{ GeV}^2$ and $3m_\eta^2 + m_\pi^2 = 0.92 \text{ GeV}^2$.

8.3.6.2 The pseudoscalar singlet: η'

If we work our way up in energy we may include more and more fields into our effective Lagrangian. The Lagrangian is constructed in the same way as before: we write all possible terms that are consistent with their transformation properties under $SU_L(3) \times SU_R(3)$. As more fields are

added, more unknown constants appear in the couplings between the fields.

Perhaps the simplest such meson to include is the remaining pseudoscalar, η_0 . η_0 does not transform at all under $SU_L(3) \times SU_R(3)$. Starting with the chiral limit, the most general $SU_L(3) \times SU_R(3)$ -invariant effective Lagrangian involving both η_0 and \mathcal{M} and containing just terms with up to two derivatives is therefore

$$\mathcal{L} = -V(\eta_0) - \frac{1}{2}G(\eta_0)\partial_\mu\eta_0\partial^\mu\eta_0 - F(\eta_0)\text{tr}[\xi^\dagger\partial_\mu\xi\xi^\dagger\partial^\mu\xi] \quad (8.69)$$

Since η_0 and \mathcal{M} are both parity odd, parity invariance of the strong interactions implies that the otherwise arbitrary real functions, V , G , and F , must be even under $\eta_0 \rightarrow -\eta_0$.

Expanding this result about the minimum of the potential, $V(\eta_0)$, and canonically normalizing the fields then gives

$$\mathcal{L} = -\frac{1}{2}\mu_0^2\eta_0^2 - \frac{1}{2}\partial_\mu\eta_0\partial^\mu\eta_0 - \frac{1}{4}\text{tr}[\partial_\mu\mathcal{M}\partial^\mu\mathcal{M}] + \dots \quad (8.70)$$

μ_0^2 here represents an unknown mass for η_0 which should, on dimensional grounds, be taken of the order of the strong interaction scale. Clearly η_0 need not be particularly light since, unlike the octet pseudoscalar mesons, its mass need not vanish in the limit that the quarks are massless. The dots represent cubic and higher interaction terms. As expected, there is no mixing in the absence of $SU_L(3) \times SU_R(3)$ breaking.

Incorporating $SU_L(3) \times SU_R(3)$ breaking by perturbing in the quark masses gives, as before,

$$\begin{aligned} \delta\mathcal{L} &= H(\eta_0)\text{tr}[M_q(\xi + \xi^\dagger)] \\ &= -\frac{c}{4}\text{tr}[M_q\mathcal{M}^2] + a\eta_0^2\text{tr}[M_q\mathcal{M}^2] + \dots \end{aligned} \quad (8.71)$$

The second equation uses parity invariance to rule out any odd powers of η_0 in H . We see that even after including leading-order symmetry-breaking terms we do not generate any mixing between η_0 and the members of the scalar-meson octet. This absence of mixing follows from the requirement that the octet mesons must become massless in the chiral-symmetry limit. We therefore expect the heavy stable pseudoscalar mass eigenstate, called the η' , to be entirely an $SU_f(3)$ singlet, and the lighter mass eigenstate, η , to be entirely octet.

8.3.6.3 The pseudovector nonet

We next apply this analysis to the nine $J^{PC} = 1^{--}$ stable mesons. These massive spin-one particles may be represented by an octet of four-vector

fields, $\tilde{\mathcal{M}}_\mu$ as in Eq. (8.36),

$$\tilde{\mathcal{M}}_\mu = \begin{pmatrix} \rho_\mu^0 + \frac{1}{\sqrt{3}}\omega_\mu^8 & \sqrt{2}\rho_\mu^+ & \sqrt{2}K_\mu^{+*} \\ \sqrt{2}\rho_\mu^- & -\rho_\mu^0 + \frac{1}{\sqrt{3}}\omega_\mu^8 & \sqrt{2}K_\mu^{0*} \\ \sqrt{2}K_\mu^{-*} & \sqrt{2}K_\mu^{0*} & -\frac{2}{\sqrt{3}}\omega_\mu^8 \end{pmatrix} \quad (8.72)$$

and a singlet vector, ω_μ^0 . Under $SU_f(3)$ transformations the octet transforms as $\tilde{\mathcal{M}}_\mu \rightarrow U\tilde{\mathcal{M}}_\mu U^\dagger$, and the singlet is invariant. Both fields may be chosen to be invariant under the spontaneously-broken axial transformations, $SU_A(3)$, by absorbing appropriate powers of ξ into the fields. Once the fields have been chosen in this way, the Lagrangian need only be constructed in an $SU_f(3)$ -invariant way. The most general mass term for these fields that would be consistent in the symmetry limit is

$$\mathcal{L}_{\text{mass}} = -\frac{\tilde{\mu}_0^2}{2}\omega_\mu^0\omega^{0\mu} - \frac{\tilde{\mu}_8^2}{4}\text{tr}[\tilde{\mathcal{M}}_\mu\tilde{\mathcal{M}}^\mu] \quad (8.73)$$

This Lagrangian represents an arbitrary mass, $\tilde{\mu}_0$, for the singlet particle, ω_μ^0 , and a common mass, $\tilde{\mu}_8$, for all of the particles in the octet.

The mass terms that are allowed may be constructed using the following trick. We need to know how the QCD quark mass term, \mathcal{L}_m of Eq. (8.23), transforms under $SU_f(3)$. The trick is to notice that although \mathcal{L}_m is not invariant under $SU_f(3)$, it would be invariant if the mass matrix, M_q , were to transform as $M_q \rightarrow U M_q U^\dagger$. The allowed mass terms are then those that are bilinear in the vector meson fields, are linear in the quark mass matrix, M_q , and would be $SU_f(3)$ -invariant if the mass matrix were to transform as $M_q \rightarrow U M_q U^\dagger$. The result is

$$\begin{aligned} \delta\mathcal{L}_{\text{mass}} &= -\frac{a}{4}\text{tr}[M_q\tilde{\mathcal{M}}_\mu\tilde{\mathcal{M}}^\mu] - \frac{b}{4}\omega_\mu^0\text{tr}[M_q\tilde{\mathcal{M}}^\mu] \\ &= -\frac{1}{2}\left(\rho_\mu^0\ \omega_\mu^8\ \omega_\mu^0\right)\left(\delta\mu_8^2\right)\begin{pmatrix} \rho^{0\mu} \\ \omega^{8\mu} \\ \omega^{0\mu} \end{pmatrix} - \frac{a}{2}(m_u+m_d)\rho_\mu^+\rho^{-\mu} \\ &\quad - \frac{a}{2}(m_u+m_s)K_\mu^{+*}K^{-*\mu} - \frac{a}{2}(m_s+m_d)\bar{K}_\mu^{0*}K^{0*\mu} \end{aligned} \quad (8.74)$$

in which the 3×3 mass matrix, $\delta\mu_8^2$, is

$$\left(\delta\mu_8^2\right) = \begin{pmatrix} \frac{a}{2}(m_u+m_d) & \frac{a}{2\sqrt{3}}(m_u-m_d) & \frac{b}{4}(m_u-m_d) \\ \frac{a}{2\sqrt{3}}(m_u-m_d) & \frac{a}{6}(m_u+m_d+4m_s) & \frac{b}{4\sqrt{3}}(m_u+m_d-2m_s) \\ \frac{b}{4}(m_u-m_d) & \frac{b}{4\sqrt{3}}(m_u+m_d-2m_s) & 0 \end{pmatrix} \quad (8.75)$$

In order to get a qualitative picture of what the spectrum should look like, it is sufficient to completely neglect m_u and m_d . This should be an even better approximation here than it was for the pseudoscalar octet, since the mass scale for the vector mesons is set by the invariant mass terms, $\tilde{\mu}_8$ and $\tilde{\mu}_0$, which should be much greater than all three of the quark masses. In this limit, we have

$$\begin{aligned} m_\rho^2 &\approx \tilde{\mu}_8^2 \\ m_{K^*}^2 &\approx \tilde{\mu}_8^2 + \frac{a}{2}m_s \end{aligned} \quad (8.76)$$

and the $\omega^8 - \omega^0$ mass matrix becomes

$$m_{0-8}^2 \approx \begin{pmatrix} \tilde{\mu}_8^2 + \frac{2a}{3}m_s & -\frac{b}{2\sqrt{3}}m_s \\ -\frac{b}{2\sqrt{3}}m_s & \tilde{\mu}_0^2 \end{pmatrix} \quad (8.77)$$

The mass eigenstates, to this approximation, are

$$\begin{aligned} \begin{pmatrix} \omega \\ \phi \end{pmatrix} &= \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \omega_8 \\ \omega_0 \end{pmatrix} \\ \tan \theta &= \frac{\sqrt{3}}{bm_s} \left[\tilde{\mu}_8^2 - \tilde{\mu}_0^2 + \frac{2a}{3}m_s + \sqrt{\left(\tilde{\mu}_8^2 - \tilde{\mu}_0^2 + \frac{2a}{3}m_s \right)^2 + \frac{b^2}{3}m_s^2} \right] \end{aligned} \quad (8.78)$$

and the corresponding eigenvalues are

$$\begin{aligned} m_\omega^2 &= \frac{1}{2} \left[\tilde{\mu}_8^2 + \tilde{\mu}_0^2 + \frac{2a}{3}m_s - \sqrt{\left(\tilde{\mu}_8^2 - \tilde{\mu}_0^2 + \frac{2a}{3}m_s \right)^2 + \frac{b^2}{3}m_s^2} \right] \\ m_\phi^2 &= \frac{1}{2} \left[\tilde{\mu}_8^2 + \tilde{\mu}_0^2 + \frac{2a}{3}m_s + \sqrt{\left(\tilde{\mu}_8^2 - \tilde{\mu}_0^2 + \frac{2a}{3}m_s \right)^2 + \frac{b^2}{3}m_s^2} \right] \end{aligned} \quad (8.79)$$

Clearly there can be a considerable amount of mixing between ω_8 and ω_0 , depending on whether m_s is large or small compared to the splitting $\tilde{\mu}_8^2 - \tilde{\mu}_0^2$. If $m_s^2 \ll |\tilde{\mu}_8^2 - \tilde{\mu}_0^2|$ then $\theta \approx 0$ and the mass eigenstates, ω and ϕ , are almost purely octet and singlet. On the other hand, if $m_s^2 \gg |\tilde{\mu}_8^2 - \tilde{\mu}_0^2|$ then the $SU_f(3)$ -invariant mass matrix is proportional to the unit matrix, and so the mixing is completely governed by the ratio of the symmetry-breaking coefficients a and b :

$$\tan \theta \approx \frac{2}{\sqrt{3}} \frac{a}{b} \left[1 + \sqrt{1 + \frac{3b^2}{4a^2}} \right], \quad \tilde{\mu}_8^2 \approx \tilde{\mu}_0^2 \quad (8.80)$$

Since there are as many free parameters as there are masses in the present approximation, it is not possible to derive a general mass relation here, or

to predict the amount of mixing between ω_8 and ω_0 . Experimentally it is found that these two states are strongly mixed, with one mass eigenstate, ϕ , being approximately purely $s\bar{s}$ and the other being $(u\bar{u} + d\bar{d})$.

8.3.6.4 Baryons

We next briefly describe the results that can be obtained for the stable octet and decuplet baryons. In these cases there is sufficient information to derive mass relations among the different baryon species. The presentation is kept fairly telegraphic since the considerations are similar to the meson systems that have already been considered.

For the octet baryons, the rules for constructing the mass terms are the same as for the pseudovector mesons. We must construct a mass term which is at most linear in the quark mass matrix, and which would be $SU_f(3)$ -invariant if the mass matrix were to transform as is described in the paragraph immediately preceding Eq. (8.73).

The most general such mass has the form of a common mass for the baryon octet, as well as two symmetry-breaking terms

$$\begin{aligned} \mathcal{L}_{\text{mass}} &= \mathcal{L}_{\text{inv}} + \delta\mathcal{L} \\ \text{with } \mathcal{L}_{\text{inv}} &= -\frac{A}{2} \text{tr}[\bar{\mathcal{B}}\mathcal{B}] \\ \text{and } \delta\mathcal{L} &= -F \text{tr}(\bar{\mathcal{B}}[M_q, \mathcal{B}]) - D \text{tr}(\bar{\mathcal{B}}[M_q, \mathcal{B}]) \end{aligned} \quad (8.81)$$

Because there are three parameters and four independent masses in the octet, we derive the following *Gell-Mann–Okubo* mass relation (in the isospin conserving limit):

$$3m_\Lambda + m_\Sigma = 2(m_N + m_\Xi) \quad (8.82)$$

N here represents the *nucleon isodoublet*, n and p . Comparison with the observed masses demonstrates that this is a very successful prediction. For instance, if we average members of each isospin multiplet, the two sides of the equations add to 4540 and 4514 MeV, respectively.

Similarly, for the decuplet there are two possible parameters in the mass term. One of these is a common $SU_f(3)$ -invariant mass for the entire multiplet, and the other is the only possible symmetry-breaking term to arise at this order. Since there are a total of four isomultiplets within the decuplet, there are two independent relations that may be derived:

$$m_\Omega - m_{\Xi^*} = m_{\Xi^*} - m_{\Sigma^*} = m_{\Sigma^*} - m_\Delta \quad (8.83)$$

These relations are also quite good; the mass differences all lie between 140 and 153 GeV.

We see, then, that even though it is at present impossible to quantitatively solve for the QCD spectrum from first principles, especially at low energies, the theory nevertheless gives a comprehensive and impressive understanding of the features of the hadron spectrum. A great deal of effort has gone into providing a more detailed understanding of the observed energy levels within the theory. We skip these here because with our current tools they all involve some level of model dependence and their proper treatment is beyond the goals of this presentation.

QCD is at least equally successful in describing the collisions and decays of these hadronic states, so it is to these that we turn in the following chapter.

8.4 Problems

[8.1] η_0 mixing

Show that the potential η_0 – \mathcal{M} mixing term:

$$\begin{aligned}\mathcal{L} &= igf^2\eta_0 \operatorname{tr}[M_q(\xi - \xi^\dagger)] \\ &= -gf\eta_0 \operatorname{tr}[M_q\mathcal{M}] + \cdots \\ &= gf\eta_0\pi^0(m_d - m_u) - \frac{gf}{\sqrt{3}}\eta_0\eta_8(m_u + m_d - 2m_s)\end{aligned}$$

violates both C and CP.

[8.2] Technical issues in Goldstone's theorem

The proof of Goldstone's theorem used the artifice of using a position-dependent symmetry parameter, $\omega(x)$, with the limit of constant ω taken at the end. This side-steps several technical issues, related to the normalization of the states which are produced by performing symmetry transformations. To illustrate these issues suppose that $R[\omega] = i\omega_0 Q$ with $Q = \int d^3x j^0$ and constant ω_0 . Show that the state $|\psi\rangle = R|\Omega\rangle$ has a norm, $\langle\psi|\psi\rangle$, which is proportional to the volume, V , of space (and so diverges) if we assume: (i) Q commutes with all Poincaré transformations; and (ii) the ground state is translation invariant $P_\mu|\Omega\rangle = 0$. For these reasons it is preferable to formulate Goldstone's theorem in terms of the matrix element $\langle G|j^\mu(x)|\Omega\rangle$ and of the commutator of R with local fields, rather than working with ill-defined quantities like $\langle G|R|\Omega\rangle = \langle G|\psi\rangle$.

[8.3] Physical interpretation of the chiral parameter c

We can find the physical interpretation for the parameter c , introduced in Eq. (8.61), by comparing the vacuum energy density, ρ_v , for QCD as computed in two equivalent ways. Using the QCD Lagrangian show that

the dependence of ρ_V on the quark masses satisfies

$$\frac{\partial \rho_V}{\partial m_q} = \langle \Omega | \bar{q} q | \Omega \rangle \quad (8.84)$$

where $|\Omega\rangle$ denotes the QCD ground state and $q(x)$ is the appropriate quark field. Compare this with the same quantity found by using the effective meson Lagrangian given by Eq. (8.58) and Eq. (8.61) in the semiclassical approximation. Show that in the limit $m_u = m_d = m$, these results imply

$$m_\pi^2 = mc \quad \text{and} \quad \langle \Omega | \bar{u} u | \Omega \rangle = \langle \Omega | \bar{d} d | \Omega \rangle = \frac{cf^2}{2} \quad (8.85)$$

9

Hadronic interactions

The last chapter reviewed the hadronic particle content of the standard model. We found that quarks and gluons are bound together, with a binding energy, $\Lambda_c \sim 400$ MeV. Nevertheless, it is possible with symmetry arguments to guess correctly the general pattern (masses, spins, and conserved or nearly conserved charges) of the bound states. However, we have not shown how to make detailed predictions; indeed, with analytical, theoretical tools alone it is not known how to make detailed predictions.

We now turn to interactions, such as scattering processes, of hadrons. Two types of scatterings should be distinguished; those of hadrons off hadrons, and those of hadrons off particles without strong interactions, such as electrons, photons, and neutrinos. It is often easier to understand the latter processes, since more of the problem (that part involving the electron, photon, or neutrino) is well under control.

Scatterings are also distinguished by the energy involved, falling roughly into four classes:

- (i) low energy scatterings, such as e^-p^+ scattering with $s - m_p^2 \ll m_p m_\pi$ so there is nowhere near enough energy to probe the structure of the proton;
- (ii) low-energy scatterings involving pions and kaons, such as $\pi\pi$ scattering or πp scattering;
- (iii) intermediate-energy scatterings, with excess kinetic energies in the range of a few hundred MeV to a few GeV; and
- (iv) high-energy scatterings, with $s > (\text{few GeV})^2$.

The first of these is quite tractable, and we will discuss it first.

In the second energy range, it is possible to make fairly accurate predictions by invoking a symmetry principle, spontaneously broken chiral

symmetry – the same symmetry which ensures the small masses of pions. We discuss this in Section 9.3.

The third case is truly complicated, because the energy is sufficiently low for the $SU_c(3)$ coupling to be strong, and yet there is enough energy that chiral symmetry is not helpful. Furthermore, at the higher energies, the set of available final states for a collision process gets larger. There are no good tools for first principles or theoretical predictions in this regime. Instead, our knowledge is based on experimental data and phenomenological modeling. In the lower energy range, scatterings are reasonably well described as occurring owing to a series of hadronic resonances (the ρ , f_0 , h_1 , Δ , N^* , etc.). There is no good first-principles understanding of the location of these resonances, though there are phenomenological models (Regge theory, etc.) which work relatively well. However, as the energy scale is raised, the resonances become broader and denser until this description loses all utility. We will not discuss this regime any further.

The last energy range is complicated, but tractable to some extent, because the large value of s sets a scale in the problem where the interactions are weak. This means that some quantities are computable, or, to be more precise, one can derive relations between what will be measured in different experiments. We discuss this at some length in Section 9.2.

9.1 Quasi-elastic scattering

Perhaps the simplest situation for which hadronic scattering can be computed quantitatively is in the limit where the energy transfer to the hadron of interest is small compared to the scale Λ_{QCD} . The case of a very heavy quark is also relatively simple, and also permits a non-relativistic approximation.

9.1.1 Non-relativistic hadrons

Consider the electromagnetic and weak interactions of stable or long-lived hadrons with photons and leptons, when the kinetic energy of the interaction ($\sqrt{s} - \sum m_i$, m_i the masses of initial-state particles) is small compared to the scale Λ_{QCD} . In this case, we can write an *effective theory* for the interactions in which we treat the hadrons as point particles (since there is not sufficient energy to probe their structure). One then writes down all terms consistent with symmetries (C, P, electromagnetic gauge symmetry, and conservation of baryon number, lepton number, strangeness, etc.) and the electromagnetic charge assignments of the various particles. Weak (P and flavor violating)

interactions can also be written down, suppressed by the Fermi coupling G_F . For instance, the renormalizable terms in the Lagrangian for the proton would be

$$\mathcal{L}_{\text{proton}} = -\bar{p}\not{D}p - m_p \bar{p}p, \quad D_\mu = \partial_\mu - ieA_\mu \quad (9.1)$$

The covariant derivative is determined by the charge, $+1$, of the proton. This is exactly the Lagrangian term used to solve for the bound state spectrum of the hydrogen atom, and it gives the correct fine structure. In general, calculations with this Lagrangian (and the analogous renormalizable terms for pions and ordinary leptons) are straightforward. One should remember that the theory is only valid in the limit where kinetic energies are small compared to hadron masses, so it is always legitimate to expand in the large mass (non-relativistic behavior) of the hadron. Also, it is forbidden to use this effective theory to consider the interaction of two or more hadrons, because the interaction energies of those hadrons can be large, even if the kinetic energy of the interaction is not.

If the energy scale of interest is at the upper end of the range of validity of this effective description, or if we desire high accuracy, one must also remember that high-dimension operators will be present. The coefficients of high-dimension operators involving electromagnetic interactions will be set by the scale where the effective description breaks down, $\Lambda_c \sim 400$ MeV. For instance, the first high-dimension term involving the proton which can appear is

$$\mathcal{L}_{\mu_a} = -\frac{i\mu_a}{4}\bar{p}[\gamma^\mu, \gamma^\nu]F_{\mu\nu}p \quad (9.2)$$

with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ the electromagnetic field strength. On dimensional grounds, together with the fact that this is an electromagnetic interaction with one power of A_μ , we can estimate that $\mu_a \sim e/\Lambda_{\text{QCD}}$. However, we cannot determine this coefficient exactly, unless we know how to solve QCD to understand in detail the structure of the proton. Therefore μ_a must be left as a parameter to be determined by experiment.

To understand the physical content of this term, consider the case where $F_{\mu\nu}$ is a static, background field, and compute the energy shift for a state containing a proton at rest,

$$\langle p(\mathbf{k} = 0, \sigma) | H_{\mu_a} | p(\mathbf{k} = 0, \sigma) \rangle \quad (9.3)$$

The \bar{p} and p operators in H annihilate the proton in the bra and ket respectively, leaving spinors $\bar{u}(\mathbf{k} = 0, \sigma)$ and $u(\mathbf{k} = 0, \sigma)$. The combination $[\gamma^0, \gamma^i]$

vanishes between these spinors, while the $[\gamma^i, \gamma^j]$ part becomes

$$-\frac{i}{4}[\gamma^i, \gamma^j]F_{ij} = \frac{1}{2}\epsilon_{ijk}\sigma^k F_{ij} = \boldsymbol{\sigma} \cdot \mathbf{B} \quad (9.4)$$

Therefore

$$H_{\mu_a} \simeq -\mu_a \bar{\mathbf{p}} \vec{\sigma} \cdot \vec{B} \quad (9.5)$$

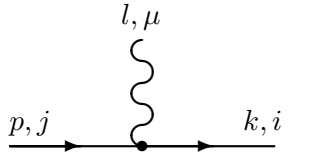
We recognize this as a spin magnetic interaction; μ_a is an anomalous contribution to the magnetic moment of the proton. The total proton magnetic moment is

$$\mu_p \equiv \frac{g_p e}{2m_p} = \frac{e}{2m_p} + \mu_a, \quad \text{or} \quad g_p = 1 + \frac{2m_p \mu_a}{e} \quad (9.6)$$

Here g_p is the *gyroscopic ratio* for the proton, defined as the ratio of the true magnetic moment to the contribution arising from the Dirac equation, $e/2m_p$.

According to our previous estimate, $\mu_a \sim e/\Lambda_{\text{QCD}}$. Since $m_p/\Lambda_{\text{QCD}} \sim 2$, we expect a large correction to the gyroscopic ratio of the proton, $g_p - 1 \sim 2$. The experimental value is $g_p = 2.793$. This correction is quite important in astrophysics, because the proton magnetic moment breaks the degeneracy between the spin-singlet and spin-triplet $1s$ states of the hydrogen atom, leading to the 21 cm hydrogen-emission line. The neutron also turns out to have a large magnetic moment, $\mu_n = (-1.913)e/2m_n$. Since the neutron is charge zero, the magnetic-moment term gives its dominant electromagnetic interaction.

To give a concrete example of the utility of this expansion, consider the photon-neutron scattering cross section, where the photon energy (in the neutron rest frame) is small, $l^0 = |\mathbf{l}| \ll \Lambda_{\text{QCD}}$. Since the electric charge is zero, at leading order only the dipole interaction need be considered. The two diagrams contributing to photon scattering are the same as in Figure 6.8, except that the vertex Feynman rule is now



$$\frac{-i\mu_n}{2}(\gamma^\nu \gamma^\mu - \gamma^\mu \gamma^\nu)_{ij} l_\nu (2\pi)^4 \delta^4(p+k+l) \quad (9.7)$$

Writing the initial and final momenta as \mathbf{l}, \mathbf{l}' for the photon and \mathbf{p}, \mathbf{p}' for the neutron, and the initial and final polarizations as ϵ and ϵ' , the matrix element is

$$\mathcal{M} = -\frac{i\mu_n^2}{4} \left\{ \bar{u}(\mathbf{p}')(\mathbf{l}' \not{\epsilon}' - \not{\epsilon}' \mathbf{l}') \frac{-i(\not{p} + \not{l}) + m}{(p+l)^2 + m^2} (\mathbf{l} \not{\epsilon} - \not{\epsilon} \mathbf{l}) u(\mathbf{p}) \right\}$$

$$+ \bar{u}(\mathbf{p}')(\not{l}\not{\epsilon} - \not{\epsilon}\not{l}) \frac{-i(\not{p} - \not{l}') + m}{(p - l')^2 + m^2} (\not{l}'\not{\epsilon}' - \not{\epsilon}'\not{l}') u(\mathbf{p}) \Big\} \quad (9.8)$$

To check that this is transverse, substitute $\epsilon_\mu \rightarrow l_\mu$; $(\not{l}\not{\epsilon} - \not{\epsilon}\not{l})$ immediately gives zero.

Squaring, and averaging/summing the initial/final spins and polarizations, gives four terms. Using

$$(\not{l}\gamma^\mu - \gamma^\mu\not{l})(-i\not{p} + m)(\not{l}\gamma^\mu - \gamma^\mu\not{l}) = -16i p \cdot l \not{l} \quad (9.9)$$

as can easily be verified by using $l^2 = 0$, the square of the first term in the matrix element reduces to

$$\overline{\mathcal{M}}_1^2 = \frac{-\mu_n^4}{64(2p \cdot l)^2} (16p \cdot l)(16p' \cdot l') \text{tr} \left[\not{l}'(-i\not{p} - i\not{l} + m)\not{l}(-i\not{p} - i\not{l} + m) \right] \quad (9.10)$$

Since $\not{l}\not{l} = l^2 = 0$ and $p^2 = -m^2$, the trace evaluates to $-8(p \cdot l)(p \cdot l')$. The square of the second term in the matrix element is the same under the substitution $l \rightarrow -l'$, so it equals the square of the first term (in the small l limit, where $(p \cdot l)^2 = (p \cdot l')^2$). After a great deal of unenlightening work, one can show that the interference terms, in the limit $l \ll p$, vanish at leading order. Therefore, using $p \cdot l \simeq -m_n l^0 \simeq p \cdot l'$, the spin-averaged matrix element squared becomes

$$\overline{\mathcal{M}}^2 = 16\mu_n^4 m_n^2 (l^0)^2 \quad (9.11)$$

The cross section is isotropic. Remembering that the relative velocity is 1 (the photon's velocity), the total cross section is (using the non-relativistic limit for the neutron energy)

$$\begin{aligned} \sigma &= \frac{1}{2m_n 2l^0} \int \frac{d^3 p' d^3 l'}{(2\pi)^6 2m_n 2l^0} (2\pi)^4 \delta^3(\mathbf{p}' + \mathbf{l}' - \mathbf{l}) \delta(|\vec{l}| - l^0) 16\mu_n^4 m_n^2 (l^0)^2 \\ &= \frac{\mu_n^4 (l^0)^2}{\pi} \end{aligned} \quad (9.12)$$

As the energy increases or more precision is needed, one must take into account still higher order interactions. For instance, at dimension 6 there is the interaction

$$\mathcal{L}_{p,6} = r_{\text{chg}}^2 \bar{p} D_\mu D^\mu \not{D} p, \quad (9.13)$$

where $r_{\text{chg}} \sim 1/\Lambda_{\text{QCD}} \sim 0.5$ fm is another *a priori* unknown parameter, traditionally called the charge radius of the proton because such a correction would arise if the proton were a ball of radius r_{chg} with its charge on the surface. Numerically, $r_{\text{chg}} = 0.87$ fm, as determined by electron scattering and energy shifts in muonic hydrogen bound states.

Weak interactions are handled similarly. To the extent that isospin is a good symmetry, the weak interactions of the proton and neutron can be handled as those of the respective isospins, except for the possibility of numerical shifts in the coefficients. The neutron and proton are isospin $1/2$, with third component of isospin $-1/2$ and $1/2$ respectively, exactly the same isospin charges as down and up quarks. The charged current interactions, *c.f.* Eq. (7.12), are given by

$$\mathcal{L}_{cc} = \frac{G_F}{\sqrt{2}} C_\mu^* C^\mu, \quad C^\mu = i\bar{e}\gamma^\mu(1 + \gamma^5)\nu_e + V_{ud}^* \bar{n}\gamma^\mu(g_V + g_A\gamma^5)p, \quad (9.14)$$

with g_V and g_A coefficients to account for possible renormalizations in going from the quark to the hadron level. The g_V term, $\bar{n}\gamma^\mu p$, is also the charged current of $SU(2)$ isospin. Therefore, to the extent that isospin is a valid symmetry, it should be a conserved current, which in particular means that there is no renormalization with respect to the value we would obtain for quarks of the same isospin; the proton-to-neutron coupling has the same coefficient as the up-to-down quark coupling. This current is therefore referred to as the *conserved vector current*. Renormalization of this current first appears at quadratic order in isospin breaking, according to the *Ademollo–Gatto theorem* (see also Subsection 9.3.3):

$$g_V = 1 + O\left(\frac{m_d^2 - m_u^2}{\Lambda_{\text{QCD}}^2}\right) = 1 + O(10^{-3}) \quad (9.15)$$

The same is not true of the axial current, which does have a substantial (though not enormous) renormalization with respect to the naive quark model value of 1;

$$g_A \simeq 1.269 \quad (9.16)$$

experimentally. Since no symmetry argument can fix this value, it must be taken from experiment.

Similarly, the weak neutral-current interactions for neutrons and protons are those of isospin $1/2$ particles with charges 0 and 1 respectively, except that the axial current has a non-negligible renormalization:

$$\mathcal{L}_{nc} = \frac{G_F}{\sqrt{2}} J_{nc}^\mu J_{nc\mu}, \quad J_{nc}^\mu = J_{nc,\text{lept}}^\mu + \bar{p}(g_{V,p} + g_{A,p}\gamma^5)p + \bar{n}(g_{V,n} + g_{A,n}\gamma^5)n \quad (9.17)$$

Again, up to isospin breaking corrections, the vector charges are as expected, $T_3 - 2Q\sin^2\theta_W$, which is $g_{V,p} = \frac{1}{2} - 2\sin^2\theta_W$ for a proton and $g_{V,n} = -\frac{1}{2}$ for a neutron. The axial couplings experience larger radiative corrections and must be determined from experiment. These interactions are relevant, for instance, in neutrino–nucleon scattering.

This effective description is an efficient way of studying interactions when the kinetic energy is $\ll 400$ MeV. However, we emphasize again two provisos. First, the higher the energy, the more important the high dimension operators become. There is no *a-priori* way of determining their coefficients; so the utility of the treatment breaks down as the energy increases. Second, the effective theory cannot be used to study interactions between hadrons, such as pp or pn scattering. This is because the energy scales involved in such scatterings need not be small, even if the kinetic energies of the widely separated particles are. At short ranges there can be large interaction energies owing to exchange of strongly interacting particles, potentially leading to bound states and strong scattering effects. This is nuclear physics, which is a complicated field not well described by a simple low-energy effective theory. An exception is when there is an electromagnetic Coulomb barrier which stops the hadrons from exploring separations $\sim 1/\Lambda_{\text{QCD}}$. This occurs for same-charge hadrons at kinetic energies $\ll \alpha \Lambda_{\text{QCD}}$ (the factor α entering to account for the weakness of the electromagnetic Coulomb interaction).

9.1.2 Non-relativistic quarks: HQET*

For scattering and decay processes involving very massive and slowly moving particles, relativistic effects are not important, and considerable simplification is obtained by expanding observables in powers of the small velocities of the massive particles (or, equivalently, the relevant momenta and energies divided by the heavy mass). This is a very instructive limit to take for the heavy quarks (c , b , and t), because it reveals new symmetries relating different quark species and spins. Furthermore, for these quarks the QCD interaction energies are small compared with their masses, and so these symmetries should be good approximations for mesons containing these quarks, up to corrections which can be computed (in principle) in inverse powers of the heavy mass. This leads to a systematic approximation scheme for understanding many of the features of strongly interacting heavy quarks.

Technically, this is accomplished by formulating the dynamics of interest in terms of an effective low-energy theory, similar in spirit to that encountered in Chapter 7, called *heavy quark effective theory* (HQET). A complete discussion of this approach goes beyond the scope of an introductory book like this one, and we content ourselves with a brief sketch of its implications here.

The main idea on which this approach is founded is that some of the symmetries which emerge in the non-relativistic limit are obscure when formulated in terms of relativistic fields, such as those we use in the previous

section for proton elastic scattering. The roots of this obscurity lie in the necessity of carrying around both particle and antiparticle within the relativistic approach, due to the appearance of both in all fields when these are expressed in terms of particle states, such as in

$$\psi(x) = \sum_{\sigma} \int \frac{d^3k}{(2\pi)^3 2k^0} \left[a_{k\sigma} u_{k\sigma} e^{ikx} + \bar{a}_{k\sigma}^* v_{k\sigma} e^{-ikx} \right] \quad (9.18)$$

Although having the underlying particle–antiparticle symmetry explicit pays dividends in relativistic applications, it is an encumbrance in non-relativistic situations, for which the influence of virtual particle–antiparticle pairs is negligible.

A simpler formulation is obtained by “integrating out” such virtual pairs, leaving an effective theory involving only the “particle” part of the field, *i.e.*, the part involving $a_{k\sigma}$ and not $\bar{a}_{k\sigma}$. The effects of virtual antiparticles can all be expressed in such a theory in terms of various effective interactions, in much the same way as the effective Fermi interaction captures the low-energy physics of virtual W exchange once these gauge bosons are integrated out. The effective interactions obtained in this way have couplings proportional to inverse powers of the relevant heavy mass, m_Q , which justifies the non-relativistic limit, and these can be characterized systematically order by order in powers of $1/m_Q$.

Our interest in what follows is in applications to things like the decays of \bar{B} mesons into final states involving D mesons, owing to the underlying heavy-quark decay $b \rightarrow c$. Since there are two heavy particles in these reactions, which can be moving (slowly) relative to one another, we phrase the non-relativistic limit without assuming we are precisely in the heavy particle’s rest frame.

Consider therefore a single Dirac fermion, ψ , describing a heavy quark having mass m_Q , whose 4-velocity is $v^\mu = p^\mu/m_Q$ (satisfying $v^\mu v_\mu = -1$). It is convenient to define the particle and antiparticle parts of ψ by

$$\psi(x) = e^{im_Q v \cdot x} \left[\chi_+(x) + \chi_-(x) \right] \quad (9.19)$$

where $\Gamma_\pm \chi_\pm = \chi_\pm$ for $\Gamma_\pm = \frac{1}{2}(1 \mp i\not{v})$. This works because $\Gamma_\pm \propto i\not{v} \mp m$, which gives zero when acting on $v_{\mathbf{k}\sigma}$ and $u_{\mathbf{k}\sigma}$ respectively, according to the Dirac equation, Eq. (1.88). The utility of the above definition for χ can be seen by inserting it into the free Lagrangian, since

$$\begin{aligned} \bar{\psi}(\not{\partial} + m_Q)\psi &= (\bar{\chi}_+ + \bar{\chi}_-)(im_Q\not{v} + m_Q + \not{\partial})(\chi_+ + \chi_-) \\ &= (\bar{\chi}_+ + \bar{\chi}_-)(2m_Q\Gamma_- + \not{\partial})(\chi_+ + \chi_-) \end{aligned} \quad (9.20)$$

Since $\Gamma_- \chi_+ = 0$, this shows that terms which diverge as $m_Q \rightarrow \infty$ drop out

of the part of the Lagrangian involving the particle part, χ_+ . It is for this reason that the large- m_Q limit is simplest to see using these variables. The next step is to integrate out the antiparticle part of the field, χ_- , and to leading order in $1/m_Q$ this amounts to eliminating it in terms of χ_+ using its equation of motion, leading to $\chi_- = -(\not{D}/2m_Q)\chi_+ + O(m_Q^{-2})$, and so

$$\psi(x) = e^{im_Q v \cdot x} \left[1 - \frac{1}{2m_Q} \not{D} \right] \chi_+(x) + O(m_Q^{-2}) \quad (9.21)$$

Using this in the Lagrangian $-\bar{\psi}(\not{D} + m_Q)\psi$ then leads to the following effective low-energy Lagrangian governing the dynamics of the slowly moving fermions:

$$\begin{aligned} \mathcal{L}_{\text{HQET}} &= -\bar{\chi}_+ \left[\not{D} - \frac{\not{D}\not{D}}{m_Q} + \frac{\not{D}\Gamma_- \not{D}}{2m_Q} \right] \chi_+ + O(m_Q^{-2}) \\ &= -\bar{\chi}_+ \left[\not{D} - \frac{\not{D}\not{D}}{2m_Q} - \frac{\not{D}\Gamma_+ \not{D}}{2m_Q} \right] \chi_+ + O(m_Q^{-2}) \\ &= \bar{\chi}_+ \left[i v^\mu D_\mu + \frac{1}{2m_Q} \left(D^\mu D_\mu - \frac{i}{2} \gamma^{\mu\nu} t_a F_{\mu\nu}^a \right) - \frac{(v^\mu D_\mu)^2}{2m_Q} \right] \chi_+ \end{aligned} \quad (9.22)$$

up to $O(m_Q^{-2})$ terms. These equations use the identities $\Gamma_+ + \Gamma_- = 1$ and $\Gamma_+ \not{D} \Gamma_+ = -i \Gamma_+ v^\mu D_\mu$, as well as $\not{D}\not{D} = D^\mu D_\mu - \frac{i}{2} \gamma^{\mu\nu} t_a F_{\mu\nu}^a$ where $\gamma^{\mu\nu} = \frac{1}{2}[\gamma^\mu, \gamma^\nu]$ and t_a is the appropriate representation matrix for the gauge generators (e.g. $t_a = eQ$ for electromagnetism and $t_a = \frac{1}{2} g_3 \lambda_a$ – with λ_a denoting the Gell-Mann matrices – for $SU_c(3)$).

9.1.2.1 Applications

The leading term in this Lagrangian is the term $i\bar{\chi} v \cdot D \chi$, and for N fields χ^i this has a large symmetry group $SU(2N)$, corresponding to unitary rotations on both the “flavor” index $i = 1, \dots, N$ and the (suppressed) 2-valued spinor index. The mixing of spin and flavor symmetries in this way is possible when $m_Q \rightarrow \infty$ because in this limit spin-orbit couplings disappear and spin essentially acts as just another internal symmetry.

To see what such a symmetry can imply, consider the decay, $\bar{B} \rightarrow D l^- \bar{\nu}$, where \bar{B} and D are mesons containing heavy b and c quarks respectively. This requires evaluating the matrix element $\langle D(p') | \bar{c} \gamma^\mu b | \bar{B}(p) \rangle$. HQET relates this to matrix elements measured in other reactions. To see this, define

$$\begin{aligned} \langle D(p') | \bar{c} \gamma^\mu c | D(p) \rangle &= f_D (p + p')^\mu \\ \langle \bar{B}(p') | \bar{b} \gamma^\mu b | \bar{B}(p) \rangle &= f_B (p + p')^\mu \\ \langle D(p') | \bar{c} \gamma^\mu b | \bar{B}(p) \rangle &= f_+ (p + p')^\mu + f_- (p - p')^\mu \end{aligned} \quad (9.23)$$

where the form factors f_D , f_B , and f_\pm are functions of the Lorentz-invariant

combination, $q^2 = -(p - p')^2$, of the appropriate meson momenta. Equivalently, since

$$\begin{aligned} q_D^2 &= -(p_D - p'_D)^2 = 2m_D^2(1 + v \cdot v') \\ q_B^2 &= -(p_B - p'_B)^2 = 2m_B^2(1 + v \cdot v') \\ q_{BD}^2 &= -(p_B - p_D)^2 = (m_B - m_D)^2 + 2m_B m_D(1 + v \cdot v') \end{aligned} \quad (9.24)$$

we may take them to be functions of the product of meson 4-velocities, $v \cdot v'$.

Now comes the main point. If the relative velocity of the D meson relative to the \bar{B} is small, then we may make use of the approximate $SU(4)$ symmetry which rotates the flavor and spin of the b and c quarks, and therefore relates the matrix element relevant in the decay to the purely B and D matrix elements. Keeping in mind the mass-dependence of our relativistic normalization for momentum states, this implies these form factors are not independent of one another:

$$f_B(v \cdot v') = f_D(v \cdot v') = f_+(v \cdot v') \sqrt{\frac{4m_B m_D}{m_B + m_D}} = -f_-(v \cdot v') \sqrt{\frac{4m_B m_D}{m_B - m_D}} \quad (9.25)$$

This implies that all four form factors (at low velocities) are determined by one unknown function, at least to leading order in $1/m_c$ and $1/m_b$. What is important here is this result is *exact* in the strong interactions since it is a consequence of general symmetry properties in the heavy-quark limit.

The reader is referred to the extensive literature for more detailed explanations and applications of these ideas.

9.2 Hard inelastic scattering: partons

After elastic collisions, perhaps the next simplest hadronic process to consider is very inelastic scattering. These kinds of collisions might be expected to be dominated by the collisions of the point-like quarks and gluons from which the hadrons are made, and these collisions should be relatively simple owing to the asymptotic freedom of the strong interactions, as discussed in Section 7.4. This property states that the strong interactions effectively become weak over short distances, and we see in this section that this ensures that many features of these collisions are well described perturbatively in the strong coupling α_3 , because the right language to describe that part of the process involving large energy exchanges should be the language of quarks and gluons (in this context collectively called *partons*). Perturbation theory in terms of quarks and gluons can then be applied to that subset of

the process. How the quark or gluon involved in the scattering process gets “found inside” the hadron is not a question perturbation theory can resolve. After being struck, the parton has enough energy to escape the hadron, but confinement subsequently requires that it produce additional partons, such that they bind off into colorless hadrons. Hard processes typically produce several such hadrons moving in the direction of the final-state parton, collectively called a *jet*. The question of how the jet forms is also not resolvable by perturbation theory. However, each of these questions is universal in a certain precise sense, meaning that once we learn how quarks and gluons are found inside of a hadron or how they turn into hadrons in one process at one energy, this information can be used to predict how it will occur in other processes and at other energies.

The key idea in doing this is called *factorization*. It is the fact, rigorously proven in some cases, that for sufficiently inclusive questions involving sufficiently high energy scattering, the problem of computing how scattering occurs divides into three parts. The hadron is described in terms of incoherent (free, independent) constituent quarks and gluons; the quarks and gluons undergo the hard scattering; and quarks and gluons produced in the scattering form into hadrons again. The middle part can be treated by perturbation theory. The first part cannot, but is universal; the description of a hadron in terms of quarks and gluons is the same for all processes and depends on energy in a predictable way. Therefore, measurements in one experiment can be used to predict the results of other experiments. The same is true for the final state part, provided one is only interested in sufficiently inclusive details of the final state. All of these statements are true only up to corrections of order $\Lambda_{\text{QCD}}^2/q^2$ with q^2 the invariant characterizing how high energy the process is.

To clarify these statements and present the details, we will begin with the simplest scattering process involving hadrons, hadron–lepton scattering, which is also where the ideas we discuss were originally tested. Then we discuss hadron–hadron collisions. Initially our discussion will be made at leading order in the strong coupling, but we end by discussing what happens beyond this order. The subject we are discussing is quite vast, and it is only our intention to introduce the relevant ideas and to present the techniques for a lowest-order treatment.

9.2.1 Deep inelastic scattering

Consider first the process $e^-p \rightarrow e^-X$, where X means any multi-hadron final state. When the energies are high and the electron makes more than

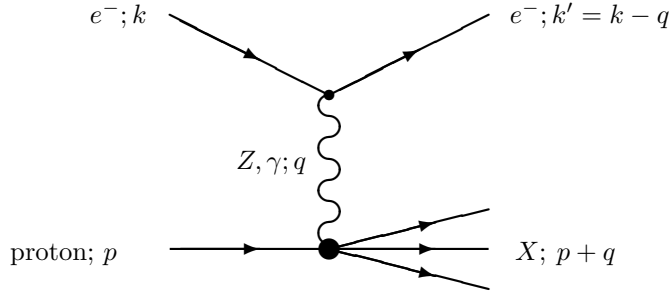


Fig. 9.1 Kinematic variables for deep inelastic scattering.

a glancing scattering from the proton, this process is called *deep inelastic scattering*.

9.2.1.1 Kinematics

The first step to analyzing the process $e^-p \rightarrow e^-X$ is to study the kinematics of the reaction. There is a standard notation for the kinematic variables which we shall adopt.

Define the incoming four-momenta to be

$$\text{electron: } k, \text{ with } k^2 = -m_e^2 \simeq 0; \quad \text{proton: } p, \text{ with } p^2 = -m_p^2 \quad (9.26)$$

and the final four-momenta to be,

$$\text{electron: } k', \text{ with } (k')^2 = -m_e^2 \simeq 0 \quad \text{Hadrons: } p + k - k' \quad (9.27)$$

as shown in Figure 9.1. We also define the momentum transfer,

$$q \equiv k - k' \quad (9.28)$$

The standard invariants for this system are

$$s \equiv -(p + k)^2, \quad \text{center of mass energy squared;} \quad (9.29)$$

$$q^2 \equiv (k - k')^2, \quad \text{virtuality of the photon;} \quad (9.30)$$

$$\nu \equiv \frac{-q \cdot p}{m_p}, \quad e^- \text{ energy loss in proton rest frame;} \quad (9.31)$$

$$x \equiv \frac{q^2}{-2p \cdot q} = \frac{q^2}{2m_p \nu}, \quad \text{see below} \quad (9.32)$$

$$y \equiv \frac{q \cdot p}{k \cdot p}, \quad e^- \text{ fractional energy loss in } p \text{ rest frame;} \quad (9.33)$$

$$W^2 = -(p + q)^2, \quad \text{mass squared of hadronic state } X \quad (9.34)$$

The interpretation of x , often called *Bjorken x* , is as follows. If we imagined

that the electron were striking a single massless quark, and that the quark scatter elastically, then x is the fraction of the proton's momentum the struck quark was carrying, evaluated in a frame where the proton is initially approaching the electron at very high energy (the infinite momentum frame or Breit frame). To see this, note that

$$\begin{aligned} x &\equiv \frac{q^2}{-2p \cdot q}, \text{ so} \\ 0 &= -2q \cdot (xp) - q^2 \\ 0 &\simeq -(xp)^2 - 2q \cdot (xp) - q^2 \\ 0 &\simeq -(xp + q)^2 \end{aligned} \quad (9.35)$$

so that, indeed, if the photon struck a quark with momentum fraction xp , then the final quark would be on shell. In passing from the second to the third line we have neglected the proton mass, which is appropriate if we are not concerned about $\Lambda_{\text{QCD}}^2/q^2$ corrections.

We will not ask for the differential cross section in terms of the specifics of the final-state X , but we will ask for the differential cross section in terms of the leptonic final state. Its final-state phase space is

$$\frac{1}{v_{\text{rel}} 2k^0 2p^0} \int \frac{d^3 k'}{(2\pi)^3 2k'^0} \quad (9.36)$$

For simplicity we will only discuss spin-averaged cross sections here, in which case there is no problem with performing the angular integration. We also systematically ignore the lepton mass, so $v_{\text{rel}} 2k^0 2p^0 = 2(s - m_p^2)$. In the proton rest frame, we may rewrite

$$\frac{1}{2(s - m_p^2)} \int \frac{d^3 k'}{(2\pi)^3 2k'^0} = \frac{1}{16\pi^2(s - m_p^2)} \int_{-1}^1 d \cos \theta_{kk'} k' dk' \quad (9.37)$$

Furthermore, in this frame, $y = k'^0/k^0$, $k^0 = (s - m_p^2)/(2m_p)$, $q^2 = -2k \cdot k' = 2(k^0)^2 y(1 - \cos \theta_{kk'})$, and $x = k^0(1 - \cos \theta_{kk'})/m_p$, so this can be re-expressed as

$$\begin{aligned} \frac{1}{2(s - m_p^2)} \int \frac{d^3 k'}{(2\pi)^3 2k'^0} &= \frac{1}{32\pi^2} \int_0^1 y dy \int_0^1 dx \\ &= \frac{1}{32\pi^2} \int_0^1 dx \int_0^{x(s - m_p^2)} \frac{q^2}{x^2(s - m_p^2)^2} dq^2 \end{aligned} \quad (9.38)$$

9.2.1.2 Structure functions

Let us now write down the most general form that the spin averaged cross section can take. This will define the “structure functions,” which parameterize our ignorance of hadronic cross sections.

The probability to scatter is given by

$$\sigma = \sum_f \left| \langle f | S | e^-(k), p(p) \rangle \right|^2 \quad (9.39)$$

where S is the S -matrix and \sum_f is a summation over all final states different from the initial state. Although we are largely ignorant of the strong coupling states and their interactions, it is certainly safe to expand perturbatively in the electromagnetic (or weak) interactions. Focusing on the photon-exchange graph in Figure 9.1, we see that S must contain at least two insertions of the electromagnetic current, of which one is the leptonic part of the current and one is the hadronic current;

$$\begin{aligned} & \langle f | S | e^-(k), p(p) \rangle \\ &= e^2 \int_{q,q'} \langle f | T [J_{\text{had}}^\mu(q) A_\mu(-q) J_{\text{lept}}^\nu(q') A_\nu(-q')] | e^-(k), p(p) \rangle \end{aligned} \quad (9.40)$$

plus terms higher order in the electromagnetic coupling. The A field operators tie off into a propagator, $\eta_{\mu\nu}/q^2(2\pi)^4\delta^4(q+q')$, and for the rest of the problem the electromagnetic and hadronic parts of the Hilbert space can be treated independently;

$$\begin{aligned} & \sum_f \left| \langle f | S | e^-(k), p(p) \rangle \right|^2 \\ &= \frac{e^4}{4v_{\text{rel}}k^0p^0} \int \frac{d^3k'}{(2\pi)^3 2k'^0} \frac{1}{q^4} \\ & \quad \times \langle e^-(k) | J^\mu(-q) | e^-(k') \rangle \langle e^-(k') | J^\nu(q) | e^-(k) \rangle \\ & \quad \times \langle p(p) | J_\mu(q) \sum_{f,\text{had}} | f, \text{had} \rangle \langle f, \text{had} | J_\nu(-q) | p(p) \rangle \\ &= \int \frac{2\pi\alpha^2 y dx dy}{q^4} L^{\mu\nu}(x, q^2, s) W_{\mu\nu}(x, q^2) \end{aligned} \quad (9.41)$$

The leptonic tensor is easily evaluated (apply crossing to Eq. (6.18)),

$$L_{\mu\nu} = 2(k_\mu k'_\nu + k'_\mu k_\nu - k \cdot k' \eta_{\mu\nu}) \quad (9.42)$$

and the hadronic tensor can at least be compactly expressed by noting that the final hadronic-state sum in Eq. (9.41) is a sum over a complete set of

hadronic final states, and is therefore the identity;

$$W^{\mu\nu} \equiv \frac{1}{2} \sum_{\sigma} \frac{1}{4\pi} \int d^4z e^{iq \cdot z} \langle p(p, \sigma) | [J_{\mu}(z), J_{\nu}(0)] | p(p, \sigma) \rangle \quad (9.43)$$

The factor of $1/4\pi$ is conventional and was taken into account in the last line of Eq. (9.41). The replacement of the ordered currents with the commutator is harmless, because when the positive frequency J operates on the proton state, it gives zero, for kinematic reasons – the energy of the proton cannot be lowered by q^0 .

The quantity $W^{\mu\nu}$ can only depend on the vectors q, p and the invariants which can be formed from them, q^2 and x . Based only on its tensorial structure, and the condition $q_{\mu} W^{\mu\nu} = 0$, we can express W in terms of three scalar functions of x and q^2 ;

$$W^{\mu\nu}(x, q^2) = \left(\eta^{\mu\nu} - \frac{q^{\mu} q^{\nu}}{q^2} \right) F_1(x, q^2) - \frac{\hat{p}^{\mu} \hat{p}^{\nu}}{p \cdot q} F_2(x, q^2) - i \epsilon^{\mu\nu\alpha\beta} \frac{q_{\alpha} p_{\beta}}{2p \cdot q} F_3(x, q^2) \quad (9.44)$$

Here $\hat{p}_{\mu} \equiv p_{\mu} - q_{\mu} p \cdot q / q^2 = p_{\mu} - (1/2x) q_{\mu}$. The quantities F_1 , F_2 , and F_3 are called *structure functions*, and they completely parameterize our ignorance of the hadronic physics involved in total (inclusive) hadronic cross sections. F_1 and F_2 are positive, while F_3 is zero for electromagnetic interactions but will be nonzero when we generalize to weak interactions (see Problem 9.4). If we were interested in spin-dependent cross sections, five additional structure functions linear in the spin vector s would be required.

For electromagnetic scattering, we find by contracting Eq. (9.44) with Eq. (9.42) that the differential cross-section is

$$\frac{d\sigma}{dx dy} = \frac{4\pi\alpha^2 y}{q^4} \left(q^2 F_1(x, q^2) + \left[(1-y) \frac{q^2}{xy^2} - x m_p^2 \right] F_2(x, q^2) \right) \quad (9.45)$$

9.2.1.3 Parton distribution functions

The hadronic current at spacetime point x , relevant in Eq. (9.43), is

$$J_{\mu}(x) \equiv - \sum_m \frac{2i}{3} \bar{u}_m(x) \gamma_{\mu} u_m(x) + \sum_m \frac{i}{3} \bar{d}_m(x) \gamma_{\mu} d_m(x) \quad (9.46)$$

with u, d the up and down type quarks and m an index running over the generations. The current operator can be re-expressed in terms of number operators for quarks; the current–current correlator, Eq. (9.43), is measuring the probability (not the amplitude, because there are two J insertions) for a quark to be in the proton, times the tensor such a quark would contribute to the scattering problem. At large momentum, the theory is weakly coupled

when described in terms of quarks and gluons, so this must be the right language for describing $W_{\mu\nu}(x, q^2)$.

With this in mind, we define the parton distribution functions (henceforth PDFs) $u_m, d_m, \bar{u}, \bar{d}, g(x, q^2)$, loosely, as follows; $u_m(x, q^2) dx$ is the probability that a probe of virtuality q^2 will find an up-type quark of generation m in the proton, with a momentum fraction between x and $x+dx$ of the full proton momentum. \bar{u}_m is the same for an up-type antiquark, g for a gluon, and so forth.

It may sound counterintuitive, but it *is* possible to find anti-quarks in a proton. Similarly, while we think of the proton as a bound state of three quarks, the gluon PDF turns out to account for most of the particles, and about half the momentum, of the proton. To see why, reconsider the collinear emission discussed in Subsection 6.7.2. An electron involved in a high energy scattering radiates off photons with probability of order $\alpha \ln(q^2/m_e^2)$. Similarly, even if we can describe a proton as a bound state of three quarks at low energy, in a hard scattering a quark will emit additional gluons with a probability of order $\alpha_3 \ln(q^2/\Lambda_{\text{QCD}}^2)$. However, Eq. (8.7) tells us that $\alpha_3(q^2) \sim 1/\ln(q^2/\Lambda_{\text{QCD}}^2)$. Therefore the gluon emission probability is order 1. The emitted gluons in turn can split into quark–antiquark pairs, also with probability $O(1)$. The PDFs at scale q^2 are defined already to include such radiations, and therefore the gluon and antiquark PDFs will be nonzero. The higher q^2 we consider, the more such radiations will occur, which means the PDFs will have q^2 dependence which can be predicted using calculations similar to those in Subsection 6.7.2. We return to this point in Subsection 9.2.3.

The splitting just described cannot change the total number of quarks minus antiquarks in any flavor. Therefore, the fact that the proton carries the flavor quantum numbers of two up quarks and a down quark is represented by a sum rule on the PDFs,

$$\int_0^1 dx \begin{Bmatrix} u(x, q^2) - \bar{u}(x, q^2) \\ d(x, q^2) - \bar{d}(x, q^2) \\ s(x, q^2) - \bar{s}(x, q^2) \end{Bmatrix} = \begin{Bmatrix} 2 \\ 1 \\ 0 \end{Bmatrix} \quad \text{in a proton} \quad (9.47)$$

To clarify, the PDFs can also be defined for any other hadron. Properly we should put an index on u_m, g_m, \dots to indicate which hadron they refer to, since the values depend on the hadron being probed. We will not do so in the following to avoid cluttering the notation. For a neutron the $u - \bar{u}$ integral gives 1 and the $d - \bar{d}$ integral gives 2, while for a π^+ they give 1 and -1 *etc.* The quarks represented by the difference $u - \bar{u}(x, q^2)$ are usually referred to as the *valence quarks*, while the residual set of quarks and anti-quarks

are called the *sea quarks*. It turns out that the valence quarks are mostly at fairly large x , while the sea dominates at small x . Also, the total momentum carried by all quarks, antiquarks, and gluons must equal the total hadron momentum, so

$$\int_0^1 dx x \left(g(x, q^2) + \sum_m \left[u_m(x, q^2) + d_m(x, q^2) + \bar{u}_m(x, q^2) + \bar{d}_m(x, q^2) \right] \right) = 1 \quad (9.48)$$

To relate the parton distribution functions to the structure functions, we need to determine the contribution of a quark with momentum fraction x' to the structure functions. Therefore, consider the case of a single quark with momentum fraction x' . The cross-section for this quark to scatter from an electron is

$$\frac{Q_q^2 e^4}{v_{\text{rel}} x' p^0 k^0} \int \frac{d^3 k' d^3 p'}{(2\pi)^6 2k'^0 2p'^0} (2\pi)^4 \delta^4(k + x'p - k' - p') \frac{1}{q^4} L^{\mu\nu} \times 2 \left(x' p_\mu p'_\nu + x' p_\nu p'_\mu - x' p \cdot p' \eta_{\mu\nu} \right) \quad (9.49)$$

where Q_q^2 is the charge squared of the quark in units of the electric charge of the proton. The final quantity in parenthesis is the result of evaluating the Dirac trace, summed on the final and averaged on the initial spins, for the quark. To evaluate this quantity, it is convenient to write

$$\int \frac{d^3 p'}{(2\pi)^3 2p'^0} = \int \frac{d^4 p'}{(2\pi)^4} 2\pi \delta(p'^2) \quad (9.50)$$

The p' integration is then performed trivially, using the energy and momentum conserving delta function; it forces $p' = (x'p + q)$. We can also treat the k' integration in the same way as in Eq. (9.38). This gives

$$\int \frac{2\pi \alpha^2 y dy dx}{q^4} \frac{1}{4\pi x'} 2\pi \delta((x'p + q)^2) L^{\mu\nu} 2Q_q^2 \left(x' p_\mu p'_\nu + x' p_\nu p'_\mu - x' p \cdot p' \eta_{\mu\nu} \right) \quad (9.51)$$

where p' is now defined to be $(x'p + q)$. The delta function can be rewritten (neglecting $x'^2 m_p^2$ next to q^2) as

$$\delta((x'p + q)^2) = \frac{1}{|2p \cdot q|} \delta \left(x' + \frac{q^2}{2p \cdot q} \right) = \frac{x}{q^2} \delta(x' - x), \quad (9.52)$$

which just enforces that x' of the initial quark and x of the leptonic kinematics are the same. The cross-section due to scattering from the quark has now become

$$\int \frac{2\pi \alpha^2 y dy dx}{q^4} L^{\mu\nu} \frac{Q_q^2}{q^2} \delta(x' - x) \left(x p_\mu p'_\nu + x p_\nu p'_\mu - x p \cdot p' \eta_{\mu\nu} \right) \quad (9.53)$$

Comparing with Eq. (9.41), and with the parameterized form of $W_{\mu\nu}$ written in Eq. (9.44), we can identify

$$\begin{aligned} & \left(\eta_{\mu\nu} - \frac{q_\mu q_\nu}{q^2} \right) F_1(x, q^2) - \frac{\hat{p}^\mu \hat{p}^\nu}{p \cdot q} F_2(x, q^2) \\ &= Q_q^2 \delta(x - x') \frac{1}{q^2} \left(x p_\mu p'_\nu + x p_\nu p'_\mu - x p \cdot p' \eta_{\mu\nu} \right) \end{aligned} \quad (9.54)$$

Now $x p \cdot p' = x p \cdot (x p + q) \simeq x p \cdot q = -q^2/2$. Further,

$$\begin{aligned} \frac{x}{q^2} (p_\mu p'_\nu + p_\nu p'_\mu) &= -\frac{1}{2 p \cdot q} (2 x p_\mu p_\nu + p_\mu q_\nu + p_\nu q_\mu) \\ &= -\frac{x}{p \cdot q} \left(p_\mu p_\nu + \frac{1}{2x} (p_\mu q_\nu + p_\nu q_\mu) + \frac{1}{4x^2} q_\mu q_\nu \right) - \frac{1}{2} \frac{q_\mu q_\nu}{q^2} \\ &= -x \frac{1}{p \cdot q} \hat{p}_\mu \hat{p}_\nu - \frac{1}{2} \frac{q_\mu q_\nu}{q^2} \end{aligned} \quad (9.55)$$

Therefore, the contribution of a quark, of momentum fraction x' , to the structure functions $F_1(x, q^2)$, $F_2(x, q^2)$ is

$$\begin{aligned} F_1(x, q^2) &= \frac{Q_q^2}{2} \delta(x - x') \\ F_2(x, q^2) &= x Q_q^2 \delta(x - x') \end{aligned} \quad (9.56)$$

More generally, in terms of the parton distribution functions defined above, the structure functions are

$$\begin{aligned} F_1(x, q^2) &= \frac{1}{2} \sum_m \left[\frac{4}{9} (u_m + \bar{u}_m)(x, q^2) + \frac{1}{9} (d_m + \bar{d}_m)(x, q^2) \right] \\ F_2(x, q^2) &= 2x F_1(x, q^2) \end{aligned} \quad (9.57)$$

Therefore the structure functions directly determine the parton distribution functions.

At this point the reader may ask whether we have just replaced one parameterization of our ignorance, in terms of F_1 and F_2 , for another, in terms of the parton distribution functions. This is not the case. For one thing, the description of the proton in terms of partons gives us the relation $F_2(x, q^2) = 2x F_1(x, q^2)$ (the Callan–Gross sum rule), which need not be true in general. This relation is obeyed by the data, up to $O(\alpha_3)$ corrections (which can in turn be computed by treating the scattering between the electron and quark at higher order in α_3 , and including scattering from gluons which first occurs at this order). Second, there are definite predictions for the q^2 dependence of the PDFs, which we present in Subsection 9.2.3. Finally, the same PDFs also apply to other processes, for instance, neutrino scattering from hadrons or hadron–hadron scattering.

Let us emphasize a few points about this calculation.

- (i) The calculation we have just outlined, for the cross section $e^-p \rightarrow e^-X$, is valid to leading order in α_3 . Corrections can be systematically computed by performing a loopwise expansion of the scattering cross section of an e^- from a parton of momentum fraction x . Such corrections are now known completely at the two-loop level, and the agreement between theory and data is very good.
- (ii) The partonic treatment is insufficient if we are interested in corrections to the cross section suppressed by powers of $\Lambda_{\text{QCD}}^2/q^2$ with respect to the leading order value we have just discussed. This is not just because we have sometimes made approximations valid in the $\Lambda_{\text{QCD}}^2 \ll q^2$ limit; the parton distribution function approach itself is insufficient to find such corrections, because it is necessary to take into account more information about the proton than its partonic structure. The PDFs describe a picture of the proton as a set of quarks and gluons, overlapping in an uncorrelated (incoherent) way. However, hadrons are bound states, so there must exist some non-perturbative correlations between the partons, describing the way in which they are bound in the hadron. The scale separation between the binding forces, of order Λ_{QCD} , and the probed scale, set by q^2 , is what ensures that these corrections are small. Such corrections are called *higher twist*.
- (iii) The description in terms of partons is useful at large q^2 and moderate x , which is more restrictive than large center-of-mass energy s . The total inelastic cross section is dominated by the $q^2 \lesssim \Lambda_{\text{QCD}}^2$ region, even at arbitrarily large s .
- (iv) It was essential that we summed over all hadronic final states in passing from Eq. (9.41) to Eq. (9.43), because that sum gives the identity which then drops out and leaves a hadronic expression depending only on the initial state. It is possible to go beyond this in a limited way, by asking either questions involving only the general kinematics of the final state, or *semi-inclusive* questions, which we will address in the next subsection.

9.2.2 Hadron-hadron collisions

Consider next the scattering of two hadrons at high energy. For processes involving a large invariant energy scale, the right description will again be in

terms of the quark and gluon constituents. In this case, each hadron must be described in terms of PDFs, so the prediction for the rate for some process will now involve an integral over the PDFs of each hadron. Again, we begin with the simplest process to consider.

9.2.2.1 The Drell–Yan process and heavy-quark production

The Drell–Yan process is the production of a lepton–antilepton pair, or *dilepton*, together with any hadronic final state X , in hadron–hadron scattering. The lepton pair $l\bar{l}$ must be produced by an electromagnetic (or weak) interaction, $e\bar{l}\gamma^\mu A_\mu l$, but since there is no photon in the initial state, the S -matrix must be quadratic in electromagnetic interactions,

$$S_{\text{Drell–Yan}} = \langle \bar{l}lX | e^2 J_{\text{lept}}^\mu A_\mu J_{\nu \text{ hadr}} A^\nu | h_1 h_2 \rangle \quad (9.58)$$

Again, the leptonic and hadronic parts of the Hilbert space factorize and the A insertions combine into a gauge-field propagator. Therefore, the total cross-section is

$$\begin{aligned} \sigma = & \int_{q,q'} \int \frac{d^3k d^3k'}{(2\pi)^6 4k^0 k'^0} \frac{e^4}{q^2 q'^2} \langle 0 | J^\mu(q) | l(k) l'(k') \rangle \langle l(k) l'(k') | J^\nu(-q') | 0 \rangle \\ & \times \sum_X \langle h_1 h_2 | J_\mu(-q) | X \rangle \langle X | J_\nu(q') | h_1 h_2 \rangle \end{aligned} \quad (9.59)$$

Evaluating the leptonic parts will give a leptonic tensor which we can find from Eq. (9.42) using crossing, and momentum conserving delta functions which force $k + k' = q = q'$.

Again, it is possible to define a current-current tensor $W_{\mu\nu}$ for the hadrons, and explore its possible structure. However, when $-q^2 \gg \Lambda_{\text{QCD}}^2$ (that is, for large invariant-mass dileptons), we can take advantage of the fact that $\alpha_3(-q^2) \ll 1$ and work in terms of quark and gluon operators. The current $J_\mu = \sum_q Q_q \bar{q} \gamma_\mu q$ probes the partonic (specifically, quark) content of the hadrons. Unlike leading-order deep inelastic scattering (DIS), \bar{q} and q must both act on partons from the initial hadrons – otherwise the process represents the creation of the $l\bar{l}$ pair from one hadron, which is not kinematically allowed.

Diagrammatically, this leading-order contribution to Drell–Yan would be drawn as in Figure 9.2. Here the blobs on the left mean the two initial-state hadrons, with the several lines leaving each representing the constituent

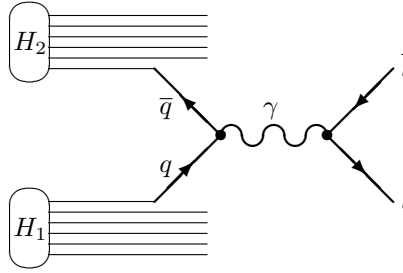


Fig. 9.2 The Drell–Yan process.

partons. One parton from each hadron – a q and a \bar{q} – interact electromagnetically to form the $\bar{l}l$ pair (with $l = e, \mu$, or τ). This picture generalizes slightly the meaning of Feynman diagrams we have used so far; when one line is extracted from a blob, it means the relevant PDF is used to find the probability of finding that parton in the hadron represented by the blob. The other lines from the blob are not terminated to indicate that the final hadronic state is summed over.

We can recycle Eq. (6.32) to find the differential rate for a quark of momentum $x_1 p$ and an antiquark of momentum $x_2 p'$ to form a dilepton pair:

$$\sigma(q\bar{q} \rightarrow \bar{l}l) = \frac{4\pi\alpha^2}{9x_1x_2s} Q_q^2 \quad (9.60)$$

Here Q_q^2 is the charge squared of the quark. x_1x_2s is the Mandelstam s for the quark–antiquark pair, as opposed to the hadron–hadron system, which we still write as s . The extra factor of $1/3$ relative to Eq. (6.32) comes about because we are averaging over the initial quark colors; the annihilation can only take place if their colors are the same.

The four-momentum of the dilepton pair system is $x_1 p^\mu + x_2 p'^\mu$. Therefore, $q^2 = -x_1x_2s$. Further, the spatial component of q transverse to p and p' in the center-of-mass frame, q_\perp , vanishes. However, the dilepton center of mass is not generally at rest in the CM frame. Instead, its motion is along the beam axis in the hadron CM frame. It is conventional to define the *rapidity* of a particle or set of particles, y , as the natural log of the boost, along the beam axis, to go from the center of mass of the hadronic system to the frame where the center of mass of the particles in question moves (if at all) only transversely to the beam axis. Thus, $y = \frac{1}{2} \ln[(q^0 + q_z)/(q^0 - q_z)]$, with q^0 and q_z the energy and momentum along the beam axis in the CM frame.

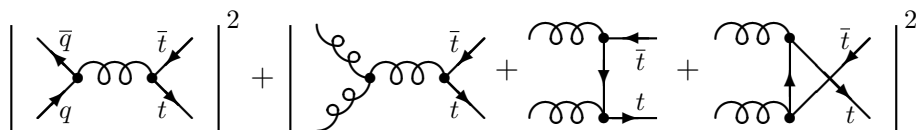


Fig. 9.3 Heavy-quark production.

For q^μ given above, the rapidity is therefore $y = \frac{1}{2} \ln(x_1/x_2)$. Combining terms, and integrating over the PDFs of both hadrons, the differential cross section in terms of q^2 and y becomes

$$\begin{aligned} \frac{d\sigma}{dq^2 dy} &= \int_0^1 dx_1 dx_2 \sum_{q=u,d,s,\dots} (\bar{q}_{h_1}(x_1) q_{h_2}(x_2) + q_{h_1}(x_1) \bar{q}_{h_2}(x_2)) \\ &\quad \times \frac{4\pi\alpha^2 Q_q^2}{9|q^2|} \delta(|q^2| - x_1 x_2 s) \delta\left(y - \frac{1}{2} \ln \frac{x_1}{x_2}\right) \end{aligned} \quad (9.61)$$

Here $q_h(x)$, $\bar{q}_h(x)$ are the quark and antiquark PDFs in hadron h . This expression predicts that

- (i) the Drell–Yan cross section and rapidity distribution can be determined simply in terms of the same PDFs measured in deep inelastic scattering, and
- (ii) Drell–Yan $\bar{l}l$ pairs have center-of-mass motion along the beam axis of the colliding hadrons.

At higher order in α_3 , the relation between Drell–Yan and DIS cross sections becomes more complicated but remains computable. Also, with $O(\alpha_3)$ probability the center-of-mass motion of the $\bar{l}l$ pair can deviate strongly from the beam axis. Again, the size of this correction is computable perturbatively.

Next, consider the production of heavy-quark pairs, for instance, $b\bar{b}$ or $t\bar{t}$ pairs, in hadron–hadron collisions. We skip straight to the analysis in terms of partons. There are two production mechanisms, shown in Figure 9.2. The spin and color-averaged differential cross section for each can be evaluated using the techniques of Chapter 6, though the gluon external lines must be treated with care – one cannot replace the polarization sum with $\eta_{\mu\nu}$ as can be done for a photon, it must be evaluated only summing over the physical transverse polarization states. The parton level differential cross

sections are

$$\frac{d\hat{\sigma}}{dt}(q\bar{q} \rightarrow q'\bar{q}') = \frac{4\pi\alpha_3^2}{9\hat{s}^4} \left[(m^2 - \hat{t})^2 + (m^2 - \hat{u})^2 + 2m^2\hat{s} \right] \quad (9.62)$$

$$\begin{aligned} \frac{d\hat{\sigma}}{dt}(gg \rightarrow q'\bar{q}') = \frac{\pi\alpha_3^2}{8\hat{s}^2} & \left[\frac{6(m^2 - \hat{t})(m^2 - \hat{u})}{\hat{s}^2} - \frac{m^2(\hat{s} - 4m^2)}{3(m^2 - \hat{t})(m^2 - \hat{u})} \right. \\ & + \frac{4(m^2 - \hat{t})(m^2 - \hat{u}) - 8m^2(m^2 + \hat{t})}{3(m^2 - \hat{t})^2} + (\hat{t} \leftrightarrow \hat{u}) \\ & \left. - 3 \frac{(m^2 - \hat{t})(m^2 - \hat{u}) + m^2(\hat{u} - \hat{t})}{\hat{s}(m^2 - \hat{t})} + (\hat{t} \leftrightarrow \hat{u}) \right] \end{aligned} \quad (9.63)$$

with \hat{s} , \hat{t} , and \hat{u} the Mandelstam variables defined in terms of x_1p and x_2p' rather than p and p' , and m^2 the mass squared of the heavy quark. Integrating over t and defining $z \equiv 4m^2/\hat{s}$ gives

$$\begin{aligned} \hat{\sigma}(q\bar{q} \rightarrow q'\bar{q}') &= \frac{4\pi\alpha_3^2}{27\hat{s}} (2 + z\sqrt{1-z}) \\ \hat{\sigma}(gg \rightarrow q'\bar{q}') &= \frac{\pi\alpha_3^2}{48\hat{s}} \left[-(28+31z)\sqrt{1-z} + (16+16z+z^2) \ln \frac{1+\sqrt{1-z}}{1-\sqrt{1-z}} \right] \end{aligned} \quad (9.64)$$

The full production cross section is

$$\begin{aligned} \sigma_{h_1 h_2 \rightarrow q' \bar{q}'} &= \int_0^1 dx_1 dx_2 \left[g_{h_1}(x_1) g_{h_2}(x_2) \hat{\sigma}_{gg \rightarrow q' \bar{q}'}(\hat{s} = x_1 x_2 s) \right. \\ & \quad \left. + (q_{h_1}(x_1) \bar{q}_{h_2}(x_2) + \bar{q}_{h_1}(x_1) q_{h_2}(x_2)) \hat{\sigma}_{q\bar{q} \rightarrow q' \bar{q}'}(\hat{s} = x_1 x_2 s) \right] \end{aligned} \quad (9.65)$$

Here we have written explicitly which hadron each PDF refers to, but we have not written the scale where they should be evaluated. The correct scale is approximately $\mu^2 = \hat{s}$.

9.2.2.2 Jets and fragmentation

We saw in Chapter 4 and Chapter 6 that electroweak processes can occur with $q\bar{q}$ pairs in the final state. Similarly, hadron–lepton and hadron–hadron scattering can produce quark or gluon final-state particles.

The differential cross sections needed in light quark and gluon production

are

$$\frac{d\hat{\sigma}}{d\hat{t}} = \frac{\pi\alpha_3^2}{\hat{s}^2} \times \left\{ \begin{array}{ll} q_1 q_2 \rightarrow q_1 q_2 & : \quad \frac{4}{9} \frac{\hat{s}^2 + \hat{u}^2}{\hat{t}^2} \\ q_1 q_1 \rightarrow q_1 q_1 & : \quad \frac{4}{9} \left(\frac{\hat{s}^2 + \hat{u}^2}{\hat{t}^2} + \frac{\hat{s}^2 + \hat{t}^2}{\hat{u}^2} \right) - \frac{8}{27} \frac{\hat{s}^2}{\hat{t}\hat{u}} \\ q_1 \bar{q}_1 \rightarrow q_2 \bar{q}_2 & : \quad \frac{4}{9} \frac{\hat{t}^2 + \hat{u}^2}{\hat{s}^2} \\ q_1 \bar{q}_1 \rightarrow q_1 \bar{q}_1 & : \quad \frac{4}{9} \left(\frac{\hat{s}^2 + \hat{u}^2}{\hat{t}^2} + \frac{\hat{u}^2 + \hat{t}^2}{\hat{s}^2} \right) - \frac{8}{27} \frac{\hat{u}^2}{\hat{s}\hat{t}} \\ q\bar{q} \rightarrow gg & : \quad \frac{32}{27} \frac{\hat{u}^2 + \hat{t}^2}{\hat{u}\hat{t}} - \frac{8}{3} \frac{\hat{u}^2 + \hat{t}^2}{\hat{s}^2} \\ gg \rightarrow q\bar{q} & : \quad \frac{1}{6} \frac{\hat{u}^2 + \hat{t}^2}{\hat{u}\hat{t}} - \frac{3}{8} \frac{\hat{u}^2 + \hat{t}^2}{\hat{s}^2} \\ qg \rightarrow qg & : \quad -\frac{4}{9} \frac{\hat{u}^2 + \hat{s}^2}{\hat{u}\hat{s}} + \frac{\hat{u}^2 + \hat{s}^2}{\hat{t}^2} \\ gg \rightarrow gg & : \quad \frac{9}{2} \left(3 - \frac{\hat{u}\hat{t}}{\hat{s}^2} - \frac{\hat{u}\hat{s}}{\hat{t}^2} - \frac{\hat{s}\hat{t}}{\hat{u}^2} \right) \end{array} \right. \quad (9.66)$$

where as before, \hat{s} , \hat{t} , and \hat{u} are parton-level Mandelstan variables, and where one must remember to integrate over only half the final state phase-space when the final particles are identical, gg or $q_1 q_1$. Here the index on q is used just to distinguish whether the quarks involved in a scattering are of the same or different species. Each particle emerges with momentum perpendicular to the beam axis of $k_\perp^2 = \hat{t}\hat{u}/\hat{s}$. Their rapidities are $\frac{1}{2} \ln(x_1 t/x_2 u)$ and $\frac{1}{2} \ln(x_1 u/x_2 t)$.

Several of the cross sections possess $1/\hat{t}^2$ or $1/\hat{u}^2$ behavior, which leads to divergent small k_\perp^2 cross-sections. However, it is k_\perp^2 , not \hat{s} , which sets the scale for determining α_3 and the PDFs. (To see this, note that diagrams with $1/\hat{t}^2$ behavior behave in this way because of t -channel gluon propagators. The square of the momentum carried by the propagator is $-(p-k)^2 = \hat{t}$, and this is what controls the scale for α_3 , not the \hat{s} of the collision.) When k_\perp^2 is not large compared to Λ_{QCD}^2 , a treatment in terms of PDFs and perturbation theory is not reliable. Note that the total hadron-hadron cross

section is dominated by small angle and small x processes, which perturbative techniques cannot reliably compute.

To compute the k_\perp and rapidity distribution of light, large k_\perp partons, one must integrate the differential cross sections given above over x_1 and x_2 with the relevant PDFs inserted. But this does not answer the question of what will be detected in an experiment. How do such final-state partons turn into hadrons? That process begins with final-state radiation. We saw in Subsection 6.7.2 that initial-state radiation occurs with probability $\alpha \log^2(q^2/m^2)$. The calculation carries through essentially unchanged for final-state radiation. In QCD the coupling α is replaced by α_3 and the logs are cut off by the QCD scale Λ_{QCD} , so the rate of final-state radiation is of order $\alpha_3 \log^2(q^2/\Lambda_{\text{QCD}}^2)$. Here one log arises from the range in angle between the final states, $\int_{\Lambda_{\text{QCD}}/q}^1 d\theta/\theta$, and the other comes from the energy fraction of the radiated state, $\int_{\Lambda_{\text{QCD}}/q}^1 dx/x$. Hard, large angle radiations are rare, $O(\alpha_3)$, and can be studied perturbatively. On the other hand, soft and collinear radiations are common. As remarked before, Eq. (8.7) can be used to infer that $\alpha_3(q^2) \sim 1/\log(q^2/\Lambda_{\text{QCD}}^2)$. Therefore there are typically several final-state radiations, which can in turn also radiate. Therefore the parton becomes several partons, most of them soft and moving in almost the same direction as the original parton. These bind off into a collection of hadrons, a process called *hadronization*. The two processes together – final state radiation and hadronization – are called *fragmentation*. The collection of hadrons, all moving in close to the direction of the original parton, are called a *jet*.

Hard (large-momentum fraction), large angle final state radiations are rare, $O(\alpha_3/\pi)$, and computable perturbatively. They have been extensively studied and there is good agreement between observation and theory. However, this subject lies beyond the scope of this book. The details of the soft and collinear radiation and hadronization involve the scale Λ_{QCD} and cannot be studied perturbatively. However, some observables are universal. The origin of the hard parton has little bearing on its fragmentation into a jet, because fragmentation occurs after the parton has moved a distance $\geq \Lambda_{\text{QCD}}^{-1}$ from the production site, and can therefore only interact through soft momentum exchange with energies $\sim \Lambda_{\text{QCD}}$. Therefore, questions involving high-energy final-state hadrons, while non-perturbative, are independent of the originating process which produced the parton, with corrections which are either perturbative and computable, or suppressed by powers of $\Lambda_{\text{QCD}}^2/k_\perp^2$.

With this in mind, we can define the fragmentation function $D_{h/q}(z, q^2) dz$, as the probability that among the hadrons produced by a quark, one will be a hadron of type h with momentum fraction between z and $z + dz$ of the originating quark's momentum. Here q^2 is the virtuality of the process which produced the quark ($-\hat{t}$ in a t -channel process, \hat{s} in an s -channel process). That is, roughly,

$$D_{h/q} = \sum_X |\langle X, h(zp) | q(p) \rangle|^2 \quad (9.67)$$

$D_{h/\bar{q}}$ and $D_{h/g}$ are defined similarly. To the extent that CP is a valid symmetry, $D_{h/q} = D_{\bar{h}/\bar{q}}$ and $D_{h/g} = D_{\bar{h}/g}$.

Like the PDFs, the fragmentation functions D cannot be evaluated perturbatively, but must be determined from experiment. Once determined in one experiment (typically e^+e^- annihilation), however, they can be used to make predictions about high energy hadron yields in other experiments. Furthermore, like the PDFs, they have weak and computable q^2 dependence. We turn to this issue next.

9.2.3 Scale dependence of parton-distribution functions

The parton-distribution functions $u_m, \bar{u}_m, d_m, \bar{d}_m, g(x, q^2)$ are *a priori* unknown functions of x . However, we *can* say something about their dependence on q^2 .

The first point is that the PDFs should be approximately scale independent. This is because, at the relevant scales involved, the strong coupling is weak, $\alpha_3 \ll 1$. If a quark were a free particle, like an electron, then the energy-scale dependence of its interactions would be trivial, just as the energy dependence of the leptonic tensor $L_{\mu\nu}$ of Eq. (9.42) is. Therefore, a quark with momentum fraction x , analyzed at one momentum q_1^2 , should, at leading order, still be a quark of momentum fraction x when analyzed at a different momentum q_2^2 . Therefore, at leading order in α_3 , the PDFs are q^2 independent;

$$u_m(x, q_1^2) = u_m(x, q_2^2) \text{ at leading order in } \alpha_3 \quad (9.68)$$

This is in good agreement with experimental data. For instance, the structure functions vary slowly with q^2 , $F_{1,2}(x, 2q^2)/F_{1,2}(x, q^2) \sim 1 + O(\alpha_3)$.

The origin of the scale dependence in the structure functions comes about because of initial-state radiation. Suppose that a proton contains a quark with momentum fraction x . Then, as we saw in Subsection 6.7.2, there is a substantial chance that, when probed with virtuality q^2 , the quark will emit collinear radiation, and be observed with a smaller momentum

fraction. In Subsection 6.7.2 we saw that, for an electron, the probability for a photon emission with energy fraction between x' and $x' + dx'$ was $\simeq \alpha/2\pi \log(q^2/m_e^2)[1 + (1-x')^2]/x' dx'$. The same calculation can be carried over to gluon radiation. The difference is that a color factor for the vertex, and a summation over the final-state colors of the quark and gluon, must be included. If the incoming quark is in color i , the coupling constants and color factors for the (final color summed) emission process can be taken over from the QED case with the replacement

$$e^2 \rightarrow g_3^2 \sum_{j\alpha} \left| \frac{\lambda_{ij}^\alpha}{2} \right|^2 = \frac{g_3^2}{4} \sum_{j\alpha} \lambda_{ij}^\alpha \lambda_{ji}^\alpha \equiv C_F g_3^2, \quad C_F = \frac{4}{3} \quad (9.69)$$

where λ_{ij}^α is the Gell-Mann matrix introduced in Eq. (1.186). The quantity C_F is called the quadratic Casimir of the fundamental representation, and is sometimes written $C_2(F)$.

The definition of the PDFs at scale q^2 is, that they have already had initial-state radiation, henceforward *splitting*, up to the scale q^2 taken into account in their definition. However, when we probe with a larger q^2 , the probability of splitting is increased. The PDF must be modified to include this additional chance that the parton has split by collinear radiation. For instance, in passing from the scale q^2 to the scale λq^2 , the distribution $u_m(x, q^2)$ would change by

$$u_m(x, \lambda q^2) = u_m(x, q^2) \left(1 - \frac{C_F \alpha_3}{2\pi} \log \left(\frac{\lambda q^2}{q^2} \right) \int dx' \frac{1 + (1-x')^2}{x'} \right) + \text{gain term} \quad (9.70)$$

where the first term is probability that a quark of momentum fraction x survives without splitting (which is 1 minus the probability of splitting), and the second “gain term” is the probability that some larger x parton undergoes a splitting which produces an up quark of momentum fraction x . If the only process to consider were $q \rightarrow qg$ splitting, then the rate at which quarks of momentum fraction x would be produced would be

$$\text{gain term}_{q \rightarrow qg} = \frac{C_F \alpha_3}{2\pi} \log \left(\frac{\lambda q^2}{q^2} \right) \int_x^1 \frac{dx'}{x'} u_m(x', q^2) \left(\frac{1 + (x/x')^2}{1 - (x/x')} \right) \quad (9.71)$$

Here x' is the momentum fraction of the quark that will radiate down to momentum fraction x . For the quark's momentum fraction to change from x' to x , it must lose $x' - x$ of the proton's momentum, which is a fraction $1 - (x/x')$ of its starting momentum. Also note that the integration measure is dx'/x' . The $1/x'$ is a Jacobian which arises because, when a quark of

momentum fraction x' loses a gluon in the range $[y, y + dy]$ of that quark's momentum, the gluon energy is in the range $[x'y, x'y + x'dy]$.

Note that both the integral over x' in the first “loss” expression, in Eq. (9.70), and over the “gain” expression, in Eq. (9.71), are logarithmically divergent. The first expression is log divergent as $x' \rightarrow 0$, the second as $x' - x \rightarrow 0$. The first divergence arises from the large probability for a quark of momentum fraction x to lose a tiny bit of momentum by a very soft gluon radiation. The second divergence arises from the large probability that a slightly higher momentum-fraction quark will lose a tiny bit of momentum to a gluon radiation, and become a momentum-fraction x quark. These divergences cancel provided that $u_m(x, q^2)$ is a smooth function. They are handled by defining

$$\frac{1}{(1-x)_+} = \frac{1}{1-x}(x < 1); \quad \int_0^1 dx \frac{f(x)}{(1-x)_+} = \int_0^1 dx \frac{f(x) - f(1)}{1-x} \quad (9.72)$$

That is, $1/(1-x)_+$ is $1/(1-x)$ everywhere but at zero, where there is a negative delta function which ensures that the second condition above will be met. In terms of this, Eq. (9.70) and Eq. (9.71) can be combined, and converted into differential form in the change of momentum scale, as (defining $y = x/x'$ of Eq. (9.71))

$$\frac{du_m(x, q^2)}{d \log(q^2)} = \frac{C_F \alpha_3}{2\pi} \int_x^1 \frac{dy}{y} \left(\frac{1+y^2}{(1-y)_+} + \frac{3}{2} \delta(1-y) \right) u_m(x/y, q^2) \quad (9.73)$$

The entire loss term, Eq. (9.70), appears in the $(1-y)_+$ condition in the denominator. In fact, this slightly over-subtracts, because the actual loss term involves $\int (1+(1-x)^2) dx/x$, and the subtraction would be appropriate if this were $\int (2) dx/x$. Therefore we had to put in the factor of $\frac{3}{2} \delta(1-y)$ to reproduce the loss term correctly. This factor $3/2$ is the same as the $3/2$ which appears in Eq. (6.88).

Because it is so easy to emit a soft gluon, the gluon PDF will diverge at least as $g(x, q^2) \sim 1/x$. The total number of gluons contained in the proton is therefore (at least) logarithmically divergent. This is not problematic, however, because no experiment counts the number of gluons in the proton, and because the total momentum fraction carried by gluons, $\int dx xg(x, q^2)$, is well behaved so long as $g(x, q^2) < o(1/x^2)$, which is the case.

To compute the scale dependence of the PDFs, we need to determine not only how the quark can break up by emitting a gluon, $q \rightarrow qg$, but also how a gluon can break up, $g \rightarrow gg$ and $g \rightarrow q\bar{q}$. The analogous expressions are

$$e \rightarrow e\gamma \text{ or } q \rightarrow qg \quad : \quad \left[C_F = \frac{4}{3} \right] \frac{\alpha}{2\pi} \frac{1 + (1-x)^2}{x}$$

$$\begin{aligned}
\gamma \rightarrow e^+e^- \text{ or } g \rightarrow q\bar{q} & : \left[\frac{C_F d_F}{d_A} = \frac{1}{2} \right] \frac{\alpha}{2\pi} (x^2 + (1-x)^2) \\
(\text{QCD only}) \quad g \rightarrow gg & : [C_A = 3] \frac{\alpha}{2\pi} \frac{1+x^4+(1-x)^4}{x(1-x)} \\
& = \frac{6\alpha}{2\pi} \left(\frac{x}{1-x} + \frac{1-x}{x} + x(1-x) \right) \quad (9.74)
\end{aligned}$$

The quantity in brackets is the result of summing over final state colors, and should be dropped for the QED processes. The new group-theoretic factors are $d_F = 3$, the number of colors of quarks, $d_A = 8$, the number of types of gluons, and $C_A = 3$, the analog of C_F but with adjoint representation objects like the gluons; $\sum_{\beta\gamma} f_{\alpha\beta\gamma} f_{\delta\beta\gamma} = C_A \delta_{\alpha\delta}$.

Each process leads to one loss term and two gain terms; for instance, the $q \rightarrow qg$ process, we already saw, leads to a loss and a gain term in the quark PDF evolution; but it also causes a gain term in the gluon PDF evolution, to account for the radiated gluon. The complete set of evolution equations for the PDFs, called the Altarelli–Parisi or DGLAP (Dokshitzer–Gribov–Lipatov–Altarelli–Parisi) equations, are (writing q for u_m, d_m)

$$\frac{dq(x, \mu^2)}{d \log \mu^2} = \frac{\alpha_3}{2\pi} \int_x^1 \frac{dy}{y} \left\{ P_{q \leftarrow q}(y) q\left(\frac{x}{y}, \mu^2\right) + P_{q \leftarrow g}(y) g\left(\frac{x}{y}, \mu^2\right) \right\} \quad (9.75)$$

$$\frac{d\bar{q}(x, \mu^2)}{d \log \mu^2} = \frac{\alpha_3}{2\pi} \int_x^1 \frac{dy}{y} \left\{ P_{q \leftarrow q}(y) \bar{q}\left(\frac{x}{y}, \mu^2\right) + P_{q \leftarrow g}(y) g\left(\frac{x}{y}, \mu^2\right) \right\} \quad (9.76)$$

$$\begin{aligned}
\frac{dg(x, \mu^2)}{d \log \mu^2} = \frac{\alpha_3}{2\pi} \int_x^1 \frac{dy}{y} \left\{ P_{g \leftarrow q}(y) \sum_q \left[q\left(\frac{x}{y}, \mu^2\right) + \bar{q}\left(\frac{x}{y}, \mu^2\right) \right] \right. \\
\left. + P_{g \leftarrow g}(y) g\left(\frac{x}{y}, \mu^2\right) \right\} \quad (9.77)
\end{aligned}$$

$$P_{q \leftarrow q}(y) = \frac{4}{3} \left[\frac{1+y^2}{(1-y)_+} + \frac{3}{2} \delta(1-y) \right] \quad (9.78)$$

$$P_{g \leftarrow q}(y) = \frac{4}{3} \left[\frac{1+(1-y)^2}{y} \right] \quad (9.79)$$

$$P_{q \leftarrow g}(y) = \frac{1}{2} \left[y^2 + (1-y)^2 \right] \quad (9.80)$$

$$P_{g \leftarrow g}(y) = 6 \left[\frac{1-y}{y} + \frac{y}{(1-y)_+} + y(1-y) + \left(\frac{11}{12} - \frac{N_f}{18} \right) \delta(1-y) \right] \quad (9.81)$$

The functions $P_{b \leftarrow a}$ are called *splitting functions*; they summarize the rate of collinear splitting of species type a to produce species type b with momentum fraction y . Here $N_f \equiv \sum_q 1$ is the number of active quark species. A quark species is active if $\mu^2 > m_q^2$, if this condition is not met then that quark's

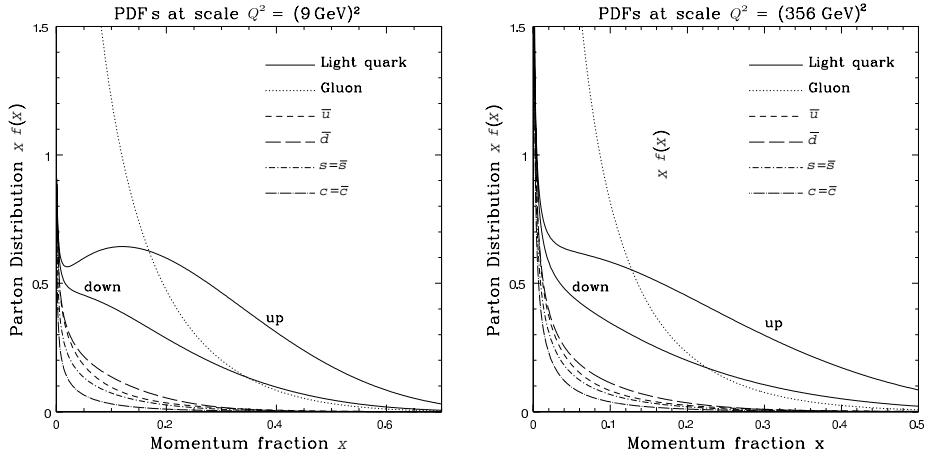


Fig. 9.4 Proton-parton distribution functions at two scales.

PDF should be taken to be zero. The value of α_3 should be evaluated at the renormalization point $\bar{\mu}^2 = \mu^2$. These equations manifestly preserve the sum rules, Eq. (9.47) and Eq. (9.48).

The values of the PDFs, and the effect of their scale dependence, is illustrated in Figure 9.4, which shows the PDFs at two scales, roughly the scales relevant for $b\bar{b}$ and $t\bar{t}$ production. Here we have plotted xf rather than f , so that the curves will fit better onto the plot. The area under a curve then represents the momentum fraction carried by that species. Note the horizontal axes, which we did not extend to $x = 1$ because the PDFs are all very small at very large x . Accurate compilations of PDFs for the proton are publicly available from collaborations (CTEQ, MRST) who perform grand fits, mostly of deep inelastic scattering data. The data for Figure 9.4 come from an MRST fit. The PDFs of the neutron are related to proton PDFs by isospin symmetry (exchange $u \leftrightarrow d$) to within small errors, smaller than the errors in the determination of the PDFs. The PDFs of other hadrons are not as well determined.

Final-state particles also emit radiation. In fact, the calculation in Subsection 6.7.2 carries through essentially unchanged for radiation from a final state. We argued there that such radiation is not important in evaluating the total cross section. However, it is important in determining the particles in the final state. In particular, the fragmentation functions $D(z, q^2)$ introduced in Subsection 9.2.2 are also q^2 dependent. For $\lambda > 1$, When a $q\bar{q}$ pair is formed from a gluon or photon of virtuality λq^2 , there is a larger opportunity for each quark to radiate off gluons than when the pair is formed at q^2 .

When such a radiation occurs, the fragmentation into a hadron must occur from one of the partonic fragments, rather than from the original hadron. This leads to scale evolution for the fragmentation functions, given by

$$\begin{aligned}
 \frac{dD_{h/q}(z, \mu^2)}{d \log \mu^2} &= \frac{\alpha_3}{2\pi} \int_z^1 \frac{dy}{y} \left\{ P_{q \leftarrow q}(y) D_{h/q}\left(\frac{z}{y}, \mu^2\right) + P_{q \leftarrow g}(y) D_{h/g}\left(\frac{z}{y}, \mu^2\right) \right\} \\
 \frac{dD_{h/\bar{q}}(z, \mu^2)}{d \log \mu^2} &= \frac{\alpha_3}{2\pi} \int_z^1 \frac{dy}{y} \left\{ P_{q \leftarrow q}(y) D_{h/\bar{q}}\left(\frac{z}{y}, \mu^2\right) + P_{q \leftarrow g}(y) D_{h/g}\left(\frac{z}{y}, \mu^2\right) \right\} \\
 \frac{dD_{h/g}(z, \mu^2)}{d \log \mu^2} &= \frac{\alpha_3}{2\pi} \int_z^1 \frac{dy}{y} \left\{ P_{g \leftarrow q}(y) \sum_q \left[D_{h/q}\left(\frac{z}{y}, \mu^2\right) + D_{h/\bar{q}}\left(\frac{z}{y}, \mu^2\right) \right] \right. \\
 &\quad \left. + P_{g \leftarrow g}(y) D_{h/g}\left(\frac{z}{y}, \mu^2\right) \right\}
 \end{aligned} \tag{9.82}$$

Here $P_{q \leftarrow q}$ etc. are the same splitting functions as for parton distributions, given already in Eq. (9.78) through Eq. (9.81).

By writing these equations in differential form, we can use them to evolve the PDFs (and splitting functions) from a scale q_1^2 to a very different scale q_2^2 , accounting not only for the $O(\alpha_3 \log(q_2^2/q_1^2))$ changes in the PDFs, but also for for *all* $(\alpha_3 \log(q_2^2/q_1^2))^n$ changes in the PDFs. In this they resemble the renormalization group equations of Subsection 7.4.1. They do not account for changes subdominant by a power of α_3 , without a log, i.e. $\alpha_3^{n+1} \log^n(q_2^2/q_1^2)$. Such corrections can be accounted for by computing the collinear splitting processes at higher order in α_3 , though this is only useful if the hard-scattering cross section is computed to a similar order. The next-to-leading-order (NLO) corrections to the splitting functions are known, as are the complete NNLO corrections and the hard-scattering cross-sections for a number of processes.

The agreement of measurements at very different values of q^2 and x with the evolution described by these equations is remarkable. A “grand fit” of QCD cross sections at many s , q^2 , and x values can be used to determine the PDFs as a function of scale, as well as the strong coupling constant. Recent fit values (2006) give $\alpha_3(\bar{\mu}^2 = M_Z^2) = 0.118 \pm 0.002$, in very good agreement with other determinations (lattice, Z pole observables, etc.).

9.3 Soft inelastic scattering: low-energy mesons

The previous section shows how the weakness of the strong interactions over short distances permits a relatively simple description of very

inelastic hadron collisions. The flip side of asymptotic freedom is that the strong interactions are strong over longer distances, and this considerably complicates the description of inelastic hadron collisions at lower energies. This section shows how it is nevertheless possible to make some quantitative predictions in this regime, by exploiting the chiral symmetries of QCD discussed in Section 8.3.

The starting point for making these predictions is the observation that the light pseudoscalar meson octet, $\xi = \exp[i\mathcal{M}/f]$, — consisting of pions, kaons and η — are pseudo-Goldstone bosons for the breaking of the approximate chiral $SU_L(3) \times SU_R(3)$ symmetry of QCD, involving the rotation of the lightest three quarks. The predictions follow because Goldstone's theorem strongly restricts the kinds of interactions which are possible for such particles at low energy. To see how this works in detail, recall that the most general low-energy Lagrangian for the pseudoscalar meson octet which is consistent with $SU_L(3) \times SU_R(3)$ is given by Eq. (8.58) and Eq. (8.61), to leading non-trivial order in the symmetry-breaking quark masses, M_q :

$$\begin{aligned}\mathcal{L}_{gb} &= \frac{f^2}{4} \text{tr}[\xi^\dagger \partial_\mu \xi \xi^\dagger \partial^\mu \xi] + \frac{cf^2}{4} \text{tr}[M_q(\xi + \xi^\dagger)] + O(M_q^2, M_q \partial^2, \partial^4) \\ &= -\frac{1}{4} \text{tr}[\partial_\mu \mathcal{M} \partial^\mu \mathcal{M}] + \frac{1}{24f^2} \text{tr}[\mathcal{M}^2 \partial_\mu \mathcal{M} \partial^\mu \mathcal{M} - \mathcal{M} \partial_\mu \mathcal{M} \mathcal{M} \partial^\mu \mathcal{M}] \\ &\quad + \frac{cf^2}{2} \text{tr} M_q - \frac{c}{4} \text{tr}[M_q \mathcal{M}^2] + \frac{c}{48f^2} \text{tr}[M_q \mathcal{M}^4] + O(\mathcal{M}^6)\end{aligned}\tag{9.83}$$

This must describe the dominant low-energy self-interactions of pions, kaons, and the η meson, and its power lies in the few parameters it involves, i.e., f and the product cM_q , compared with the number of observables which it predicts. In particular, Section 8.3 shows that the matrix cM_q is proportional to the meson squared-mass matrix, and so it is completely determined by the pion, kaon, and η masses. The only quantity left undetermined by the meson spectrum is therefore the dimensionful parameter f , which we next show can be determined from the measured meson decay lifetimes. Once this is done, the meson scattering cross sections are completely determined, and so these provide quantitative predictions of the standard model against which experiments may be compared.

9.3.1 Meson decays

Meson decays into leptons are described by supplementing the QCD interactions with the charged-current weak-interaction term,

$$\mathcal{L} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{weak}} \quad (9.84)$$

and working perturbatively in $\mathcal{L}_{\text{weak}}$. Because our interest is in energies comparable to the meson masses, we *cannot* perturb in the strong interactions, although this does not invalidate the expansion in powers of $\mathcal{L}_{\text{weak}}$. In this approximation the lowest-order energy eigenstates are the hadrons themselves, and we do not try to resolve these into strongly-interacting bound states of quarks or leptons. Since the individual quark flavors are conserved in the absence of the weak interactions, the charged members of the pseudoscalar octet of light mesons are stable in the limit where $\mathcal{L}_{\text{weak}}$ is turned off, and their decay rate may be computed perturbatively in the weak couplings.

9.3.1.1 Semileptonic decays

The simplest decays to consider are those for which there is at most one hadron amongst the final decay products. For the pseudoscalar mesons, such decays are called *semileptonic* and they include both the dominant decays, such as $\pi^+ \rightarrow \mu^+ \nu_\mu$ or $K^+ \rightarrow \mu^+ \nu_\mu$, and subdominant decays such as $K^+ \rightarrow \pi^0 e^+ \nu_e$ etc. For these decays the relevant part of the charged-current weak interactions, Eq. (7.12), is

$$\mathcal{L}_{\text{weak}} = \frac{G_F V_{uj}}{\sqrt{2}} \sum_{\ell=e,\mu} [\bar{u} \gamma^\nu (1 + \gamma_5) d_j] [\bar{\ell} \gamma_\nu (1 + \gamma_5) \nu_\ell] + \text{h.c.} \quad (9.85)$$

where the implied sum on j is over the two relevant down-type quarks: $d_1 = d$ and $d_2 = s$. Here $G_F = g_2^2/4M_W^2 \sqrt{2}$ is the usual Fermi coupling constant introduced in Chapter 5, and V_{ud} and V_{us} are the relevant CKM matrix elements. As discussed in Subsection 2.4.3, these are well approximated by $V_{ud} = \cos \theta_C$ and $V_{us} = \sin \theta_C$, with θ_C the Cabbibo angle. Numerically, $\cos \theta_C = 0.9753(6)$.

Consider now the decay rate for the reaction $\pi^+ \rightarrow \mu^+ \nu_\mu$, which is overwhelmingly the most common π^+ decay channel (making up 99.987 70(4)% of all π^+ decays). In order to compute this rate we require the matrix element

$$\langle \mu^+ \nu_\mu | \mathcal{L}_{\text{weak}} | \pi^+ \rangle \quad (9.86)$$

which is difficult to compute reliably from first principles because it involves the strong-interaction matrix element $\langle \Omega | \bar{d} \gamma^\mu (1 + \gamma_5) u | \pi^+ \rangle$, where $|\Omega\rangle$ is the QCD ground state.

For later purposes it is convenient to rewrite the quark combination appearing in $\mathcal{L}_{\text{weak}}$ as a linear combination of the conserved Noether currents for the symmetry $SU_L(2) \times SU_R(2)$:

$$\begin{aligned} i\bar{u}\gamma^\mu(1+\gamma_5)d &= i\bar{q}\gamma^\mu(1+\gamma_5)\left(\frac{\tau_1+i\tau_2}{2}\right)q \\ &= \left[(j_I)_1^\mu + i(j_I)_2^\mu\right] + \left[(j_A)_1^\mu + i(j_A)_2^\mu\right] \end{aligned} \quad (9.87)$$

where

$$(j_I)_k^\mu = \frac{i}{2}\bar{q}\gamma^\mu\tau_k q \quad \text{and} \quad (j_A)_k^\mu = \frac{i}{2}\bar{q}\gamma^\mu\gamma_5\tau_k q \quad (9.88)$$

Here τ_k are the three Pauli matrices and we re-express the left- and right-handed currents of $SU_L(2) \times SU_R(2)$ in terms of axial and vector currents using $j_L^\mu = \frac{1}{2}(j_I^\mu + j_A^\mu)$.

Some progress can be made using symmetries. Since the pion is a pseudo-scalar particle, the parity invariance of the strong interactions implies

$$\langle\Omega|(j_I)_k^\mu|\pi^+\rangle = 0 \quad (9.89)$$

Similarly, the most general form for the matrix element of the axial current which is consistent with Poincaré and isospin invariance is given by

$$\langle\Omega|(j_A)_k^\mu|\pi_l(q)\rangle = iF_\pi q^\mu e^{iqx}\delta_{kl} \quad (9.90)$$

where π_l are the isospin eigenstates for the pi meson, with $\pi^\pm = (\pi_1 \mp i\pi_2)/\sqrt{2}$. Therefore, the matrix element of the pion which we need is

$$\langle\Omega|\bar{d}\gamma^\mu(1+\gamma_5)u|\pi^+(q)\rangle = i\sqrt{2}F_\pi q^\mu e^{iqx} \quad (9.91)$$

This allows us to complete the computation of the rest frame decay rate,

$$\frac{1}{\tau_{\text{th}}} = \frac{G_F^2 \cos^2 \theta_C F_\pi^2 m_\mu^2 m_\pi}{4\pi} \left(1 - \frac{m_\mu^2}{m_\pi^2}\right)^2 \quad (9.92)$$

which may be compared with the observed lifetime, $\tau_{\text{exp}} = 2.6030(24) \times 10^{-8}$ s, to determine that $F_\pi = 92$ MeV.

Our goal is to use this to determine the value of the parameter f , which we do by repeating the decay-rate calculation in the low-energy effective meson theory, obtained by supplementing the interactions of Eq. (8.58) and Eq. (8.61) with an expression for the charged-current weak interaction in terms of the meson field \mathcal{M} : $\mathcal{L} = \mathcal{L}_{\text{meson}} + \mathcal{L}_{\text{weak}}(\mathcal{M})$. We saw, in Eq. (9.85) through Eq. (9.87), that the weak-interaction Lagrangian can be written in terms of the $SU_L(2) \times SU_R(2)$ Noether currents. Now we need to find those Noether currents in the effective theory describing the pions.

According to Eq. (1.147) in Chapter 1, we need two things to determine these Noether currents; the variation of the fields with respect to the symmetry transformations, and the canonical momenta of the fields. The symmetry transformations of the fields are given in Eq. (8.55) and Eq. (8.57). The canonical momenta can be found from the Lagrangian, Eq. (8.58). Keeping only the terms of lowest order in pion fields and derivatives gives

$$(j_1)_k^\mu = -\epsilon_{klm}\pi_l\partial^\mu\pi_m + \cdots \quad \text{and} \quad (j_A)_k^\mu = f\partial^\mu\pi_k + \cdots \quad (9.93)$$

where the ellipses denote an infinite number of higher order terms in these currents involving more derivatives (or powers of quark masses) and/or more powers of π_k .

Using Eq. (9.93) to evaluate $\langle\Omega|(j_A)_k^\mu|\pi_l\rangle$ immediately gives

$$\langle\Omega|(j_A)_k^\mu|\pi_l(q)\rangle = ifq^\mu e^{iqx}\delta_{kl} \quad (9.94)$$

Comparing with Eq. (9.90) determines f to be simply

$$f = F_\pi = 92 \text{ MeV} \quad (9.95)$$

With this constant in hand, the low-energy form of the meson-meson interaction is completely specified, and so may be used to predict the cross section for low-energy pion scattering.

9.3.1.2 Non-leptonic decays

Before turning to meson scattering we pause briefly to discuss the case of *non-leptonic* meson decays, defined as those for which only hadrons appear in the final state. Important examples of such decays include $K^\pm \rightarrow \pi^\pm\pi^0$ and $K^0 \rightarrow \pi^+\pi^-$ or $K^0 \rightarrow \pi^+\pi^-\pi^0$. If we focus on those terms which change the strangeness quantum number by ± 1 unit, then the relevant part of the charged-current weak-interaction Lagrangian is

$$\mathcal{L}_{\text{weak}} = \frac{G_F V_{is} V_{jd}^*}{\sqrt{2}} [\bar{u}_i \gamma^\nu (1 + \gamma_5) s] [\bar{d} \gamma_\nu (1 + \gamma_5) u_j] + \text{h.c.} \quad (9.96)$$

where the sum on i and j is over the three species of up-type quarks $u_i = \{u, c, t\}$. (Even though the light hadrons involve only valence u , d , and s quarks, the terms involving c and t quarks can be relevant to precision calculations due to the virtual effects of these quarks in the hadronic quark sea.)

It is a more complicated proposition to use this Lagrangian to compute non-leptonic decay rates, as may be seen in two complementary ways. For instance, from the microscopic (quark) point of view the decay rate for

$K^+ \rightarrow \pi^+\pi^0$ requires the matrix element

$$\langle \pi^+\pi^0 | [\bar{u}\gamma^\nu(1+\gamma_5)s][\bar{d}\gamma_\nu(1+\gamma_5)u] | K^+ \rangle \quad (9.97)$$

and much less than before can be said about this matrix element purely on symmetry grounds. For instance, the parity invariance of the strong interactions implies that only the pseudoscalar part of the operator has a nonzero matrix element involving three pseudoscalar mesons.

The implications of the vector-like $SU_V(3)$ invariance similarly follow using a bit of group theory for $SU(3)$. The starting point is the recognition that all of the currents and particle states appearing in Eq. (9.97) transform as octets of $SU_V(3)$. This implies that the product of the two currents in Eq. (9.97) transforms like a symmetric product of two octets, $(\mathbf{8} \otimes \mathbf{8})_{\text{sym}} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{27}$, as does the two-pion state (due to Bose statistics for the pions). Now, $SU_V(3)$ invariance implies that the matrix element of two octet currents between octet meson states must be built using only the three invariant $SU(3)$ tensors: $\delta_{ab} \propto \text{Tr}(\lambda_a \lambda_b)$; the completely antisymmetric structure constants, $f_{abc} \propto \text{Tr}([\lambda_a, \lambda_b]\lambda_c)$; and the completely symmetric constants $d_{abc} \propto \text{Tr}(\{\lambda_a, \lambda_b\}\lambda_c)$ (where λ_a here denotes the Gell-Mann matrices introduced in Chapter 1). It follows that the most general form possible for the matrix element of two currents in the $SU_V(3)$ -invariant limit is

$$\begin{aligned} \langle \mathcal{M}_a(p)\mathcal{M}_b(q) | [\bar{\mathcal{Q}}\gamma^\nu(1+\gamma_5)\lambda_c\mathcal{Q}][\bar{\mathcal{Q}}\gamma_\nu(1+\gamma_5)\lambda_d\mathcal{Q}] | \mathcal{M}_e(r) \rangle \\ = f_1\delta_{ab}d_{cde} + f_2(\delta_{de}d_{abc} + \delta_{ce}d_{abd}) \\ + f_3(\delta_{ac}f_{bde} + \delta_{bc}f_{ade} + \delta_{ad}f_{bce} + \delta_{bd}f_{ace}) \\ + f_4\delta_{cd}d_{abe} + f_5(\delta_{ae}d_{bcd} + \delta_{be}d_{acd}) \end{aligned} \quad (9.98)$$

where the indices a, \dots, e run from 1 to 8, and Lorentz invariance requires the form factors f_k to be functions only of Lorentz-invariant combinations of the external momenta. Given that four-momentum conservation implies $p^\mu + q^\mu = r^\mu$, this requires the f_k to be functions only of the invariants p^2 , q^2 , and r^2 , and so they must be constants which are functions only of the meson masses. Furthermore, for the $\Delta S = \pm 1$ decays involving the operator $[\bar{u}\gamma^\mu(1+\gamma_5)s][\bar{d}\gamma_\mu(1+\gamma_5)u]$ we have $c = 4$ and $d = 1$, and for $K \rightarrow \pi\pi$ transitions inspection of Eq. (8.34) shows that $a, b = 1, 2, 3$ and $e = 4, 5, 6, 7$, and so only the three constants f_1 through f_3 contribute. We see that in the $SU_V(3)$ -invariant limit the $\Delta S = \pm 1$ transitions from this operator may be parameterized by three independent constants, f_1 , f_2 , and f_3 , all of which must themselves be inferred from experiments.

This same counting can also be seen from the point of view of the low-energy effective theory involving only mesons. In this case the complications

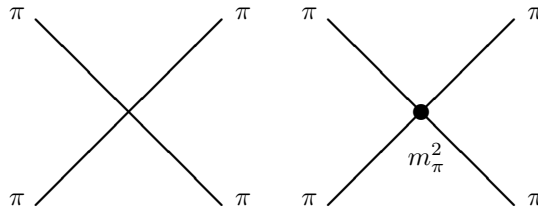
arise once we try to express $\mathcal{L}_{\text{weak}}$ in terms of the meson field \mathcal{M} , since the weak interactions are no longer simply linear combinations of Noether currents. Instead we must construct $\mathcal{L}_{\text{weak}}$ as a function of the meson field in such a way that it involves the fewest derivatives (or powers of quark masses) and transforms under all symmetries in the same way as does $\mathcal{L}_{\text{weak}}$ when expressed in terms of quark fields. For the $\Delta S = \pm 1$ interaction relevant to Eq. (9.97) we have seen that the product of currents transforms under $SU_L(3) \times SU_R(3)$ as the product $(\mathbf{8}, \mathbf{1}) \oplus (\mathbf{27}, \mathbf{1})$. That is, $\mathcal{L}_{\text{weak}}$ is a singlet under $SU_R(3)$ because it involves only left-handed quark fields, and it transforms like the symmetric part of the product of two octets (without the singlet contribution, since $\Delta S \neq 0$) with respect to $SU_L(3)$.

Keeping in mind the transformation rule, Eq. (8.53), for $\xi = \exp[i\mathcal{M}/f]$ and focusing on terms involving only two derivatives leads to the following three possible effective interactions:

$$\begin{aligned} \mathcal{L}_{\text{weak}} = & \frac{G_F V_{us} V_{ud}^*}{\sqrt{2}} \left[g_8 \text{Tr} \left(\lambda_6 \partial_\mu \xi \partial^\mu \xi^\dagger \right) + g_{27}^{(1/2)} c_{1/2}^{ab} \text{Tr} \left(\lambda_a \partial_\mu \xi \xi^\dagger \lambda_b \partial^\mu \xi \xi^\dagger \right) \right. \\ & \left. + g_{27}^{(3/2)} c_{3/2}^{ab} \text{Tr} \left(\lambda_a \partial_\mu \xi \xi^\dagger \lambda_b \partial^\mu \xi \xi^\dagger \right) \right] + \dots \end{aligned} \quad (9.99)$$

where the first term represents the left-handed octet while the second two terms transform as left-handed **27**s. These last two terms are distinguished by how they transform under isospin, with the first transforming as isospin-1/2 and the last as isospin-3/2, with $c_{1/2}^{ab}$ and $c_{3/2}^{ab}$ being the appropriate Clebsch–Gordan coefficients. In principle there is also an effective operator involving quark masses which transforms properly to appear in $\mathcal{L}_{\text{weak}}$, consisting of the operator $\tilde{g}_8 \text{Tr}[\lambda_6 M_q \xi] + \text{h.c.}$ However, this operator has the same form as the lowest-order meson mass term, Eq. (8.61), and so cannot generate transitions between meson flavors once the total meson-mass term is diagonalized. It therefore does not contribute at all to meson decays, and simply corrects the formulae for the meson masses in terms of the quark masses.

We see that in the end the number of undetermined constants arising in $K \rightarrow \pi\pi$ decays is three, corresponding to the three effective couplings, g_8 , $g_{27}^{(1/2)}$ and $g_{27}^{(3/2)}$, which must be inferred by comparison with experiments. This agrees precisely with the number of constants obtained by counting the independent form factors which symmetries allow when $\mathcal{L}_{\text{weak}}$ is expressed in terms of quarks. There is a long-standing puzzle about the size of the constants that are obtained by fitting to the observed decay rates, because these fits show that $|g_8| \gg |g_{27}^{(1/2)}|, |g_{27}^{(3/2)}|$, a result which is contained in the phenomenological “ $\Delta I = 1/2$ -rule.” It is not yet understood why QCD should produce this kind of hierarchy amongst these couplings, although the

Fig. 9.5 Feynman graphs which dominate $\pi\pi$ scattering.

resolution of this puzzle may have to await the eventual reliable calculation of these coefficients from first principles within QCD, calculations which remain at present beyond our reach.

The bottom line is that predictivity is reduced for non-leptonic decays compared with semi-leptonic decays because the hadronic operator appearing in $\mathcal{L}_{\text{weak}}$ is not a conserved current, and so involves effective coupling constants which are independent of the constants appearing in the rest of $\mathcal{L}_{\text{meson}}$. As a result all of these constants must be obtained from experiment before any testable predictions can be made. Although considerable effort has been expended along these lines, we here abandon the pursuit on the grounds that it takes us too far beyond our main line of development.

9.3.2 Meson–meson scattering

We return now to our main line of argument, which is to use the effective theory given by Eq. (9.83) to obtain predictions for low-energy scattering amongst the pseudoscalar octet mesons. To do so we evaluate the Feynman graphs of Figure 9.5, using the Feynman rules obtained from the Lagrangian of Eq. (9.83). The first graph in the figure involves two derivatives; the second has no derivatives, but one insertion of the pion mass. For simplicity we focus on low-energy pion–pion scattering when doing so, and simply quote the corresponding results for kaons and the η meson.

A straightforward calculation gives the following matrix element for the scattering $\pi_a\pi_b \rightarrow \pi_c\pi_d$ (which we temporarily call \mathcal{A} rather than \mathcal{M} , to avoid confusion with the pion field \mathcal{M}):

$$\mathcal{A}_{ab,cd} = \frac{1}{F_\pi^2} \left[\delta_{ab}\delta_{cd}(s - m_\pi^2) + \delta_{ac}\delta_{bd}(t - m_\pi^2) + \delta_{ad}\delta_{bc}(u - m_\pi^2) \right] \quad (9.100)$$

In the CM frame s , t , and u have simple expressions in terms of the pion energy, E , and three-momentum, q : $s = 4E^2$, $t = -2q^2(1 - \cos \vartheta)$, and $u = -2q^2(1 + \cos \vartheta)$. Here ϑ denotes the scattering angle, also in the CM frame.

Comparison with the data is made using channels having definite angular momentum and isospin. If we decompose $\mathcal{A}_{ab,cd}$ into combinations, $\mathcal{A}^{(i)}$, having definite initial isospin,

$$\begin{aligned} \mathcal{A}_{ab,cd} = & \mathcal{A}^{(0)} \frac{1}{3} \delta_{ab} \delta_{cd} + \mathcal{A}^{(1)} \frac{1}{2} (\delta_{ac} \delta_{bd} - \delta_{ad} \delta_{bc}) \\ & + \mathcal{A}^{(2)} \left[\frac{1}{2} (\delta_{ac} \delta_{bd} + \delta_{ad} \delta_{bc}) - \frac{1}{3} \delta_{ab} \delta_{cd} \right] \end{aligned} \quad (9.101)$$

then

$$\mathcal{A}^{(0)} = \frac{2s - m_\pi^2}{F_\pi^2}, \quad \mathcal{A}^{(1)} = \frac{t - u}{F_\pi^2}, \quad \mathcal{A}^{(2)} = -\frac{s - 2m_\pi^2}{F_\pi^2} \quad (9.102)$$

The next step is to resolve these amplitudes into partial waves,

$$\mathcal{A}_\ell^{(i)} \equiv \frac{1}{64\pi} \int_{-1}^1 d \cos \vartheta P_\ell(\cos \vartheta) \mathcal{A}^{(i)} \quad (9.103)$$

where $P_\ell(\cos \vartheta)$, as usual, denote the Legendre polynomials (so $P_0(x) = 1$ and $P_1(x) = x$). Since all of the dependence on ϑ appears through the variables t and u , and since Eq. (9.102) gives $\mathcal{A}^{(0)}$ and $\mathcal{A}^{(2)}$ as functions of s only, it is clear that only the partial wave $\ell = 0$ is predicted at lowest order for the even isospin configurations. Also, since $\mathcal{A}^{(1)}$ is strictly linear in $\cos \vartheta$, it only involves the partial wave $\ell = 1$.

The actual comparison with the data is made by expanding the (real part of) $\mathcal{A}_\ell^{(i)}$ in powers of the squared pion momentum: $q^2/m_\pi^2 = E^2/m_\pi^2 - 1 = (s - 4m_\pi^2)/4m_\pi^2$. That is, writing

$$\mathcal{A}_\ell^{(i)} = \left(\frac{q^2}{m_\pi^2} \right)^\ell \left(a_\ell^i + b_\ell^i \frac{q^2}{m_\pi^2} + \cdots \right) \quad (9.104)$$

defines the pion scattering lengths, a_ℓ^i , and slopes, b_ℓ^i . Applying these definitions to Eq. (9.99) gives the predictions of the second and third columns of Table 9.1. Column three gives the numerical value corresponding to the analytic expression which is given in column two. The predictions including the next-order terms in the q^2 expansion have also been worked out, and are given in the fourth column of this table.

Comparison of these predictions with experiment is not straightforward, since it is not feasible to directly perform pion–pion scattering experiments. Instead, the pion–pion scattering amplitudes at low energies are inferred

Table 9.1 *Theory vs. experiment for low-energy pion scattering*

Parameter	Leading order		Next order	Experiment
a_0^0	$7m_\pi^2/32\pi F_\pi^2$	0.16	0.20	0.26(5)
b_0^0	$m_\pi^2/4\pi F_\pi^2$	0.18	0.26	0.25(3)
a_1^1	$m_\pi^2/24\pi F_\pi^2$	0.030	0.036	0.038(2)
a_0^2	$-m_\pi^2/16\pi F_\pi^2$	-0.044	-0.041	-0.028(12)
b_0^2	$-m_\pi^2/8\pi F_\pi^2$	-0.089	-0.070	-0.082(8)

from their influence on the final state in other processes, such as $K \rightarrow \pi\pi e\nu_e$ or $\pi N \rightarrow \pi\pi N$. The experimental results, as obtained from kaon decays, for those quantities which are predicted to be nonzero at lowest order are listed in the right-hand-most column of Table 9.1. Data also exist for other partial waves which are predicted to vanish at lowest order, such as $I = 0, \ell = 2$, and these are found to be in good agreement with the nonzero predictions which arise at next-to-leading order in the low-energy expansion.

This example nicely illustrates the predictive power which is possible with a low-energy effective Lagrangian, even if it is impossible to predict the values for the couplings of this Lagrangian in terms of an underlying theory. This predictive power arises because many observables – e.g. the pion scattering lengths and slopes – are all parameterized in terms of a single constant – the decay constant, F_π – which can be extracted directly from experiment. We emphasize that this predictive power holds regardless of the renormalizability of the effective theory. Computing to higher orders involves the introduction of more parameters, but predictions remain possible provided that more observables are computed than there are parameters to fix from experiment. The information underlying these predictions comes from the symmetries of the underlying theory, as well as the restrictions owing to the comparatively small number of possible interactions which can appear at low orders of the low-energy expansion.

9.3.3 Including nucleons

The interpretation of the light pseudoscalar mesons as pseudo-Goldstone bosons also constrains the kinds of interactions they can have with other kinds of particles at low energies. We pause briefly to illustrate these constraints using the example of pion–nucleon couplings, which for simplicity we examine in the isospin-invariant limit: $m_u = m_d$. The logic proceeds in the same way as in the previous section, wherein we use the measured neutron decay rate to infer the value of an effective coupling which governs

the low-energy effective pion–nucleon interactions. Once this has been done, the result may be used to make predictions for the properties of low-energy pion–nucleon scattering, which succeed when they are compared with experiment.

The couplings between nucleons and pions which dominate at lowest order in the derivative expansion involve only one derivative. It can be shown that the most general form for these which is consistent with both $SU_L(2) \times SU_R(2)$ invariance and parity invariance is:

$$\begin{aligned} \mathcal{L}_{\pi NN} = & -\bar{N}(\not{\partial} + m_N)N - \frac{ig}{2f} \left(\bar{N} \gamma^\mu \gamma_5 \tau_k N \right) \cdot \partial_\mu \pi_k \\ & - \frac{i}{4f^2} \epsilon_{klm} \left(\bar{N} \gamma^\mu \tau_k N \right) \cdot (\pi_l \times \partial_\mu \pi_m) + \cdots \quad (9.105) \end{aligned}$$

where $N = \begin{pmatrix} p \\ n \end{pmatrix}$ is the nucleon isodoublet. The ellipses here represent terms which involve either three or more powers of the pion field, more than two powers of the nucleon field, or more than one derivative.

What is important is that *all* of the interactions involving arbitrary powers of the pion field but only one derivative are completely dictated by the symmetries, and have strengths which are completely determined in terms of the two constants f and g . In particular, those terms involving even powers of the pion field, π_k/f , do not depend on g , while those involving odd powers of π_k/f are strictly proportional to g . The constant f here is also the same one which governs the meson self-interactions, which we have seen is fixed by the charged-pion lifetime to be $f = F_\pi = 92$ MeV. Clearly, only the one additional constant, g , is required to determine the low-energy couplings of pions to nucleons in the limit of exact $SU_L(2) \times SU_R(2)$ symmetry, and we infer the value of this constant below using the measured decay lifetime of free neutrons.

Of course $SU_L(2) \times SU_R(2)$ is not an exact symmetry, so non-derivative interactions involving the quark masses must also be included. Such terms generate two effects; a mass splitting between the neutron and proton, and nucleon–pion interactions. However, the $m_N - m_P$ splitting has comparable electromagnetic contributions which cannot be easily disentangled, and the $NN\pi$ interactions are suppressed by a power of m_u or m_d , and so are much smaller than the interactions present in Eq. (9.105). Therefore we will neglect these terms in what follows.

9.3.3.1 Nucleon decays

In order to proceed we must first determine the value of g , which we do by computing the lifetime for neutron decay through its dominant channel

$n \rightarrow p e \bar{\nu}_e$. From the microscopic (quark) point of view this proceeds through the weak interaction Lagrangian, so we require the matrix element

$$\begin{aligned}\langle N(p, \sigma) | (j_I)_k^\mu | N'(p', \sigma') \rangle &= \frac{i}{2} e^{iqx} \bar{u} \left[F_1(q^2) \gamma^\mu + F_2(q^2) \gamma^{\mu\nu} q_\nu \right] \tau_k u' \\ \langle N(p, \sigma) | (j_A)_k^\mu | N'(p', \sigma') \rangle &= \frac{i}{2} e^{iqx} \bar{u} \left[G_1(q^2) \gamma^\mu \gamma_5 + G_2(q^2) \gamma_5 q^\mu \right] \tau_k u'\end{aligned}\quad (9.106)$$

Here, p_μ and p'_μ are the four-momenta of the initial and final nucleons – and $q^\mu = (p - p')^\mu$ is their difference – while $\sigma, \sigma' = \pm 1/2$ similarly represent the polarizations of the initial and final nucleons. $u = u(p, \sigma)$ and $u' = u(p', \sigma')$ are the Dirac spinors for free spin-1/2 particles having the indicated momentum and polarization σ , with the dispersion relations $p^2 + m_N^2 = 0$ and $(p')^2 + m_{N'}^2 = 0$. Finally, $\gamma^{\mu\nu}$ stands for the commutator $[\gamma^\mu, \gamma^\nu]/2$.

Lorentz invariance requires the four form factors, F_1, F_2, G_1 , and G_2 , to be functions of the invariant momentum transfer, q^2 . Furthermore, these functions are subject to a constraint which expresses the fact that we are working in a limit where $SU_L(2) \times SU_R(2)$ is taken to be a symmetry of the QCD Lagrangian. The implications of this symmetry follow from the conditions of current conservation, $\partial_\mu (j_I)_k^\mu = \partial_\mu (j_A)_k^\mu = 0$. This implies

$$(m_N - m_{N'}) F_1(q^2) = 0 \quad \text{and} \quad i(m_N + m_{N'}) G_1(q^2) = q^2 G_2(q^2), \quad (9.107)$$

where m_N and $m_{N'}$ are respectively the masses of the nucleon which appears on the left- and right-hand side of the matrix element.

To a very good approximation, only the quantities $F_1(0)$ and $G_1(0)$ are required when evaluating the neutron decay rate, and this is because the components of the momentum transfer, q^μ , are at most of order 1 MeV in the rest frame of the decaying neutron. This justifies the neglect of q^μ in the matrix element because it is much smaller than the typical strong-interaction scale, $\Lambda \sim m_N \sim 1$ GeV, over which the form factors vary appreciably. The same reasoning also allows the neglect of the neutron–proton mass difference, $m_n - m_p$, in the matrix element since this is also of order an MeV.

Symmetry arguments provide the further information that $F(0) = 1$ because of the following argument. Recall that the Noether currents $(j_I)_k^\mu$ generate the isospin charges once they are integrated over all space, according to

$$Q_k = \int d^3\mathbf{r} (j_I)_k^0 \quad (9.108)$$

But we also know that the matrix elements of these generators only depend

on the isospin transformation properties of the states whose matrix elements are taken, since they may be computed purely using the commutation relations of $SU_I(2)$. Since integrating over all space corresponds to taking $q^\mu \rightarrow 0$ in Fourier space, this implies that the form factor $F_1(0)$ should also depend only on the isospin transformation properties of the nucleons. But since both quarks and nucleons are isodoublets, and since inspection of Eq. (9.106) shows that taking the same matrix element between quark states would give $F_1(0) = 1$, it follows that the same must be true for nucleons.

This is a special case of a general result, which states that the $q^\mu \rightarrow 0$ part of the matrix element of a conserved current is never renormalized. As applied to the electromagnetic current this theorem is exact, and is what underlies the statement that the electric charge of the proton or neutron is simply obtained by summing the charges of the constituent quarks. As applied to the isospin current it is only approximately valid, since isospin is only an approximate symmetry of the strong interactions. Once symmetry-breaking effects are included, the amount by which $F_1(0)$ differs from unity is given by the *Ademollo–Gatto* theorem, which states that the first contribution arises at second order in the quark masses

$$F_1(0) = 1 + O\left(\frac{m_d^2 - m_u^2}{\Lambda_{\text{QCD}}^2}\right) = 1 + O(10^{-3}) \quad (9.109)$$

One way to prove this theorem is to show that $F_1(0) = 1$ in the most general low-energy effective theory for the mesons.

The same argument does not hold for the axial current because this is a current for a symmetry which is spontaneously broken. As Problem 8.2 shows, this turns out to imply that the corresponding conserved charge is not well defined when acting on particle states, and so $G_1(0)$ need not be unity even in the limit of exact symmetry.

Finally, using the matrix element (9.106) with $q^\mu \rightarrow 0$, together with $F_1(0) = 1$, gives the rate for $n \rightarrow pe\bar{\nu}$ in terms of the one unknown constant $G_1(0)$, and comparison of the result with the measured neutron lifetime, $\tau_{\text{exp}} = 885.7(8)$ s, allows the inference $G_1(0) = 1.269$. In order to relate this to the unknown coupling constant g we next compare this with the same calculation performed within the low-energy pion–nucleon theory.

In order to make this comparison we compute the low-energy limit of the matrix elements of the isospin current using the low-energy theory. This is done by computing the vector and axial-vector Noether currents starting with the pion–nucleon effective Lagrangian given by Eq. (9.105), to obtain

$$(J_I)_k^\mu = -\epsilon_{klm}\pi_l\partial^\mu\pi_m + \frac{i}{2}\bar{N}\gamma^\mu\tau_k N + \cdots$$

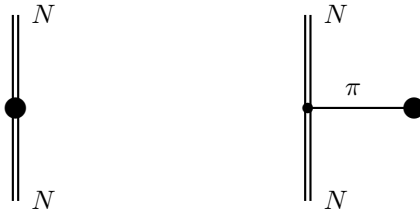


Fig. 9.6 Feynman graphs dominating nucleon Noether currents.

$$(j_A)_k^\mu = f\partial^\mu\pi_k + \frac{ig}{2}\bar{N}\gamma^\mu\gamma_5\tau_k N + \cdots \quad (9.110)$$

where the purely pionic terms were obtained earlier using the pion effective Lagrangian. As usual, the ellipses denote an infinite number of higher-order terms corresponding to the infinite number of interactions in the effective pion-nucleon Lagrangian. All of the terms not written explicitly above involve additional factors of the fields π_k or N , or involve more derivatives of these fields, than do the terms displayed.

We next evaluate the matrix elements of these currents within the effective theory, by evaluating the two Feynman graphs of Figure 9.6. The blob in the figure represents the current, double lines represent nucleons, and single lines represent pions. The first graph gives the direct matrix element of Eq. (9.110), and contributes to the form factors F_1 and G_1 . The second graph uses the $NN\pi$ interaction of the effective Lagrangian, Eq. (9.105), together with the vacuum-pion matrix element of Eq. (9.90). It contributes only to the form factor G_2 . Evaluating these graphs, we find

$$F_1 = 1, \quad G_1 = g, \quad \text{and} \quad G_2 = \frac{2igm_N}{q^2} \quad (9.111)$$

from which we see $F_1(0) = 1$ and $G_1(0) = g$. The factor $1/q^2$ in G_2 comes from the pion propagator in the second diagram of Figure 9.6, which is massless in the symmetry limit in which we work. Notice that this result for G_2 is precisely what is required to satisfy the current-conservation condition of Eq. (9.107).

We may finally determine the value of g . The neutron decay rate is completely determined by the constants $F_1(0) = 1$ and $G_1(0) = g$, and the measured neutron mean life therefore implies $g = 1.269$.

9.3.3.2 Nucleon-pion scattering

Having determined from experiment the values taken by f and g , we are now in a position to use the effective pion–nucleon Lagrangian to predict their low-energy properties. Historically the phenomenology of pion–nucleon scattering has been described in terms of a trilinear $NN\pi$ Yukawa interaction, with no derivatives:

$$\mathcal{L}_{NN\pi} = ig_{NN\pi} (\bar{N} \gamma_5 \tau_k N) \cdot \pi_k \quad (9.112)$$

with the constant $g_{NN\pi}$ found from phenomenological studies to be close to 14. But the value of this constant can be predicted in terms of the constant g , and this prediction serves as a test of the success of the low-energy pion–nucleon Lagrangian.

The prediction starts with the trilinear $NN\pi$ interaction of Eq. (9.105):

$$\mathcal{L}_{NN\pi} = -\frac{ig}{2f} (\bar{N} \gamma^\mu \gamma_5 \tau_k N) \cdot \partial_\mu \pi_k \quad (9.113)$$

and performs an integration by parts to move the derivative to the nucleon fields. Using the lowest-order nucleon equations of motion $-(\not{\partial} + m_N)N = 0$ – then gives a result of the form of Eq. (9.112), but with

$$g_{NN\pi} = \frac{gm_N}{f} \quad (9.114)$$

Using the experimental values, $g = 1.269$, $m_N = 940$ MeV, and $f = 92$ MeV, gives the prediction $g_{NN\pi} = 13.0$, which agrees well with the phenomenologically inferred value. The successful relation, Eq. (9.114), is known as the *Goldberger–Treiman relation*.

9.3.4 Anomalies and π^0 decay

Historically, the understanding of pions as Goldstone bosons provided an important stimulus for the understanding of anomalous symmetries (i.e. symmetries of the classical action which do not survive quantization). The criterion for when symmetries are anomalous is described in detail in Subsection 2.5.3, where anomalies provide an important consistency requirement for the standard model itself. Anomalies are also crucial for being able to describe the low-energy properties of pions, because they underlie the theoretical understanding of the π^0 meson decay rate.

The dominant decay process for neutral pi mesons is the decay into two photons, $\pi^0 \rightarrow \gamma\gamma$, which tells us that the dominant π^0 decay must be electromagnetic. Consider, therefore, the limit of the standard model where weak

interactions are ignored, but both electromagnetic and strong interactions are kept. We have seen that in the absence of electromagnetic interactions the strong dynamics of QCD ensures that low-energy pion interactions enjoy an $SU_L(2) \times SU_R(2)$ invariance up to small corrections proportional to m_d and m_u . But only those elements of this symmetry which commute with the u - and d -quark electric-charge matrix

$$Q_{\text{em}} = \begin{pmatrix} \frac{2}{3} & \\ & -\frac{1}{3} \end{pmatrix} \quad (9.115)$$

are also symmetries of the electromagnetic interactions, and so only the subgroup $U_{\text{em}}(1) \times U_{3A}(1)$ generated by Q_{em} and

$$T_{3A} = \begin{pmatrix} \frac{1}{2} & \\ & -\frac{1}{2} \end{pmatrix} \gamma_5 \quad (9.116)$$

survive as symmetries of the combined strong and electromagnetic theory (up to small quark-mass corrections).

Since $U_{3A}(1) \in SU_A(2)$ is an axial symmetry, it is spontaneously broken by the ground state, with the result that its Goldstone boson – the π^0 meson – must be exactly massless to all orders in the electromagnetic interactions (in the limit of vanishing quark masses). Since the rest of $SU_A(2)$ does not commute with electromagnetic interactions, the symmetries for which the π^\pm mesons are the Goldstone bosons are broken by electromagnetism, thereby allowing electromagnetic interactions to contribute a nonzero shift to the charged pi-meson mass.

There is a potential puzzle buried in the neutral-pion decay rate. As is shown in Problem 9.9, the decay rate for the process $\pi_0 \rightarrow \gamma\gamma$ is well-described by a pion-photon interaction term of the form

$$\mathcal{L} = \frac{e^2}{32\pi^2 F_\pi} \pi^0 F^{\mu\nu} F^{\lambda\rho} \epsilon_{\mu\nu\lambda\rho} \quad (9.117)$$

The problem with this Lagrangian is that the π^0 field it contains is not differentiated, and so as a result it is *not* invariant with respect to the $U_{3A}(1)$ symmetry for which π^0 is the Goldstone boson. Instead, since the $U_{3A}(1)$ transformation rule is $\delta\pi^0 = \omega F_\pi$, for constant infinitesimal parameter ω , we see that

$$\delta\mathcal{L} = \frac{e^2}{32\pi^2} F^{\mu\nu} F^{\lambda\rho} \epsilon_{\mu\nu\lambda\rho} \omega \quad (9.118)$$

Invariant terms involving the fewest derivatives, such as

$$\mathcal{L}' = \frac{e^2}{32\pi^2 F_\pi} \left(\frac{\square\pi^0}{\Lambda^2} \right) F^{\mu\nu} F^{\lambda\rho} \epsilon_{\mu\nu\lambda\rho} \quad (9.119)$$

with $\Lambda \sim 4\pi F_\pi \sim 1$ GeV (or its equivalent, with $\square\pi^0 \rightarrow m_\pi^2\pi^0$) give decay rates which are $O(m_\pi^4/\Lambda^4)$ smaller, which would be too small (for any Λ large enough to justify the validity of the low-energy field theory description in terms of pion fields). The puzzle is how to reconcile the observed neutral-pion decay rate with the $U_{3A}(1)$ symmetry of the pion Lagrangian.

As the section title suggests, anomalies are crucial to the resolution of this puzzle. In particular, there is an anomaly for the axial $U_{3A}(1)$ generator, T_{3A} , taken together with two electromagnetic generators, Q_{em} , since the trace over left-handed quarks of the quantity $\mathcal{A} = 2\text{tr}(T_{3A}Q_{\text{em}}^2)$ is nonzero. Its nonzero value is

$$\mathcal{A} = N_c \left[\left(\frac{2}{3}\right)^2 - \left(-\frac{1}{3}\right)^2 \right] = \frac{N_c}{3} \quad (9.120)$$

where the only contributions come from the u and d quarks, and the overall prefactor of $1/2$ arises because we take the eigenvalues of the $U_{3A}(1)$ generator to be $\pm 1/2$.

This anomaly is important for the pion Lagrangian because it says that the theory is *not* really $U_{3A}(1)$ invariant, even though it appears to be classically. Furthermore, it is possible to compute how the standard model action transforms under the anomalous $U_{3A}(1)$ symmetry, leading to the result

$$\delta\mathcal{L}_{SM} = \frac{e^2}{32\pi^2} \mathcal{A} F^{\mu\nu} F^{\lambda\rho} \epsilon_{\mu\nu\lambda\rho} \omega \quad (9.121)$$

Using Eq. (9.120) for \mathcal{A} we see from this that Eq. (9.118) precisely reproduces the transformation property which the anomaly requires, but only if there are precisely three colors: $N_c = 3$. The agreement between the pion decay rate computed with Eq. (9.117) and observations therefore provides direct evidence for the relevance of anomalies, and that there are precisely three colors.

This success is a special case of a more general principle, called *anomaly matching*, which simply states that the transformation properties of the action (invariance or anomalous transformation) should be the same, regardless of the scale at which it is examined. The consequences of this for the low-energy pion (or pseudoscalar meson) Lagrangian can be extended beyond the relatively simple implications of the $U_{3A}(1)$ symmetry to the complete case of $SU_L(3) \times SU_R(3)$. In this more general case the known standard model anomalies for the axial symmetries imply the existence of a more complicated symmetry-breaking term, called the *Wess–Zumino term*, for the low-energy meson effective action. These terms encode for the mesons the anomalous transformation properties of the standard model action.

It goes beyond the scope of our presentation to record here the explicit

form for the entire anomalous action, for the details of which we instead refer the reader to the literature. We do give the leading term it generates once it is expanded in powers of the meson field \mathcal{M} , which has the form

$$\mathcal{L}_{WZW} = \frac{N_c}{240\pi^2 F_\pi^5} \epsilon^{\mu\nu\lambda\rho} \text{tr} \left[\mathcal{M} \partial_\mu \mathcal{M} \partial_\nu \mathcal{M} \partial_\lambda \mathcal{M} \partial_\rho \mathcal{M} \right] + \cdots \quad (9.122)$$

where the ellipses represent terms involving more powers of \mathcal{M} . As above, the coefficient is fixed by demanding that its transformation under $SU_L(3) \times SU_R(3)$ reproduces the anomalies of the underlying quarks, and the resulting value is already used in Eq. (9.122). As usual $N_c = 3$ here denotes the number of quark colors. Although this interaction turns out to vanish if it is restricted to the $SU_L(2) \times SU_R(2)$ subgroup, it does have physical implications such as through its contribution to reactions like $K^+ K^- \rightarrow \pi^+ \pi^- \pi^0$.

9.4 Neutral-meson mixing

Neutral pseudoscalar mesons like $K^0 = d\bar{s}$ and $\bar{K}^0 = s\bar{d}$ (or $\bar{D}^0 = \bar{c}u$, $D^0 = \bar{u}c$, $\bar{B}_d^0 = b\bar{d}$, $B_d^0 = d\bar{b}$, and $\bar{B}_s^0 = b\bar{s}$, $B_s^0 = s\bar{b}$)[†] deserve special study because they turn out to provide particularly precise laboratories for testing the standard model. They do so because they (possibly together with the neutrinos) are particles which are not distinguished from their antiparticles by any exactly conserved charge, and so in principle the *mass* eigenstates can be linear combinations of the *flavor* eigenstates. This possibility gives rise to a variety of mixing phenomena whose properties turn out to depend sensitively on fairly detailed properties of the standard model.

9.4.1 Kinematics of meson mixing

Before describing the K and B systems in detail, we start with a general discussion of the special features of neutral-meson kinematics. To this end we first consider the effective Lagrangian which can be expected to govern the free propagation of a neutral spinless meson which differs from its antiparticle but which is not distinguished from it by a conserved charge. Such a meson can be described by a complex scalar field, φ , but one for which the transformation $\varphi \rightarrow e^{i\alpha} \varphi$ is *not* a symmetry.

Our interest in what follows is in effects which change flavor by two units, which first arise at order G_F^2 in the weak interactions. At this order it is

[†] Notice that for historical reasons for down-type quarks it is the meson containing the *antiquark* which is called meson rather than the antimeson. This convention is not used for mesons involving up-type quarks, such as the D system.

inconsistent to neglect the loss of probability associated with meson decays, whose rate is also $\Gamma \propto G_F^2$. As in earlier sections we do so by including a small imaginary part in the meson mass, $\mu = m - i\Gamma/2$, leading to a meson energy of the form

$$E_k = \sqrt{k^2 + \mu^2} \approx \sqrt{k^2 + m^2} - \frac{i\Gamma(k)}{2} \quad (9.123)$$

where $\Gamma(k) = m\Gamma/\sqrt{k^2 + m^2}$ is the decay rate for a moving particle, including relativistic time dilation.

Because this energy is complex, it is necessary to distinguish the charge-conjugate field,

$$\bar{\varphi}(x) = \int \frac{d^3k}{(2\pi)^3 2k^0} [\bar{a}_{\mathbf{k}} e^{ikx} + a_{\mathbf{k}}^* e^{-ikx}] \quad (9.124)$$

from the complex conjugate of the field

$$\varphi(x) = \int \frac{d^3k}{(2\pi)^3 2k^0} [a_{\mathbf{k}} e^{ikx} + \bar{a}_{\mathbf{k}}^* e^{-ikx}] \quad (9.125)$$

since $\bar{\varphi} \neq \varphi^*$ due to the energy which appears in the inner product $kx = -E_k t + \mathbf{k} \cdot \mathbf{x}$. Here $a_{\mathbf{k}}$ and $\bar{a}_{\mathbf{k}}$ are the destruction operators for the corresponding particle and antiparticle.

With this distinction in mind, and using the freedom to redefine fields to put the kinetic energy terms into canonical form, the most general quadratic (free) Lagrangian density for a complex scalar field is

$$\begin{aligned} \mathcal{L} &= -\partial_\mu \bar{\varphi} \partial^\mu \varphi - A \bar{\varphi} \varphi - \frac{1}{2} (B \varphi^2 + C \bar{\varphi}^2) \\ &= -\partial_\mu \bar{\varphi} \partial^\mu \varphi - \frac{1}{2} \begin{pmatrix} \varphi \\ \bar{\varphi} \end{pmatrix}^T \begin{pmatrix} B & A \\ A & C \end{pmatrix} \begin{pmatrix} \varphi \\ \bar{\varphi} \end{pmatrix} \end{aligned} \quad (9.126)$$

where we take A , B , and C to be complex constants. There are three important special cases of this Lagrangian density which are of interest for the present purposes.

Flavor conservation: The free propagation of φ particles preserves the $U(1)$ symmetry $\varphi \rightarrow e^{i\alpha} \varphi$ and $\bar{\varphi} \rightarrow e^{-i\alpha} \bar{\varphi}$ if and only if $B = C = 0$.

CP conservation: If we choose our phase conventions so that CP takes $\varphi \rightarrow \pm \bar{\varphi}$ then the fields $\varphi_+ = (\varphi + \bar{\varphi})/\sqrt{2}$ and $\varphi_- = i(\bar{\varphi} - \varphi)/\sqrt{2}$ are CP eigenstates and CP invariance for φ propagation requires $B = C$.

Unitary evolution: The propagation of φ is unitary if $\varphi^* = \bar{\varphi}$ and $\mathcal{L} = \mathcal{L}^*$, and this is so if and only if $A = A^*$ and $B = C^*$.

Although the first condition (flavor conservation) becomes exact in the limit that the weak interactions are turned off, none of the above three strictly holds when the weak interactions are turned on. They fail because the weak interactions break both the quark flavor symmetries and CP, and cause non-unitary evolution because the mesons also can decay into other particles through the weak interactions.

The mass term in Eq. (9.126) is diagonalized by the field redefinition

$$\varphi = z\psi, \bar{\varphi} = z^{-1}\bar{\psi} \quad \text{with} \quad z = \left(\frac{C}{B}\right)^{1/4} \quad (9.127)$$

Note that z is a pure phase in the unitary case. With these choices the propagation eigenstates are

$$\begin{aligned} \psi_+ &= \frac{1}{\sqrt{2}}(\bar{\psi} + \psi) = \frac{1}{\sqrt{2}}(z\bar{\varphi} + z^{-1}\varphi) \\ \text{and } \psi_- &= \frac{i}{\sqrt{2}}(\bar{\psi} - \psi) = \frac{i}{\sqrt{2}}(z\bar{\varphi} - z^{-1}\varphi) \end{aligned} \quad (9.128)$$

and the corresponding masses and decay rates are given by

$$\mu_{\pm}^2 \approx m_{\pm}^2 - im_{\pm}\Gamma_{\pm} = A \pm \sqrt{BC} \quad (9.129)$$

and so if we write $A = |A|e^{-i\alpha}$, $B = |B|e^{-i\beta}$, and $C = |C|e^{-i\gamma}$, then

$$\begin{aligned} m_{\pm}^2 &= |A| \cos \alpha \pm \sqrt{|BC|} \cos \left[\frac{1}{2}(\beta + \gamma) \right] \\ \text{and } m_{\pm}\Gamma_{\pm} &= |A| \sin \alpha \pm \sqrt{|BC|} \sin \left[\frac{1}{2}(\beta + \gamma) \right] \end{aligned} \quad (9.130)$$

Notice that $\Gamma_{\pm} = 0$ when $\bar{\varphi} = \varphi^*$ and $\mathcal{L} = \mathcal{L}^*$ since in this case $\alpha = \beta + \gamma = 0$.

Defining the states, $|\varphi\rangle$ and $|\bar{\varphi}\rangle$, destroyed by the fields $\varphi(x)$ and $\bar{\varphi}(x)$, together with $|\psi_{\pm}\rangle$ destroyed by the fields $\psi_{\pm}(x)$, we see these are related by

$$\begin{aligned} |\psi_+\rangle &= \frac{1}{\sqrt{|p|^2 + |q|^2}} [p|\varphi\rangle + q|\bar{\varphi}\rangle] \\ |\psi_-\rangle &= \frac{i}{\sqrt{|p|^2 + |q|^2}} [p|\varphi\rangle - q|\bar{\varphi}\rangle] \end{aligned} \quad (9.131)$$

where

$$\frac{p}{q} = z^2 = \left(\frac{C}{B}\right)^{1/2} \quad (9.132)$$

It is useful to relate these states to those which would be obtained if CP

were a symmetry. In this case $B = C$ and so $z = 1$, and so the propagation eigenstates would instead be

$$\varphi_+ = \frac{\varphi + \bar{\varphi}}{\sqrt{2}} \quad \text{and} \quad \varphi_- = \frac{i(\bar{\varphi} - \varphi)}{\sqrt{2}} \quad (\text{CP-preserving}) \quad (9.133)$$

with mass eigenvalues $\mu_{\pm}^2 \simeq m_{\pm}^2 - im\Gamma_{\pm} = A \pm B$. The eigenstates in the general case are related to these states, $|\varphi_{\pm}\rangle$, by

$$\begin{aligned} |\psi_+\rangle &= \frac{1}{\sqrt{1+|\hat{\epsilon}|^2}} [|\varphi_+\rangle - i\hat{\epsilon}|\varphi_-\rangle] \\ |\psi_-\rangle &= \frac{1}{\sqrt{1+|\hat{\epsilon}|^2}} [|\varphi_-\rangle + i\hat{\epsilon}|\varphi_+\rangle] \end{aligned} \quad (9.134)$$

where

$$\hat{\epsilon} = \frac{(z^2 - 1)}{1 + z^2} = \frac{p - q}{p + q} \quad (9.135)$$

The parameters z , m_{\pm} , and Γ_{\pm} are measured in oscillation phenomena, such as when a state is initially prepared as $|\varphi\rangle$ and then remeasured to be $|\varphi\rangle$ or $|\bar{\varphi}\rangle$ at a later time t . The probability amplitude for such a transition is given by expanding these states into propagation eigenstates and evolving these forward in time, to get

$$\langle\varphi|\varphi\rangle_t = \frac{1}{2} [e^{-iE_+t} + e^{-iE_-t}] \quad \langle\bar{\varphi}|\varphi\rangle_t = \frac{q}{2p} [e^{-iE_-t} - e^{-iE_+t}] \quad (9.136)$$

giving probabilities

$$\begin{aligned} P_t[\varphi(k) \rightarrow \varphi(k)] &= \frac{1}{4} \left[e^{-\Gamma_+(k)t} + e^{-\Gamma_-(k)t} + 2e^{-\frac{[\Gamma_+(k)+\Gamma_-(k)]t}{2}} \cos \Omega_k t \right] \\ P_t[\varphi(k) \rightarrow \bar{\varphi}(k)] &= \frac{1}{4} \left| \frac{q}{p} \right|^2 \left[e^{-\Gamma_+(k)t} + e^{-\Gamma_-(k)t} - 2e^{-\frac{[\Gamma_+(k)+\Gamma_-(k)]t}{2}} \cos \Omega_k t \right] \end{aligned} \quad (9.137)$$

where

$$\begin{aligned} \Omega_k &= \sqrt{k^2 + m_+^2} - \sqrt{k^2 + m_-^2} \approx (m_+ - m_-) \quad (\text{if } k \ll m_{\pm}) \\ &\approx \frac{m_+^2 - m_-^2}{2k} \quad (\text{if } k \gg m_{\pm}) \end{aligned} \quad (9.138)$$

9.4.2 Phenomenology of $K-\bar{K}$ mixing

Our first application of this formalism will be to the strange $K^0(=d\bar{s})$ and $\bar{K}^0(=\bar{d}s)$ mesons. In the absence of the weak interactions both strangeness

number (for which $S|K\rangle = +|K\rangle$ and $S|\bar{K}\rangle = -|\bar{K}\rangle$) and CP are conserved, ensuring that both K^0 and \bar{K}^0 are stable in this limit, and the effective description of the propagation of neutral K mesons would be described by Eq. (9.126), with $B = C = 0$ and $A = A^* = m_K^2$.

The possibility for oscillations arises once the weak interactions are turned on, since these generate both the $\Delta S = \pm 2$ terms in Eq. (9.126) parameterized by B and C , and an imaginary part for A describing the probability loss due to weak decays. Strange particles are typically produced through strong interactions between a beam and a target, and so are produced in pairs because of the strangeness conservation of the strong interactions. Neutral kaons are often produced amongst the decay products of the strange particles which arise in this way, and they start off in strangeness eigenstates because of their origin in the decay of hadrons having definite strangeness.

Once produced, the propagation eigenstates having definite mass and lifetime are denoted by $|K_S\rangle = |\psi_+\rangle$ and $|K_L\rangle = |\psi_-\rangle$, where the subscripts “S” and “L” denote “short” and “long” since $|K_L\rangle$ is very long-lived compared with the short-lived $|K_S\rangle$. These two states have very different lifetimes due to a combination of two effects: approximate CP conservation and the accident that m_K is not much bigger than $3m_\pi$.

As we show in more detail below, it turns out that for K mesons the weak interactions conserve CP to within a good approximation, and so in particular $|\epsilon| \ll 1$ in Eq. (9.134). Consequently, we can take to a good approximation $|K_L\rangle \approx |K_-\rangle = |\varphi_-\rangle$ and $|K_S\rangle \approx |K_+\rangle = |\varphi_+\rangle$, where $CP|K_\pm\rangle = \pm|K_\pm\rangle$.[†] This is significant because the main decay channel for neutral kaons is into pions, and the number of pions into which a kaon can decay depends upon its CP quantum numbers in a CP-conserving world. Since pions are CP odd, a two-pion state having angular momentum ℓ satisfies $CP|\pi\pi\rangle = (-)^2(-)^\ell|\pi\pi\rangle$. Since the K is spinless, the final state has no angular momentum; since the pion is also spinless, we must have $\ell = 0$. Therefore, only the CP-even state $|K_+\rangle$ can decay into two pions if CP is conserved. It is indeed true that by far the most common decay mode for K_S is into two pions, with $K_S \rightarrow \pi^+\pi^-$ (68.95 ± 0.14)% of the time and with $K_S \rightarrow \pi^0\pi^0$ (31.05 ± 0.14)% of the time.

Conversely, a purely pionic decay of the CP-odd state $|K_-\rangle$ must involve at least three pions, and so approximate CP-conservation would explain the observation that the purely pionic decay of the K_L is dominantly into three pions, with $K_L \rightarrow \pi^+\pi^-\pi^0$ in (12.59 ± 0.19)% of decays and $K_L \rightarrow \pi^0\pi^0\pi^0$ (21.05 ± 0.23)% of the time. (The remaining K_L decays are mostly

[†] Do not confuse K_\pm with the charged kaons, K^\pm .

semi-leptonic, into $\pi^\pm e^\mp \nu$ in $(38.81 \pm 0.27)\%$ of decays and into $\pi^\pm \mu^\mp \nu$ $(27.19 \pm 0.25)\%$ of the time.) Furthermore, because $m_K = 498$ MeV is just a little bigger than $3m_\pi = 3 \times 140 = 420$ MeV there is much less phase space available for the decay $K_L \rightarrow 3\pi$ than there is for the decay $K_S \rightarrow 2\pi$, with the result that the decay rates for these processes should differ considerably even if their underlying decay amplitudes are similar in size. Indeed, their observed mean lifetimes – $\tau_S = \Gamma_S^{-1} = 8.96 \times 10^{-11}$ s and $\tau_L = \Gamma_L^{-1} = 5.2 \times 10^{-8}$ s – imply that $\Gamma_S \approx 580\Gamma_L$.

Using $\Gamma_+ \gg \Gamma_-$ and assuming slowly-moving kaons leads to the following approximate form for the oscillation probabilities:

$$\begin{aligned} P_t[K(k) \rightarrow K(k)] &\approx \frac{1}{4} e^{-\Gamma_L(k)t} \left[1 + 2 e^{-[\Gamma_S(k) - \Gamma_L(k)]t/2} \cos(\Delta m t) \right] \\ P_t[K(k) \rightarrow \bar{K}(k)] &\approx \frac{1}{4} \left| \frac{q}{p} \right|^2 e^{-\Gamma_L(k)t} \left[1 - 2 e^{-[\Gamma_S(k) - \Gamma_L(k)]t/2} \cos(\Delta m t) \right], \end{aligned} \quad (9.139)$$

where $\Delta m = m_{K_L} - m_{K_S}$. These oscillations may be observed, such as by tagging the decaying K^0 or \bar{K}^0 through the charge of the lepton which is produced in the semileptonic decay, $K_L \rightarrow \pi^\pm \ell^\mp \nu$ (with $\ell = e, \mu$). Since this decay proceeds at the quark level either from $s \rightarrow u \ell^- \bar{\nu}$ or $\bar{s} \rightarrow \bar{u} \ell^+ \nu$, we see that the charge of the final lepton reveals what component of the kaon state is K^0 or \bar{K}^0 at the instant of decay.

Measurements of the number of positively and negatively charged leptons produced by a neutral kaon beam as a function of distance traveled by the kaons reveals these oscillations, with a frequency which is found to be comparable to the K_S lifetime. This frequency corresponds to the remarkably small mass difference $m_{K_L} - m_{K_S} = (3.483 \pm 0.006) \times 10^{-12}$ MeV. Such a small mass difference is only measurable at all because of the striking oscillation phenomenon to which it gives rise.

The sign of the mass difference is measurable because of the phenomenon of kaon *regeneration* in matter. When a beam of neutral kaons passes through matter, it interacts with atomic nuclei through the strong interactions. But because the basic reactions involved amount to quark rearrangements, K^0 mesons interact very differently than do \bar{K}^0 s. For example, the reaction $\bar{K}^0 p \rightarrow \Sigma^0 \pi^+$ can happen through the interchange of an s and u quark between the incident meson and the target baryon, but similar processes involving K^0 are impossible, since there is no light baryon containing an \bar{s} antiquark[†] and matter contains no antibaryons.

[†] The closest would be the Θ^+ pentaquark of Subsection 8.3.4; but even assuming it exists, it is heavier than $m_K + m_p$.

As a result of this difference, the \bar{K}^0 component of a neutral kaon beam can be rapidly absorbed by nuclear reactions as the beam passes through matter. Suppose the initial kaon beam had been in flight long enough before encountering the matter for its K_S component to completely decay (as can be observed by the disappearance of the tell-tale decays of the beam into two-pion final states). Because passage through matter removes the \bar{K}^0 component of the beam it has the effect of regenerating the beam's K_S component, leading to the reappearance of the two-pion decays immediately after the beam's passage through matter.

9.4.2.1 CP-violation

We have seen that if CP were conserved then it would forbid the long-lived neutral kaon from decaying into two pions. And since $K_L \rightarrow \pi^+\pi^-$ is observed in $(2.090 \pm 0.025) \times 10^{-3}$ of its decays (and into $\pi^0\pi^0$ in $(9.32 \pm 0.12) \times 10^{-4}$ of decays), CP cannot be exactly conserved. Historically, it was the observation of these rare K_L decays which led to the discovery of CP-violation in nature.

Suppose we denote by \mathcal{L}_w the effective term in the $\Delta S = \pm 1$ weak interaction Lagrangian density at low energies which is responsible for this decay. Then we know that the matrix element $\langle \pi\pi | \mathcal{L}_w | K_L \rangle$ is nonzero. In principle there are two ways in which CP-violation can lead to the decay $K_L \rightarrow \pi\pi$:

- (i) \mathcal{L}_w can be CP-invariant but K_L contains a small admixture of K_+ (i.e. $\hat{\epsilon} \neq 0$ in Eq. (9.134)), allowing the decay to proceed through the CP-preserving matrix element $\langle \pi\pi | \mathcal{L}_w | K_+ \rangle$; or
- (ii) \mathcal{L}_w itself breaks CP, allowing the decay to proceed due to the matrix element $\langle \pi\pi | \mathcal{L}_w | K_- \rangle$

These options respectively correspond to the cases where the dominant source of CP-violation has the selection rule $\Delta S = \pm 2$ or ± 1 . Naturally, they may both be present.

Which of these options actually arises in nature can be quantified by comparing the sizes of several different CP-violating observables in neutral kaon decays. Three such observables which can be compared in this way are

$$\eta_{00} = \frac{\langle \pi^0\pi^0 | \mathcal{L}_w | K_L \rangle}{\langle \pi^0\pi^0 | \mathcal{L}_w | K_S \rangle}, \quad \eta_{+-} = \frac{\langle \pi^+\pi^- | \mathcal{L}_w | K_L \rangle}{\langle \pi^+\pi^- | \mathcal{L}_w | K_S \rangle} \quad (9.140)$$

and

$$\delta_L = \frac{\Gamma(K_L \rightarrow \pi^+\ell^-\nu) - \Gamma(K_L \rightarrow \pi^-\ell^+\nu)}{\Gamma(K_L \rightarrow \pi^+\ell^-\nu) + \Gamma(K_L \rightarrow \pi^-\ell^+\nu)} \quad (9.141)$$

These would all be given in terms of the single parameter $\hat{\epsilon}$ if mixing were

the sole source of CP-violation: $\eta_{00} = \eta_{+-} = -\hat{\epsilon}$ and $\delta_L = (|q|^2 - |p|^2)/(|q|^2 + |p|^2) = -(\hat{\epsilon} + \hat{\epsilon}^*)/(1 + |\hat{\epsilon}|^2)$.

More generally these quantities depend on more parameters. For instance suppose that both $\hat{\epsilon} \neq 0$ and $\langle \pi\pi | \mathcal{L}_w | K_- \rangle \neq 0$. In order to express how η_{00} and η_{+-} can differ from one another in this case it is useful first to decompose the $K \rightarrow \pi\pi$ amplitude into its isospin components. Recall to this end that neutral kaons are isodoublets and pions are isotriplets, and so a spinless state involving two pions must have isospin 0 or 2. We define the isospin 0 and 2 amplitudes by

$$\begin{aligned} \langle \pi^+ \pi^- | \mathcal{L}_w | K^0 \rangle &= A_0 e^{i\xi_0} + \frac{A_2 e^{i\xi_2}}{\sqrt{2}} \\ \text{and} \quad \langle \pi^+ \pi^- | \mathcal{L}_w | \bar{K}^0 \rangle &= A_0 e^{-i\xi_0} + \frac{A_2 e^{-i\xi_2}}{\sqrt{2}} \end{aligned} \quad (9.142)$$

where A_0 and A_2 are the relevant (complex) CP conserving strong interaction matrix elements for each pionic isospin channel, while ξ_0 and ξ_2 are the CP-violating phases which arise due to CP-violation within \mathcal{L}_w .

Notice that since the CP-violating phases only enter into physical decay rates through the interference term in $|A_0 e^{i\xi_0} + A_2 e^{i\xi_2}|^2$, in the absence of any other sources of CP-violation it is only the *relative* CP-violating phase, $\xi_2 - \xi_0$, which can have physical consequences for neutral K decays. With these definitions we therefore have

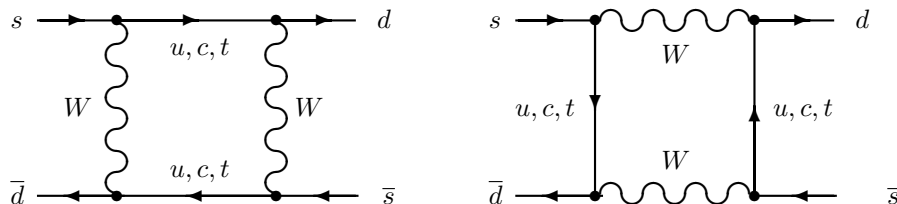
$$\eta_{+-} = \epsilon + \epsilon' \quad \eta_{00} = \epsilon - 2\epsilon' \quad \text{and} \quad \delta_L = \frac{2 \operatorname{Re} \epsilon}{1 + |\epsilon|^2} \quad (9.143)$$

where, assuming $\hat{\epsilon}, \xi_0, \xi_2 \ll 1$, we define

$$\begin{aligned} \epsilon &= -\hat{\epsilon} + \xi_0 \\ \text{and} \quad \epsilon' &= \left(\frac{A_2}{A_0 + A_2} \right) (\xi_2 - \xi_0) \end{aligned} \quad (9.144)$$

The definitions of ϵ and ϵ' here are made in such a way that they are invariant under any rephasing of the states K^0 and \bar{K}^0 . In particular, ϵ' gives the measure of “direct” CP-violation by \mathcal{L}_w , depending as it does only on the phase difference, $\xi_2 - \xi_0$. At the time of this writing the survey of current measurements given by the Particle Data Group gives evidence that both ϵ and ϵ' are nonzero,

$$\begin{aligned} \operatorname{Re} \epsilon &= (1.657 \pm 0.021) \times 10^{-3} \\ \operatorname{Im} \epsilon &= (1.572 \pm 0.022) \times 10^{-3} \\ \operatorname{Re} \left(\frac{\epsilon'}{\epsilon} \right) &= (1.67 \pm 0.26) \times 10^{-3} \end{aligned} \quad (9.145)$$

Fig. 9.7 Box diagrams for $\Delta S = \pm 2$ interactions.

although the evidence for $\epsilon' \neq 0$ has involved some controversy due to there being conflicting results from competing experiments. This reveals a pattern wherein (i) CP-violation exists but is a small effect, and (ii) direct CP-violation also exists but for neutral kaons is suppressed relative to the CP-violating contribution from K^0 - \bar{K}^0 oscillations.

9.4.2.2 Standard model predictions

We next turn to what the standard model predicts for these neutral-kaon parameters, and the extent to which accurate predictions can be made. As we shall see, the extreme precision of the measured value for $m_{K_L} - m_{K_S}$ imposes strong constraints on the model.

Our interest is in the low-energy effective theory defined at energies below the W mass but above the QCD scale. In this theory the W and Z bosons do not appear since they are “integrated out,” and strongly-interacting particles are still described by quarks rather than the hadronic bound states. We discussed this effective theory extensively in Chapter 7. As we saw there, at the renormalizable level each quark number is separately conserved; so both mixing and decay of the K mesons is forbidden. Therefore we must investigate high dimension operators, specifically dimension-6, four-fermion operators. The lowest-order operators encountered in Chapter 7 all change strangeness by at most 1 unit, and so cannot contribute to $K\bar{K}$ mixing. They also conserve CP, and so cannot contribute to direct CP-violation. Therefore we must consider dimension-6 operators which arise from loops.

The leading $\Delta S = \pm 2$ contribution to the low-energy Lagrangian arises from the diagrams of Figure 9.7. Its contribution to the Lagrangian takes the form

$$\mathcal{L}_w^{(|\Delta S|=2)} = \mathcal{C}[\bar{d}\gamma^\mu P_L s][\bar{d}\gamma_\mu P_L s] + \text{c.c.} \quad (9.146)$$

Naively, since the diagrams in Figure 9.7 are one loop and since the result

is a dimension-6 operator, we would expect the coefficient to be of order

$$\mathcal{C} \sim \frac{g_2^4}{4\pi^2 M_W^2} \quad (9.147)$$

This estimate would lead to a mass splitting of order $m_K^3 g_2^4 / (4\pi^2 M_W^2) \sim 10^{-4}$ MeV, which is drastically larger than the observed value.

In fact, details of the structure of the diagrams make the contribution much smaller, as we now explore. Each diagram contains two virtual quark lines, which can be any of the up-type quarks, u, c, t . We must sum over each possibility; each gives the same answer, except that the CKM angles appearing in the vertices differ, and the masses appearing on the propagators also differ. Therefore, a parametrization of the coefficient \mathcal{C} is,

$$\mathcal{C} = \frac{G_F^2}{16\pi^2} \sum_{i,j=u,c,t} V_{id}^* V_{is} V_{jd}^* V_{js} f(m_i^2, m_j^2, m_s^2, m_d^2, M_W^2, \alpha_3) \quad (9.148)$$

The function f has dimension $(\text{mass})^2$ and is obtained by evaluating the lowest-order graphs contributing to $\Delta S = \pm 2$ transitions, given by the “box” diagrams of Figure 9.7. At the level discussed so far, f is α_3 independent; but because α_3 is not very small, especially at low momentum scales, one should include corrections arising from dressing these graphs in all possible ways with gluon lines.

The key observation is that, if all of the quarks in the loop had the same mass, the function f would be the same for each i and j . Therefore, the coefficient \mathcal{C} would be proportional to $\sum_i V_{id}^* V_{is}$. However, this combination of CKM elements actually vanishes: $\sum_i V_{id}^* V_{is} = (V^\dagger V)_{ds} = \delta_{ds} = 0$. This implies that the result for this graph must be further suppressed by factors of small quark masses. In particular, because of the massive W -boson propagators in the diagram, the typical momentum “flowing” in the loop is $\sim M_W$; therefore, the masses on the fermion lines represent $O(m^2/M_W^2)$ corrections to these propagators, and so the function f above must depend weakly on small quark masses, $f(m) = f(0)$ up to m_i^2/M_W^2 corrections. Therefore, the cancellation we have just found should be valid up to m_i^2/M_W^2 corrections. This argument obviously does not apply to the top quark, but in fact the top-quark contribution is suppressed by small mixing angles, $|V_{td}V_{ts}|^2 \sim (0.2)^{10} \approx 6 \times 10^{-8}$. Cancellations such as these, in which loop-generated flavor-changing neutral currents are suppressed by small mass ratios due to the unitarity of the KM matrix, are known as the *GIM mechanism*, after its discoverers Glashow, Iliopoulos, and Maiani.

These general observations are borne out by the explicit evaluation of the

diagrams of Figure 9.7. Neglecting m_u^2/M_W^2 in these diagrams gives

$$\begin{aligned} \mathcal{C} = & \frac{G_F^2}{16\pi^2} \left[(V_{td}^* V_{ts})^2 m_t^2 \eta_t h_1(x_t) + (V_{cd}^* V_{cs})^2 m_c^2 \eta_c h_1(x_c) \right. \\ & \left. - 2 V_{td}^* V_{ts} V_{cd}^* V_{cs} \left(\frac{m_c^2 m_t^2}{M_W^2} \right) \eta_{tc} h_2(x_c, x_t) \right] \end{aligned} \quad (9.149)$$

where $x_i = m_i^2/M_W^2$ and

$$\begin{aligned} h_1(x) &= \frac{1}{4} + \frac{9}{4(1-x)} - \frac{3}{2(1-x)^2} - \frac{3x^2}{2(1-x)^3} \ln x \\ h_2(x, y) &= \left\{ \left[\frac{1}{4} + \frac{3}{2(1-x)} - \frac{3}{4(1-x)^2} \right] \frac{\ln x}{y-x} + (x \leftrightarrow y) \right\} \\ &\quad + \frac{3}{4(1-x)(1-y)} \end{aligned} \quad (9.150)$$

The QCD correction factors, η_t , η_c , and η_{tc} , are all unity if only the box graphs are evaluated, but become non-trivial once gluon lines are included. For a QCD scale of 200 MeV, these QCD correction factors become $\eta_t \approx 0.62$, $\eta_c \approx 0.85$, and $\eta_{tc} \approx 0.36$.

We may simplify the expression for \mathcal{C} using $x_c \ll 1$ and $x_t \gg 1$, which implies $h_1(x_t) \approx \frac{1}{4} + O[(\ln x_t)/x_t]$, $h_1(x_c) \approx 1 + O(x_c)$ and

$$h_2(x_c, x_t) \approx \frac{1}{x_t} \left(-\frac{3}{4} + \ln x_c - \frac{\ln x_t}{4} \right) \quad (9.151)$$

In order to see the relative size of the three terms in Eq. (9.149) it is useful to adopt the Wolfenstein parameterization, Eq. (2.92), of the CKM matrix, for which $V_{cs} \approx 1$, $V_{cd} \approx -\lambda$, $V_{ts} \approx -A\lambda^2$, and $V_{td} \approx A\lambda^3(1 - \rho - i\eta)$, where A , ρ , and η are at most $O(1)$ and the small quantity $\lambda \approx 0.2$ expresses the amount of suppression of flavor-changing in the charged-current weak interactions. Ignoring the $O(1)$ factors, the relative size of the three terms contributing to \mathcal{C} are of order

$$\lambda^{10} m_t^2 : \lambda^2 m_c^2 : \lambda^6 m_c^2 \approx 0.04 : 1 : 0.002 \quad (9.152)$$

which shows that the small off-diagonal entries of the CKM matrix overwhelm the large t -quark mass to leave the c -quark graph as dominant.

It is the matrix elements of this $\Delta S = \pm 2$ effective interaction which govern the size of the K^0 - \bar{K}^0 mixing matrix elements, B and C . If we estimate the size of the matrix element $\langle K^0 | [\bar{d}\gamma^\mu P_L s] [\bar{d}\gamma_\mu P_L s] | \bar{K}^0 \rangle$ by the appropriate power of the strong-interaction scale, m_K , then we obtain the following size for the $\Delta S = \pm 2$ terms in the effective neutral-kaon mass

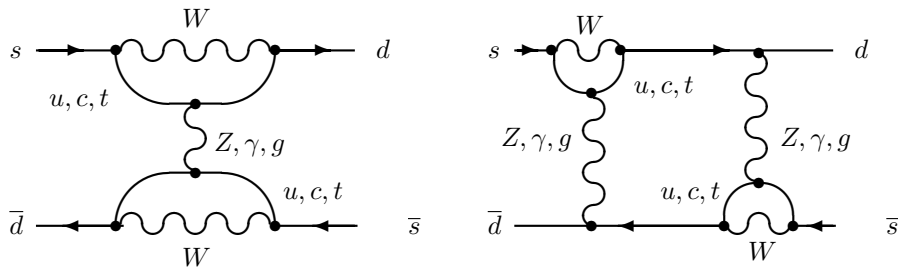


Fig. 9.8 Potentially large higher-order diagrams for $\Delta S = \pm 2$ interactions.

matrix,

$$|B|, |C| \sim \mathcal{C} m_K^4 \sim \frac{G_F^2 \lambda^2 m_c^2 m_K^4}{16\pi^2} \sim 4 \times 10^{-15} \text{ GeV}^2 \quad (9.153)$$

This leads to the following remarkably accurate estimate for the K_L – K_S mass difference: $\Delta m \sim 10^{-12} \text{ MeV}$.

“Long-distance” corrections

We see from this that the standard model can account for the size of K^0 – \bar{K}^0 mixing, but only because large cancellations amongst the contributions of heavy up-type quarks imply a result which is strongly suppressed by small quark masses and/or small CKM mixing angles. Since this makes the result smaller than a generic one-loop result, it is necessary to check that the result is not smaller than contributions which are nominally higher-order in the gauge interactions but less suppressed by masses and mixings. One might worry in particular that QCD corrections could be large, since for these α_3 need not be very small (as we have seen in the numerical values quoted above for the QCD correction factors, η_t , η_c , and η_{tc}).

Two examples of such graphs are given in Figure 9.8. Of these, the left-hand graph involves an additional two gauge coupling constants while the right-hand graph is an iteration of the loop-corrected effective $\Delta S = \pm 1$ interaction (more about which below). Both involve an additional loop, but when the virtual quarks and gluons have low energies the additional coupling factor, $\alpha_3/4\pi$, need not be small. These kinds of contribution turn out to generate an effective $\Delta S = \pm 2$ coupling whose coefficient differs from that computed above in that it need not be proportional to a squared quark mass. This is possible because for these graphs the GIM cancelation is *logarithmic*,

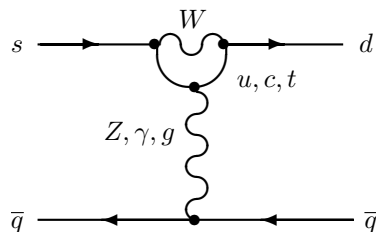


Fig. 9.9 Diagrams introducing CP-violation into the $\Delta S = \pm 1$ interactions.

involving virtual-quark sums of the form

$$\sum_{i=u,c,t} V_{id}^* V_{is} \ln \left(\frac{m_i^2}{M_W^2} \right) \quad (9.154)$$

As a result, such higher-order contributions are typically suppressed in neutral kaon mixing, relative to that computed above, by powers of a strong-interaction scale over the relevant heavy quark mass, such as by m_K/m_c . Although these do not ruin the order-of-magnitude agreement between the standard model and experiment, they do seriously complicate making a more precise comparison between theory and experiment for neutral kaon mixing. As we shall see, these same complications are less important for mixing amongst mesons involving heavier quarks, such as for neutral B mesons.

CP-violation

The standard model provides a similar qualitative (but not yet quantitatively precise) understanding of the size of direct and indirect CP-violation, as is measured by $\epsilon \sim 10^{-3}$ and $\epsilon' \sim 10^{-6}$ for neutral kaons. For instance, ϵ measures the strength of CP-violation in kaon mixing relative to the CP-preserving terms, and since the effects of CP-violation can be banished (such as in the Wolfenstein parameterization) to those terms of the CKM matrix which involve third-generation quarks, a phase can only enter into the effective $\Delta S = \pm 2$ interactions through the graphs involving virtual t -quarks. However, we have seen that, for neutral kaons, the virtual t -quark contribution is suppressed by small mixing angles relative to the dominant c -quark term, leading to the reasonably successful estimate $\epsilon \sim \lambda^8 (m_t/m_c)^2 \sim 10^{-2}$.

Attempts to similarly estimate the parameter ϵ' are more difficult due to bigger uncertainties to do with the strong interactions. In particular, as the definition, Eq. (9.144), shows, ϵ' is only sensitive to the *relative* phase

between the parts of the interaction which change isospin by $3/2$ and by $1/2$. An important complication on the theoretical side is the fact that the amplitude of $\Delta I = 1/2$ transitions are observed experimentally to be about 22 times stronger in K decays than are the $\Delta I = 3/2$ ones, as discussed in Subsection 9.3.1.

Since the tree-level contributions to kaon decays do not involve the third generation, they cannot violate CP. However, direct CP-violation can enter at one loop in the effective $\Delta S = \pm 1$ four-quark interactions through the graphs of Figure 9.9 – known for historical reasons as “penguin” diagrams. Although the gluon-exchange diagrams would naively be expected to dominate those involving the Z or photon, this need not be true for their contributions to ϵ' because gluon exchange contributes only to the $\Delta I = 1/2$ transition.

The upshot is that ϵ' is expected to be nonzero and small within the standard model, but a precise comparison with experiments remains difficult due to the uncertainties associated with the strong-interaction matrix elements.

9.4.3 Phenomenology of B – \bar{B} mixing

Oscillations and CP-violation have also been observed for neutral mesons involving b quarks. In the standard model, these phenomena arise for the same reasons as in the K meson system; the standard model therefore makes unambiguous predictions for what should be seen for B mesons, and because the b quark is much heavier than the s quark, these predictions are often less clouded by matrix-element uncertainties than for K mesons. This potential for meaningful comparison between theory and experiment has stimulated considerable experimental effort towards measuring the properties of B mesons, including the construction of “ B factories,” where the copious production of B mesons allows unprecedented precision.

The relatively large b quark mass compared with the QCD scale provides the main difference between the analysis of B and K mixing. This mass has several effects.

- First, the much larger phase space available in B decays allows many more decay channels and removes the coincidence in masses which made the lifetime of the K_L so different from that of the K_S . Consequently the oscillation kinematics can be simplified by neglecting the difference between the decay rates of the two propagation eigenstates relative to the decay rates themselves $\Delta\Gamma \ll \Gamma$.

- Second, the B -meson decays are more heavily CKM suppressed than K -meson decays (by λ^4 rather than λ^2). This makes the B -meson decay lifetimes comparatively longer than would otherwise be true for such a heavy meson. For instance the B^\pm and neutral B_d lifetimes are of order 10^{-12} s, which is just long enough for the mesons to travel observable distances (up to a millimeter) within detectors before they decay. Since both Δm and Γ are $O(G_F^2)$, we generically expect $\Gamma \sim \Delta m \ll m$ for B mesons.
- Finally, the large matrix element $V_{tb} \approx 1$ implies there is no suppression of the virtual t -quark in loops, and so no trade-off between small mixing angles and enhancement by powers of m_t in mixing- and CP-violating parameters. This allows the t -quark contribution to dominate, and makes the mixing parameters depend strongly on m_t .

The propagation eigenstates for B mesons are denoted B_H and B_L , with the “heavy” meson (B_H) defined to be the more massive of the two. For most purposes the transition probability for oscillating from B^0 to \bar{B}^0 is given by specializing Eq. (9.137) to the case $\Gamma_H \approx \Gamma_L = \Gamma$, and taking $\Delta m = m_{B_H} - m_{B_L}$ comparable in size to Γ . Assuming non-relativistic B s leads to

$$\begin{aligned} P_t[B^0 \rightarrow B^0] &= e^{-\Gamma t} \cos^2\left(\frac{\Delta m t}{2}\right) \\ P_t[B^0 \rightarrow \bar{B}^0] &= \left|\frac{q}{p}\right|^2 e^{-\Gamma t} \sin^2\left(\frac{\Delta m t}{2}\right) \end{aligned} \quad (9.155)$$

which shows that the time-dependence of $B\bar{B}$ oscillations is controlled by the dimensionless ratio $\xi = \Delta m/\Gamma$. The non-relativistic approximation is appropriate for B factories, where they are produced by pair-production in the reaction $e^+e^- \rightarrow B\bar{B}$ with the beam energy adjusted to run at or near to a resonance corresponding to a bound state in the $b\bar{b}$ system – the $\Upsilon(4S)$ state. Because the mass of this state is just above the threshold for $B\bar{B}$ production, the final-state B and \bar{B} are produced almost at rest.

In practice the comparatively short distance traveled by the B meson before it decays makes it more difficult to measure the time, t , elapsed since the state was a pure B^0 eigenstate (compared with what is possible for kaons). Since the B s are made through electron–positron annihilations *via* a virtual photon: $e^+e^- \rightarrow \gamma^* \rightarrow B\bar{B}$, the produced B mesons start off in a state with relative angular momentum $\ell = 1$ (p -wave) and with $CP|B\bar{B}\rangle = -|B\bar{B}\rangle$. This implies the initial state necessarily has entanglement between

the B -meson momentum and flavor:

$$|B\bar{B}\rangle_{t=0} = \frac{1}{\sqrt{2}} \left[|B(k)\bar{B}(-k)\rangle - |B(-k)\bar{B}(k)\rangle \right] \quad (9.156)$$

Starting with this state, oscillations can be observed by tagging the decays of the B and \bar{B} mesons, such as by observing the lepton charge in their semileptonic decays, $\bar{B}^0 \rightarrow X\ell^-\nu$ and $B^0 \rightarrow X'\ell^+\nu$, where X denotes an arbitrary hadronic state. These decays represent about 10% of the B meson decay width per lepton flavor (e and μ). If the tracks of the daughter particles are precisely measured, the point where these decays take place can be precisely determined. Then the above formulae give the probability of seeing semileptonic decays into same-sign or opposite-sign leptons as a function of the distance (and so also time elapsed) between the decays.

Observations of B_d - \bar{B}_d oscillations show that the mixing parameter, ξ_d , for these decays is given by

$$\begin{aligned} \xi_d &= \left(\frac{\Delta m}{\Gamma} \right)_{B_d} = 0.771 \pm 0.012 \\ \text{or } \Delta m_{B_d} &= (3.304 \pm 0.046) \times 10^{-10} \text{ MeV} \end{aligned} \quad (9.157)$$

Very recent (2006) preliminary results on B_s - \bar{B}_s oscillations indicate

$$\begin{aligned} \xi_s &= \left(\frac{\Delta m}{\Gamma} \right)_{B_s} \simeq 25 \\ \Delta m_{B_s} &\simeq 110 \times 10^{-10} \text{ MeV} \end{aligned} \quad (9.158)$$

Standard-model predictions

The standard-model predictions for B - \bar{B} mixing are given by the same box graphs, Figure 9.7, as for kaon mixing, whose evaluation gives the following $\Delta B = \pm 2$ contributions to the low-energy Lagrangian,

$$\mathcal{L}_w^{(|\Delta B|=2)} = \mathcal{C}_d [\bar{d}\gamma^\mu P_L b] [\bar{d}\gamma_\mu P_L b] + \mathcal{C}_s [\bar{s}\gamma^\mu P_L b] [\bar{s}\gamma_\mu P_L b] + \text{c.c.} \quad (9.159)$$

where the coefficients \mathcal{C}_d and \mathcal{C}_s are given by

$$\begin{aligned} \mathcal{C}_q &= \frac{G_F^2}{16\pi^2} \left[(V_{tq}^* V_{tb})^2 m_t^2 \eta_t h_1(x_t) + (V_{cq}^* V_{cb})^2 m_c^2 \eta_c h_1(x_c) \right. \\ &\quad \left. - 2V_{tq}^* V_{tb} V_{cq}^* V_{cb} \left(\frac{m_c^2 m_t^2}{M_W^2} \right) \eta_{tc} h_2(x_c, x_t) \right] \end{aligned} \quad (9.160)$$

Here $q = d, s$ and we define as before $x_i = m_i^2/M_W^2$, with the functions h_1 and h_2 as given in Eq. (9.150). The above expression also uses $x_u \approx 0$.

In this case the relative size of the charm and top-quark terms in \mathcal{C}_q are of order

$$\begin{aligned} \lambda^6 m_t^2 : \lambda^6 m_c^2 : \lambda^6 m_c^2 & \quad (\mathcal{C}_d) \\ \lambda^4 m_t^2 : \lambda^4 m_c^2 : \lambda^4 m_c^2 & \quad (\mathcal{C}_s) \end{aligned} \quad (9.161)$$

which shows that the t -quark contribution dominates in both cases. This dominance also guarantees that the contributions of graphs like those of Figure 9.8, involving higher-order QCD corrections, are less important than for kaon mixing, being suppressed by powers of $m_B^2/m_t^2 \ll 1$.

Neglecting all but the t -quark contribution, the prediction for the relative size of Δm for the B_d and B_s systems becomes

$$\frac{(\Delta m)_{B_d}}{(\Delta m)_{B_s}} = \frac{\mathcal{A}_d |V_{td}|^2}{\mathcal{A}_s |V_{ts}|^2} = \left[\frac{\mathcal{A}_d}{\mathcal{A}_s} \right] \lambda^2 |1 - \rho - i\eta|^2 \quad (9.162)$$

where $\mathcal{A}_q = \langle B_q^0 | [\bar{q}\gamma^\mu P_L b] [\bar{q}\gamma_\mu P_L b] | \bar{B}_q^0 \rangle$ is the required strong-interaction matrix element. Assuming the matrix elements involved to be of the same order of magnitude, we see that mixing effects for the B_s system should be enhanced relative to the B_d system because of the factor $\lambda^{-2} \sim 25$.

The mixing parameter p/q may be similarly obtained. To the extent that we may neglect $\Delta\Gamma$, the parameters B and C of the effective neutral meson Lagrangian, Eq. (9.126), are simply given by $B = \mathcal{C} \mathcal{A}$ and $C = \mathcal{C}^* \bar{\mathcal{A}}$, where \mathcal{A} and $\bar{\mathcal{A}}$ denote the matrix elements

$$\mathcal{A} = \langle B^0 | [\bar{q}\gamma^\mu P_L b] [\bar{q}\gamma_\mu P_L b] | \bar{B}^0 \rangle \quad \text{and} \quad \bar{\mathcal{A}} = \langle \bar{B}^0 | [\bar{b}\gamma^\mu P_L q] [\bar{b}\gamma_\mu P_L q] | B^0 \rangle \quad (9.163)$$

The CP-invariance of the strong interactions implies $\bar{\mathcal{A}} = \mathcal{A}$, and so using this in Eq. (9.132) we obtain (when $\Delta\Gamma \approx 0$) the following result for the mixing parameter: $p/q = (C/B)^{1/2} \approx (\mathcal{C}^*/\mathcal{C})^{1/2}$.

This shows that for B mesons we expect $|p/q| = 1$ to good accuracy. This agrees well with the measured value $|q/p| = 0.998 \pm 0.007$. The phase of p/q depends only on the phase of the relevant CKM matrix elements (up to an accuracy of m_c^2/m_t^2 , using only the t -quark graph):

$$\begin{aligned} \left(\frac{p}{q} \right)_{B_d} &= \frac{V_{td} V_{tb}^*}{V_{td}^* V_{tb}} \approx \frac{1 - \rho - i\eta}{1 - \rho + i\eta} \\ \left(\frac{p}{q} \right)_{B_s} &= \frac{V_{ts} V_{tb}^*}{V_{ts}^* V_{tb}} \approx 1 + O(\lambda^2) \end{aligned} \quad (9.164)$$

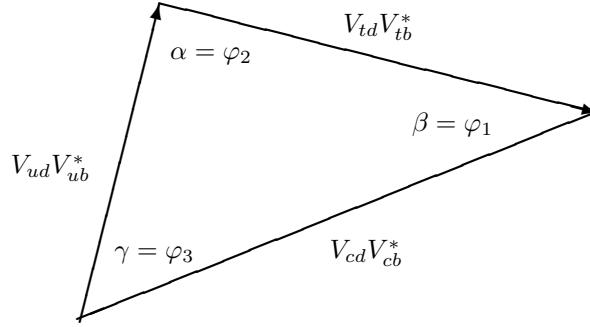


Fig. 9.10 The “unitarity triangle.”

CP-violation

The standard model’s central prediction is that all flavor-changing and CP-violating physics is due to the four parameters of the unitary CKM matrix, and the B -meson system provides a multitude of observables against which this may be tested. One approach is to infer the values of the appropriate CKM matrix elements from various observables and then to check whether the results agree with the unitarity conditions $\sum_i V_{in} V_{im}^* = \delta_{mn}$. As of this writing, all such tests support these standard model predictions.

Many B -physics experiments probe in particular the special case $m = b$ and $n = d$, for which unitarity implies

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \quad (9.165)$$

This condition is often called the “unitarity triangle” because it is the statement that the sum of three complex numbers vanish, and so can be visualized as saying that three vectors in the complex plane sum to zero (and so form a triangle, see Figure 9.10).

One can define a similar triangle for any pair of distinct rows or columns of the CKM matrix, and it turns out that all of these triangles have the same area, conventionally denoted $J/2$. Within the Wolfenstein parameterization $J \approx A^2 \lambda^6 \eta$, and so $J \rightarrow 0$ in the limit where CP is preserved ($\eta \rightarrow 0$). Since the triangle therefore degenerates into a line when CP is unbroken, the angles of the unitarity triangle are good measures of CP-violation. These angles are conventionally denoted α , β , and γ (or alternatively $\varphi_1 = \beta$, $\varphi_2 = \alpha$, and $\varphi_3 = \gamma$), and are defined in terms of the CKM matrix elements by

$$\alpha = \arg \left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*} \right), \quad \beta = \arg \left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*} \right), \quad \text{and} \quad \gamma = \arg \left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right) \quad (9.166)$$

A particularly clean measure of CP-violation in the B_d system compares the rates with which B_d^0 and \bar{B}_d^0 decay into a CP eigenstate, like ψK_S , ψK_L , $\pi^+\pi^-$, and so on. (Here ψ represents the long-lived $c\bar{c}$ bound state.) For example, suppose at $t = t_1$ one member of the $B\bar{B}$ pair is tagged as a B^0 by observing the charge of the lepton emitted in its semileptonic decay, and at a later time $t = t_2$ the other meson decays into ψK_S . We now calculate how the difference between these rates depends on the above angles, and so compute

$$D(t_1, t_2) = \Gamma[(\ell^+ X \nu)_{t_1}; (\psi K_S)_{t_2}] - \Gamma[(\ell^- X \nu)_{t_1}; (\psi K_S)_{t_2}] \quad (9.167)$$

Since we have seen that $|p/q| \approx 1$ for B mixing, we may write $p/q = e^{i\omega}$, where $\omega = \arg[(V_{td}V_{tb}^*)/(V_{td}^*V_{tb})]$. The direct CP-violation from the $\Delta B = \pm 1$ terms, \mathcal{L}_w , in the low-energy weak interactions are similarly parameterized by writing

$$\langle \psi K_S | \mathcal{L}_w | B^0 \rangle = A e^{i\xi/2} \quad \text{and} \quad \langle \psi K_S | \mathcal{L}_w | \bar{B}^0 \rangle = A e^{-i\xi/2} \quad (9.168)$$

Here A can be complex but is CP-preserving, and ξ represents the direct CP-violating phase due to the CKM matrix elements in this transition. For the decays $B^0 \rightarrow \psi K_S$ and $\bar{B}^0 \rightarrow \psi K_S$, the dominant decay process is believed to be the direct quark decays $b \rightarrow c\bar{c}s$ and $\bar{b} \rightarrow \bar{c}c\bar{s}$, and so $\xi = \arg[(V_{cs}^*V_{cb})/(V_{cs}V_{cb}^*)] \approx \arg[(V_{cd}^*V_{cb})/(V_{cd}V_{cb}^*)]$. The last equality uses that V_{cs} and V_{cd} are both real to within the accuracy required.

Using these expressions gives

$$\begin{aligned} D(t_1, t_2) &= 2|A|^2 e^{-\Gamma(t_1+t_2)} \sin(\omega + \xi) \sin[\Delta m(t_2 - t_1)] \\ &= -2|A|^2 e^{-\Gamma(t_1+t_2)} \sin(2\beta) \sin[\Delta m(t_2 - t_1)] \end{aligned} \quad (9.169)$$

where we use

$$e^{i(\omega+\xi)} \approx \left[\frac{V_{cd}^*V_{cb}}{V_{cd}V_{cb}^*} \right] \left[\frac{V_{td}V_{tb}^*}{V_{td}^*V_{tb}} \right] \approx \left[\frac{V_{cd}^*V_{cb}}{V_{td}^*V_{tb}} \right] \left[\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} \right] = e^{-2i\beta} \quad (9.170)$$

Notice that $D(t_1, t_2)$ is odd under $t_1 \leftrightarrow t_2$, and so vanishes once it is integrated over all times. This is a general consequence of the entanglement in the initial meson state, and makes the measurement of this asymmetry more complicated. It is in order to circumvent this kind of vanishing that B factories are often built asymmetrically, so that the initial e^+e^- pair collide with a net overall momentum. This ensures that the center of mass of the $B\bar{B}$ pair moves, thereby separating the decay products and so allowing the measurement of the time difference $t_1 - t_2$.

An enormous theoretical and experimental effort has been devoted to understanding in detail the decays and asymmetries in a great variety of rare

and common B decays, and how best to make contact with the standard-model predictions. We do not try to describe this literature since it would force us to stray too far from our main line of development. Suffice it to say that at present all measurements are consistent with the standard-model predictions.

9.4.4 $D-\bar{D}$ mixing

Electrically neutral charmed mesons, $D^0 = (\bar{u}c)$ and $\bar{D}^0 = (\bar{c}u)$, are also expected to oscillate. Within the standard model, we expect these oscillations to be unobservably small. The reason is that the decay process is not suppressed by small CKM angles, but the mass splitting is CKM suppressed, and GIM suppression makes it proportional only to m_b^2 or m_s^2 , rather than the much larger m_t^2 . Therefore, a rough estimate of the ratio of mass splitting to decay rate would be,

$$\frac{\Delta m_D^2}{\Gamma_D} \sim \frac{\lambda^{10} m_b^2}{m_c^2}, \frac{\lambda^2 m_s^2}{m_c^2} \quad (9.171)$$

with the former contribution from the b quark and the latter from the s quark in the loop. Each term results in an estimate much less than 10^{-3} , and indeed, D oscillation phenomena have not been observed, and are experimentally constrained down to the 10^{-3} level.

It is nonetheless of experimental interest to look for $D-\bar{D}$ mixing, because it is possible in some extensions of the standard model. It therefore represents a sensitive probe of physics beyond the standard model, and any detection at the present level of sensitivity would signal the standard model's failure.

9.5 Problems

[9.1] Low-energy photon–proton scattering

Consider the scattering of a photon from a proton.

- [9.1.1] Including only the leading interaction, $-ieA_\mu \bar{p}\gamma^\mu p$, find the spin and polarization averaged differential scattering cross section, in the proton rest frame and in the small photon energy limit, as a function of angle θ . Compare with the classical electromagnetism result (essentially the same as Thompson scattering),

$$\frac{d\sigma}{d\cos\theta} = \frac{\pi\alpha^2}{m_p^2} (1 + \cos^2\theta)$$

[9.1.2] Now consider corrections arising from the magnetic-dipole interaction. Show that the linear in μ_p correction to the integrated cross section (arising from interference between a scattering with one appearance of μ_p and with no appearances), vanishes. Is the same true for the angular distribution? At what order in μ_p do corrections to the integrated and differential scattering cross sections first appear? (If you are wise, you will attempt this problem using classical electromagnetic theory, rather than Feynman graphs.)

[9.2] **Color factors**

Verify the color factors,

$$\sum_{j,\alpha} \lambda_{ij}^\alpha \lambda_{jk}^\alpha = 4\delta_{ij} C_F$$

$$\sum_{i,j} \lambda_{ij}^\alpha \lambda_{ji}^\beta = 4\delta^{\alpha\beta} \frac{1}{2}$$

$$\sum_{\beta,\gamma} f_{\alpha\beta\gamma} f_{\delta\beta\gamma} = C_A \delta_{\alpha\delta}$$

with $C_F = 4/3$ and $C_A = 3$ for QCD.

[9.3] **DGLAP kernel functions**

Derive Eq. (9.74) by considering the collinear splitting of a gluon into quarks and the radiation of a gluon from a gluon. Work in a frame where the gluon momentum $p = (p, 0, 0, p)$ is large, and treat the components of the outgoing gluon momenta, transverse to the z direction, as small compared to the longitudinal momentum. To avoid problems with gauge invariance, it is *essential* to sum over the two transverse polarization states of all external gluons at all times, rather than using $\eta_{\mu\nu}$ to compute the spin summed external states. (You can verify this by doing it the wrong way.)

[9.4] **Pascos–Wolfenstein ratio**

Consider the scattering of a muon neutrino from a proton. Four processes are possible, $\nu_\mu p \rightarrow \nu_\mu X$, $\nu_\mu p \rightarrow \mu X$, $\bar{\nu}_\mu p \rightarrow \bar{\nu}_\mu X$, and $\bar{\nu}_\mu p \rightarrow \bar{\mu} X$.

[9.4.1] Write down the leptonic tensors $L_{\mu\nu}$ which replace Eq. (9.42), for each case. Work out the cross section in terms of the structure functions $F_{1,\text{cc}}^\pm$, $F_{2,\text{cc}}^\pm$, $F_{3,\text{cc}}^\pm$, and $F_{1,\text{nc}}$, $F_{2,\text{nc}}$, $F_{3,\text{nc}}$ (the charged-current and neutral-current structure functions are in general independent functions, and we must distinguish the charged-current structure function in which the proton gains charge, from the one where it loses charge).

These are defined similarly to the electromagnetic current, except using J_{cc} , J_{cc}^* , and J_{nc} rather than J_{em} in Eq. (9.43).

- [9.4.2] Express the charged-current and neutral-current structure functions, $F_{cc}^\pm(x, q^2)$, $F_{nc}(x, q^2)$, in terms of the quark and anti-quark PDFs, $u_m(x, q^2)$, $\bar{u}_m(x, q^2)$, $d_m(x, q^2)$, and $\bar{d}_m(x, q^2)$, in analogy with the electromagnetic case, Eq. (9.57). To do this, consider a quark of momentum fraction x , and study what tensor arises from its squared matrix element to interact with a W or Z boson.

What specific PDF's does each scattering process probe? Is there any relation between the F_1 , F_2 , and F_3 structure functions, akin to the Callan–Gross relation?

- [9.4.3] Consider muons and antimuons impinging on a target of equal quantities of neutrons and protons—for instance, deuterium, D —so that the target-averaged $u(x, q^2) = d(x, q^2)$ and $\bar{u}(x, q^2) = \bar{d}(x, q^2)$ (up to small isospin breaking effects, which we will ignore). Also assume that, for all other quark types (s , c , etc.) the quark and antiquark distributions are the same, $s(x, q^2) = \bar{s}(x, q^2)$.

Show that the Paschos–Wolfenstein ratio,

$$R_{PW} \equiv \frac{\sigma(\nu_\mu D \rightarrow \nu_\mu X) - \sigma(\bar{\nu}_\mu D \rightarrow \bar{\nu}_\mu X)}{\sigma(\nu_\mu D \rightarrow \mu X) - \sigma(\bar{\nu}_\mu D \rightarrow \bar{\mu} X)}$$

satisfies

$$R_{PW} = \frac{1}{2} - \sin^2 \theta_w$$

In particular, each component cross section has complicated dependence on the structure functions, but the ν , $\bar{\nu}$ differences only depend on the differences $u - \bar{u}$, $d - \bar{d}$, and the neutral-to-charged current ratio has a common x dependence, so the neutral-to-charged interaction coefficient can be factored out, and the detailed dependence on the structure functions cancels in the ratio. This presents a very convenient tool for determining $\sin^2 \theta_w$.

[9.5] Light scalar quark

Suppose that there existed a scalar quark, \tilde{q} , meaning a scalar field in the fundamental representation of $SU_c(3)$. Write its electromagnetic charge as Q . Also imagine that the mass of this quark were small, $m \sim 100$ MeV, so that it could appear in long-lived bound states.

Completing this problem will require a few Feynman rules for such a scalar quark. The propagator for the scalar quark, at momentum k , is $-i/k^2$, and the vertex rules are the same as Eq. (5.65) and Eq. (5.75), but with γ^μ replaced with $i(k - p)^\mu$.

- [9.5.1] Write down the lightest new meson-like states which could result as bound states of u and d quarks with the hypothetical scalar quark. What are the expected quantum numbers of these states – particularly under spin, parity, and charge conjugation? Give a vague estimate of the mass of such a state.
- [9.5.2] The scalar quark must be included as a partonic species in the analysis of deep inelastic scattering. What are the new contributions to the structure functions $F_1(x, q^2)$, $F_2(x, q^2)$ due to the scalar quark parton distribution function $\tilde{q}(x, q^2)$, $\bar{\tilde{q}}(x, q^2)$? Is the Callan–Gross sum rule, $F_2(x, q^2) = 2xF_1(x, q^2)$, satisfied by the scalar-quark contribution? (The spin dependence of the Callan–Gross sum rule was one of the early pieces of evidence for the spin-half nature of quarks.)
- [9.5.3] Revisit the calculation of collinear gluon radiation, but now considering radiation from the scalar quark \tilde{q} . Find the appropriate replacements for Eq. (6.88) and Eq. (9.75), for the scalar quark. Does the quark or the scalar quark lose its large x contribution faster, as q^2 is increased?

[9.6] **Top-quark production**

Consider $t\bar{t}$ (top–antitop) production at a hadron collider. Observing this process was one of the principal motivations of the Tevatron collider at Fermilab.

- [9.6.1] Find the squared matrix elements for the two production processes, $q\bar{q} \rightarrow t\bar{t}$ and $gg \rightarrow t\bar{t}$, at leading order in α_3 , namely, $O(\alpha_3^2)$. Average over initial-state spins and colors, sum over final-state spins and colors. Note that the calculation of the gg process is quite tricky; there are contributions from multiple diagrams, and their relative sizes are dependent on how the polarization sums are done. The safest method is to sum over the two transverse gluon polarizations in the center of mass frame of the system. Systematically neglect all masses other than the top quark mass (which *cannot* be neglected), and work only to leading order in the strong coupling α_3 , neglecting α_2 and α_1 .
- [9.6.2] Evaluate the $t\bar{t}$ production cross section, integrated over angles, as a function of the center of mass energy squared s of the initial particles. If possible find a closed form expression in terms of s , m_t , and α_3 .
- [9.6.3] Express the total production rate for $t\bar{t}$ pairs as an integral over x_1 and x_2 , in terms of $q(x)$, $\bar{q}(x)$, and $g(x)$ the parton distribution functions of the proton, for pp collisions. Now, compute for $p\bar{p}$ collisions, using $q_p(x) = \bar{q}_{\bar{p}}(x)$, $\bar{q}_p(x) = q_{\bar{p}}(x)$. What part of the answer is different *vis-à-vis* the pp case?

Looking at the plot of parton distribution functions, do you see why the Tevatron (running at 1800 GeV center of mass energy) was designed as a $p\bar{p}$ machine, rather than a pp machine (despite the added complications of creating an antiproton beam)?

[9.7] Isospin transitions

Consider the matrix element

$$\langle A | J_{\text{had}}^\mu(x=0) | B \rangle \approx \sqrt{(i - i_3)(i + i_3 + 1)}(p_A^\mu + p_B^\mu)V_{ud}$$

for two spin-zero states, $|A\rangle$ and $|B\rangle$, with isospin i and $i_3 = i_3(B) = i_3(A) - 1$. This type of matrix element arises, for example, in the decay $\pi^- \rightarrow \pi^0 + \text{leptons}$ or in nuclear decays such as ${}^6\text{C}^{14} \rightarrow {}^7\text{N}^{14}$. In both of these examples we would have $i = 1$, $i_3(B) = -1$, and $i_3(A) = 0$.

[9.7.1] For the particular process: $\pi^- \rightarrow \pi^0 + e^- + \bar{\nu}_e$, compute the decay rate, $\frac{d\Gamma}{dE}(E)$, as a function of the energy of the emitted electron as seen in the rest frame of the π^- . Sum over all spins and integrate over the neutrino momentum and electron direction. Feel free to also neglect the electron mass.

[9.7.2] What is the π^- decay lifetime in its rest frame, via the process $\pi^- \rightarrow \pi^0 e^- \bar{\nu}_e$?

[9.8] π^+ decays

Given the hadronic matrix element

$$\langle \text{vacuum} | J_{\text{had}}^\mu | \pi^+ \rangle = \sqrt{2} F_\pi p^\mu V_{ud}$$

[9.8.1] Compute the total decay rate for the purely leptonic decay $\pi^+ \rightarrow \ell^+ \nu_\ell$.

[9.8.2] Using $F_\pi = 93$ MeV and the results of the previous problem calculate the ratio of the decay rates for the two processes $\Gamma(\pi^+ \rightarrow \pi^0 e^+ \nu_e) = \Gamma(\pi^- \rightarrow \pi^0 e^- \bar{\nu}_e)$ and $\Gamma(\pi^+ \rightarrow e^+ \nu_e)$. Compare your rates to those found in the Particle Data Book.

[9.9] Neutral-pion decay

Suppose the interaction Hamiltonian involves a term

$$\mathcal{H}_{\text{int}} = \frac{e^2}{32\pi^2 F_\pi} \left(\frac{N_c}{3} \right) \pi^0(x) F^{\mu\nu}(x) F^{\lambda\rho}(x) \epsilon_{\mu\nu\lambda\rho}$$

that mediates the decay of a π^0 into two photons. e in this expression is the proton charge, F_π is the pion decay constant, $F_\pi = 93$ MeV, and $N_c = 3$ is the number of colors. $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ denotes, as usual, the electromagnetic-field strength.

[9.9.1] Show that the field decompositions

$$A_\mu = \sum_{\lambda=\pm 1} \int \frac{d^3\mathbf{p}}{2p^0(2\pi)^3} \left[\epsilon_\mu(\mathbf{p}, \lambda) e^{ip \cdot x} a_{\mathbf{p}\lambda} + \text{c.c.} \right]$$

$$\text{and } \pi^0 = \int \frac{d^3\mathbf{k}}{2k^0(2\pi)^3} \left[e^{ikx} b_{\mathbf{k}} + \text{c.c.} \right]$$

imply the following matrix elements:

$$\begin{aligned} & \langle \gamma(\mathbf{p}, \sigma) \gamma(\mathbf{q}, \lambda) | \mathcal{H}_{\text{int}} | \pi^0(\mathbf{k}) \rangle \\ &= \frac{e^2}{32\pi^2 F_\pi} \left(\frac{N_c}{3} \right) \epsilon_{\mu\nu\lambda\rho} 4(-ip^\mu) \epsilon^\nu(\mathbf{p}, \sigma) (-iq^\lambda) \epsilon^\rho(\mathbf{q}, \lambda) \end{aligned}$$

[9.9.2] Use the following polarization vectors:

$$\begin{aligned} \epsilon^\mu(\mathbf{p}, \sigma = \pm 1) &= \frac{1}{\sqrt{2}}(0, 1, \pm i, 0) = \epsilon^\mu(-\mathbf{p}, \sigma = \mp 1) \\ &= \epsilon^{*\mu}(\mathbf{p}, \sigma = \mp 1) \end{aligned}$$

in the frame for which the photon momentum has the form

$$p^\mu = (\omega, 0, 0, \omega)$$

to show that the total rate for $\pi^0 \rightarrow 2\gamma$ decay in the π^0 rest frame is

$$\Gamma(\pi^0 \rightarrow 2\gamma) = \frac{\alpha^2 m_\pi^3}{(4\pi)^3 F_\pi^2} \left(\frac{N_c}{3} \right)^2$$

$\alpha \equiv \frac{e^2}{4\pi}$ is the usual fine-structure constant. Compare the value you find with that in the Particle Data Book. Notice that this result gives an experimental check on the number of colors: $N_c = 3$.

Part IV

Beyond the standard model

10

Neutrino masses

The previous chapters have revealed the standard model as a remarkable theory which successfully describes the vast majority of all of modern particle physics. Better yet, its implications are very robust since it is the most general theory which is consistent with general principles (like Lorentz invariance and unitarity) plus the two assumptions: (i) renormalizability and (ii) the given particle content – all of which are seen in experiments (modulo confinement for the case of quarks and gluons) save the as-yet-undiscovered Higgs scalar.

In this chapter we encounter the first (and, as of this writing, only) known case where there is good evidence that the standard model does *not* provide a good description: the phenomenon of neutrino oscillations. After describing the conceptual issues and the evidence for the standard model’s failure, we briefly describe what can be said at present about which of the two core assumptions must be relinquished.

The standard-model prediction which has gone sour – encountered in Subsection 2.5.2 – asserts the separate conservations (up to negligible corrections involving the electroweak anomaly) of the three lepton numbers, L_e , L_μ , and L_τ . Even considering the electroweak anomaly, the quantities $L_e - L_\mu$ and $L_\mu - L_\tau$ are anomaly free and so should be exactly conserved. As a consequence the theory predicts exactly massless and stable neutrinos, ν_e , ν_μ , ν_τ , with ν_ℓ only participating in charged-current weak interactions together with its corresponding charged lepton, ℓ^- .

There are now several kinds of experiments involving neutrinos of relatively low energy, propagating over long distances, which indicate that the last of these properties is not true. In particular, the nuclear reactions which power the Sun are known to produce neutrinos through charged-current reactions involving only electrons. These “solar” neutrinos escape the Sun and

some make their way to the Earth where they can be detected, and some are observed no longer to interact through reactions involving electrons (they instead are detected through their neutral-current interactions). Evidence for a similar effect is also seen when electron (anti)neutrinos are both prepared and detected on Earth.

Similarly, the Earth is under constant bombardment from space by very energetic cosmic rays (mostly protons), which strongly interact with nitrogen and oxygen nuclei when they hit the upper atmosphere. π^\pm particles are copiously produced in these reactions and the neutrinos generated when these decay in flight are seen in detectors on and under the Earth's surface. Since, for example, a π^+ decays almost 100% of the time into $\mu^+\nu_\mu$ and the muon decays in turn into $e^+\nu_e\bar{\nu}_\mu$, a $\nu_\mu : \nu_e$ ratio of close to 2 : 1 was expected for the “atmospheric” neutrinos detected in this way (even after more careful simulations of the decay process in the atmosphere). Instead, this ratio is seen only for downward moving neutrinos; upward moving neutrinos (those originating from showers on the opposite side of the Earth) exhibit a ratio close to 1 : 1. It appears that about half of the expected ν_μ neutrinos are missing. Again, evidence for the same effect is also seen for muon neutrinos that are both produced and detected on Earth.

Neither of these effects is consistent with the separate conservation of L_e , L_μ , and L_τ , and this provides the first real evidence that the standard model as described in previous chapters is an incomplete theory. One of the two key assumptions – particle content and/or renormalizability – is breaking down, but which one? And if it is not new particles that are at root, what would a failure of renormalizability mean? We take up these questions in more detail after first digressing to describe the kinematics of massive neutrinos.

In essence, the message of this chapter is that it is possible to amend the standard model in several ways to permit neutrino masses that can describe the observed neutrino oscillations. However, we do not yet know which amendment is correct.

10.1 The kinematics of massive neutrinos

Neutrinos resemble the neutral mesons encountered in Chapter 9, in that they both can participate in oscillations. However, the situation for neutrinos differs from that described earlier for neutral mesons in two important ways: the nonzero neutrino spin, and the possible participation of at least three species of neutrinos in the oscillations.

10.1.1 Dirac vs. Majorana

Within the standard model the three massless neutrinos are labeled by their flavor, momentum, and helicity, $|\nu_i(\mathbf{p}, h)\rangle$, with helicity defined as the projection of the neutrino spin along its direction of motion: $h = \mathbf{s} \cdot \hat{\mathbf{p}}$, where $\hat{\mathbf{p}} = \mathbf{p}/|\mathbf{p}|$. Helicity is a sensible label for massless particles because all observers can agree on its value, regardless of their Lorentz frame. Since only left-handed neutrinos appear in the standard model, $h = -\frac{1}{2}$ for all three neutrino states.

In general, however, Lorentz invariance and locality imply that chiral perturbation theory (CPT) is a symmetry, and so we immediately infer the existence of three more states:

$$|\bar{\nu}_i(\bar{\mathbf{p}}, \bar{h})\rangle = CPT|\nu_i(\mathbf{p}, h)\rangle \quad (10.1)$$

where $\bar{\mathbf{p}} = \mathbf{p}$ and $\bar{h} = -h$. The interesting feature here is that CPT changes the sign of helicity, and so the states obtained in this way must differ from those with which we started. Within the standard model these new states are the right-handed antineutrinos, and they are distinguished from the neutrinos both by their helicity and their eigenvalues of the lepton numbers, L_e , L_μ , and L_τ .

Things are different if neutrinos are massive, however, because in this case helicity is no longer a good quantum number. It fails as a quantum number because different Lorentz observers can disagree on its value if they are moving relative to one another. For instance, massive neutrinos possess a rest frame for which the direction, $\hat{\mathbf{p}}$, of neutrino momentum is undefined. Similarly, two observers moving in the same direction as the neutrino – with one moving more slowly than the neutrino and the other moving more quickly – will disagree on the direction of $\hat{\mathbf{p}}$, with one finding it points along the direction of their mutual motion and the other finding it pointing in the opposite direction. For this reason we instead label massive spin-half states by their spin direction in the particle rest frame, $|\nu_i(\mathbf{p}, \sigma)\rangle$, and any massive spin-half particle must have *both* eigenvalues $\sigma = \pm 1/2$.

Applying CPT to such a state again changes the sign of σ , but because the original particle has both signs in any case it is no longer automatic that $|\bar{\nu}_i\rangle \equiv CPT|\nu_i\rangle \neq |\nu_i\rangle$. In this case it is only the existence of another conserved charge, such as lepton number, which can decisively distinguish particle from antiparticle by assigning them opposite charge. If such a charge exists, then the neutrino is said to be a “Dirac” neutrino, and the six states (two-spin times three-flavor), $|\nu_i\rangle$, are physically distinct from their CPT conjugates, $|\bar{\nu}_i\rangle = CPT|\nu_i\rangle$. If such a charge does not exist, the neutrino is

said to be a “Majorana” neutrino, for which the action of CPT returns the same states we started with: $|\nu_i\rangle$.

The names Dirac and Majorana are taken from the fields which are used to describe such particles. The minimal two states, $|\nu(\mathbf{p}, \sigma)\rangle$ and its CPT conjugate $|\bar{\nu}(\mathbf{p}, -\sigma)\rangle$, of a spin-half particle can always be described by Majorana spinor fields, and a collection of N such particles can be represented without loss of generality by N Majorana fields. However, when there is a $U(1)$ symmetry (corresponding to an additive conserved charge) acting on the spin-half states it is very convenient to group the pair of Majorana spinors, ψ_1 and ψ_2 , which describe the distinct particle and antiparticle – $|\nu(\mathbf{p}, \sigma)\rangle$ and $|\bar{\nu}(\mathbf{p}, \sigma)\rangle$ plus their respective CPT conjugates – together into a Dirac spinor: $\psi_D = P_L\psi_1 + P_R\psi_2$, with the $U(1)$ acting as $\psi_D \rightarrow e^{i\alpha}\psi_D$.

This grouping into Dirac spinors was done in previous chapters for all of the fermions of the standard model besides neutrinos because for these particles electric charge distinguishes them from their antiparticles. Since neutrinos are not electrically charged (and don’t carry baryon number) the question of whether they are distinct from their antiparticle (i.e. Majorana vs. Dirac) is equivalent to the question whether there is a conserved lepton number whose charge they do carry. The observed neutrino oscillations do not settle this issue because although they show that L_e , L_μ , and L_τ are not all separately conserved, they do not rule out the conservation of total lepton number, $L = L_e + L_\mu + L_\tau$, or $B - L$.

10.2 Neutrino oscillations

Before trying to describe neutrino masses and oscillations in terms of all of the standard-model fields, we first put aside issues of $SU_L(2) \times U_Y(1)$ gauge invariance and summarize them in the simpler context of three neutrinos with a phenomenological mass matrix, interacting with other particles through their couplings with the W and Z bosons. This simplified description is all that is needed to describe the observations, and we return to the issue of gauge invariance in later sections.

The basic choice to be made in describing neutrino oscillations is whether or not to add any new neutrino states above and beyond the three states which we already know exist. (Once we reintroduce gauge invariance in later sections, we shall see that this choice is equivalent to the choice of modifying the standard model by adding new particles or by adding nonrenormalizable interactions.) In the case where new neutrinos are added there is a subsidiary question as to whether or not the couplings of these new particles conserve any form of overall lepton number, L .

We consider each case in turn, starting first with the minimal assumption that no new neutrino states exist.

10.2.1 Three neutrinos only

Suppose first that there are no more neutrino states beyond those which are already present in the standard model. In this case, by assumption, we know that CPT takes the existing neutrinos into themselves and so neutrinos must be their own antiparticles. This precludes in turn lepton number from being an approximate symmetry, and we indeed find in this case that neutrino masses in this scenario necessarily break L conservation in addition to breaking separately L_e , L_μ , and L_τ .

To describe this situation phenomenologically, imagine supplementing the standard model Lagrangian by an explicit neutrino mass term:

$$\mathcal{L} = \mathcal{L}_{\text{SM}} - \frac{1}{2} \left[m_{ij} (\bar{\nu}_i P_L \nu_j) + \text{c.c.} \right] \quad (10.2)$$

where m_{ij} is an arbitrary complex symmetric 3×3 matrix. We work in a basis for which the charged-lepton mass matrix has already been diagonalized. Since the standard model contains a grand total of six neutrino and antineutrino states (including spin labels) we may take the three neutrino fields to be Majorana without loss of generality.

This mass term, which couples the fields, ν_a , to themselves, differs in several ways from the quark and charged-lepton mass terms, which couple two different fields (*e.g.* L and E or Q and U or D) to one another. In terms of the two component left-handed component χ of the Majorana spinor $\nu = [\chi \epsilon \chi^*]^T$, the mass term reads

$$\mathcal{L}_\nu \text{ mass} = \frac{1}{2} \left(m_{ij}^* \chi_i^T \epsilon \chi_j + \text{h.c.} \right) \quad (10.3)$$

Such a mass term, where a particle is coupled to its own transpose rather than a distinct right-handed state, is called a Majorana mass. Unlike the mass matrix of earlier chapters, the Majorana mass matrix, m_{ij} , must be symmetric and so has fewer independent entries than does the corresponding charged-lepton matrix.

As advertised, this neutrino mass term breaks all of the lepton-number symmetries of the standard model, for which the left-handed fields rotate by $P_L \nu_j \rightarrow e^{i\omega_j} P_L \nu_j$, with a separate parameter ω_j for each neutrino flavor “ j .” But the Majorana condition for ν_i then implies the right-handed fields rotate oppositely, $P_R \nu_j \rightarrow e^{-i\omega_j} P_R \nu_j$ (as must any antiparticle), and so $\bar{\nu}_j P_L \rightarrow e^{i\omega_j} \bar{\nu}_j P_L$. Since the mass matrix transforms as $m_{jk} \rightarrow m_{jk} e^{i(\omega_j + \omega_k)}$, it is

invariant under all three transformations only if $m_{jk} = 0$. The same is also true for the overall lepton symmetry, for which $\omega_1 = \omega_2 = \omega_3$. The only type of lepton number which Eq. (10.2) can preserve is the difference between two lepton numbers; e.g. $L_{12} = L_1 - L_2$ would be a symmetry if the only nonzero mass matrix elements were m_{33} and $m_{12} = m_{21}$. In such a case the neutrinos ν_1 and ν_2 would group into a Dirac neutrino (with antiparticle distinguished by the eigenvalue of L_{12}), with the third neutrino being necessarily Majorana if $m_{33} \neq 0$.

10.2.1.1 The PMNS mixing matrix

In general the neutrino mass matrix may be diagonalized by redefining the neutrino fields, $P_L \nu_i = V_{ij} P_L \nu_j$, with unitary V_{ij} , following the arguments of Chapters 1 and 2 with only minor modifications. The main difference is that the mass matrix, m_{jk} , has fewer independent entries than did the quark-mass matrix, g_{mn} , because the matrix m_{jk} must be symmetric. On the other hand, since the neutrino matrix does not relate different left- and right-handed fields, there are fewer fields which can be independently rotated to accomplish the diagonalization. This changes the counting of the number of independent physical parameters which these rotations can introduce into the charged-current weak interactions, leading to more CP-violating phases for leptons than for quarks (as we now describe).

After transforming to the neutrino mass basis in this way, the neutrino part of the Lagrangian becomes

$$\mathcal{L} = -\frac{1}{2} \bar{\nu}_i (\not{\partial} + m_i) \nu_i + \mathcal{L}_{\text{nc}} + \mathcal{L}_{\text{cc}} \quad (10.4)$$

where the m_i are real and non-negative. The unitarity of V ensures the neutral-current interaction \mathcal{L}_{nc} is unchanged from the standard model result (the leptonic GIM mechanism), and the charged-current interaction picks up a CKM-like mixing matrix,

$$\mathcal{L}_{\text{cc}} = \frac{igV_{ai}}{\sqrt{2}} W_\mu (\bar{\ell}_a \gamma^\mu \gamma_L \nu_i) + \text{h.c.} \quad (10.5)$$

Here we use $i = 1, 2, 3$ to label the three neutrino types and $a = 1, 2, 3$ to label the charged leptons: $\{\ell_1, \ell_2, \ell_3\} = \{e, \mu, \tau\}$. We use different letters to do so in anticipation of the next section where we have more neutrinos than charged leptons. Notice that the convention for leptons is to use the charged leptons to label the rows and the neutrinos to label the columns, which is opposite to the convention used for quarks.

The mixing matrix, V_{ai} , may be parameterized in terms of mixing angles

and phases as $V = U K$, with $K = \text{Diag}(e^{i\alpha_1/2}, e^{i\alpha_2/2}, e^{i\alpha_3/2})$ and

$$\begin{aligned}
 U &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13} e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13} e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13} e^{-i\delta} \\ -c_{23}s_{12} - s_{23}c_{12}s_{13} e^{i\delta} & c_{23}c_{12} - s_{23}s_{12}s_{13} e^{i\delta} & s_{23}c_{13} \\ s_{23}s_{12} - c_{23}c_{12}s_{13} e^{i\delta} & -s_{23}c_{12} - c_{23}s_{12}s_{13} e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{10.6}
 \end{aligned}$$

where $c_{ij} = \cos \theta_{ij}$ and $s_{ij} = \sin \theta_{ij}$. The CKM-like matrix, U_{ai} , is called the PMNS matrix – for Pontecorvo, Maki, Nakagawa, and Sakata – and describes the amplitude with which the neutrino type “ i ” participates in charged-current reactions with the charged lepton “ a .”

The matrices U and V differ because of the appearance of the extra phases, α_i , in the matrix K . Such phases are conventionally removed from the CKM matrix by performing phase rotations of the first- and second-generation quarks, and this is possible within the standard model because the rest of the Lagrangian conserves flavor and so is unchanged by this re-phasing. The same rotations also would have removed the phases $e^{i\alpha_j}$ if neutrinos had been Dirac particles. However, we have seen that the neutrino-mass term is not invariant under lepton-number transformations and so is not preserved by re-phasings of the left-handed neutrino states.

The phases δ and α_i can have physical implications because they introduce CP-violation into neutrino physics. Since the phase δ is the direct analogue of the CP-violating phase in the CKM matrix, its effects disappear in the limit $\theta_{13} \rightarrow 0$. One of the phases α_i can be rotated away by making a common phase rotation of the charged leptons (conventionally α_3 is removed), and the other two are not observable in processes which conserve total lepton number L .

10.2.1.2 Neutrino oscillations

Since neutrinos are usually created by charged-current weak interactions, they are produced in association with a charged lepton, and so start off in a flavor eigenstate, ν_a . For instance, μ^- decay produces a ν_μ and a $\bar{\nu}_e$, while the reaction $e^- X \rightarrow X' \nu_e$ produces an electron neutrino. But the subsequent time evolution of the produced neutrino simply involves time-dependent phases for the neutrino-mass eigenstates, ν_i . If neutrinos are detected by observing the charged lepton in another charged-current reaction then the detector again probes the final flavor state, which need not be the same as

the initial flavor if the mass and flavor eigenstates differ from one another. This difference between flavor and mass bases therefore allows the possibility of complicated time-dependent oscillation phenomena, along the lines we have already encountered for neutral mesons. Most of our information about neutrino masses and mixings arises from such neutrino oscillations, as we now explore.

To this end suppose a neutrino is produced as a flavor eigenstate, ν_a . It then propagates to a detector at a distance \mathbf{x} , where it is measured to have flavor ν_b . The amplitude for this process is

$$\langle \nu_b(\mathbf{x}, t) | \nu_a(0, 0) \rangle = \langle \nu_b | \exp(-i\hat{H}t + i\hat{\mathbf{P}} \cdot \mathbf{x}) | \nu_a \rangle \quad (10.7)$$

Here \hat{H} and $\hat{\mathbf{P}}$ are the time-evolution and translation operators, and t is the time of propagation. Since neutrinos are almost massless, $t \simeq |\mathbf{x}|$ to high precision.

The t - and \mathbf{x} -dependence of this amplitude is most conveniently evaluated by inserting a complete basis of mass eigenstates, giving

$$\langle \nu_b(\mathbf{x}, t) | \nu_a(0, 0) \rangle = \sum_i \sum_\sigma \int d^3\mathbf{k} e^{-iE_i(k)t + i\mathbf{k} \cdot \mathbf{x}} \langle \nu_b | \nu_i(\mathbf{k}, \sigma) \rangle \langle \nu_i(\mathbf{k}, \sigma) | \nu_a \rangle \quad (10.8)$$

Were $|x| = t$ and $E = |\mathbf{p}|$ for the neutrino, then the phase from the temporal and spatial propagation would exactly cancel. A small failure in $|x| = t$ leads to a species-independent phase, which is not important for the observables of interest. The *species-dependent* part of the phase arises because different kinds of neutrinos have slightly different dispersion relations: $E_i(k) = [\mathbf{k}^2 + m_i^2]^{1/2}$.

Because the source and detector are by assumption of finite size, a range of momenta can contribute to the process even if the energy were perfectly well measured. In typical applications, the neutrino energy is measured accurately enough that for our purposes we can take E to be known in computing the difference $E - |\mathbf{k}|$. Given the neutrino energy, E , and using the (very good) approximation that neutrinos are ultra-relativistic, implies $|\mathbf{k}| \approx E - m_i^2/E$, and so $E - |\mathbf{k}| \approx m_i^2/E$. This allows the species-dependent part of the phase to be written $e^{-im_i^2|\mathbf{x}|/(2E)}$, and so the desired amplitude becomes

$$\begin{aligned} \langle \nu_b(\mathbf{x}, t) | \nu_a(0, 0) \rangle &= e^{i\xi} \sum_i e^{-im_i^2 L/(2E)} \langle \nu_b | \nu_i \rangle \langle \nu_i | \nu_a \rangle \\ &= e^{i\xi} \sum_i e^{-im_i^2 L/(2E)} V_{bi} V_{ai}^* \end{aligned} \quad (10.9)$$

where ξ is an uninteresting overall phase. We see that each neutrino energy

eigenstate picks up a relative phase of $e^{-im_i^2 L/(2E)}$. By making E modest and L very large, the different terms in the amplitude can acquire large phase differences even if their squared-mass differences, $\Delta m^2 = m_i^2 - m_j^2$, are very small.

The probability of the neutrino being produced as flavor eigenstate ν_a and being detected in flavor eigenstate ν_b is therefore

$$\begin{aligned} P_{\nu_a \rightarrow \nu_b}(E, L) &= \left| \langle \nu_b(L, t) | \nu_a(0, 0) \rangle \right|^2 \\ &= \sum_{ij} e^{-i(m_i^2 - m_j^2)L/(2E)} U_{bi} U_{bj}^* U_{aj} U_{ai}^* \end{aligned} \quad (10.10)$$

Notice that the Majorana phases, $e^{i\alpha_j}$, cancel in this expression, since the same phase appears in both V_{ai} and V_{bi} , and this is what justifies replacing V_{ai} with U_{ai} in the above. This means that one cannot distinguish between the possibilities of Dirac and Majorana neutrinos purely on the basis of neutrino-oscillation experiments.

On the other hand, the rotation angles θ_{12} , θ_{13} , and θ_{23} do not cancel, and so their values can be inferred from oscillation experiments. Notice also that the oscillation probability involves only the difference of squares of masses, and not the masses directly. Furthermore, in the limit $L \ll \lambda(E)$, with the “oscillation length” defined by $\lambda(E) = 2E/\Delta m^2$, the unitarity of the PMNS matrix ensures that the probability becomes diagonal,

$$P_{\nu_a \rightarrow \nu_b}(E, 0) = \delta_{ab} \quad (10.11)$$

Since

$$\lambda(E) = 500 \text{ m} \left(\frac{E}{1 \text{ GeV}} \right) \left(\frac{1 \text{ eV}^2}{\Delta m^2} \right) \quad (10.12)$$

and because the observed neutrino mass splittings turn out to be very small – $\Delta m^2 \leq 5 \times 10^{-3} \text{ eV}^2$ – in practice most neutrino experiments only probe the small- L limit. This is part of the reason why evidence for oscillations has been so long in coming. The sensitivity to mass splitting improves as the length between source and detection point increases and as the energy of the neutrinos decreases. These observations have driven the construction of long-baseline experiments, having a very large separation between source and detector, several of which have been running as of this writing.

It is often the case that only two species of neutrino have mass differences which permit them to appreciably oscillate in a given experiment. In this case it suffices to consider the two-neutrino special case, for which the 2×2

unitary matrix U_{ai} may be chosen to be

$$U = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \quad (10.13)$$

In such a case the probability, Eq. (10.10), is particularly simple:

$$\begin{aligned} P_{\nu_a \rightarrow \nu_a}(E, L) &\simeq 1 - \sin^2(2\theta) \sin^2\left(\frac{\Delta m^2 L}{4E}\right) \\ P_{\nu_a \rightarrow \nu_b \neq \nu_a}(E, L) &\simeq \sin^2(2\theta) \sin^2\left(\frac{\Delta m^2 L}{4E}\right) \end{aligned} \quad (10.14)$$

and for this reason experimental results are often quoted in terms of $\sin^2(2\theta)$ for an effective mixing angle for the oscillation relevant to the experiment.

10.2.1.3 Experimental values

At the time of writing two distinct types of neutrino oscillations have been unambiguously detected. The first of these – solar neutrino oscillations – correspond to observations of the disappearance of ν_e neutrinos coming from the Sun, as well as the disappearance of $\bar{\nu}_e$ neutrinos coming from reactors. If interpreted in terms of two-flavor oscillations (as we shall see, a good approximation even in the three-flavor case), these measurements point to

$$\begin{aligned} 0.40 \leq \tan^2 \theta_\odot \leq 0.50 \quad (30^\circ < \sin \theta_\odot < 38^\circ, 99\% \text{ CL}) \\ \Delta m_\odot^2 = |m_2^2 - m_1^2| = 8.0(3) \times 10^{-5} \text{ eV}^2 \quad (7.1\text{--}8.9, 99\% \text{ CL}) \end{aligned} \quad (10.15)$$

For historical reasons this range for solar neutrinos is called the large-mixing-angle (LMA) solution, and is the only one which accounts for both the experiments involving neutrinos produced on the Sun and on Earth.

The second type of observed oscillations involves either the disappearance of ν_μ and $\bar{\nu}_\mu$ neutrinos generated by cosmic-ray showers in the upper atmosphere, or of ν_μ s from long-baseline beam experiments on Earth. Interpreted in terms of two-neutrino oscillations (again a good approximation), these observations point to the parameters

$$\begin{aligned} \sin^2 2\theta_{\text{atm}} = 1.02(4) \quad (36^\circ < \theta_{\text{atm}} < 54^\circ, 99\% \text{ CL}) \\ \Delta m_{\text{atm}}^2 = |m_3^2 - m_2^2| = 2.5(3) \times 10^{-3} \text{ eV}^2 \quad (1.7\text{--}3.3, 99\% \text{ CL}) \end{aligned} \quad (10.16)$$

Notice that these oscillations are close to maximal, with $\sin^2(2\theta_{\text{atm}})$ consistent with unity.

We see that within a three-neutrino picture the splitting between the two neutrinos responsible for solar-neutrino oscillations must be much smaller than their common splitting from the third neutrino, and in the above we

follow common practice by choosing our labeling for the neutrino-mass eigenstates so that the small splitting is between ν_1 and ν_2 . Since the small mass difference governs solar-neutrino oscillations, and these involve the disappearance of ν_e neutrinos, either ν_1 or ν_2 must have a significant overlap with ν_e , and we conventionally choose ν_1 to be the state which has this large overlap.

In themselves neutrino oscillations do not fix the absolute values of each of the masses separately, leaving open two possibilities. If the nearly degenerate pair, ν_1 and ν_2 , are less massive than ν_3 , then the mass pattern is known as a “normal” hierarchy, while if $m_1, m_2 > m_3$ the hierarchy is “inverted.”

In addition to these two kinds of observed oscillations, there are also a large number of searches using electron neutrinos, but probing a range of Δm^2 comparable to or larger than that involved in atmospheric mixing. In our conventions for which it is $\nu_1 - \nu_2$ oscillations which are seen in solar neutrino experiments, and for which it is ν_1 which strongly overlaps ν_e , these null reactor-neutrino searches constrain the strength of any $\nu_1 - \nu_3$ oscillations which are possible. They imply the 3σ limit

$$|U_{e3}|^2 = s_{13}^2 < 0.067 \quad (10.17)$$

Because this implies $c_{13} > 0.966$, we have $1 > |U_{e1}|^2 + |U_{e2}|^2 = c_{13}^2 > 0.933$ and so it is convenient to write $|U_{e1}| = c_{12}c_{13} \simeq \cos\theta_\odot$ and $|U_{e2}| = s_{12}c_{13} \simeq \sin\theta_\odot$ when inferring the constraints on U_{ai} from solar oscillation measurements, and so $\theta_{12} \simeq \theta_\odot$. From the third column of the PMNS matrix it is similarly useful to write $|U_{\mu 3}| = s_{23}c_{13} \simeq \sin\theta_{\text{atm}}$ and $|U_{\tau 3}| = c_{23}c_{13} \simeq \cos\theta_{\text{atm}}$, in which case $\theta_{\text{atm}} \simeq \theta_{23}$.

The qualitative features of the PMNS matrix which are emerging from experiment can be summarized by the following approximate form, which is obtained by setting $\theta_{13} = 0$ and taking maximal mixing for atmospheric neutrinos, $\theta_{23} = \pi/4$ (i.e. $s_{13} = 0$, $c_{13} = 1$, and $s_{23} = c_{23} = 1/\sqrt{2}$):

$$U \approx \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12}/\sqrt{2} & c_{12}/\sqrt{2} & 1/\sqrt{2} \\ s_{12}/\sqrt{2} & -c_{12}/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \quad (10.18)$$

It is clear that this mixing matrix does not resemble at all the CKM matrix in the quark sector.

Since the experimental accuracy with which the neutrino-mass splittings and phases have been measured is evolving rapidly, the reader should consult recent literature for more up-to-date numbers and errors.

10.2.1.4 Direct mass searches

Neutrino oscillation experiments do not tell us the values of the neutrino masses, only mass squared differences. However, for three neutrinos taking the lightest neutrino to be massless gives a *lower* bound on the mass of the heaviest neutrino of 0.05 eV. How does this compare with direct searches for neutrino masses?

The best kinematic laboratory limit on neutrino masses puts an upper limit on any neutrino having significant overlap with ν_e , and so which can be produced in the beta decay $n \rightarrow pe^-\nu$ in tritium. Although some experiments give slightly better results, the Particle Data Group quotes an upper limit of $m_{\nu_e} \lesssim 3$ eV due to difficulties in understanding the precise shape of the electron spectrum at its endpoint. Weaker limits apply for neutrinos which are dominantly ν_μ , with the best coming from measurements of the muon spectrum in charged pion decay, which gives $m_{\nu_\mu} < 0.19$ MeV (90% C.L.).

The strongest upper bound comes from cosmology in a three-neutrino world, for which (as of 2006) the sum of neutrino masses must satisfy $\sum_i m_i < 0.7$ eV. This result assumes the simplest viable cosmological model, and is weaker under more general assumptions about cosmology. Such cosmological bounds are also evolving rather rapidly and the reader is advised to consult the recent literature for up-to-date results.

10.2.1.5 The MSW effect

The treatment presented above assumes that the neutrino propagates in vacuum between the source and the detector. Since neutrino interactions are very weak, this is often a valid approximation; but not always. Just as light traveling through a transparent medium has its speed of propagation altered by the index of refraction, so also a neutrino traveling through matter experiences medium effects. These effects can change the way that neutrinos oscillate – an effect called the MSW (Mikheyev–Smirnov–Wolfenstein) effect – in a way which appears to be important in practice for neutrinos coming to us from the Sun. Although a complete treatment of this effect goes beyond the scope of this book, we include here a brief discussion due to its practical importance for neutrino experiments.

Consider therefore neutrinos propagating in an environment like the interior of the Sun. If fluctuations in this environment are sufficiently small then the propagation through it of neutrinos is well described by the effective Lagrangian obtained by averaging the standard-model Lagrangian over

the environment,

$$\mathcal{L}_{\text{env}} \approx \langle \mathcal{L}_{\text{SM}} \rangle_{\text{env}} = \text{tr}[\rho_{\text{env}} \mathcal{L}_{\text{SM}}] \quad (10.19)$$

where ρ_{env} here denotes the density matrix which describes the state through which the neutrinos move. Within a solar environment this trace is over the nuclei and electrons present within the plasma inside the Sun, and the resulting Lagrangian describes the average effects on neutrino propagation caused by the charged- and neutral-current interactions, Eq. (7.12) and Eq. (7.15), between the neutrinos and these particles. Detailed studies show that this mean-field picture appears to provide a good description of the solar environment, at least deep within the solar radiative zone where neutrino oscillations take place.

For example, consider the part of the charged-current interaction which governs the interactions between electrons and neutrinos, which, when written in the flavor basis, is

$$\mathcal{L}_{\text{cc}} = \frac{G_{\text{F}}}{\sqrt{2}} [\bar{\nu}_e \gamma^\mu (1 + \gamma_5) \nu_e] [\bar{e} \gamma_\mu (1 + \gamma_5) e] \quad (10.20)$$

In order to perform the average of this term over the environment it is useful to use the Fierz re-arrangement theorem (see Problem 1.9), which allows the above term to be rewritten as

$$\mathcal{L}_{\text{cc}} = \frac{G_{\text{F}}}{\sqrt{2}} [\bar{\nu}_e \gamma^\mu (1 + \gamma_5) \nu_e] [\bar{e} \gamma_\mu (1 + \gamma_5) e] + \dots \quad (10.21)$$

What makes this rearrangement useful is the observation that $\bar{e} \gamma^\mu e$ is nothing but the electron current operator, J_e^μ , and so in a medium full of electrons its mean value is simply the 4-current of electrons in that medium, which in the medium's rest frame is $\langle J_e^\mu \rangle_{\text{env}} = n_e \delta_0^\mu$ with n_e denoting the local electron number density. The $\bar{e} \gamma^\mu \gamma_5 e$ term similarly gives the axial electron current, which vanishes when averaged over any parity-invariant environment, which we assume the solar environment to be.

Using this to perform the average of \mathcal{L}_{cc} in the environment corresponds to replacing the electron current in Eq. (10.21) with this average, in which case the charged-current interactions with the medium introduce the following neutrino-propagation term into the effective Lagrangian:

$$\delta \mathcal{L}_{\text{env}} = - \frac{G_{\text{F}} n_e}{\sqrt{2}} [\bar{\nu}_e \gamma^0 (1 + \gamma_5) \nu_e] \quad (10.22)$$

The appearance of this term in the Lagrangian shifts neutrino energies by $G_{\text{F}} n_e \sqrt{2}$, and shifts antiparticle energies by the opposite amount.

None of the other terms in the charged-current neutrino interactions give a nonzero result (to leading order in the weak interactions) when averaged over

the solar environment, given that this environment contains no muons or taus. The same is not true for averages over the neutral-current interactions, which do produce similar neutrino energy shifts due to their interactions with the electrons, protons, and neutrons in the solar environment. However, an important feature in all of these other nonzero medium-dependent interactions is that they are diagonal in neutrino flavor space, and so shift all neutrino types up (and antineutrino types down) by precisely the same amount. As such they do not contribute at all to the oscillation effects of interest in what follows.

Let us now follow how these medium-dependent terms affect the propagation of an individual neutrino. For simplicity we take $\theta_{13} \approx 0$ when doing so, so that neutrino oscillations can be treated in the two-flavor limit. Since the new term has the effect of shifting only the electron-type neutrino energy its contribution to the single-particle neutrino Hamiltonian is

$$\delta \hat{H}_{\text{MSW}} = \sqrt{2} G_F n_e \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + (\text{diagonal}) \quad (\text{flavor basis}) \quad (10.23)$$

To this must be added the single-particle Hamiltonian which describes neutrino propagation *in vacuo*, which is diagonal in the mass basis but when rotated to the flavor basis becomes

$$\hat{H}_{\text{vac}} = E + \frac{m_{\text{av}}^2}{2E} + \frac{\Delta m^2}{4E} \begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix} \quad (\text{flavor basis}) \quad (10.24)$$

Here we write the squared mass eigenvalues as $m_{\pm}^2 = m_{\text{av}}^2 \pm \Delta m^2/2$.

Adding Eq. (10.23) and Eq. (10.24) and dropping all terms proportional to the unit matrix yields the following effective mass matrix for neutrinos within the solar environment:

$$\delta \hat{H} = \frac{G_F n_e}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \frac{\Delta m^2}{4E} \begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix} \quad (10.25)$$

This can be suggestively re-expressed in terms of an effective medium-dependent mass splitting and mixing angle as follows:

$$\delta H = \Delta' \begin{pmatrix} -\cos 2\theta' & \sin 2\theta' \\ \sin 2\theta' & \cos 2\theta' \end{pmatrix} \quad (10.26)$$

where

$$\begin{aligned} \Delta' &\equiv \left[\left(\frac{G_F n_e}{\sqrt{2}} \right)^2 - \left(\frac{G_F n_e}{\sqrt{2}} \right) \frac{\Delta m^2 \cos 2\theta}{2E} + \left(\frac{\Delta m^2}{4E} \right)^2 \right]^{1/2} \\ \sin 2\theta' &\equiv \left(\frac{\Delta m^2}{4E} \right) \frac{\sin 2\theta}{\Delta'} \end{aligned} \quad (10.27)$$

Notice in particular that the medium-dependent mixing becomes maximal, $\sin 2\theta' = 1$, if

$$\frac{G_F n_e}{\sqrt{2}} = \left(\frac{\Delta m^2}{4E} \right) \cos(2\theta) \quad (10.28)$$

a condition which defines what are called “resonant” oscillations.

We see from this that the problem of finding the neutrino propagation eigenstates in the Sun is an exercise in degenerate perturbation theory. In the absence of masses and medium-dependent interactions the two states, $|\nu_e\rangle$ and $|\nu_\mu\rangle$, are precisely degenerate in energy, but this degeneracy is broken by the sum of two small effects: the neutrino mass matrix and the small medium-dependent term. As a result, *large* mixings can be generated even by small mixing angles and medium-dependent effects, provided only that these two small effects are comparable in size. In particular, the resonance condition, Eq. (10.28), is the case where the diagonal entries of the effective medium-dependent Hamiltonian, Eq. (10.25), vanish in the flavor basis.

Now imagine following a neutrino as it travels out from the central regions of the Sun towards its surface. Such a neutrino sees a *time-dependent* Hamiltonian, because the electron density, n_e , through which it passes falls monotonically with increasing distance from the solar center. As a result, the overlap with flavor eigenstates with which such a neutrino emerges from the Sun depends on the integrated history of how it responds to this changing density profile, and in particular whether it ever passes through a region within the Sun for which the resonance condition, Eq. (10.28), is satisfied.

In practice only the most energetic neutrinos which leave the Sun pass through such a resonance, but since it happens that these are the ones to which terrestrial detectors are most sensitive, this resonance has dramatic implications for the observed signal. How much conversion happens at the resonant point depends crucially on whether the transition through the resonant regime is abrupt or adiabatic – that is, on whether the electron density experienced by the neutrino in the resonant region changes significantly over the typical oscillation time at resonance.

Consider the case of adiabatic evolution, which is not too bad an approximation for neutrinos in the Sun. The electron is created as purely electron type by a nuclear reaction deep in the solar core, and so must be expressed in terms of the mass eigenstates using the medium-dependent mixing angle, θ'_p , which is relevant at its production point. The amplitude for this

neutrino to be produced in the heavier mass eigenstate is then $\sin \theta'_p$, and the amplitude to be in the lighter state is $\cos \theta'_p$.

Depending on its production site and energy, the neutrino may pass through the resonance region on its way out. (More energetic neutrinos have a larger MSW effect, see Eq. (10.25), and also happen to be produced deeper in the core, where n_e is larger.) During adiabatic evolution through the resonance region, the amplitude for the electron to be in each mass state stays the same, while such a large phase develops between them that the interference between mass states can be ignored.

Once they arrive at the Earth, the appropriate mixing angle is the vacuum one, and so the amplitude for the mass eigenstates to be measured as electron type is controlled by $\sin \theta$ and $\cos \theta$. The total probability for the neutrino to be produced and detected as ν_e is then well approximated by the sum of probabilities for each mass state, weighted by the probability for the electron to be in that mass state. That is:

$$P(\nu_{e\odot} \rightarrow \nu_{e\oplus}) = \sin^2 \theta'_p \sin^2 \theta + \cos^2 \theta'_p \cos^2 \theta \quad (10.29)$$

For instance, for the lowest energy neutrinos we have $\Delta m^2/(4E) \gg G_F n_e/\sqrt{2}$ at the solar center, and so can take $\theta'_p = \theta$. In this case the survival probability becomes $\sin^4 \theta + \cos^4 \theta = 1 - \frac{1}{2} \sin^2 2\theta$, corresponding to the length average of Eq. (10.14). Numerically it is about 5/9. By contrast, for the highest energy neutrinos $G_F n_e/\sqrt{2} \gg \Delta m^2/(4E)$ at the center of the Sun; so at the point of production we have $\theta' \simeq \pi/2$. The probability of observing an electron neutrino at the Earth is then approximately $\sin^2 \theta$, or about 1/3.

These numbers provide a reasonable description of the experimental data. There are three kinds of experiments, each of which is sensitive to a different energy range. The Gallium experiments (SAGE, GALLEX, and GNO), which can detect only the ν_e component and are sensitive primarily to the more numerous low-energy neutrinos, say about (50–60)% of the expected ν_e flux. By contrast, water Cherenkov detectors are only sensitive to the most energetic neutrinos ($E > 5$ MeV), and of these the SNO detector can detect both charged-current and neutral-current reactions and so is capable of measuring the flux of all three neutrinos. Indeed, these detectors see about 1/3 as many electron neutrinos as expected, and the SNO detector sees a compensating increase in the number of non- ν_e events. The whole picture is well described by resonant neutrino oscillations amongst three neutrino species.

10.2.2 Sterile neutrinos

Next consider the possibility that the three standard-model neutrino states are supplemented by additional new neutrinos. In general we might imagine adding N Majorana neutrino fields, $s_x, x = 1 \dots N$, in addition to the three standard-model fields, ν_a . We again start by writing down masses for these neutrinos without worrying about consistency with the standard-model gauge interactions, and return in later sections to show how the choices made here can be reconciled with gauge invariance.

There is considerably more freedom in the kinds of interactions which can be entertained when there are additional neutrino fields. We restrict these options by assuming these new fermions couple only to standard-model fields through their mixing in the fermion-mass matrix with the three standard-model neutrinos, and so they typically interact much more weakly with matter than do ordinary neutrinos. Any fermion which mixes with ordinary neutrinos, but otherwise does not couple to standard-model matter, is known for this reason as a “sterile” neutrino.

These assumptions lead to the following Lagrangian:

$$\mathcal{L} = \mathcal{L}_{\text{SM}} - \frac{1}{2} \bar{s}_x \not{\partial} s_x - \frac{1}{2} \left[M_{xy} (\bar{s}_x P_L s_y) + m_{ab} (\bar{\nu}_a P_L \nu_b) + 2\mu_{ax} (\bar{\nu}_a P_L s_x) + \text{c.c.} \right] \quad (10.30)$$

where the total left-handed neutrino-mass matrix

$$\begin{pmatrix} m & \mu \\ \mu^T & M \end{pmatrix} \quad (10.31)$$

is an arbitrary complex and symmetric $(3 + N) \times (3 + N)$ matrix. As before we work in a basis for which the charged-lepton mass matrix has already been diagonalized.

The freedom to choose the neutrino-mass matrices m , μ , and M introduces considerable latitude for generating different kinds of neutrino physics, which we cannot hope to explore systematically here. We instead describe only a few important subclasses of models which illustrate important alternatives, one being the case where neutrino masses preserve overall lepton number. In the next sections we consider this case separately before summarizing some of the other possibilities and constraints on the existence of sterile neutrinos.

10.2.2.1 Dirac neutrinos

In this section we examine an illustrative example where the lepton number is unbroken, corresponding to the case with $N = 3$ new neutrino states which transform in the following way under the lepton symmetry: $P_L \nu_a \rightarrow e^{i\omega} P_L \nu_a$

and $P_L s_a \rightarrow e^{-i\omega} P_L s_a$. In this case the mass term is invariant for any μ , provided the diagonal blocks vanish: $m = M = 0$. The existence of this symmetry makes it convenient to group the six Majorana fields, ν_a, s_a , into three Dirac fields, $\psi_a = P_L \nu_a + P_R s_a$, for which $\psi_a \rightarrow e^{i\omega} \psi_a$.

The diagonalization of the neutrino-mass matrix proceeds along the same lines as did the diagonalization of the quark-mass matrix in the standard model. The neutrino spectrum consists of three massive states which in this scenario are distinguished from their antiparticles by their eigenvalues of the conserved lepton charge. The three distinct masses are given by the positive square roots of the eigenvalues of the matrix $\mu\mu^\dagger$.

The mismatch between the unitary rotation required to diagonalize the neutrino and charged-lepton masses generates in the usual way a 3×3 charged-current PMNS mixing matrix:

$$\mathcal{L}_{cc} = \frac{igU_{ai}}{\sqrt{2}} W_\mu (\bar{\ell}_a \gamma^\mu \gamma_L \psi_i) + \text{h.c.} \quad (10.32)$$

where U_{ai} is as given in Eq. (10.6). The lepton symmetry precludes the appearance of left-handed antiparticles, $P_L s_a$, in the charged current interaction. Consequently, the only difference from the case of three Majorana neutrinos discussed above is the omission of the three CP-violating phases, $e^{i\alpha_j}$, since these can be rotated away by re-phasing the various lepton fields. No off-diagonal neutrino couplings appear in the neutral-current interactions due to the leptonic version of the GIM mechanism. In particular this implies that the sterile neutrinos do not couple to the Z boson, and so do not contribute to the well-measured invisible Z width.

Because neutrino oscillations do not see the CP-violating phases, oscillation phenomenology in this model is precisely the same as it was for three Majorana neutrinos: the PMNS matrix describes 3-flavor oscillations because lepton-number conservation ensures that there are only three distinct mass eigenvalues amongst which oscillations can take place. In particular there are no new observable oscillation effects between the neutrinos and their antiparticles, since these are guaranteed to have exactly equal masses by the lepton symmetry. (Dirac and Majorana neutrinos do differ in their implications for double-beta decay experiments, however, as we describe in more detail below.) At the time of writing this remains a phenomenologically viable picture of neutrino oscillations.

It is counter-intuitive that it should be possible to duplicate the low-energy neutrino spectrum by replacing Majorana neutrinos with Dirac neutrinos, and yet still not be able to detect these new light-neutrino states in scattering experiments. We now elaborate in more detail why this is so, by directly

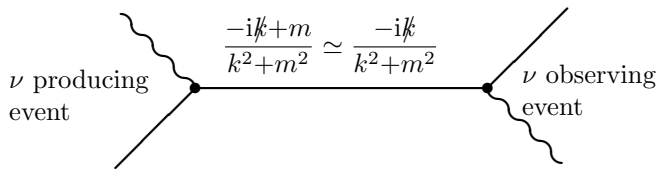


Fig. 10.1 Helicity suppression in neutrino physics.

computing the production amplitude for the sterile, right-handed component of the Dirac neutrino, without making use of the oscillation formalism. Our purpose in so doing is to show that this production amplitude is *helicity-suppressed*, and so is proportional to powers of the small neutrino mass.

To see this, consider a neutrino which propagates from source to detector as in Figure 10.1. In this Feynman graph the exchanged neutrino is a (very slightly) virtual particle, whose evolution is described by its propagator. Evaluating the graph gives the following matrix element for the production and subsequent scattering of such a neutrino,

$$\mathcal{M} = \frac{G_F^2}{2} J_{\text{prod}}^\mu J_{\text{det}}^\nu \left\{ (\dots) \gamma_\mu (1 + \gamma_5) \left[\frac{-i\not{k} + m}{k^2 + m^2} \right] \gamma_\nu (1 \pm \gamma_5) (\dots) \right\} \quad (10.33)$$

where k is the neutrino momentum and (\dots) refers to initial- and final-state factors whose form is not important for the argument we are about to make. Since neutrinos are produced through their charged-current weak interactions, the factor of $\gamma_\mu (1 + \gamma_5) J_{\text{prod}}^\mu$ describes the interaction at the production point which brings the neutrino into being. The corresponding current at the detection point is $\gamma_\nu (1 \pm \gamma_5) J_{\text{det}}^\nu$, and the sign which appears depends on whether or not it is the charged-current or some other interaction which appears here.

For experiments which detect the sterile component, s_a , of the Dirac neutrino, $\psi_a = P_L \nu_a + P_R s_a$, we must choose the sign $(1 - \gamma_5)$ in Eq. (10.33), while detection of the active component, ν_a , requires using the factor $(1 + \gamma_5)$. Consequently only the m term in the numerator of the neutrino propagator can contribute to s_a -detection amplitudes while only the \not{k} part contributes to the detection of the ν_a component. The ratio of these two amplitudes is therefore typically of order k^0/m , evaluated in a frame determined by the production and detection currents. Since most experiments involve neutrinos with lab-frame energies $\gtrsim 1$ MeV, the amplitude for ν_a detection is at

least six orders of magnitude larger than that for s_a scattering, and this must then be squared to find the probability.

The conclusion is that almost none of the interaction rate of the final neutrino is “lost” due to the small likelihood of the neutrino arriving in a right-handed, sterile helicity state. This is responsible for the paradoxical result that even though half of available neutrino states have essentially no electroweak interactions, this has virtually no bearing on neutrino phenomenology.

10.2.2.2 Generic sterile neutrinos

In the generic case the diagonalization of the neutrino and charged-lepton mass matrices leads to a complicated pattern of neutrino masses, which may or may not preserve any of the flavor symmetries carried by the $(3 + N)$ neutrinos. We restrict ourselves here to listing three of the main possibilities, which are distinguished by the relative size of the three kinds of matrix elements, m_{ab} , μ_{ax} , and M_{xy} . When comparing these matrices we assume for simplicity that the square matrices, m and M , have eigenvalues which are all nonzero and roughly the same size.

The diagonalization of the mass matrix has two main effects for the weak interactions, once these are expressed in terms of mass eigenstates. First, it generates a PMNS matrix, U_{ai} , for the charged-current interactions, but this matrix need no longer be square since it has three rows ($a = 1, 2, 3$, for each charged lepton) but $3 + N$ columns ($i = 1, \dots, 3 + N$ if the three active neutrinos are supplemented by N sterile counterparts). The existence of these new mixing elements implies that the new neutrinos participate in the charged-current weak interactions, and so can contribute to oscillation phenomena. No such effect has been observed to date, and this constrains the size of the PMNS matrix elements which involve sterile neutrinos. Some of these observational bounds are summarized in the next section.

A second implication of the diagonalization process is the generation of an off-diagonal mixing matrix for the neutral-current weak interactions, once these are expressed in terms of mass eigenstates, \mathcal{N}_u , $i = 1 \dots (3 + N)$. These have the generic form

$$\mathcal{L}_{\text{nc}} = \frac{ie}{s_w c_w} Z_\mu [i \overline{\mathcal{N}}_u \gamma^\mu (H_{uv} P_L + H'_{uv} P_R) \mathcal{N}_v] \quad (10.34)$$

where $c_w = \cos \theta_w$ and $s_w = \sin \theta_w$ denote the weak mixing angle as usual and H_{uv} and H'_{uv} denote the new neutral-current mixing matrices.

We distinguish the following four illustrative regimes.

$\mu \ll m, M$: In this case, the sterile neutrinos do not significantly mix with

the standard-model neutrinos, and so for practical purposes they do not couple at all to observable particles. In this scenario the observed neutrino oscillations are completely described by the Majorana mass matrix, m , and the potential existence of the sterile fields are irrelevant for experiments. They may have implications for cosmology depending on the details of the thermal history of the universe.

Pseudo-Dirac neutrinos: $\mu \gg m, M$. If $m = M = 0$ this reduces to the case of Dirac neutrinos described above in the case where there are $N = 3$ sterile neutrinos. If $N > 3$ then this scenario leads to three Dirac neutrinos whose squared masses are the eigenvalues of $\mu\mu^\dagger$, plus $(N - 3)$ massless sterile neutrinos which do not participate in the weak interactions.

If M and m are not exactly zero then the neutrino spectrum is a small perturbation of the Dirac case just described. In this case there are six massive neutrino states which participate in the weak interactions and come in almost-degenerate pairs whose squared masses are close to the three eigenvalues of the matrix $\mu\mu^\dagger$, plus $(N - 3)$ massless sterile states. The almost-degenerate pairs of neutrinos are close in mass because they must become the particle and antiparticle parts of the Dirac neutrinos in the limit $m, M \rightarrow 0$. For the same reason the mixing angle between these nearly-degenerate states is approximately maximal, $\theta \approx \pi/4$, since it must become precisely maximal in the limit $m, M \rightarrow 0$. This kind of maximally-mixed and almost-degenerate neutrino is referred to as *pseudo-Dirac*.

Because the two components of a pseudo-Dirac neutrino are close to maximally mixed, strong oscillations can develop between these states if distances $L \approx 2E/\Delta m^2$ can be probed. As we see in more detail below, for those neutrinos participating in solar neutrino oscillations the lack of evidence for such oscillations rules out their being pseudo-Dirac neutrinos with m, M greater than $\sim 10^{-9}$ eV.

Light sterile neutrinos: $m \sim \mu \sim M$. In such a scenario all of the mass matrices are comparable in size, generically leading to $(3 + N)$ Majorana neutrino states which all have masses comparable in size. Successful description of the observed neutrino oscillations requires these masses to be of order 10^{-2} eV. In this case the mass eigenstates are generically complicated mixtures of the standard model ν_a and the singlets s_x , typically leading to large oscillation effects amongst all of the neutrino eigenstates. This *light sterile neutrino* scenario is therefore generically disfavored by the absence of evidence (described below) for oscillations into sterile states, although models which avoid these bounds are also possible.

Seesaw neutrinos: $m \ll \mu \ll M$. In this case there are N heavy mass

eigenstates whose masses are given by the square roots of the eigenvalues of the $N \times N$ matrix $M^\dagger M$, plus three mass eigenstates whose squared masses are given by the eigenvalues of the 3×3 matrix $\mathcal{M}^\dagger \mathcal{M}$ with $\mathcal{M} = \mu M^{-1} \mu^T + m$. The heavy eigenstates are almost purely sterile, with mixing angles with standard-model neutrinos which are of order μ/M , where this means the ratio of combinations of matrix elements of these matrices. The light eigenstates are similarly almost pure ν_a , and although their mixing with the sterile neutrinos is also $O(\mu/M)$, their mixings amongst themselves need not be small.

This way of obtaining neutrino masses is called the *seesaw mechanism*, because of the way that \mathcal{M} gets lighter (goes down) as M gets heavier (goes up). It clearly provides a very attractive picture of neutrino masses particularly if the matrix M is extremely large. Large M is attractive because it would both explain why the sterile neutrinos have not yet been discovered (they are too heavy), and why the observed neutrino masses are so light (they are $O(\mu^2/M)$). We shall see in later sections that this hierarchy of masses is also natural to expect once we embed this picture into an $SU_L(2) \times U_Y(1)$ framework.

10.2.2.3 Constraints on sterile neutrinos

Many of the phenomenological constraints on sterile neutrinos are based on the absence of oscillations between these neutrinos and the usual active ones. Since sterile neutrinos are impossible to detect experimentally such oscillations act to drain probability away from the known three neutrinos, and bounds can be obtained only if the total population of the known three neutrino species can be determined. This can either be done by testing the unitarity of the PMNS matrix, assuming it is only 3×3 in size, or by looking for the physical consequences of draining significant amounts of energy into an invisible channel.

Since the oscillation probability is a strong function of the mass difference, $\Delta m_{jk}^2 = m_j^2 - m_k^2$, between the relevant sterile and active neutrino states, as well as of the active–sterile mixing parameter, $|U_{ax}|$, the strength of the bound obtained also depends sensitively on these parameters. For this reason it is unwieldy to try to bound a general sterile-neutrino model, and so most of the known bounds are made under the assumption that there is only a single sterile neutrino, s , corresponding to a fourth mass eigenstate ν_4 .

We now list some of the strongest such bounds.

Solar neutrinos: The SNO experiment can detect neutrinos using both the charged-current and neutral-current reactions, and so can measure the flux

from the Sun of all three active neutrino species. The sum of these agrees with the known rates for nuclear reactions in the Sun and this agreement constrains the amount of flux which could have been lost into sterile states. For generic neutrino- mass differences larger than 10^{-12} eV² these bounds constrain the sterile-active neutrino mixing to be smaller than $|U_{ex}| \lesssim 0.1$, where the precise bound depends on the mass difference. This bound improves to $|U_{ex}| \lesssim 0.001$ for mass splittings in a narrow range around $\Delta m_{14}^2 = 10^{-4}$ eV², and to $|U_{ex}| \lesssim 0.01$ for 10^{-8} eV² $\lesssim \Delta m_{14}^2 \lesssim 10^{-6}$ eV², due to the presence in these cases of resonant active-sterile MSW oscillations within the Sun.

Reactor neutrinos: The absence of neutrino-flux disappearance from reactor-generated neutrinos constrains $|U_{ex}| \lesssim 0.1$ for $\Delta m_{14}^2 \gtrsim 10^{-3}$ eV².

Atmospheric neutrinos: Active-sterile neutrino oscillations can be excluded for $|U_{ex}| \lesssim 0.2$ for the range $10^{-4} \lesssim \Delta m_{14}^2 \lesssim 10^{-2}$ eV².

Supernova neutrinos: Supernovae are exploding stars, and they start off so hot and dense that ordinary neutrinos scatter so frequently that they can be trapped inside and come to equilibrium with other types of particles. Nevertheless the emission of neutrinos is a supernova's most efficient channel for cooling, and the predicted neutrino fluxes agree with those that were detected on Earth coming from Supernova SN1987a. Sterile neutrinos can ruin this agreement because they can escape more easily than ordinary neutrinos from a supernova environment, and so excessively large active-sterile antineutrino conversions within supernovae can be ruled out if they drain too much flux from the supernova neutrino signal.

For resonant $\nu_e - \nu_s$ oscillations within the supernova interior, this can constrain mixings down to $|U_{ax}| \lesssim 0.01$ for mass splittings, Δm_{14}^2 , which are larger than around 10 eV². However, these bounds carry some uncertainty since they rely on our present imperfect theoretical understanding of supernovae, and because they are computed neglecting the feedback of significant changes to the neutrino densities on the dynamics of the supernova and the strength of the resonant oscillations.

Nucleosynthesis: The dominant cosmological constraint for sterile neutrinos whose masses are smaller than 1 eV comes from the requirement that there not be too many light degrees of freedom contributing to the universal expansion during Big-Bang nucleosynthesis (BBN). Even if the primordial abundance of sterile neutrinos is small, this constrains the strength of active-sterile oscillations because these can cause too efficient production of sterile neutrinos from the equilibrated active neutrinos. This latter condition can

constrain neutrino abundances down to mass differences of order 10^{-8} eV^2 , with the strongest constraints (coming for the largest mass differences) being of order $|U_{ax}| \lesssim 0.1$.

Cosmic microwave background: Measurements of CMB temperature fluctuations constrain the energy density which can be present in neutrinos at the epoch of recombination, leading to the constraint $\Omega_\nu h^2 < 10^{-2}$, where $\Omega_\nu = \rho_\nu / \rho_c$ is the present day mass-energy density of neutrinos in units of the critical density, $\rho_c \sim 10^{-29} \text{ g/cm}^3$. Since this constrains the energy density present in neutrinos, the mixing angles to which this allows constraints are larger for heavier neutrinos, being sensitive to mixing angles as small as $|U_{ax}| \sim 0.001$ for sterile neutrino masses of order 10 eV.

In summary, there is at present no experimental evidence for the existence of any neutrinos beyond the three which come as standard equipment with the standard model, although neither can we rule out all of the models of this type which are of theoretical interest.

10.3 Neutrinoless double-beta decay

We have seen that a central property which distinguishes different categories of neutrino-oscillation models is the question of whether sterile-neutrino states exist, with the related issue of whether or not neutrino masses preserve an unbroken lepton-number invariance (i.e. whether neutrinos are Majorana or Dirac in character). We have seen in particular that although neutrino oscillations provide considerable information about neutrino masses, they can be equally well described by the mixing of three Majorana or three Dirac neutrinos. Is there a way to determine experimentally whether neutrinos are Dirac or Majorana? The answer is “yes,” in principle, and this section describes the experiments which are potentially the most decisive.

Neutrinos are Dirac (Majorana) if they are distinct from (identical to) their antiparticles, and so the central issue is whether or not there exists a conserved lepton number whose eigenvalues can distinguish a neutrino from an antineutrino. At present neutrinoless double-beta decay experiments provide the most sensitive tests of overall lepton-number conservation, and so it is the features of these experiments which we must describe.

Certain nuclei, such as $^{76}_{32}\text{Ge}$, $^{100}_{42}\text{Mo}$, $^{130}_{52}\text{Te}$, and $^{136}_{54}\text{Xe}$, cannot undergo ordinary (single) beta decay, such as

$$^{76}_{32}\text{Ge} \not\rightarrow ^{76}_{33}\text{As} + e^- + \bar{\nu}_e \quad (10.35)$$

because the $^{76}_{33}\text{As}$ nucleus is heavier than is the $^{76}_{32}\text{Ge}$ nucleus. They are

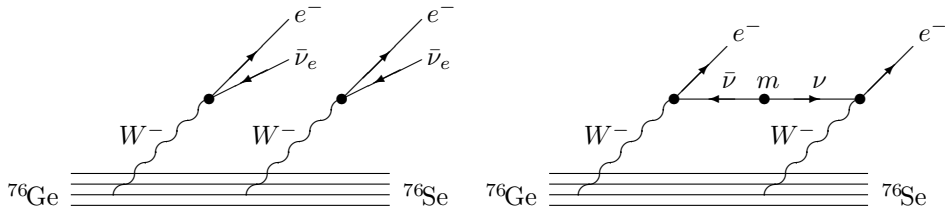


Fig. 10.2 Two neutrino and neutrinoless double-beta decay.

nonetheless radioactively unstable, because the ${}^{76}_{34}\text{Se}$ nucleus *is* lighter, and so allows the (very rare) process

$${}^{76}_{32}\text{Ge} \rightarrow {}^{76}_{34}\text{Se} + 2e^- + 2\bar{\nu}_e \quad (10.36)$$

This process is very rare because it arises only at second order in the weak interactions, corresponding to the basic decay, $n \rightarrow p e^- \bar{\nu}_e$, happening simultaneously for *two* neutrons at once. The left-hand panel of Figure 10.2 gives a Feynman diagram for this decay, wherein a complex nucleus (the lines at the bottom) radiates two virtual W bosons, which both subsequently dissociate into $e^- \bar{\nu}_e$ pairs. The requirement for two W -boson exchanges implies an enormous penalty: a suppression in the rate of order $(Q/M_W)^8 \sim 10^{-40}$, where $Q \sim 2$ MeV is a typical energy available in the decay. The rate is further suppressed because the 4 particle final state phase space is very small. The lifetime should therefore be extremely long, $\tau \sim 10^{21}$ years; and yet this process is nevertheless observed (so far for a total of 12 different nuclides). In ${}^{76}_{32}\text{Ge}$ the half-life for this process is measured to be 1.5×10^{21} years.

A related reaction which is forbidden in the standard model is the process illustrated in the right-hand panel of Figure 10.2: *neutrinoless* double beta decay. This decay is very much like the observed double beta-decay reaction, with the difference that the final-state neutrinos mutually annihilate and so do not appear amongst the final daughter particles. This process is forbidden in the standard model for two separate reasons: (i) lepton number conservation, which forbids the creation of two electrons without the concomitant creation of two antineutrinos, and (ii) helicity conservation, which forbids the right-handed antineutrino produced in the first vertex from being absorbed as a left-handed neutrino at the second vertex. Although the helicity-conservation obstacle evaporates if neutrinos have a mass, the real

significance of this decay is that its observation would establish the failure of lepton-number conservation.

To see this in more detail consider the matrix element for the decay, which from Figure 10.2 is

$$\mathcal{M} \propto \frac{G_F^2}{2} W^{\mu\nu} \left\{ \bar{u} \gamma_\mu (1 + \gamma_5) \left[\frac{-i\not{p} + m}{p^2 + m^2} \right] (1 + \gamma_5) \gamma_\nu^T \bar{u}^T \right\} \quad (10.37)$$

where \bar{u} and \bar{u}^T are the external-state electron spinors, and $W^{\mu\nu}$ denotes the relevant nuclear matrix element of the two weak currents, which depends on details of the structure of the nucleus. The last factors arise as $[\bar{u} \gamma_\nu (1 + \gamma_5)]^T$ because this part of the diagram is describing an antiparticle. Because $(1 + \gamma_5) \gamma^\mu (1 + \gamma_5) = 0$, the $-i\not{p}$ part of the propagator does not contribute; only the (helicity-flipping, lepton-number violating) mass term contributes.

For small neutrino masses (compared with the MeV energies of the decay) the decay rate is therefore suppressed by a factor of $|U_{ej}^2 m_j|^2 = |m_{ee}|^2$. Even for $m_j \sim 1$ eV and $Q \sim 1$ MeV, $(m_j/Q)^2$ represents 12 orders of magnitude in suppression. This is partially compensated by the much larger phase space available for the neutrinoless decay, as well as by differences in the nuclear-physics part of the matrix element, but the rate is still expected to be suppressed by orders of magnitude with respect to the observed $2\nu\beta\beta$ rate.

The neutrinoless mode can be distinguished experimentally from the two-neutrino mode by using the energy spectrum of the daughter electrons. Since four light leptons emerge in $2\nu\beta\beta$ decays, the sum of the electron energies is continuously distributed because some energy is lost to the neutrinos: $E_1 + E_2 \leq Q$. For the neutrinoless decay, all of the available decay energy goes into the two electrons, leading to the unique result $E_1 + E_2 = Q$ for the sum of the two electron energies. Distinguishing these alternatives requires a detector capable of examining the electron spectrum near its endpoint with high energy resolution.

Since neutrinoless double beta decay probes the combination

$$|m_{ee}|^2 = \left| \sum_j U_{ej}^2 m_j \right|^2 \quad (10.38)$$

it is also sensitive to differences in the phases, $e^{i\alpha_j}$, which appear in the lepton charged current for Majorana neutrinos. At the time of this writing, no compelling evidence for neutrinoless double beta decay exists, and the best limit on the decay lifetime for ${}^{76}_{32}\text{Ge} \rightarrow {}_{0\nu\beta\beta} {}^{76}_{34}\text{Se} - 1.9 \times 10^{25}$ years — leads to the constraint $|\sum_j U_{ej}^2 m_j| < 0.35$ eV. The precise limit on m_ν can be weaker if differences $\alpha_j - \alpha_k$ are close to π , due to potential cancellations

in the sum appearing in Eq. (10.38). The limit also depends on nuclear matrix elements for which there is considerable uncertainty.

10.4 Gauge-invariant formulations

We now return to the issue of how to embed the neutrino models described above in a way which respects the $SU_L(2) \times U_Y(1)$ gauge symmetry of the standard model. As the introductory paragraphs in this chapter make clear, there are two broad ways to proceed: either introduce additional light degrees of freedom (and so modify the assumed particle content of the model), or relax the assumption of renormalizability. Each of these alternatives has its merits, and we examine both in turn in the next two sections.

10.4.1 Sterile neutrinos

We start by entertaining the possibility that the set of fermionic fields chosen for the standard model in Eq. (2.7) is incomplete, and in particular that what is missing there is a right-handed lepton field, N_m , for each generation, in direct analogy with what was done for the quark sector. We choose the quantum numbers of this field to be what is required in order to allow a Yukawa interaction with the lepton doublet, L , required to generate a new mass term. Repeating the arguments of Chapter 2 in this case shows that the new field must be an $SU_L(2)$ singlet and to have hypercharge $Y = 0$ in order to allow the gauge-invariant interaction of the schematic form $(LN)\tilde{\phi}$.

Since the field N must be a singlet under *all* $SU_c(3) \times SU_L(2) \times U_Y(1)$ interactions, it couples to ordinary matter only through the Yukawa coupling which gives it a mass. We see in this way why the additional neutrino states must be sterile. The new renormalizable terms which can be introduced into the standard model Lagrangian given this new field are

$$\mathcal{L}_N = -\frac{1}{2}\bar{N}_m \not{\partial} N_m - \frac{1}{2}M_m \bar{N}_m N_m - (k_{mn}\bar{L}_m P_R N_n \tilde{\phi} + \text{h.c.}) \quad (10.39)$$

The Yukawa couplings between the lepton doublet, L , and N preserve total lepton number, provided we assign lepton number $L = +1$ to the right-handed field, $P_R N$. The same Yukawa interactions violate the separate lepton numbers, L_e , L_μ , and L_τ , so long as k_{mn} is not diagonal.

The Majorana mass term $M_m \bar{N}_m N_m$ has not been encountered before for the standard-model fields, since for these a direct mass term of this type cannot be gauge invariant. This term breaks each of the lepton numbers, L_m , so long as the corresponding mass term, M_m , is nonzero. In particular, any nonzero M_n breaks the total lepton number $L = L_e + L_\mu + L_\tau$, by two

units, as may be seen because $P_R N$ and $\bar{N} P_R$ carry lepton number $+1$ and $P_L N$ and $\bar{N} P_L$ carry lepton number -1 .

The implications for neutrino masses in this scenario are found by replacing the Higgs doublet by its vacuum expectation value. Once this is done we recognize that this model is equivalent to the sterile-neutrino model of Eq. (10.30), with $N = 3$ sterile neutrinos: $N_n \rightarrow s_a$; and mass matrices

$$m_{ab} = 0, \quad \mu_{ab} = \frac{k_{ab} v}{\sqrt{2}} \quad \text{and} \quad M_{ab} = M_a \delta_{ab} \quad (10.40)$$

We find in this way the gauge-invariant extension of the sterile-neutrino models of earlier sections. Having a gauge-invariant formulation allows some further insights into the possible neutrino-mass scenarios which were discussed earlier, some of which we now record.

- In principle, gauge invariance implies relationships amongst different couplings, and as usual we find that these relationships relate mass terms to Higgs couplings. Unfortunately these cannot be tested until the Higgs boson is found.
- The embedding into a gauge-invariant formulation provides a simple explanation for why the new sterile-neutrino states, s_x , introduced in Subsection 10.2.2 do not couple to other standard-model particles apart from through their mixings with neutrinos. They do not do so because no such renormalizable couplings are possible given that participation in the neutrino-mass term requires the new particles to be singlets under the standard-model gauge group.
- The gauge-invariant formulation also provides a useful insight into how large the various mass terms might reasonably be expected to be. It implies in particular that it is natural to choose $m_{ab} = 0$, because no renormalizable interaction can give this term given the assumed sterile-neutrino field content. The Dirac mass term, μ_{xa} , is seen to be proportional to the Higgs *v.e.v.*, and so is naturally at most a few hundred GeV and possibly much less than this if the dimensionless Yukawa couplings, k_{mn} , should be very small. By contrast, the Majorana mass terms, M_m , need not vanish in the limit where the electroweak symmetry is unbroken, and so (as is argued in more detail in the next chapter) would naturally be expected to be the weak scale or larger. This naturally points us towards the see-saw scenario of neutrino masses discussed in Subsection 10.2.2, for which $m \ll \mu \ll M$.
- A problem with any scenario for which μ_{ax} dominates or is comparable with the observed neutrino masses (like Dirac or pseudo-Dirac neutrinos or the light-sterile-neutrino scenario) is that extremely small Yukawa

couplings are required. This is because direct laboratory searches for neutrino masses require that all neutrinos which mix appreciably with ν_e have masses satisfying $m < 2$ eV. Cosmological considerations also suggest $m_\nu < 0.3$ eV for all of the neutrinos which are abundant during the cosmologically observable universe, and in particular for those which overlap significantly with the three active neutrinos.

Requiring $k_{mn}v$ to be this small requires the eigenvalues of the matrix k_{mn} must be $< 10^{-11}$. Indeed, the cosmological bound constrains the sum of eigenvalues, and so requires $\text{tr } k_{mn} < 4 \times 10^{-12}$. This is to be contrasted with the other Yukawa coupling matrices, for which $\text{tr } f_{mn} \simeq 10^{-2}$, $\text{tr } h_{mn} \simeq 2.5 \times 10^{-2}$, and $\text{tr } g_{mn} \simeq 1$. Of course, we do not understand why any of the fermion masses are what they are, but one must always pause when arbitrarily setting dimensionless couplings to be zero to more than ten decimal places.

- These same models similarly require extremely small choices to be made for the dimensionful Majorana mass parameters, M_n . For instance, for the Dirac-neutrino scenario to be correct the mass, M_m , must satisfy the extraordinary bound $M_m < 10^{-20}\mu$, where μ is the mass parameter appearing in the Higgs potential, see Eq. (2.26), and indeed is the only other mass parameter at all in the standard model. By contrast, the normal situation in the standard model is that all couplings which can be present on the grounds of symmetry and renormalizability are actually measured to be nonzero. The sole exception is the Θ_3 term in Eq. (2.13), which satisfies $|\Theta_3| < 10^{-9}$. (The puzzle as to why this is so is discussed in Section 11.4.) A Dirac-neutrino scenario must explain why the same should be true (with extraordinary accuracy) for the Majorana neutrino masses.

There is a symmetry principle which could enforce the vanishing of M_n ; if we demand, besides the gauge symmetries, that the standard model satisfy an exact $B-L$ symmetry, then the vanishing of M_m is automatic. As noted in Subsection 2.5.3, such a symmetry has vanishing gauge and gravitational anomalies; so this possibility is at least consistent. At our present level of understanding such a proposal has two drawbacks, however. The first drawback of exact $B-L$ conservation is that it is difficult to reconcile such a symmetry with the observed matter-antimatter asymmetry in the universe today. Since a proper discussion of this is beyond the scope of this book, we do not discuss it further.

The second drawback is that such a theory requires stepping back from the successful standard-model understanding of the charge assignments of the various matter fields (and so for understanding the exact charge

neutrality of the neutron, for example). In the standard model without right-handed neutrinos, the hypercharges of all particles are determined uniquely (up to overall normalization) by gauge-anomaly cancellation, as we saw in Subsection 2.5.3. This guarantees that the quark hypercharge assignments ensure that the neutron electric charge is precisely zero. Adding an N field to each generation adds a new contribution to Eq. (2.131), which turns out to make it hold automatically. In the theory with an N particle added, the anomaly-cancellation conditions remain satisfied even if we shift each particle's hypercharge y by δ times its $B - L$ charge, for any δ . Such a shift would change the electric charge of the neutron by δ . Since the neutron's electric charge is measured to differ from zero by no more than 10^{-21} , a theory with an N field and exact $B - L$ conservation requires an extraordinary tuning of particle hypercharges. (On the other hand, when a Majorana-mass term like $M\bar{N}N$ is present, gauge invariance of that term demands $\delta = 0$, and so again ensures the exact neutrality of the neutron.)

10.4.2 Dimension-5 interactions

Suppose next that we do not enlarge the field content of the standard model, as would be appropriate if there are no new particles at the energies accessible in current experiments. It is still possible to account for neutrino oscillations in this case, provided that we allow the introduction of non-renormalizable interactions. Before exploring the implications of this approach for neutrino physics, we first address what is implied physically when we relinquish the condition of renormalizability.

10.4.2.1 The significance of non-renormalizable physics

Up to this point, we have considered the standard model to be a renormalizable theory. That is, we have considered the most general theory which consists of the particle content presented in Chapter 2, together with all terms in the Lagrangian consistent with the gauge symmetries which only having couplings with dimension $(\text{mass})^p$, with $p \geq 0$. This led to an experimentally tremendously successful model of modern particle physics, which passes with flying colors an enormous number of high-precision experimental tests. Its one failure is its inconsistency with the observed neutrino-oscillation phenomena discussed above. Yet a single failure is sufficient to be fatal and so we now entertain the possibility of extending the standard model by non-renormalizable interactions.

To see what this choice means physically we revisit the discussion of

Chapter 7, where we encountered non-renormalizable interactions in the context of the Fermi theory of the low-energy weak interactions. There we saw that non-renormalizable four-fermion interactions were required to describe the effects of virtual W and Z exchange within the effective theory describing physics at energies well below the W mass. These effective interactions had couplings proportional to $G_F \propto M_W^{-2}$, and this is why they were non-renormalizable. But such couplings were also physically reasonable because they expressed the large penalty which must be paid (because of the uncertainty principle) at low energy in order to produce a high-energy state like a virtual W or Z . If we never try to probe energy scales as large as the W mass, the Fermi theory provides a completely adequate description of the weak interactions at low energies.

Non-renormalizable interactions were also encountered in the effective Lagrangians discussed in Chapter 9, which were claimed to describe the dynamics of strongly-interacting mesons and baryons at energies well below typical strong-interaction scales. In this case the non-renormalizable interactions were either proportional to powers of $(4\pi F_\pi)^{-1}$ or the inverse of a similar strong-interaction scale. Again this expressed the fact that the effective theory was meant only to capture the physics of the strong interactions at energies small compared with $4\pi F_\pi \sim 1$ GeV.

The potential problem with non-renormalizable interactions is that there are in principle an infinite number of them, corresponding to arbitrary powers of the fields and their derivatives, and so the admission of so many interactions into the theory threatens to destroy its predictivity. It is the low-energy approximation which saves the day in this case, since there are only a limited number of effective interactions possible whose coefficients have a given power of an inverse heavy mass. For instance, if our interest is only in contributions to observables which are proportional to one power of M_W^{-1} , we need consider only interactions whose dimension (counting only derivatives and fields) is $(\text{mass})^5$ or lower. Similarly, allowing contributions up to M_W^{-2} requires keeping interactions whose operator dimension is $(\text{mass})^6$ or lower. Effective interactions having operator dimension $(\text{mass})^p$ are called “dimension- p ” interactions. We may restrict our attention to low-dimension interactions because it is the lowest-dimension interactions which should be the most important at low energies.

Since renormalizable interactions are defined to be those having dimension 4 or less, they are the only ones which do not involve any inverse powers of a heavy mass like M_W . From this point of view keeping only renormalizable interactions is the appropriate treatment if we wish to neglect all of the virtual effects of higher-energy particles. This is what would be appropriate if

we believe that these particles are very much more massive than the energies for which we wish to make predictions. Applied to the standard model, we see that renormalizability is appropriate if we believe the standard model contains all of the particles which are relevant up to energies much higher than those of present experimental interest. Indeed, because the standard model is such a successful theory we expect that the mass scale associated with any high-dimension operators is likely to be much larger than M_W .

What would it mean to use non-renormalizable interactions to describe neutrino masses? It would mean that there is some new physics at energies much higher than the weak scale which violates the conservation of the three lepton numbers, and it is the virtual effects of these particles which underlies the physics of neutrino mass. If such an approach is successful, the mass scale in the denominator of the relevant effective coupling gives an indication of just how massive the lepton-violating physics is likely to be.

10.4.2.2 Lowest-dimension non-renormalizable operators

Setting aside for the moment the desire to understand neutrino oscillations, we instead first adopt the most conservative and constructive approach and ask: what are the lowest-dimension non-renormalizable interactions which are possible? Since the lowest-dimension interactions should have the largest effects at low energies, the catalog of lowest-dimension interactions should point to the kinds of observables in which new physics should first appear.

In this section we perform this exercise for the standard model and find that there is a unique dimension-5 interaction which is allowed by the standard-model particle content and gauge symmetries, and so if any new physics is very heavy this should be the interaction which first shows deviations from standard-model predictions. What is most remarkable is that this unique term is precisely what gives Majorana masses to the three standard-model neutrinos!

We now construct this term. Effective interactions must be Lorentz scalars and gauge invariant. The fields available with which to build them are scalars, ϕ , fermions, ψ , and gauge bosons (or, equivalently, covariant derivatives, $D = \partial - ieA$). Notice that gauge-field strengths need not be counted separately because they can be constructed as the commutator of two covariant derivatives, $[D_\mu, D_\nu] = -ieF_{\mu\nu}$. Since interactions having dimension 4 or less are renormalizable, the leading non-renormalizable interactions have dimension 5. Such an operator can be made up of the given fields and derivatives in the following possible schematic forms:

$$\phi^5, D^1\phi^4, D^2\phi^3, D^3\phi^2, \bar{\psi}D^2\psi, \phi D\bar{\psi}\psi, \phi^2\bar{\psi}\psi \quad (10.41)$$

Starting with these, and requiring Lorentz and gauge invariance, it is straightforward to show that the standard model field content allows only one dimension-5 interaction, namely (see Problem 10.1)

$$\mathcal{L}_{\text{eff}} = -\tilde{k}_{mn}\tilde{\phi}_\alpha(\bar{L}_m^\alpha P_R L_n^\beta)\tilde{\phi}_\beta + \text{h.c.} \quad (10.42)$$

where we write out the $SU_L(2)$ indices explicitly. Notice that this interaction violates total lepton number, L , by two units. The matrix of coefficients \tilde{k}_{mn} must be complex and symmetric. Since this interaction has dimension 5, these couplings have mass dimension -1 , and so can be written $\tilde{k}_{mn} = c_{mn}/\Lambda$, with the dimensionless coefficients $c_{mn} \lesssim 1$, and where Λ is an energy scale which must satisfy $\Lambda \gg M_W$. (If it did not then the virtual heavy particles whose physics this interaction describes would not be heavy, and it would not be a good approximation to neglect dimension 6, dimension 7 and higher effective interactions.) Otherwise, we impose no *a priori* constraints on the \tilde{k}_{mn} , and try to determine them from experiment.

The physical implications of this interaction are most easily seen by going to unitary gauge, for which it becomes

$$\mathcal{L}_{\text{eff}} = -\frac{1}{2}\tilde{k}_{mn}(\bar{\nu}_m P_R \nu_n)(v + H)^2 + \text{h.c.} \quad (10.43)$$

where as before $v = 246$ GeV and H is the physical Higgs boson. We see that it describes a left-handed Majorana-neutrino mass of size

$$m_{ab} = \tilde{k}_{ab}^* v^2 = c_{ab}^* \frac{v^2}{\Lambda} \quad (10.44)$$

plus $\nu\nu H$ and $\nu\nu HH$ neutrino–Higgs interactions. Because the observation of the implication of the νH interactions must await the discovery of the Higgs, the only observable implication of this unique dimension-5 interaction is an arbitrary Majorana neutrino mass, of order v^2/Λ . The larger the scale Λ the smaller the neutrino masses which result. Neutrino masses of order $m \sim 50$ meV (the order of magnitude of the observed neutrino masses), require a mass scale $\Lambda \sim 10^{14}$ GeV. The enormity of this scale gives us good faith that it is legitimate to neglect still higher-dimension interactions. It also means that if there are only three light neutrinos it is not hard to understand theoretically why their masses should be so small, since this is a natural consequence of the scale of the new lepton-number-violating physics being very large.

10.4.2.3 Possible origins of the dimension-5 operator

The previous sections argue that if the standard model is an incomplete description of physics at very high energies, then the low-energy consequence

is that it must be augmented by high-dimension operators, the first of which is a unique dimension-5 operator. This operator induces neutrino masses, which are in fact observed. If this is correct then we know very little about the nature of the physics at the scale Λ whose virtual effects are the physical origin of the dimension-5 operator of Eq. (10.42).

The purpose of this section is to furnish a few speculations as to what this physics might be. We show, first, that rather simple extensions of the standard model, incorporating new heavy degrees of freedom, can generate the dimension-5 operator. However, by providing several possibilities we show that even if low-energy experiments completely determine the coefficients \tilde{k}_{mn} of the dimension-5 operator, this is still insufficient information to reconstruct completely the high-energy physics which is responsible for it.

We emphasize that this section is by no means a complete discussion of the scenarios put forward to explain neutrino masses. On the contrary, the models described here are the tip of the proverbial iceberg made up of the innumerable models presented in the literature over the years.

The seesaw mechanism

The first model we examine is one we have already encountered, wherein the new heavy physics consists of a collection of heavy sterile neutrinos, N_m , which have vanishing hypercharge and are singlets under $SU_L(2)$ and $SU_C(3)$. The new terms which can be added to the Lagrangian once such fields are included are

$$\mathcal{L}_N = -\frac{1}{2}\bar{N}_m \not{\partial} N_m - \frac{1}{2}M_{mn}\bar{N}_m N_n - (k_{mn}\bar{L}_m P_R N_n \tilde{\phi} + \text{h.c.}) \quad (10.45)$$

The mass matrix M_{mn} can be taken diagonal and real as discussed in Subsection 1.3.1.

There is no principle which forbids the eigenvalues of $M^\dagger M$ from being much larger than the other mass scales of the standard model, like M_W^2 or the Higgs-boson potential parameter μ^2 . In fact, as we saw in Subsection 10.2.2, taking M to be large leads to the very attractive seesaw picture of neutrino masses, wherein the sterile neutrinos have masses of order M while the three standard-model neutrinos acquire a small mass matrix of order $(kM^{-1}k^T)v^2$.

We therefore consider the possibility that the Majorana masses, M_m , are much larger than all of the electroweak scales, $M_m \gg |\mu| \sim 100$ GeV. In this case the sterile neutrinos we have included are actually so heavy that they can never be produced at the energies available in experiments, and so they can only influence observables through their virtual effects. We should

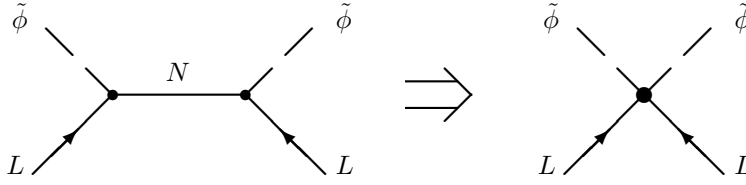


Fig. 10.3 Integrating out the right-handed neutrino.

therefore be able to capture their effects to leading order in M_n^{-1} using a non-renormalizable effective interaction.

In order to demonstrate that this is true we can follow the steps taken in Chapter 7 for the W boson. To this end consider the Feynman graph of Figure 10.3, through which virtual sterile-neutrino exchange generates an interaction between two standard-model lepton and Higgs doublets. The matrix element for the diagram on the left in Figure 10.3 is (writing out the $SU_L(2)$ indices explicitly)

$$k_{mo} k_{pn}^T \tilde{\phi}_\alpha \left\{ \bar{L}_m^\alpha P_R \left[\frac{M - i\not{p}}{M^2 + p^2} \right]_{op} P_R L_n^\beta \right\} \tilde{\phi}_\beta + \text{h.c.} \quad (10.46)$$

With the idea that M is large, we now take the large-mass limit, $M \gg p$, and expand this amplitude in powers of M^{-1} . To leading order the resulting amplitude is reproduced by taking the matrix element of the following effective interaction

$$\mathcal{L}_{\text{eff}} = - \left(-k M^{-1} k^T \right)_{mn} \tilde{\phi}_\alpha \bar{L}_m^\alpha P_R L_n^\beta \tilde{\phi}_\beta + \text{h.c.} \quad (10.47)$$

There are other interactions generated by integrating out the N field; but the one we have written is the only interaction suppressed by only a single power of the mass M . We recognize this interaction as exactly the one written down in Eq. (10.42), with the identification

$$\tilde{k}_{mn} \equiv - \left(k M^{-1} k^T \right)_{mn} \quad (10.48)$$

We see in this way that very heavy sterile neutrinos can induce in a simple way the dimension-5 operator responsible for Majorana neutrino masses. Conversely, since the dependence of light neutrino masses obtained from the dimension-5 interaction is generically $m \sim v^2/M$, this interaction expresses the seesaw relation between neutrino masses and heavy scales in its most general context.

It is clear from this example that not all of the information about the heavy right-handed neutrino and its interactions survives in the effective couplings, \tilde{k}_{mn} . In particular, the original Yukawa coupling matrix, k_{mn} , has 18 independent real parameters, and the heavy Majorana mass matrix M has three more. Unless the mass matrix has degenerate eigenvalues, the only reparameterization freedom we are allowed to redefine couplings are the three phase rotations of the three components of N_m ; so altogether there are 18 parameters needed to describe the sterile neutrino interactions in the full theory. On the other hand, the effective coupling matrix, $\tilde{k}_{mn} \equiv k_{mo} M_{op}^{-1} k_{pn}^T$, is symmetric, and so has only 12 real parameters. At least six of the sterile-neutrino parameters therefore cannot be encoded in the effective couplings of the dimension-5 operator in the low-energy theory.

Triplet scalar fields

Another way to generate the same dimension-5 interaction at low energies is by modifying the Higgs sector. To this end suppose that the standard model were supplemented by a new complex Higgs scalar, Φ , which transforms under the gauge group in precisely the same way as does the pair of Higgs doublets appearing in Eq. (10.42). That is, suppose Φ carries hypercharge -1 , is a color singlet and is a triplet of $SU_L(2)$:

$$\Phi \quad \text{transforms as} \quad (\mathbf{1}, \mathbf{3}, -1) \quad (10.49)$$

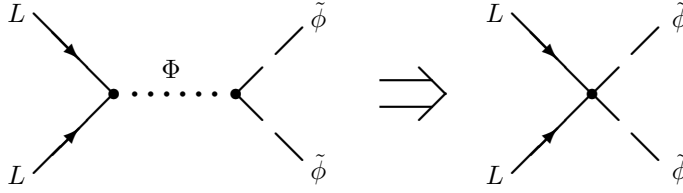
Such a field could also generate a neutrino mass through a new kind of Yukawa coupling.

The new terms which appear in a renormalizable extension of the standard model to include such a field are

$$\begin{aligned} \mathcal{L}_\Phi = & -(D_\mu \Phi)^\dagger D^\mu \Phi - M_\Phi^2 \Phi^\dagger \Phi - \lambda_\Phi (\Phi^\dagger \Phi)^2 - \lambda' \Phi^\dagger \Phi \phi^\dagger \phi \\ & - [y_{mn} \bar{L}_m (\Phi_a \tau^a \epsilon) P_R L_n + c \phi^\dagger \Phi_a^* \tau^a \tilde{\phi} + \text{h.c.}] \end{aligned} \quad (10.50)$$

Here τ^a are the Pauli matrices acting in the space of $SU_L(2)$ indices, and $\epsilon = i\tau^2$ also acts on $SU_L(2)$ indices. The last terms, in brackets, are the most important for our purposes. Because of these terms, it is impossible to assign lepton numbers so that L is a good symmetry. Either Φ must carry lepton number 2, in which case the c term is lepton number violating, or Φ carries no lepton number, in which case the y_{mn} term is lepton number violating.

We take the Φ mass parameter, M_Φ^2 , to be large and positive. There are two sound reasons for doing so. One of these is a direct bound, since the Φ

Fig. 10.4 Integrating out the Φ field.

particle cannot be lighter than ~ 100 GeV without having been observed in Higgsstrahlung events at LEP II. Furthermore, the experimentally successful mass relation $M_Z \cos \theta_W = M_W$ is ruined if the Φ -field *v.e.v.* is too large compared with that of the Higgs doublet, and large M_Φ suppresses the size of $\langle \Phi \rangle$. To see why large M_Φ^2 keeps the Φ *v.e.v.* small but nonzero consider the expansion of the scalar potential near $\Phi = 0$:

$$\begin{aligned}
 V(\Phi) = & \left(M_\Phi^2 + \frac{\lambda' v^2}{2} \right) \Phi^\dagger \Phi \\
 & + \left\{ c \begin{pmatrix} 0 & v/\sqrt{2} \end{pmatrix} \begin{pmatrix} \Phi_3^* & \Phi_1^* - i\Phi_2^* \\ \Phi_1^* + i\Phi_2^* & \Phi_3^* \end{pmatrix} \begin{pmatrix} v/\sqrt{2} \\ 0 \end{pmatrix} + \text{h.c.} \right\} \\
 & + O(\Phi^4)
 \end{aligned} \tag{10.51}$$

Once the Higgs doublet acquires its expectation value, we see that it generates a linear term in the Φ potential due to the c interaction, leading to

$$\langle \Phi_1 \rangle = -i\langle \Phi_2 \rangle = -\frac{cv^2}{2M_\Phi^2 + \lambda'v^2} \tag{10.52}$$

Given the choice $M_\Phi^2 \gg v^2$, from here on we drop the $\lambda'v^2$ term in this last expression.

Substituting this result into the y term in the Lagrangian leads to the neutrino mass

$$y_{mn} \bar{L}_m (\Phi_a \tau^a \epsilon) P_R L_n + \text{h.c.} \rightarrow \left(\frac{y_{mn} c v^2}{M_\Phi^2} \right) \bar{\nu}_m P_R \nu_n + \text{h.c.} \tag{10.53}$$

a result which depends on the dimensionless coupling, y_{mn} , the dimensionful coupling c , the ϕ expectation value, and the Φ mass. Most importantly, it varies inversely with M_Φ , and so the heavier M_Φ is, the smaller the neutrino masses become.

This is another example where the particle which is responsible for

generating neutrino masses is very heavy, and so can only contribute to low-energy physics through its virtual effects. Consequently its implications at low energy – and for neutrino masses in particular – should be captured through an effective non-renormalizable interaction, as we now compute.

To this end we evaluate the Feynman diagram of Figure 10.4, through which virtual Φ -exchange generates a coupling between the standard-model lepton and Higgs doublets. Using the relationship

$$\tau_{ij}^a \tau_{kl}^a = 2\delta_{il}\delta_{jk} - \delta_{ij}\delta_{kl} \quad (10.54)$$

and the fact that $\phi^\dagger \tilde{\phi} = \phi^{+*} \phi^{0*} - \phi^{0*} \phi^{+*} = 0$, we can evaluate the diagram to obtain

$$\mathcal{L}_{\text{eff}} = - \left(\frac{2k_{mn}c}{M_\Phi^2} \right) \tilde{\phi}_\alpha [\bar{L}_m^\alpha P_R L_n^\beta] \tilde{\phi}_\beta + \text{h.c.} \quad (10.55)$$

which is again our dimension-5 operator, with the effective coupling computed to be $\tilde{k}_{mn} = 2k_{mn}c/M_\Phi^2$. Using this value of \tilde{k}_{mn} to compute the neutrino masses gives the same neutrino mass as Eq. (10.53). To identify how large M_Φ must be we choose the coefficients k_{mn} to be $O(1)$, and choose $c \sim v \sim 100$ GeV. In this case the resulting neutrino mass is in the 50 meV range when $M_\Phi \sim 10^8$ GeV. Alternatively, if $c \sim M_\Phi$, then for $k_{mn} \sim 1$, we need $M_\Phi \sim 10^{14}$ GeV, the same scale we found in the seesaw mechanism.

Again it is clear that knowledge of the effective coupling \tilde{k}_{mn} does not suffice to determine the higher-energy physics. In particular, the values of c and M_Φ^2 appear only in a particular combination which cannot be separated purely from \tilde{k}_{mn} . Furthermore, the coefficients λ_Φ and λ' of the scalar-triplet model do not appear at all in the dimension-5 interaction, and so have no consequence (to leading order) for low-energy physics.

The final message is this: the observation of neutrino oscillations forces a change to be made to the standard model, and this implies either that we are seeing the influence of new particle states at low energies, or the virtual effects of heavy particles whose implications are captured by non-renormalizable interactions. Phenomenologically successful models of both types can be built, and there is at present insufficient information to distinguish which option is right.

The absence of any evidence for sterile-neutrino oscillations provides some evidence in favor of the heavy-physics option, and in this case the unique effective interaction which would be expected to dominate at low energies is precisely the kind of operator which would generate a Majorana neutrino mass of the kind required to describe neutrino oscillations. If this is the

correct picture then the small size of the observed neutrino masses is explained by the very high mass of the lepton-violating physics which is responsible for these masses. Unfortunately, in this scenario we are also unlikely to learn more from experiments at accessible energies about the details of this physics.

10.5 Problems

[10.1] Dimension-5 interactions

Starting with the schematic expressions, Eq. (10.41), prove that the dimension-5 neutrino-mass term, Eq. (10.42), is the unique dimension-5 interaction consistent with Lorentz invariance, the standard model field content and gauge invariance.

[10.2] Neutrino decay

Suppose that the neutrino were massive and, for some reason, it could decay, with a lifetime of τ in its rest frame and invisible decay products.

Compute the survival probability for solar neutrinos as a function of the neutrino energy. Show that the energy dependence of this rate is opposite to what is seen in solar neutrino data; it is the low-energy neutrinos which are the most depleted and the high energy neutrinos which are most likely to survive.

[10.3] MSW effect and sterile neutrinos

In discussing the MSW effect, we discarded corrections which arise due to neutral current interactions, since they are generation blind. However, if neutrino oscillations involve a sterile neutrino, then this contribution is relevant, since the sterile neutrino does not participate in neutral current interactions.

Find the medium-induced correction to the neutrino effective Lagrangian due to neutral-current weak interactions, analogous to that found in Eq. (10.22). In doing so, treat the neutron and proton currents as $i\bar{p}\gamma^\mu(T_3P_L - Q\sin^2\theta_w)p$ and similarly with n , where p, n are taken to be an isospin doublet, $T_3(p) = 1/2$ and $T_3(n) = -1/2$. How much of the neutral current contribution arises from electrons and how much from nuclei?

[10.4] Dimension-5 operators with squarks

Find all renormalizable and dimension-5 operators for a theory in which the standard model is supplemented by the color triplet scalar introduced in Problem 2.7.

11

Open questions, proposed solutions

The standard model – augmented (say) with dimension-5 operators to account for neutrino oscillations – explains all particle physics experiments performed to date (2006). Yet there are a number of reasons to believe that it is incomplete, and should be regarded at best as being the effective theory describing particle physics at the energy scales which have been probed experimentally (roughly several hundred GeV).

This chapter aims to summarize these reasons, with an eye to identifying the main themes which govern the searches for the standard model’s replacement. These themes typically revolve about “puzzles,” which either center around attempts to explain the values of some of the standard model’s couplings, or around speculations about what kinds of new particles might exist at very large masses, and what their implications might be for experiments at accessible energies. Our goal in this summary is not to be exhaustive, but rather to provide a conceptual framework for further reading of the many research directions within the literature on the broad topic of physics “beyond the standard model.”

The organizing theme for our discussion is the assumption that any particles which have not yet been discovered must be heavy compared with the energies to which we presently have experimental access. This assumption has three motivations, not least of which is the outstanding success of the standard model itself. After all, given that it is itself the most general theory consistent with general principles and the assumed particle content, its success would automatically follow if any new particles (beyond the Higgs scalar) were quite heavy. A second motivation is the great success of this approach (see Subsection 10.4.2) when it is brought to bear on the problem of neutrino oscillations. Finally, the approach is motivated by the fact that it is effective; this assumption allows us to organize more systematically the

discussion of the puzzles and open questions of the standard model in terms of the kinds of low-dimension effective interactions which are possible.

11.1 Effective theories (again)

It is an experimental fact that nature comes to us with an enormous variety of energy scales, ranging from the size of the entire visible universe down to the smallest distances we can presently measure. Science progresses because long-distance physics can be understood without needing to know in detail what is going on at shorter distances, and it can do so because the physics of long distances does not depend strongly on very many features of shorter-distance phenomena. For instance, atomic energy levels can be understood in detail without knowing more about the nucleus than its total mass, spin, magnetic moment, and electric charge.

Effective field theory provides a mathematical expression for this remarkable fact, by showing us how to identify efficiently which interactions are important at long distances (and so low energies) in any given physical situation. When the microscopic (high-energy) theory is itself understood, it is often possible to explicitly construct the relevant effective theory governing the low-energy fields, ℓ_i , by “integrating out” the heavy degrees of freedom, h_a . We have seen some explicit examples of this kind of calculation already in this book, in Chapters 7 and 10.

What such calculations show is that the relevant low-energy theory may be written as the sum over low-dimension interactions constructed purely from the light fields, of the form

$$\mathcal{L}_{\text{eff}} = \sum_I c_I \mathcal{O}_I(\ell_i) \quad (11.1)$$

where the effective interactions, \mathcal{O}_I , are built from powers of the light fields and their derivatives, and are organized in order of increasing dimension. If \mathcal{O}_I is a dimension- d operator, then its effective coupling, c_I , has dimension $(\text{mass})^{4-d}$ and may be computed once the dynamics of the heavy fields, h_a , is known.

In the examples of these kinds of calculations performed in Chapters 7 and 10 it was found that the virtual effects of a heavy particle of mass M could be reproduced by effective dimension-5 (or dimension-6) interactions with effective couplings $c \sim \tilde{c}/M$ (or $c \sim \tilde{c}/M^2$), with the dimensionless constants \tilde{c} being order 1 (or perhaps involving a loop factor, like $\tilde{c} \sim \alpha/4\pi$, or a logarithmic dependence on mass ratios). More general calculations verify that it is generically true that integrating out a heavy particle of mass M

generates a host of effective interactions whose couplings depend on M as is required on essentially dimensional grounds: dimension- d operators \mathcal{O}_I acquire effective couplings with $c_I \sim \tilde{c}_I M^{4-d}$.

11.1.1 Naturalness and UV sensitivity

This heavy-mass dependence of effective couplings – $c_I \propto M^{4-d_I}$ for interactions \mathcal{O}_I having operator dimension d_I – naturally ensures the suppression of non-renormalizable interactions (i.e. those with $d_I > 4$) by powers of $1/M$, expressing how the uncertainty principle suppresses the quantum fluctuations which generate the virtual effects of very energetic states. The fact that these couplings vanish as $M \rightarrow \infty$ expresses the experimental fact that high-energy (short-distance) physics “decouples” from low-energy (long-distance) physics.

But what about those operators whose dimension is $d_I \leq 4$? For these, the same rule would lead us to expect *enhancement* by the heavy mass M , with $c_I \sim \ln M$ for $d_I = 4$, $c_I \sim M$ for $d_I = 3$, or $c_I \sim M^2$ for $d_I = 2$. This is indeed what explicit calculations reveal, some of which are described in subsequent sections. Generally speaking, the possibility of obtaining such low-dimension interactions expresses the appearance of positive powers of the heavy masses in observables of the full theory. (Indeed, the factor of m_t^2 in the ρ -parameter of Chapter 7, or in $K\bar{K}$ mixing, can be interpreted in this way within the effective theory obtained from the standard model by integrating out the t -quark.) These low-dimension interactions are therefore important, because they can identify when observables are more sensitive than usual to very heavy mass scales.

This observation leads to several puzzles if the standard model itself is regarded as being an effective theory, obtained once some still-heavier particles are integrated out. This is because there is some evidence that nature includes physics involving mass scales much larger than a few hundred GeV, and yet the standard model *does* contain some low-dimension operators whose effective couplings cannot also be similarly large.

There are, at present, two kinds of evidence for the existence of physics above the electroweak scale. The strongest of these comes from gravity, which is characterized by a single dimensionful quantity: Newton’s constant, G_N . Newton’s constant defines an enormous mass scale, $M_p = G_N^{-1/2} \simeq 1.22 \times 10^{19}$ GeV, which is characteristic of gravitational physics. Indeed gravity is by far the weakest force – e.g. $G_N/G_F \simeq 10^{-33}$ – precisely because this mass scale is so large compared with the other mass scales we know in physics. Since Newton’s constant is an inverse power of mass the

gravitational interactions are non-renormalizable, and so are likely the low-energy effective theory for some new physics which intervenes at some energy $v \ll E \lesssim M_{\text{p}}$.

A second indication that nature involves physics at scales larger than $v = 246 \text{ GeV}$ comes from the analysis of neutrino masses in terms of dimension-5 interactions (or in terms of the seesaw mechanism). In this picture neutrino masses are light because they are inversely related to a very heavy mass, $m \sim \tilde{k}v^2/M$. Neglecting dimensionless factors, this implies a new scale $M \sim \tilde{k} \times 10^{14} \text{ GeV}$. (As we see below, a third hint that there may be new physics at these energies also comes from the observation that under certain circumstances the standard-model gauge couplings can run until all three take a common value at very high energies, of order 10^{16} GeV .)

In viewing the standard model as an effective theory and asking how it may be extended, we are already making a key assumption – that the Higgs mechanism is correct, and the Higgs boson will be found with the properties expected within the model. This assumption should give us pause, but there are good grounds for making it. First, the theory *must* be extended to include a Higgs boson or some other degrees of freedom which perform its function. For instance, as discussed in Problem 7.3, certain cross sections become unphysically large at high energies if there is no Higgs boson. Further, precision tests are now sensitive to radiative (loop) effects from Higgs bosons, and the fits favor a light Higgs boson. We will, however, contemplate the possibility that there are multiple Higgs fields, or a more complicated implementation of electroweak symmetry breaking, in Subsection 11.3.2.

11.2 Dimension zero: cosmological-constant problem

The lowest possible dimension for an operator is dimension 0, corresponding to no dependence at all on any fields or derivatives. Such an operator can appear in the standard model, corresponding to adding an arbitrary constant, ρ_0 , to the Lagrangian density. Indeed we encountered such a term in Chapter 1 – see Eq. (1.56) – where it was required to renormalize the quantum zero-point energy, and was interpreted as the energy density of the vacuum.

11.2.1 The cosmological-constant problem

For a particle physicist the vacuum energy is normally irrelevant, since its value cannot be measured; it is only the *difference* of energies for various states which is normally measured, and the vacuum contribution always

cancels in this difference. This is the reason we have not discussed this term in any detail to this point in this book.

However, the vacuum energy *is* measurable, just not by particle physicists. It can be measured because gravity couples to all energies regardless of their source, and so in particular gravity responds to the energy density of the vacuum.[†] For gravity a nonzero vacuum energy is equivalent to including the presence of Einstein’s “cosmological constant” into the gravitational field equations, and it can lead to measurable effects for the expansion history of the universe (but not to observable effects on smaller distances, like within the solar system). In fact, current observations of the universal expansion suggest that the universe does indeed have a net vacuum energy density, of size

$$\rho_{\text{cc}} \sim (2.4 \times 10^{-3} \text{ eV})^4 \quad (11.2)$$

The problem with this measurement is that it is so small, since the arguments of the previous section would lead us to expect on dimensional grounds that the contribution to ρ_0 [‡] of a particle of mass m is (including a loop factor)

$$\delta\rho_0 \sim O\left(\frac{m^4}{16\pi^2}\right) \quad (11.3)$$

which is much larger than the measured value for pretty much any particle we know of (except possibly for neutrinos).

This expectation is borne out by the explicit calculation of the zero-point fluctuations of a scalar and fermionic field, presented in Eq. (1.71) and Eq. (1.114) respectively. Naively, these calculations actually gave an ultraviolet-divergent result, $\delta\rho_0 = c_0\Lambda^4 + c_2m^2\Lambda^2 + c_4m^4\log(\Lambda/m) + \dots$, with appropriate constants c_0, c_1, c_2 , etc., but when integrating out a particle we are instructed to define the effective interactions such that they encode the difference between the theory with the heavy particle and without, *after* regularization. The issue of whether the Λ^4 and Λ^2m^2 terms exist and how to understand them is a question about the ultraviolet regularization of the theory, that is, about quantum gravity, which we leave aside. The existence of an m^4 contribution is an infrared effect which is independent of the regularization and of quantum gravity issues. It must be taken seriously.

If there is physics at scale $M \sim M_{\text{p}} \sim 10^{19} \text{ GeV}$ contributing $\delta\rho_0 \simeq (M_{\text{p}}^2/4\pi)^2$, then this dominates the contributions from physics at lower energies. In this case the prediction for ρ_0 would differ from observations by

[†] In technical terms, once gravity is included the dimension-0 term is no longer independent of fields because general covariance requires it to depend on the spacetime metric.

[‡] Do not confuse ρ_0 in this section with the ρ parameter of the previous one!

some 122 orders of magnitude! However, we saw in Chapter 1 that bosons and fermions can cancel in their contributions to ρ_0 , and so it could happen that such a cancellation is at work for physics at scales beyond our current experimental reach. For instance, this cancellation might be expected if a new symmetry between bosons and fermions – supersymmetry – should emerge at these higher energies (more about which below).

However, a serious problem remains even if the physics above the electroweak scale happens to cancel in the vacuum energy. This is because we believe that we know the spectrum of particles having masses below the electroweak scale, and these do *not* cancel in their contributions, leading to $\delta\rho_0 \sim (m_t^2/4\pi)^2$. This is 54 orders of magnitude larger than the scale actually observed in nature. Apparently, the constant term in the Lagrangian must be chosen to cancel the zero-point fluctuation contribution (and other quantum contributions) to a precision of more than 50 decimal places, unless some other mechanism exists to balance off these energy densities.

At present there is no theoretical proposal which is generally accepted by the community for modifying the standard model in a way which may be able to explain the tiny size of the vacuum energy in a natural way.[†] What makes such proposals difficult to find is the requirement that they must already modify the predictions for the *electron's* contribution to the vacuum energy, since $\delta\rho_0 \sim (m_e^2/4\pi)^2$ is also much too large a contribution. But it is very difficult to modify our theories at such low energies without running into conflicts with other experiments. The cosmological-constant problem is arguably the most severe theoretical problem in high-energy physics today, as measured by both the difference between observations and theoretical predictions, and by the lack of convincing theoretical ideas which address it.

11.3 Dimension two: hierarchy problem

The lowest-dimension effective interaction possible which explicitly involves the standard model fields is the quadratic term in the Higgs potential

$$\mathcal{L}_2 = \mu^2 \phi^\dagger \phi \tag{11.4}$$

whose coupling we've seen is related to the Higgs boson mass by $m_H^2 = 2\mu^2$, and to the Higgs expectation value by $v^2 = \mu^2/\lambda$. Since $v = 246$ GeV is measured, these relations show that μ cannot be made large without also increasing λ to compensate. Since the validity of perturbation theory requires

[†] At the time of writing the “fine-tuning” option, in which the constant term precisely cancels against the contributions of virtual particles, is enjoying a minor revival, motivated by ideas from string theory.

$\lambda/16\pi^2 \lesssim 1$, we see control over our approximations keeps $\mu \lesssim 4\pi v = 3 \text{ TeV}$.

11.3.1 The hierarchy problem

We see that phenomenology requires the sole dimensionful parameter of the standard model to satisfy $\mu \lesssim 3 \text{ TeV}$. But we have also argued that virtual effects involving very heavy new particles having mass M generically contribute to μ an amount of order

$$\delta\mu^2 \sim \frac{gM^2}{16\pi^2} \quad (11.5)$$

where g is a dimensionless measure of the coupling strength between the heavy virtual particle and the Higgs doublet. The denominator is a typical one-loop factor, which we include since this is the order at which many heavy particles contribute. Strictly speaking (see Problem 11.2) explicit calculations give results which diverge quadratically in the ultraviolet, leading to expressions of the form $\delta\mu^2 = c_2\Lambda^2 + c_0M^2 \ln(\Lambda^2/M^2) + \dots$. However, the contribution to the effective coupling, $\delta\mu^2$, is defined to reproduce the difference between the theory with and without the massive particle, after renormalization, and this leads to an estimate of the form of Eq. (11.5).

Clearly, if all else is equal then the particle with the largest mass contributes the largest amount to $\delta\mu^2$. And the result is clearly much too large – for any reasonable value of coupling g – if M is as high as $M_p \sim 10^{19} \text{ GeV}$, or as 10^{14} GeV , as might be indicated by the dimension-5 description of neutrino masses. This enormous mismatch between the “observed” value of μ and its natural value once the standard model is embedded into a more fundamental theory at much higher energies is known as the *hierarchy problem*.

Although it is logically possible that the “bare” term, μ_0^2 , appearing in the Lagrangian precisely cancels these types of contributions for all such massive particles, the required cancellation is fantastically accurate for M extremely large – to within 33 decimal places if $M \sim M_p$. A cancellation this accurate is generally known as a *fine tuning*, and it is widely considered unlikely that such a cancellation is the explanation. Rather, it is generally believed that some new physics must be involved at an energy scale not much above the weak scale, which explains why the Higgs mass parameter does not depend sensitively on any very massive particles which may exist.

Support for this point of view can be found in other areas of physics, since this sort of problem is generic to any physical systems described by effective

theories which contain scalar fields. What happens in these other, better understood, cases? In some cases the scalar is very light, but the scalar field is either a Goldstone boson for a spontaneously-broken exact symmetry (e.g. phonons, magnons, spin density waves), or a pseudo-Goldstone boson for a spontaneously-broken approximate symmetry (such as the pion). Otherwise the scalar mass is generically near the heavy scale which defines the limits of validity of the effective description. Sometimes such a state can be made anomalously light through the explicit tuning of a control parameter of the system, such as by adjusting the temperature (or another variable) to be close to a critical point, often related to a phase transition. Examples of this sort occur within a superconductor close to the superconducting transition, or by tuning the temperature and pressure of a liquid-gas system to be close to the critical point at the end of the phase-coexistence curve.

11.3.2 What to do?

Once the fine-tuning alternative is set aside, what are the other options for the new physics just above the electroweak scale? There are several proposals for what the physics could be, and at the time of writing (2006) it is not clear which, if any, is correct. We now briefly summarize some of the main options, and while so doing dwell a little longer on a particular one (supersymmetry) than on the others.

The goal is to find modifications of physics at the scale $\Lambda \sim 4\pi m_H/g$ which can explain the absence of $O(M^2)$ contributions to the Higgs mass from virtual very massive particles. Broadly speaking there are three kinds of approaches:

The radical break: The most radical approach is to assert that the fundamental scale of physics really is the electroweak scale, and to regard the evidence presented above for higher scales to be misleading. The gravitational evidence is the trickiest to dispose of in this way, because gravity exists and the Planck mass, $M_P = G_N^{-1/2}$, is well measured.

There is nonetheless some room for manoeuvre, however. For instance, an estimate of the W boson mass using only the Fermi constant would have led to $M_W \sim G_F^{-1/2} \sim 300$ GeV, rather than its true value, $M_W = 80$ GeV. This difference is due to dimensionless factors, like gauge couplings, and one could hope for similar factors to make the “real” scale of gravity much smaller than the effective scale obtained on dimensional grounds from Newton’s constant. What is remarkable is that this can be possible in models for which there are more dimensions of space than the usual three, and where some of these dimensions are quite large, because for such models

the Planck mass is related to the scale of gravity in the higher dimensions, M_g , by $M_p = (M_g r)^{n/2} M_g$, where n is the number of extra dimensions and $V = r^n$ is their volume.

As of this writing it is not clear whether this idea successfully deals with all of the new naturalness issues which arise within this new context; however, the remarkable thing is that any proposals are possible at all along these lines. The proposal also requires relinquishing a seesaw description of neutrino masses, which might be done by introducing sterile neutrinos.

Composite models: A second possibility asserts that the Higgs doublet is really not an elementary field at all, but is a composite built from other more fundamental degrees of freedom. In this picture very massive particles do not make too large a contribution to μ^2 because in the new physics which intervenes at the electroweak scale there is no elementary Higgs for which there is a coupling to modify.

In many ways this is the historically conservative idea, for which the Higgs is qualitatively similar to the other scalar fields appearing in particle physics, i.e., the mesons (π , K , etc.) that are bound states in QCD. Models based on this idea generally involve additional gauge interactions, which form bound states in analogy with the mesons of QCD. The would-be Goldstone bosons which get eaten by the W and Z are then some of these mesons. Because of the analogy to QCD, the new gauge field is called *technicolor*, and models based on this idea are called *technicolor models*. In this case, the breaking of electroweak symmetry is closely analogous to the breaking of chiral symmetry in QCD.

Simple technicolor models generally predict a Higgs boson which is both very heavy and strongly self-coupled, often leading to effective theories for which the spontaneously-broken gauge symmetries are not linearly realized, along the lines of those considered in Subsection 7.5.1. As such, they can give comparatively large contributions to quantities like the oblique parameters S , T , and U , and their simpler variants are therefore often ruled out by precision electroweak measurements. More elaborate variants (walking technicolor, topcolor), which predict a light Higgs mass and small custodial symmetry breaking corrections, have been proposed, and some of these may be consistent with current experimental results. A proper discussion of these models goes well beyond the scope of this book.

Symmetry-driven cancellations: The third option differs from the previous two in that it grants that new physics is likely to exist at energies well above the electroweak scale, and that the Higgs is elementary and appears

in the as-yet-unknown, more fundamental theories which govern the physics of these higher scales. In this proposal the contributions of all of the high-energy particles nonetheless cancel in their contributions to μ^2 , up to terms of order m_{H}^2 , by virtue of a symmetry of the more fundamental theory.

The symmetry in question is called supersymmetry (often abbreviated SUSY), and it relates bosons to fermions. Relating bosons to fermions is useful because these contribute oppositely in sign to $\delta\mu^2$, and so can cancel if their masses and couplings are related to one another. Tantalizingly, this same symmetry also causes similar cancelations between bosons and fermions in their contributions to the cosmological constant, as was mentioned in passing in Subsection 11.2.1.

Although at the time of writing there is no direct experimental evidence in favor of supersymmetry, there are some fairly indirect hints in its favor, and it is at present the choice regarded as the most likely explanation of the hierarchy problem by a majority of the community. Therefore we give a slightly less brief account of this idea, although still one which just shallowly skims the surface.

11.3.3 Proposal: supersymmetry

Supersymmetry (SUSY) is a symmetry whose transformations relate particles having different spin to one another. It is not an *internal* symmetry, in the sense that it leaves the Lagrangian density, \mathcal{L} , invariant. Rather, an infinitesimal SUSY transformation changes the Lagrangian by a total derivative, $\delta\mathcal{L} = \partial_\mu V^\mu$ for some vector V^μ , and as a consequence leaves the action, $S = \int d^4x \mathcal{L}$ invariant. In this sense SUSY is more like the spacetime symmetries: rotations, translations, and boosts.

11.3.3.1 Supermultiplets

Like the other spacetime symmetries, supersymmetry acts differently on massless- and massive-particle states. Acting on a massless-particle state having helicity h , a SUSY transformation gives a new particle state having helicity $h \pm \frac{1}{2}$. Because it changes spin by half a unit, it follows that the supersymmetry charge must be a Weyl spinor, Q_a , with $a = \pm\frac{1}{2}$. If Q_a raises the helicity by $\frac{1}{2}$ then Q_a^* lowers it by the same amount. In general there could be more than one such supersymmetry charge, up to a maximum of $N = 8$ generators. But we restrict our attention to the case where $N = 1$, which turns out to offer the best prospects for particle physics.

The Q_a s satisfy the algebra $\{Q_a, Q_b\} = 0$ and so if a supersymmetry transformation is applied twice to any given state it gives zero. Because

of this, successive SUSY transformations do not continue to generate new states, and so massless *supermultiplets* come with pairs of helicities, (h_1, h_2) , with $h_2 = h_1 + \frac{1}{2}$. The ones of physical interest are

$$(h_1, h_2) = \left(0, \frac{1}{2}\right), \left(\frac{1}{2}, 1\right), \left(1, \frac{3}{2}\right) \quad \text{and} \quad \left(\frac{3}{2}, 2\right) \quad (11.6)$$

together with their counterparts of opposite helicity. For non-gravitational physics it is only the first two of these that are of interest.

In a field theory the particles must come together with their CPT conjugates, and since CPT reverses helicity the basic multiplets become:

$$\begin{array}{ll} \text{matter multiplets:} & \left(0, \frac{1}{2}\right) \oplus \left(-\frac{1}{2}, 0\right) \\ \text{and} \quad \text{gauge multiplets:} & \left(\frac{1}{2}, 1\right) \oplus \left(-1, -\frac{1}{2}\right) \end{array} \quad (11.7)$$

Given the particle content, it is clear that a matter multiplet can be represented by a complex scalar field and a Weyl (or Majorana) spinor field – φ, ψ_L say – while a gauge multiplet is represented by a real gauge field and a Weyl (or Majorana) spinor field – A_μ, λ_L . By construction, each multiplet contains equal numbers of bosons (integer spins) and fermions (half-odd-integer spins).

SUSY transformations commute with the Hamiltonian, $[H, Q_a] = 0$, and so the particle states related by these transformations have exactly the same mass, provided supersymmetry is not spontaneously broken. Spontaneous SUSY breaking occurs if the vacuum is not invariant, $Q_a|0\rangle \neq 0$. If the vacuum is not supersymmetric, then this is normally indicated by the existence of some field, F , having a nonzero expectation value, $\langle F \rangle \neq 0$. In this case the particles in a supermultiplet need no longer have the same mass, and their difference is typically of order $\Delta m^2 \sim g M_s^2$, where M_s is the scale set by $\langle F \rangle$, and g is a dimensionless measure of the strength of the coupling of the multiplet of interest to F .

11.3.3.2 The minimal supersymmetric standard model

The minimal field content of a supersymmetric version of the standard model is obtained by adding a *superpartner* for each of the standard model's particle types, leading to what is called the minimal supersymmetric standard model (MSSM). Assuming supersymmetry commutes with the gauge symmetries (as it must for $N = 1$ supersymmetry) the gauge transformation properties of these superpartners are precisely the same as for the original standard model particles.

For the gauge bosons this leads to adding spin-half *gauginos* for each of

the gauge generators, transforming in the adjoint representation of the gauge group: the *gluinos* transform under $SU_c(3) \times SU_L(2) \times U_Y(1)$ as $(\mathbf{8}, \mathbf{1}, 0)$; the *winos* transform as $(\mathbf{1}, \mathbf{3}, 0)$; and the *bino* transforms as $(\mathbf{1}, \mathbf{1}, 0)$.

Similarly, for each standard model fermion we must add a new collection of complex scalar superpartners, each of which has the same quantum numbers as the corresponding left-handed fermion. For each generation the complete list of these is: a doublet of *sleptons*, transforming as $(\mathbf{1}, \mathbf{2}, -1/2)$; singlet sleptons, transforming as $(\mathbf{1}, \mathbf{1}, 1)$; doublet *squarks*, transforming as $(\mathbf{3}, \mathbf{2}, 1/6)$; and up- and down-type singlet squarks transforming as $(\mathbf{\bar{3}}, \mathbf{1}, -2/3)$ and $(\mathbf{\bar{3}}, \mathbf{1}, 1/3)$.

The supersymmetric expansion of the Higgs sector is a bit more complicated, because the rules for supersymmetric interactions preclude generating a mass for all of the above particles using only a single Higgs doublet. Gauge anomalies would also not cancel if only a single Higgs scalar (and its superpartner) were included. Consequently the minimal Higgs sector requires two Higgs doublets – denoted H_U and H_D – both transforming as $(\mathbf{1}, \mathbf{2})$, but differing in their hypercharge assignments, with H_U having $y = \frac{1}{2}$ and H_D having $y = -\frac{1}{2}$; plus their superpartner *Higgsinos*, whose left-handed components transform in the same way.

The nomenclature for these particles is to label the spin-half superpartners of a standard model boson with the ending “-ino,” and the spinless superpartners of standard model fermions with the prefix “s-.” Symbolically, the superpartners are all denoted with the same letter as their standard model partner, overscored with a tilde (e.g., \tilde{W} , \tilde{L} , \tilde{U} , etc.).

11.3.3.3 Supersymmetric interactions

What restrictions does supersymmetry place on the interactions amongst these particles? Since only particles with spins zero, half, and one are present, the only renormalizable interactions possible are Yukawa couplings, gauge interactions, and a scalar potential which contains terms up to quartic in the fields.

Since the gauge interactions are dictated by the gauge transformation properties given above, they need not be spelled out here in detail. For the self-interactions of gauge bosons and gauginos these interactions are already supersymmetric. For matter–gauge interactions supersymmetry supplements the standard model gauge couplings

$$(t_a)_j^i [\bar{\psi}_i \gamma^\mu P_L \psi^j] A_\mu^a + \text{c.c.} \quad (11.8)$$

with new Yukawa interactions and new terms in the scalar potential. Here t_a is the appropriate gauge generator for left-handed fermions, which includes

a factor of the associated gauge coupling, g . The new Yukawa interaction is obtained from this by replacing two of these fields by their superpartners:

$$\sqrt{2}(t_a)_j^i [\bar{\psi}_i P_L \lambda^a] \varphi^j + \text{c.c.} \quad (11.9)$$

where φ^i is the scalar partner for ψ_L^i and λ^a is the appropriate gaugino.

The new term in the scalar potential which is related to the gauge couplings by supersymmetry is called (for historical reasons) a “ D -term,” and is given by

$$V_D = \frac{1}{2} \sum_a \left[\varphi_i^* (t_a)_j^i \varphi^j \right]^2 \quad (11.10)$$

In each of these expressions, the sum on i, j is over both the gauge index and over field types (for which t_a is diagonal). Notice that supersymmetry dictates that both of these new terms have the same normalization as the original gauge interaction, implying their strength is set by the corresponding gauge coupling.

The structure of the non-gauge, matter self-interactions are all given in terms of a function, W , called the *superpotential*. $W = W(\varphi)$ is a holomorphic function – that is, depending on φ but not on φ^* – of all of the complex scalars, φ^i , which appear within the matter supermultiplets. It must be gauge-invariant, and for renormalizable theories W must be at most a cubic polynomial of its arguments.

This function specifies two kinds of terms in the Lagrangian which are related to one another by supersymmetry: the Yukawa couplings involving only matter fields, and a set of scalar self-interactions. (For historical reasons the scalar potentials, V_F , describing these self-interactions are known as “ F -terms.”) These terms are given in terms of $W(\varphi)$ by

$$\begin{aligned} \mathcal{L}_{\text{yuk}} &= -\frac{1}{2} W_{ij} [\bar{\psi}^i P_L \psi^j] + \text{c.c.} \\ V_F &= \bar{W}^i W_i \end{aligned} \quad (11.11)$$

where $W_i = \partial W / \partial \varphi^i$, $\bar{W}^i = (W_i)^*$, and $W_{ij} = \partial^2 W / \partial \varphi^i \partial \varphi^j$. For example, a mass term for a single matter multiplet corresponds to the choice $W = \frac{1}{2} m \varphi^2$, in which case $\mathcal{L}_{\text{yuk}} = -\frac{1}{2} m [\bar{\psi} P_L \psi] + \text{c.c.}$ and $V_F = |m \varphi|^2$. Notice that supersymmetry in this case dictates that the scalars and fermions all share the same mass, $|m|$. The total scalar potential is the sum of the matter and gauge contributions: $V = V_F + V_D$.

The MSSM is defined (up to SUSY-breaking terms, to be discussed below) by its superpotential, and requiring the Yukawa couplings to include the

standard model mass terms leads to the following form,

$$W_{\text{MSSM}} = f_{mn} H_D L_m E_n + g_{mn} H_D Q_m D_n + h_{mn} H_U Q_m U_n + \mu H_U H_D \quad (11.12)$$

with f_{mn} , g_{mn} , h_{mn} being the dimensionless complex Yukawa matrices we have already seen, while μ is a complex parameter having dimensions of (mass), called the *mu parameter*. We see here why a second Higgs doublet is required; unlike for the standard model, we are not free to use the complex conjugate, H_D^* , to construct a mass term for the up-type quarks. We cannot do so because supersymmetry requires W to be a strictly holomorphic function, which cannot depend on the complex-conjugate fields.

Notice that Eq. (11.12) is *not* the most general gauge-invariant cubic superpotential which can be built using only the assumed MSSM matter content. W_{MSSM} excludes several kinds of terms which violate baryon number and lepton number, such as

$$\begin{aligned} W_{\text{BL}} = & \hat{\mu}_m \epsilon_{ij} L_m^i H_U^j + u_{mnp} \epsilon_{ij} Q_{\alpha m}^i L_n^j D_p^\alpha + v_{mnp} \epsilon_{ij} L_m^i L_n^j E_p \\ & + w_{mnp} \epsilon_{\alpha\beta\gamma} D_m^\alpha D_n^\beta U_p^\gamma \end{aligned} \quad (11.13)$$

Here we label the complex scalar field using the associated fermion, and all indices have been made explicit. $\alpha, \beta, \gamma = 1, 2, 3$ are color indices, $i, j = 1, 2$ are $SU_L(2)$ doublet indices and $m, n, p = 1, 2, 3$ are generation indices. The dimensionful coefficient $\hat{\mu}_m$ and the dimensionless coefficients u_{mnp} , v_{mnp} , and w_{mnp} are unconstrained by gauge invariance or supersymmetry. If all such terms are present in addition to those of Eq. (11.12) then baryon and lepton number cannot be conserved, leading to catastrophic predictions such as for rapid proton decay.

Recall that an attractive feature of the standard model is its economical explanation of why baryon number and lepton number seem to be conserved to a good approximation in experiments. This is understood in the standard model as being an accidental consequence of the assumed particle content and renormalizability, since all gauge-invariant renormalizable interactions automatically conserve both B and L . The existence of the baryon-number and lepton-number non-conserving terms in the superpotential, Eq. (11.13), shows that the MSSM does not have this same attractive feature. The reason the MSSM does not have this property can be traced to the existence of the superpartners, since the presence of these new fields allows us to construct renormalizable B - and L -violating interactions, such as Yukawa couplings between quarks and sleptons.

If we wish to exclude the possibility of including the terms in Eq. (11.13) on symmetry grounds, we must assume the existence of a symmetry in addition to gauge invariance. Of course the required symmetry could simply be $B-L$, since the anomalies for this could easily be arranged to cancel. (See, however, the discussion of Subsection 10.4.1 for potential drawbacks of having this as an exact symmetry.) A more minimal proposal which can forbid the terms in Eq. (11.13) is conservation of a multiplicative quantum number called R -parity. This can be defined to be $R = (-)^{F+3B-L}$, where F denotes fermion number and B and L are the usual lepton and baryon numbers (with the same B and L charges assigned to all of the particles within a supermultiplet). Since $F = 0$ for any term in a Lorentz-invariant Lagrangian, such a symmetry clearly forbids any interactions for which $3\Delta B \neq \Delta L \pmod{2}$. Although this precludes writing down renormalizable B - and L -violating interactions given the MSSM field content, it does not do so at a non-renormalizable level (as we shall see) and so can potentially be consistent with the existence of rare processes which don't conserve B and/or L .

We remark in passing that if R -parity is a symmetry, it has important implications for the phenomenology of the MSSM. It does so because its definition ensures that $R = +$ for all of the ordinary standard model particles (plus the new Higgs), but $R = -$ for all of the new superpartners. Among the important implications of this is the requirement that superpartners must be pair-produced in accelerators, since all initial-state particles are R even. Furthermore, the lightest superpartner (often called the LSP for short) must be stable in an R -parity conserving world, since there are no daughters which can carry off the initial particle's odd R -parity. LSP's must therefore be copiously generated in reactions which produce superpartners, and once produced they cannot decay. If the LSP is also electrically neutral and a color singlet (as if often true), then it would escape invisibly from a detector, and so would appear to experimenters simply as missing energy and momentum. Conservation of R -parity underlies the robustness of missing-energy signals for supersymmetry for a wide range of the underlying parameters.

A stable weakly-interacting LSP having a mass of order M_w could have important implications for cosmology. Since such particles cannot decay, a residual abundance would be left over from early, hot cosmological epochs. It turns out that an electroweak scale mass and electroweak strength interactions predict an abundance leading to an energy density comparable to the energy density of the universe today. Such an LSP therefore provides a very natural candidate to explain the *dark matter* which is actually observed to gravitationally dominate galaxies and clusters of galaxies, but whose makeup is otherwise completely unknown.

11.3.3.4 Supersymmetry breaking and naturalness

The MSSM as described to this point has two very attractive features: its prediction for the vacuum energy is precisely zero, and its prediction for the corrections to the dimensionful Higgs parameter μ also vanish. Both of these predictions vanish because of the precise cancellation which occurs between the contributions of the bosons and fermions within any given supermultiplet circulating inside the loop, along the lines of that discussed for the vacuum energy in Subsection 1.3.1 and Subsection 1.3.2. It is essential for this cancellation that the bosons and fermions within any supermultiplet share the same masses and have couplings which are related to one another as described above. Better yet, it can be shown that these cancellations apply to all orders in perturbation theory and are not artifacts of the one-loop calculation.

The bad news is that the MSSM as described to this point does not break supersymmetry spontaneously, and this is why all of the particles within a given supermultiplet have exactly the same masses, and couplings that are related to one another in the desired way. Unfortunately, this is not a good description of experiments, since, among other things, it would predict the existence of a spinless electron in addition to (and having the same mass as) the observed one. This would alter atomic physics, not to mention collider physics. In fact, the absence of evidence for such particles in e^+e^- collisions implies a lower limit on the selectron mass of 73 GeV (95% CL). Any successful phenomenology must supplement the above picture with some source of supersymmetry breaking.

Unfortunately, once SUSY is broken the cancellations in ρ_c and μ are no longer perfect, although the failure to cancel must vanish if the typical splitting in mass in a supermultiplet, $\Delta m^2 \sim M_s^2$, is taken to zero. This means that the resulting corrections to μ are typically $\delta\mu^2 \sim M_s^2$, even if the overall mass of the supermultiplet within the loop is $M \gg M_s$. As a result they can be small enough to be natural provided M_s is near enough to the electroweak scale. On the other hand, for M_s this large the partial nature of the cancellations lead to contributions to ρ_c which are at least as large as[†] $\delta\rho_c \sim M_s^4$. This is mixed news because it implies that breaking supersymmetry at the weak scale can provide a solution to the hierarchy problem (provided M_s is not too large), but is unable in itself to resolve the cosmological constant problem.

How does supersymmetry break? This is normally done by coupling the

[†] In many models the corrections to ρ_c can be larger, since they can be of order $\delta\rho_c \sim M_s^2 M^2$, where M is the scale of the heavy particles being integrated out.

MSSM as described above to some new degrees of freedom which are chosen so that their ground state breaks supersymmetry. This is not at all trivial to do, but there are a number of extant proposals for which it is possible to split the masses within a supermultiplet by an amount M_s which can easily be several hundred GeV.

Happily, although it goes beyond the scope of this book to show when and why, it turns out that it is often true that the implications of supersymmetry breaking for the MSSM particles can be expressed in terms of a collection of low-dimension effective interactions amongst the MSSM particles which explicitly break supersymmetry. These interactions are meant to capture the effects for the MSSM of the virtual effects of the particles in the supersymmetry-breaking sector, which are assumed themselves not to be directly observed. The resulting SUSY-breaking interactions are general enough to capture the implications of coupling the MSSM to a broad class of specific supersymmetry-breaking sectors. By parameterizing our ignorance about supersymmetry breaking in this way we can see how its effects can affect phenomenology without having to commit to a particular model.

There are three kinds of terms which arise as effective SUSY-breaking interactions: (i) generic scalar masses; (ii) generic gaugino masses; and (iii) holomorphic trilinear scalar self-interactions: $\mathcal{L}_{\text{SB}} = \mathcal{L}_s + \mathcal{L}_g + \mathcal{L}_3$, where

$$\begin{aligned}\mathcal{L}_s &= -(\mathcal{M}^2)_j^i \varphi_i^* \varphi^j \\ \mathcal{L}_g &= -\frac{1}{2} m_{ab} [\bar{\lambda}^a R_L \lambda^b] + \text{c.c.} \\ \mathcal{L}_3 &= \mathcal{W}(\varphi) + \text{c.c.}\end{aligned}\tag{11.14}$$

where φ^i generically represents all of the matter scalars, λ^a similarly denotes the gauginos, and the holomorphic function \mathcal{W} is an arbitrary cubic polynomial of the scalar fields consistent with the assumed symmetries. Notice that gauge invariance implies there are at most three independent gaugino masses, m_1 , m_2 , and m_3 , one for each factor of the standard-model gauge group. Similarly, the definition of the function \mathcal{W} is very similar to that of the superpotential, W , and so we expect to find in \mathcal{W} any term which is present in W . For instance, if R -parity is conserved, this means that \mathcal{W} has the same form as in Eq. (11.12), although coupling constants $-\hat{f}_{mn}$, \hat{g}_{mn} , \hat{h}_{mn} , and $\hat{\mu}$ – need not be related to the couplings $-f_{mn}$, g_{mn} , h_{mn} , and μ – which appear in W . (It is nevertheless common to assume these to be proportional to one another, taking $\hat{f}_{mn} = A_{mn} f_{mn}$ (no implied sum on m, n), etc., and for this reason these trilinear interactions are often called “ A -terms.”)

It is clear that supersymmetry breaking introduces a great number of new

effective couplings, and this considerably complicates the comparison of the MSSM with experiment. Notice however that it is generic that it is the superpartners and not the standard model particles which get masses when supersymmetry breaks, and so it would be natural to find that the superpartners should be much heavier (and so not yet discovered in experiments). The reason for this is that, besides the higgsinos (which receive a mass from the μ term in the superpotential even before SUSY breaking), these new particles are either fermions in the adjoint representation or scalars, and so for them gauge invariance allows mass terms. This is in contrast to the standard-model quarks, leptons, and gauge bosons, which cannot get a mass until the gauge group $SU_L(2) \times U_Y(1)$ spontaneously breaks, and so whose masses therefore remain suppressed by powers of the Higgs *v.e.v.*, v , and small Yukawa couplings, even after supersymmetry breaking.

11.3.3.5 Phenomenology of the MSSM

A complete discussion of the phenomenology of the MSSM would require a second volume to this book. We restrict ourselves here to a discussion of two aspects of the model. First we describe a particular sector of the MSSM – the Higgs sector – which we choose to illustrate some of the features which are generic to (broken) supersymmetric extensions of the standard model. Then we briefly describe some of the potential problems to which the generic parameterization of SUSY-breaking interactions can be prone. We make no attempt to be complete or to provide full details.

The Higgs sector

The first requirement for successful phenomenology is for the scalar potential to have minima which break the standard-model gauge group in the desired way, with only Higgs doublets getting expectation values and not squarks or sleptons (say). This can be more difficult than one might think, due to the large number of scalars present and because of the restrictions which supersymmetry puts on the form of its potential. It turns out that the soft supersymmetry-breaking terms can help to ensure the proper symmetry-breaking pattern arises, in an interesting way.

An important feature of the supersymmetry-breaking effective couplings is that they vary with scale, with all of the effective couplings running in a calculable way. This running can be important if the supersymmetry-breaking sector should involve very massive particles (as is often the case), since the models then predict the couplings at an ultraviolet scale and they must be run down to the electroweak scale in order to be used in observables

measured at these lower energies. But scalar self-couplings tend to induce negative corrections to the Higgs mass parameters (see Problem 11.2), and the Higgs doublets couple quite strongly to the squark fields due to the large values of the top and bottom Yukawa couplings which appear in the supersymmetric part, V_F , of the scalar potential. These couplings therefore tend to drive the H_U mass-squared to negative values, as is required to generate a nonzero $\langle H_U \rangle$. Within this picture the large top-quark mass makes it very natural that the Higgs fields are the only ones to have negative squared masses, while all other scalars are stabilized at zero field by positive squared-masses which can easily be larger than M_W^2 .

An important distinction between the Higgs sector of the MSSM and of the standard model is the requirement in the MSSM for two Higgs fields, H_U and H_D . Both are required to acquire nonzero vacuum expectation values, which we denote by v_u and v_d , in order to give masses to all fermions. These must satisfy $v_u^2 + v_d^2 = v^2 = (246 \text{ GeV})^2$ in order to properly reproduce the value of the gauge boson masses. However, a two-Higgs model can have a problem, because the generic symmetry-breaking pattern when two Higgs doublets acquire *v.e.v.s* is to completely break the $SU_L(2) \times U_Y(1)$ gauge symmetry, including the electromagnetic subgroup. To see why, choose a gauge such that the *v.e.v.* of H_D is in the lower component,

$$\langle H_D \rangle = \begin{pmatrix} 0 \\ v_d \end{pmatrix} \quad (11.15)$$

In this case, any *v.e.v.* in the lower component of H_U will break electromagnetism. The photon is only massless if H_U carries a *v.e.v.* only in its top component,

$$\langle H_U \rangle = \begin{pmatrix} v_u \\ 0 \end{pmatrix} \quad (11.16)$$

This kind of *vacuum alignment* is automatic in the standard model, because H_D and H_U are then replaced by a single field, ϕ , and its conjugate, $\tilde{\phi}$.

It happens that the MSSM does tend to generate the proper expectation values, and the reason for this is that its scalar potential is not that of a generic two-Higgs model. In particular, the quartic terms of the potential are quite restricted because they all arise only from the supersymmetric terms in the Lagrangian, and in particular the D -term potential provides an energetic bias in favor of the proper vacuum alignment. The relevant term comes from the contribution of the diagonal $SU_L(2)$ generator in the sum

appearing in V_D :

$$V_D = \frac{g_2^2}{2} \left(H_U^\dagger t_3 H_U + H_D^\dagger t_3 H_D \right)^2 + \cdots \quad (11.17)$$

where $t_3 = \text{Diag} \left(\frac{1}{2}, -\frac{1}{2} \right)$. This term prefers the fields H_U and H_D to align so that they take opposite eigenvalues for t_3 , as is required.

Another important difference between the standard model and a two-Higgs model like the MSSM concerns the size of the Yukawa couplings which can be inferred from the observed pattern of fermion masses. The inferred values of these couplings can differ because they depend on the relative size of the *v.e.v.s*, v_u and v_d , since the fermions only couple to either v_u or v_d and so $m_t \propto v_u$ while $m_b \propto v_d$. The size of the Yukawa matrices within the MSSM compared to the standard model therefore becomes

$$g_{mn}(\text{MSSM}) = g_{mn}(\text{SM})v/v_u, \quad h_{mn}(\text{MSSM}) = h_{mn}(\text{SM})v/v_d \quad (11.18)$$

In particular, for two-Higgs models it could happen that the largest eigenvalues of f_{mn} , g_{mn} , and h_{mn} are all comparable, with the large difference between the t mass and the b and τ masses being due to a large value for v_u/v_d .

It is conventional to write this ratio of *v.e.v.s* as the tangent of an angle,

$$\frac{v_u}{v_d} \equiv \tan \beta, \quad \frac{v_u}{v} = \cos \beta, \quad \frac{v_d}{v} = \sin \beta \quad (11.19)$$

Although β can be predicted in principle given the scalar potential, in the MSSM this prediction turns out to depend on the details of the poorly-known supersymmetry-breaking contributions. The size of β can be probed to a certain degree by comparing the MSSM radiative corrections with precision electroweak measurements. The best electroweak fits at present favor a rather large value: $\tan \beta \sim 5\text{--}40$, perhaps favoring this kind of explanation for the large ratio m_t/m_b .

The two Higgs fields of the MSSM contain eight real fields, of which three are “eaten” to become the longitudinal components of the W and Z bosons, as in the standard model. This leaves five physical scalars rather than the single one of the standard model. These scalars correspond to two charged and three neutral particles, usually written as H^\pm , h , H , and A . The convention is to denote by h and H the lighter and heavier neutral scalars, while A is a neutral pseudoscalar. It is possible to make more specific predictions for the masses of these particles in the MSSM than in a generic two-Higgs doublet model. This predictivity arises because the quartic interactions of the MSSM Higgs potential are completely determined by the supersymmetric

part of the theory, and so involve couplings (like gauge couplings) which can be measured elsewhere. It turns out that the masses for all of these scalars are determined in terms of three unknown mass parameters, subject to two constraints coming from the known value of v and the observational constraints on $\tan\beta$.

The pattern of masses which emerges from the analysis of the classical Higgs potential includes the relations

$$\begin{aligned} m_H^2 &= \frac{1}{2} [m_A^2 + M_Z^2 + \Delta] \\ m_h^2 &= \frac{1}{2} [m_A^2 + M_Z^2 - \Delta] \end{aligned} \quad (11.20)$$

where $\Delta = [(m_A^2 - M_Z^2)^2 + 4m_A^2 M_Z^2 \sin^2(2\beta)]^{1/2}$. This implies that the mass of the h particle is smaller than the lighter of the A and Z masses. If $\tan\beta$ is large (so β is near $\pi/2$) then we further expect m_h to equal either m_A or M_Z , whichever is lighter. Since current bounds indicate $m_A > M_Z$ this leads to the tree-level prediction that m_h is close to the Z boson mass. Radiative corrections to the scalar potential can be important, and tend to raise m_h compared with these predictions. They leave the general picture that these scalars should not be much beyond the reach of current accelerators.

Flavor and CP problems

We close this section with a brief discussion of a potentially serious difficulty with the MSSM. This difficulty is related to the Pandora's Box of parameters which enter into its physical predictions once the soft supersymmetry-breaking interactions are admitted. These parameters raise two separate kinds of potential problems for the MSSM.

Generally, below the mass scale of superpartners, the phenomenology of the MSSM is very similar to that of the standard model. This is because R -parity causes superpartners to appear an even number of times in any vertex, which precludes them from inducing tree-level effective operators; they can only appear in loops, and so their effects are suppressed by an extra factor of $O(\alpha/\pi)$ relative to the generic prediction of particles in their mass range. Therefore, for instance, the corrections to the ρ parameter and other electroweak observables owing to SUSY particles are generally within experimental bounds. However, there are some types of experimental results which are so precise that the added MSSM particles can conflict with experiment.

The first problem arises due to the complicated flavor structure which the

soft SUSY-breaking terms allow. These terms introduce potentially complicated flavor-changing interactions into the gauge and gaugino interactions of the scalars. But we have seen that these flavor-changing interactions are measured to be very small, with the success of the standard-model description of $K-\bar{K}$ mixing relying crucially on delicate GIM cancelations. However, the effective $\Delta S = \pm 2$ interactions which are generated by loops involving the flavor-changing scalar interactions generically do not similarly cancel, leading to predictions which are too large compared with the observed mixing. Agreement between theory and experiment requires special properties, such as generation diagonal or generation blind mass matrices.

The same is true for CP-violation in the MSSM, because there are an enormous number of phases which are possible within the many complex parameters of the soft SUSY-breaking terms. These generically give unacceptably large effects for CP-violation in B mesons and nuclear dipole moments unless these phases are chosen to be small, or the scalar masses have special features.

Although it is difficult to draw a definitive conclusion without knowing the kind of soft-breaking terms which are produced by the SUSY-breaking sector, any successful SUSY-breaking proposal must explain why these CP-violating and flavor-changing effects are acceptably small.

11.4 Dimension four: triviality, θ_{QCD} , flavor problems

All of the interactions within the standard model apart from the two just described – dimensions 0 and 2 – have dimension-4, and so have dimensionless couplings. On one hand they can be sensitive to ultraviolet scales and so it is useful to look for cases where phenomenology requires us to make unusual choices for these couplings. On the other hand, the ultraviolet sensitivity is only logarithmic, and so it is likely that only contributions from physics at extremely large scales can appear in an unnatural way. We describe here three kinds of puzzles which are driven by the observed pattern of dimension-4 interactions in the standard model.

11.4.1 Triviality

We saw in Chapter 7 that the electromagnetic and strong couplings vary with scale. As the energy scale considered becomes larger, the strong coupling becomes weaker, see Eq. (7.42), a property called asymptotic freedom. The electromagnetic coupling, on the contrary, becomes larger as we explore more ultraviolet scales, see Eq. (7.41).

There is a potential problem with a theory where the coupling grows when it is run up to higher energies. It is possible that the coupling could become infinite at some finite energy scale, Λ . Indeed, if trusted, Eq. (7.41) would indicate that α , considered as a function of $\ln \mu$, diverges as μ approaches an ultraviolet scale Λ_L , called the Landau pole;

$$\begin{aligned} \frac{1}{\alpha(\mu)} &= \frac{1}{\alpha(\mu')} + b_{\text{em}} \log \left(\frac{\mu'^2}{\mu^2} \right) \quad \text{implies} \\ \frac{1}{\alpha(\Lambda_L)} &= 0 \quad \text{for} \quad \Lambda_L = \mu \exp \left(\frac{1}{2b_{\text{em}}\alpha(\mu)} \right) \end{aligned} \quad (11.21)$$

with $\Lambda_L > \mu$ because $b_{\text{em}} > 0$.

Of course, just because perturbation theory may indicate that a coupling diverges in this way doesn't mean that it really does diverge, because the coupling becomes large before becoming infinite and at some point perturbative tools are no longer reliable for studying its evolution. However, for scalar field theories and abelian gauge theories there are numerical, strong-coupling calculations which give sound reasons for believing that these couplings really do diverge at approximately the scale implied by a perturbative calculation like the one above.

Theories for which couplings diverge in this way are often referred to as “trivial” – hence the name *triviality problem* – because the only way they can make sense, unmodified, to arbitrarily high energy scales is if the renormalized coupling is zero (hence, a trivial theory). But the nature of the problem is different within the point of view that the theory of interest is an effective theory which is only meant to describe the low-energy limit of some more fundamental, as-yet-undiscovered, higher-energy physics. Since a low-energy theory with a triviality problem cannot make sense up to arbitrarily high energies, it must be replaced at higher energies by an *ultraviolet completion*: that is, some other kind of theory involving new degrees of freedom and interactions, whose presence changes the physics in some way which prevents the coupling of interest from diverging. (Alternatively, the coupling might not be defined at all, such as would happen if the particles whose couplings it describes were bound states which don't appear in the new theory of their constituents which supersedes at higher energies). From this point of view triviality problems are interesting because they provide a clue as to what goes on at higher energies, by indicating an upper limit on the energy where this must be.

There are similarities and differences between this situation and the apparent divergence of an asymptotically-free coupling, like that of QCD. In both cases the development of a strong coupling at a particular scale points

to the appearance there of interesting, possibly strongly-interacting, physics. In both cases it can happen that the coupling of interest does not even appear in the new effective theory that intervenes at this scale. The main difference is that for theories like QCD this new theory intervenes at low energies, and so the degrees of freedom it involves have their roots in the particles of the asymptotically-free microscopic theory which underlies it (e.g. the hadrons are built as bound states out of the quarks and gluons of QCD, with a known dynamics). Much less can be said for couplings which diverge in the ultraviolet, since so much less is known about the new degrees of freedom which might arise there, apart from the scale above which they must appear.

Within the standard model we can ask the perturbative part of the problem: is there a scale above which the perturbative running of the dimensionless couplings drives us beyond the domain of perturbation theory? We can already see from Eq. (7.41) that this does occur for QED, although for QED we know that the standard model itself intervenes well before the Landau pole is reached. It also occurs in the standard model itself, however, as we now see by computing the relevant running explicitly. There are three kinds of dimensionless couplings whose evolution must be followed: the gauge couplings, the Yukawa couplings, and the dimensionless Higgs self-coupling. The gauge couplings run according to the simple equations discussed in Chapter 7,

$$16\pi^2\mu^2\frac{dg_1}{d\mu^2} = \frac{41}{12}g_1^3, \quad 16\pi^2\mu^2\frac{dg_2}{d\mu^2} = -\frac{19}{12}g_2^3, \quad \text{and} \quad 16\pi^2\mu^2\frac{dg_3}{d\mu^2} = -\frac{7}{2}g_3^2 \quad (11.22)$$

The running of the Yukawa couplings is described by the more complicated relations

$$16\pi^2\mu^2\frac{df_{mn}}{d\mu^2} = \frac{3}{4}f_{mp}f_{qp}^*f_{qn} + \frac{1}{2}f_{mn}\left(Y^2 - \frac{9}{4}g_1^2 - \frac{9}{4}g_2^2\right) \quad (11.23)$$

$$16\pi^2\mu^2\frac{dg_{mn}}{d\mu^2} = \frac{3}{4}g_{mp}g_{qp}^*g_{qn} - \frac{3}{4}g_{mp}h_{qp}^*h_{qn} + \frac{1}{2}g_{mn}\left(Y^2 - \frac{17}{12}g_1^2 - \frac{9}{4}g_2^2 - 8g_3^2\right) \quad (11.24)$$

$$16\pi^2\mu^2\frac{dh_{mn}}{d\mu^2} = \frac{3}{4}h_{mp}h_{qp}^*h_{qn} - \frac{3}{4}h_{mp}g_{qp}^*g_{qn} + \frac{1}{2}h_{mn}\left(Y^2 - \frac{5}{12}g_1^2 - \frac{9}{4}g_2^2 - 8g_3^2\right) \\ \text{where } Y^2 \equiv 3h_{pq}^*h_{pq} + 3g_{pq}^*g_{pq} + f_{pq}^*f_{pq} \quad (11.25)$$

and the Higgs self-coupling satisfies

$$\begin{aligned}
 16\pi^2\mu^2 \frac{d\lambda}{d\mu^2} = & \left(-\frac{9g_2^2}{2} - \frac{3g_1^2}{2} + 6h_{mn}h_{mn}^* + 6g_{mn}g_{mn}^* + 2f_{mn}f_{mn}^* \right) \lambda \\
 & + 12\lambda^2 + \frac{9}{4}g_2^2 + \frac{9}{8}(g_2^2 + g_1^2)^2 - 3h_{mn}h_{pn}^*h_{pq}h_{mq}^* \\
 & - 3g_{mn}g_{pn}^*g_{pq}g_{mq}^* - f_{mn}f_{pn}^*f_{pq}f_{mq}^* \quad (11.26)
 \end{aligned}$$

Although we write here the complete matrix structure for the Yukawa coupling matrices f_{mn} , g_{mn} , and h_{mn} , in practice one can neglect all Yukawa couplings except those appearing in m_t , m_b , and m_τ , and if this is done then each can be considered as a single parameter.

No general closed-form analytic solution is known for these equations, but they are relatively straightforward to evolve numerically, and we can say a good deal about their behavior even without doing this. First, we see that g_2 and g_3 are asymptotically free. There is no danger that they will grow at ultraviolet scales. Also, g_1 is fairly small, and the lepton and down-type quark Yukawa couplings, f_{mn} and h_{mn} , are tiny; since $dg_1/d\ln\mu \propto g_1^3$ and $df_{mn}/d\ln\mu \propto f_{mn}$ and similarly for h_{mn} , this prevents these couplings from diverging unless we consider extremely ultraviolet scales, $\mu \gg M_p \approx 10^{19}$ GeV. But we expect new physics to intervene in any case before these scales because of the non-renormalizability of the gravitational interactions. Therefore, we should not be concerned about the UV behavior of these couplings. Since the top-quark Yukawa coupling, g_{33} , is quite large, it is not obvious that the same holds for it; but the negative $g_3^2 g_{mn}$ term turns out to hold this coupling in check, so g_{33} does not grow divergently before the gravitational scale. Provided new physics enters somewhere below the gravitational scale, triviality has no implications for these couplings.

The potentially problematic coupling is the Higgs scalar self-coupling, λ . The numerical value of this coupling, which is related to the Higgs mass, is not known; indeed it is not even known experimentally whether a single Higgs field is the correct description of how electroweak symmetry-breaking occurs in nature. We nevertheless press on and assume that it is, in order to find what we can learn about the Higgs mass from the assumption that the standard model provides a good description well into the ultraviolet.

The running of λ introduces two potential dangers. Besides checking to see if λ gets driven to be large at high energies – the triviality problem – we must also check to see if it changes sign, since $\lambda > 0$ is crucial for obtaining a Higgs *v.e.v.* $v \sim \mu/\lambda$. Negative λ instead indicates that *both* the $\phi^\dagger\phi$ and $(\phi^\dagger\phi)^2$ terms in the Higgs potential prefer to drive $\phi^\dagger\phi$ to large values. (For a quartic potential positive λ also is required to keep the energy bounded

from below, but we do not know that the potential need be strictly quartic at large fields once loop corrections are large enough to change the sign of λ .) If λ becomes negative at any renormalization scale $\mu = M$, then we expect that the true vacuum has a Higgs expectation value, v , of order the largest scale where we reliably know $\lambda < 0$, which would imply the unacceptable result $v \gg 246$ GeV.

Both of these possible dangers arise in practice. Indeed, if λ is small then the presence of the negative term proportional to $-g_{mn}^4$ in the standard-model running can cause λ to become negative. Since this occurs within the weak-coupling regime, its description using the equations governing perturbative running should remain valid. Numerical study of the renormalization group equations show that, if we require the standard model to be valid right up to the scale of $\mu = M_p \sim 10^{19}$ GeV, then the Higgs boson mass must satisfy $m_H > 130$ GeV to prevent λ from becoming negative at some ultraviolet scale below M_p . This constraint weakens if new physics is allowed to intervene at scales lower than M_p .

If, on the other hand, λ is large, then the dominant term in its evolution is the λ^2 term, which causes it to grow with scale. This can lead to a runaway behavior in which λ leaves the domain of perturbation theory at some finite renormalization scale, Λ_L . If we neglect the other terms in favor of the λ^2 term, we find

$$\mu^2 \frac{d\lambda}{d\mu^2} = \frac{12\lambda^2}{16\pi^2} \quad \Rightarrow \quad \frac{1}{\lambda(\mu)} = \frac{3}{4\pi^2} \log \left(\frac{\Lambda_L^2}{\mu^2} \right) \quad (11.27)$$

which shows that perturbation theory indicates that new physics must intervene at energies lower than $\Lambda_L \sim \mu \exp[2\pi^2/3\lambda(\mu)]$. While this result is based on a perturbative treatment which breaks down when λ is large, because only λ is involved the large- λ regime should be well-described by non-perturbative (lattice) studies of purely scalar field theories, and these indicate that the coupling really does diverge. If we therefore demand that the standard model be a complete theory all the way up to $M_p \sim 10^{19}$ GeV, so that $\Lambda_L > M_p$, the Higgs mass must have an *upper* bound $m_H \lesssim 180$ GeV. Again, this constraint weakens if the standard model is only required to be a complete theory up to some less ultraviolet scale.

In summary, the standard model's triviality problem brings both good news and bad news. The good news is that it shows that the standard model *cannot* be a complete theory at all scales, because there is then no value of the Higgs self-coupling for which λ is always finite and positive. The bad news is that the scale above which triviality implies it must fail is exponentially large compared with the electroweak scale, and it could be

as large as M_p , before which we know in any case that something has to happen because of the non-renormalizability of gravity. How much this tells us depends on how heavy the Higgs proves to be, once it is discovered. If $130 \text{ GeV} < m_H < 180 \text{ GeV}$ it is logically possible that the first new scale which is encountered is the lepton-number-violating scale to which neutrino masses appear to point. However, the larger m_H is, the lower is the triviality scale, with eventually m_H meeting Λ_L slightly below 1 TeV. If the Higgs is sufficiently heavy, then new physics is likely nearby.

11.4.2 The strong CP problem*

Every renormalizable interaction which could appear in the standard model is actually present with a nonzero coefficient, except for one; the parameter Θ_3 , introduced in Eq. (2.13). Present constraints on the size of the neutron electric dipole moment imply that it must be smaller than $\sim 10^{-9}$. This remarkably small magnitude ensures that low-energy QCD is approximately P and CP conserving; the puzzle of why the value should be so small is called the *strong CP problem*.

One might ask: what is wrong with simply choosing $\Theta_3 = 0$? Within the standard model this choice must be made at a particular scale, for instance somewhere above the Z - or t -quark mass. Since the weak interactions also break C and CP, Θ_3 will renormalize (or run) due to loops involving standard model particles, and so will be nonzero at the low energies where neutron dipole moments are measured. Within the standard model the resulting nonzero value is less than 10^{-15} , however, and so is too small to be detected in existing experiments. This shows that it *is* consistent simply to choose the value of Θ_3 to be small within the standard model, and that phenomenologically-small values for Θ_3 do not pose a naturalness problem for the standard model itself.

Such a small value for Θ_3 is nevertheless rather surprising. Furthermore, the small size of Θ_3 often does become a naturalness problem once the standard model is embedded into a more microscopic theory, such as within the supersymmetric extensions discussed above. For this reason we describe in this section a fairly mild modification of the standard model for which the small size of Θ_3 could be made to “relax” to zero. It can do so because in the modified theory Θ_3 is related to the expectation value for an almost-massless new scalar field (called the *axion*), for which the dynamics favors configurations with $\Theta_3 = 0$. Although the particle associated with this axion field is very light, it could have escaped detection so far because it couples

to the other standard model fields only through small non-renormalizable interactions.

Another reason we examine the strong CP problem is that its discussion provides a vehicle for describing some important features of the QCD ground state, which are interesting in their own right. However this discussion involves more advanced techniques in quantum field theory than does the rest of this book, and for this reason this section is somewhat technical and can be skipped on a first reading. Limitations of space also only permit us to sketch some of the main ideas without giving complete details.

11.4.2.1 The QCD vacuum angle

The physical implications of a term in the Lagrangian of form

$$\mathcal{L}_\Theta = \Theta_3 \frac{g_3^2}{64\pi^2} \epsilon^{\mu\nu\lambda\beta} G_{\mu\nu}^\alpha G_{\lambda\beta}^\alpha \quad (11.28)$$

are not immediately obvious, and require some discussion. The first puzzle is that the quantity Θ_3 multiplies is in fact a total derivative, which one might therefore expect not to matter at all. Explicitly, one may write

$$\begin{aligned} \frac{g_3^2}{64\pi^2} \epsilon^{\mu\nu\lambda\beta} G_{\mu\nu}^\alpha G_{\lambda\beta}^\alpha &= \partial_\mu K^\mu \\ K^\mu &\equiv \frac{g_3^2}{32\pi^2} \epsilon^{\mu\nu\lambda\beta} \left(G_\nu^\alpha G_{\lambda\beta}^\alpha - \frac{g}{3} f_{\alpha\beta\gamma} G_\nu^\alpha G_\alpha^\beta G_\beta^\gamma \right) \end{aligned} \quad (11.29)$$

Because of this property, the spacetime integral of the term multiplying Θ_3 is determined by the change in the charge $\int d^3x K^0 \equiv N_{\text{CS}}$ between initial and final surfaces;

$$\begin{aligned} \frac{g_3^2}{64\pi^2} \int d^3x \int_{t_i}^{t_f} dt \epsilon^{\mu\nu\lambda\beta} G_{\mu\nu}^\alpha G_{\lambda\beta}^\alpha &= \int d^3x K^0(x, t_f) - \int d^3x K^0(x, t_i) \\ &\equiv N_{\text{CS}}(t_f) - N_{\text{CS}}(t_i) \end{aligned} \quad (11.30)$$

where we assume boundary conditions which ensure the vanishing of contributions from spatial infinity. The quantity N_{CS} is called the *Chern–Simons number* of the gauge-field configuration.

Even though Θ_3 multiplies a total derivative, it can have physical implications, because the change in N_{CS} need not be zero in a vacuum-to-vacuum process. This is because the QCD vacuum is *topologically non-trivial*; there exist gauge-field configurations for which $G_{\mu\nu}^\alpha = 0$, and yet $N_{\text{CS}} \neq 0$. The

easiest way to see this is simply to present an example. The field

$$A_i^a(x) = \frac{2}{g_3} \left(\frac{d\theta}{dr} \hat{x}_i \hat{x}_a + \frac{\sin 2\theta}{2r} (\delta_{ia} - \hat{x}_i \hat{x}_a) - \frac{\sin^2 \theta}{r} \epsilon_{iaj} \hat{x}_j \right), \quad \theta \equiv \frac{\pi \rho^2}{\rho^2 + r^2} \quad (11.31)$$

(with $A_i^{a=4,\dots,8} = 0$ and ρ some length scale) satisfies $N_{\text{CS}} = 1$ even though $G_{\mu\nu}^\alpha = 0$. Since the classical Hamiltonian vanishes when evaluated at configurations having vanishing field strength, such fields are associated with the vacuum semiclassically. Topology enters by ensuring that N_{CS} does not take arbitrary values; there is a topological theorem which shows that N_{CS} is an integer whenever it is evaluated for a configuration for which $G_{\mu\nu}^\alpha(x) = 0$ everywhere on a fixed time surface. A 4-dimensional gauge-field configuration, $A_\mu^\alpha(x, t)$, with the property that $N_{\text{CS}}(t \gg 0) - N_{\text{CS}}(t \ll 0) = 1$, is called an *instanton*.

These observations have the following physical implication. The existence of the charge N_{CS} implies that the QCD ground state is not unique, but instead consists of a collection of degenerate states which are distinguished by their eigenvalue for N_{CS} . The significance of \mathcal{L}_Θ is that it assigns a phase to any transition between two vacuum states having different values for N_{CS} . Namely, if $|a\rangle$ has eigenvalue $N_{\text{CS}}(a)$ and $|b\rangle$ has eigenvalue $N_{\text{CS}}(b)$, then

$$\langle b | e^{-\beta H} | a \rangle = \langle b | e^{-\beta H} | a \rangle_{\Theta_3=0} \exp [i\Theta_3(N_{\text{CS}}(b) - N_{\text{CS}}(a))] \quad (11.32)$$

It is clear from this that Θ_3 is only defined modulo 2π , since a shift by this amount has no physical significance. Furthermore, this phase shares the symmetry properties of N_{CS} , and because $\epsilon^{\mu\nu\lambda\beta}$ is P and T odd, so is N_{CS} . N_{CS} is C even because $G_{\mu\nu}^\alpha$ appears within it an even number of times. The presence of \mathcal{L}_Θ associates a phase, Θ_3 , to the QCD ground state, and the physics of this phase breaks P and CP but not C (or CPT).

For later purposes we note in passing that, strictly speaking, the existence of the nonzero matrix element, Eq. (11.32), lifts the degeneracy between the states $|a\rangle$ and $|b\rangle$, leading to a ground state which is instead the following linear combination of these states (up to overall normalization)

$$|0_\theta\rangle = \sum_n \exp(i\theta n) |n\rangle \quad (11.33)$$

where $N_{\text{CS}}|n\rangle = n|n\rangle$. Such a vacuum is called a *theta vacuum*, and has an energy which depends on θ . Notice that we can absorb the parameter θ into a shift in Θ_3 , in which case the vacuum becomes $|0\rangle = \sum_n |n\rangle$. We implicitly make this shift in what follows.

11.4.2.2 The axial anomaly (again)

When fermions (like quarks) are present Θ_3 does not, in general, give the only phase associated with a $\Delta N_{\text{CS}} \neq 0$ transition. There can also be phases associated with the way fermions couple to the instanton which mediates the transition. This fermionic phase is associated with the anomaly in the approximate axial symmetries of QCD, which were first encountered in Subsection 2.5.3.

As we saw in Section 2.5, and discussed in much more detail in Section 8.3, the QCD Lagrangian has approximate axial symmetries under which the left- and right-handed components of the light quark fields are rotated in opposite directions. For instance, for the up quark, the Lagrangian is almost symmetric on the rotation, $u \rightarrow \exp(i\gamma_5\theta_u)u$. The associated axial current

$$J_{u5}^\mu \equiv i\bar{u}\gamma^\mu\gamma_5u \quad (11.34)$$

would be exactly conserved were it not for two things: the small up-quark mass, which breaks this symmetry, and the axial anomaly. The current's non-trivial gauge anomaly is $2\text{tr}(T_5\{T_a, T_b\}) = \delta_{ab}$, and so if we ignore the quark mass Eq. (2.126) implies the axial current is not conserved by the amount

$$\partial_\mu J_{u5}^\mu = \frac{g_3^2}{32\pi^2} \epsilon^{\mu\nu\lambda\beta} G_{\mu\nu}^\alpha G_{\lambda\beta}^\alpha \quad (11.35)$$

The relevance of these facts for the present discussion is revealed once both sides of this last expression are integrated over space and time, since this shows that the axial charge, $Q_{u5} \equiv \int d^3x J_{u5}^0$, changes by an amount

$$Q_{u5}(t_f) - Q_{u5}(t_i) = 2[N_{\text{CS}}(t_f) - N_{\text{CS}}(t_i)] \quad (11.36)$$

This shows that an instanton process also changes the axial-quark number by two units (for massless quarks). By considering separately the contribution from the right- and left-handed components of the current, $i\bar{u}\gamma^\mu P_L u$ and $-i\bar{u}\gamma^\mu P_R u$, one can show that this arises through the creation of a left-handed particle (or destruction of its antiparticle) and the destruction of a right-handed particle (or creation of its antiparticle).

The above reasoning shows that there would be no vacuum-to-vacuum instanton processes if any of the quark masses vanished, because it proves that these processes would instead produce a quark and an antiquark, and so would therefore not take the vacuum into the vacuum. For massive quarks, the mass term in the Lagrangian allows these particles to annihilate one other, so the instanton process remains vacuum-to-vacuum after all.

Therefore, when all quarks are massive, instanton processes can relate

ground states. However, if the quark masses in the Lagrangian have nonzero phases (if m_q in the Lagrangian is not a real parameter), this will introduce an extra phase into the amplitude of Eq. (11.32). The full phase associated with an instanton process is

$$\tilde{\Theta}_3 \equiv \Theta_3 + \sum_q \text{Arg } m_q = \Theta_3 + \text{Arg Det}[g h] \quad (11.37)$$

with \sum_q running over all quark flavors, and g and h denoting the Yukawa-coupling matrices g_{mn} , h_{mn} of Eq. (2.25).

This expression requires some explanation because to this point we have generally chosen the mass parameters for all quarks to be real, in which case $\text{Arg } m_q = 0$. If a quark-mass parameter in the Lagrangian were not initially real, we were always free to make it real (and non-negative) by performing a field redefinition under which the relevant quark field was rotated by an axial transformation. Within the standard model the only term which was changed by this rotation was the charged-current weak interaction, and as a consequence this redefinition pushed any phases into the CKM matrix. However, since this argument proceeds at the classical level, it ignores the axial anomaly. Once the axial anomaly is taken into account, it turns out that the redefinition which removes phases from the quark-mass matrix also shifts the Θ_3 term in the QCD Lagrangian, in such a way that the quantity $\tilde{\Theta}_3$ remains unchanged. Because of this, only the quantity $\tilde{\Theta}_3$, not Θ_3 , is independent of field redefinitions, and so is a physical quantity which determines whether P and CP are violated in QCD. When $\tilde{\Theta}_3 = 0$, CP is only violated by the phase in the CKM matrix, and so the strong CP problem really asks why $\tilde{\Theta}_3$ is small, not Θ_3 .

11.4.2.3 Neutron electric-dipole moment

It is a complicated problem to compute how a nonzero value for $\tilde{\Theta}_3$ causes CP violation in physical observables. We restrict ourselves here to summarizing more detailed calculations of how $\tilde{\Theta}_3$ contributes to the neutron electric-dipole moment, since this furnishes the strongest constraint.

Section 9.1 argued that the low-energy behavior of hadrons can be described by an effective theory involving hadron fields rather than quarks. If CP is not a good symmetry, this Lagrangian can have a term

$$\mathcal{L}_{\text{edm}} = -\frac{id_n}{8} \epsilon_{\mu\nu\alpha\beta} F^{\mu\nu} \bar{n} [\gamma^\alpha, \gamma^\beta] n \quad (11.38)$$

with $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ being the usual electromagnetic field strength. This term resembles the magnetic-dipole term of Eq. (9.2), except for the appearance of the antisymmetric tensor $\epsilon_{\mu\nu\alpha\beta}$. Following the reasoning which

leads to Eq. (9.5) now gives an interaction with an slowly-varying background electromagnetic field of the form

$$H_d = -d_n(\bar{n}\sigma n) \cdot \mathbf{E} \quad (11.39)$$

which describes an *electric* dipole moment for the neutron.

A rough estimate for the size implied for d_n when $\tilde{\Theta}_3 \neq 0$ can be found on dimensional grounds, giving $d_n \sim e\Lambda_{\text{QCD}}^{-1}\tilde{\Theta}_3$. In fact this turns out to be an overestimate, since we saw that the physical implications of the angle $\tilde{\Theta}_3$ can be removed by a harmless field redefinition if any quarks become massless. This implies that d_n must vanish if either $m_u = 0$ or $m_d = 0$. Since both are smaller than the scale Λ_{QCD} , a better estimate would be

$$\delta_n \sim \frac{em_um_d}{(m_u + m_d)\Lambda_{\text{QCD}}^2}\tilde{\Theta}_3 \quad (11.40)$$

This is to be compared with the experimental value

$$|\delta_n| < 6 \times 10^{-26} e \text{ cm} \quad (11.41)$$

leading to the constraint $\tilde{\Theta}_3 < 5 \times 10^{-10}$. More careful treatments give approximately the same answer. The neutron electric dipole moment currently provides the strongest constraint on $\tilde{\Theta}_3$.

Is it reasonable to expect the combination $\tilde{\Theta}_3$ to be exactly zero? We know relatively little about the origin of the Yukawa matrices g_{mn} , h_{mn} ; from the point of view of the standard model, they are parameters to be determined from experiment. We do know that these matrices are in general complex, and when we diagonalize them as in Subsection 2.3.3 and Eq (2.90), the resulting CKM matrix is complex, with a phase δ_{13} of Eq. (2.90) which is not small. This shows there is no evidence that the Yukawa matrices should not be complex, with entries having $O(1)$ phases. We therefore expect that, even if for some reason $\Theta_3 = 0$, the phases in the Yukawa matrices should lead to a $\tilde{\Theta}_3$ which is $O(1)$. Although within the standard model it is technically natural to have $\tilde{\Theta}_3$ small, it would be preferable if there were a physical mechanism which makes it so.

11.4.3 Proposal: axions

The axion proposal posits that a new real scalar field a is added to the standard model. This field is assumed to be a Goldstone boson for an approximate axial symmetry, carried by at least one of the quarks. If this symmetry were to be exact, the axion would therefore couple exclusively through derivatives, and so possess a shift symmetry, $a \rightarrow a + \epsilon$, for constant

parameter ϵ . It is assumed that this symmetry is broken by standard-model anomalies, which make this shift symmetry only approximate.

With these assumptions, the allowed Lagrangian terms are

$$\begin{aligned} \mathcal{L}_a = & -\frac{1}{2}\partial^\mu a \partial_\mu a + J^\mu \partial_\mu a - \frac{a}{64\pi^2 f_a} \left[g_3^2 \epsilon^{\mu\nu\alpha\beta} G_{\mu\nu}^b G_{\alpha\beta}^b \right. \\ & \left. - k_2 g_2^2 \epsilon^{\mu\nu\alpha\beta} W_{\mu\nu}^b W_{\alpha\beta}^b - k_1 g_1^2 \epsilon^{\mu\nu\alpha\beta} B_{\mu\nu} B_{\alpha\beta} \right] \end{aligned} \quad (11.42)$$

Here f_a is an unknown constant with dimensions of mass, and

$$f_a J^\mu = i\bar{\psi}_{i,m} \gamma^\mu (X_{imn} + iY_{imn} \gamma_5) \psi_{i,n} + X_\phi (\phi^\dagger iD^\mu \phi + \text{h.c.}) \quad (11.43)$$

denotes the part of the Noether current for the axion symmetry which depends on standard-model fields. The dimensionless coefficients X_{imn} , Y_{imn} , and X_ϕ are model-dependent and depend on how the symmetry acts on the various fields. The index i runs over $[Q, U, D, E, L]$, and m and n are generation indices. As written, the coefficients X_{imn} are antisymmetric and Y_{imn} are symmetric under $m \leftrightarrow n$. The constants k_1 and k_2 express the $SU_L(2) \times U_Y(1)$ part of the anomaly for the axion symmetry, and the corresponding quantity for QCD has been assumed to be nonzero and absorbed into the definition of f_a .

Notice that all of the interaction terms written are dimension-5, and lower-dimension interactions are forbidden by the assumed properties of the axion shift symmetry. The quantity f_a parameterizes the scale at which this symmetry is spontaneously broken (as must occur in order to have a Goldstone boson) and this scale is assumed to be large compared with standard model scales. Although the specific form for k_1 , k_2 , X_{imn} , Y_{imn} , and X_ϕ , may be important in studying the phenomenology of the production and detection of axions, our interest here is in the interaction term with the gluon fields since these are what allow the axion to solve the strong CP problem.

The introduction of the field a is relevant to the strong CP problem because its expectation value, $\langle a \rangle$, in a Poincaré-invariant vacuum is a constant which does not vary throughout space, and when evaluated at such a configuration \mathcal{L}_a is exactly like \mathcal{L}_Θ , corresponding to a shift in the value of the Θ angle by $\tilde{\Theta}_{3,\text{eff}} = \tilde{\Theta}_3 + \langle a \rangle / f_a$. But expectation values are determined dynamically, by minimizing energies, and if this minimization should choose $\langle a \rangle = -f_a \tilde{\Theta}_3$, then the effect of the axion field will be to exactly cancel the CP violating effect of $\tilde{\Theta}_3$, leading to P- and CP-invariant strong interactions.

Now, if the shift symmetry, $a \rightarrow a + \epsilon$, were exact, then the axion would be exactly massless and any value for $\langle a \rangle$ would be equally acceptable from an energetic point of view. In this case we have gained nothing, since the

strong CP problem is changed from why $\tilde{\Theta}_3$ is so small, to why the axion field should take a special value of $-f_a\tilde{\Theta}_3$ in its vacuum.

However, as we have seen, the term \mathcal{L}_Θ can have physical significance in the absence of massless quarks, since it appears in the matrix elements which relate vacua having different values of N_{CS} . It follows that in these circumstances a shift in a *does* have physical effects, indicating that the corresponding symmetry is broken. Since these show up through effects involving N_{CS} -changing processes (instantons), these are also the effects which induce a potential, $V(a)$, for the field a , with the true vacuum being the one in which this potential is minimized.

In order to compute $V(a)$ we must calculate the energy per unit volume associated with the field a taking on a constant value throughout space. To this end, consider comparing the energies of two would-be vacua, $|0, a_1\rangle$ and $|0, a_2\rangle$, having different values of a . When computing their energy it is convenient to assume the volume of space to take a finite but large value, Ω , and to compute the following expectation value

$$\langle 0, a | \exp(-H\tau) | 0, a \rangle = \exp[-E(a)\tau] \quad (11.44)$$

The energy difference between the two configurations a_1 and a_2 is then

$$V(a_1) - V(a_2) = -\frac{1}{\Omega\tau} \left(\ln \langle 0, a_1 | e^{-H\tau} | 0, a_1 \rangle - \ln \langle 0, a_2 | e^{-H\tau} | 0, a_2 \rangle \right) \quad (11.45)$$

This energy depends on a for the same reason that the energy of the θ -vacuum depends on θ : $\exp(-H\tau)$ has an a -dependent matrix element between states which have different eigenvalues for N_{CS} . Using the freedom, as discussed above, to write the vacuum as $|0, a\rangle = \sum_n |n, a\rangle$, and using that the phase of the matrix element $\langle n, a | \exp[-H\tau] | m, a \rangle$ is $\exp[i(\tilde{\Theta}_3 + a/f_a)(n - m)]$, leads to

$$\begin{aligned} \exp(-\tau\Omega V(a)) &= \sum_{nm} \langle n, a | \exp(-H\tau) | m, a \rangle \\ &= \sum_{nm} V_{mn} \exp[i(m - n)(\tilde{\Theta}_3 + a/f_a)] \end{aligned} \quad (11.46)$$

where

$$V_{mn} = \langle n, a = 0, \tilde{\Theta}_3 = 0 | \exp[-H\tau] | m, a = 0, \tilde{\Theta}_3 = 0 \rangle \quad (11.47)$$

which is real and positive.[†]

Now comes the main point: this expression is maximized – and so $V(a)$

[†] This quantity is real because the Hamiltonian is purely real when the Θ_3 term is absent (modulo a tiny correction due to the KM matrix). Its positivity follows from unitarity, which ensures that Euclidean correlation functions such as this satisfy a property called *reflection positivity*. Further discussion on this point lies outside the scope of this book.

is minimized – when the phases are all zero, which occurs precisely at the CP-preserving value $a = -f_a \tilde{\Theta}_3$. Therefore, the correct vacuum value of a , obtained by minimizing the energy, is automatically the one which ensures the P and CP conservation of QCD.

In order to compute the mass of the particle associated with the scalar field, a , we must compute the curvature of the effective potential, $V(a)$, near its minimum;

$$m_a^2 = \left(\frac{\partial^2 V(a)}{\partial^2 a} \right)_{\min} \quad (11.48)$$

This requires an evaluation of V_{mn} , which can be shown (using a field-theoretic calculation we do not present) to be given in the large- Ω limit by

$$V_{mn} \sim \exp \left[-\frac{(n-m)^2}{2\chi\Omega\tau} \right] \quad (11.49)$$

where χ is called the *topological susceptibility*, in terms of which the mean-squared change in N_{CS} is given by $\chi\Omega\tau$, with $\chi \sim m_u m_d \Lambda_{\text{QCD}}^3 / (m_u + m_d)$. χ can be computed in detail using lattice gauge-theory techniques.

For notational convenience we denote the difference between a and its value at the minimum of V by $\tilde{a} = a - a_{\min}$. Expanding $V(a)$ for small \tilde{a} gives

$$\begin{aligned} \exp(-\Omega\tau V(a)) &\propto \sum_{nm} e^{i(m-n)\tilde{a}/f_a} \langle n | \exp(-H\tau) | m \rangle \\ &\propto \sum_{nm} e^{i(m-n)\tilde{a}/f_a} \exp \left[-\frac{(n-m)^2}{2\chi\Omega\tau} \right] \\ &= \sum_n e^{in\tilde{a}/f_a} \exp \left(-\frac{n^2}{2\chi\Omega\tau} \right) \\ &\sim \int dn \exp \left(\frac{in\tilde{a}}{f_a} - \frac{n^2}{2\chi\Omega\tau} \right) \\ &\propto \exp \left(-\frac{\tilde{a}^2 \chi \Omega \tau}{2f_a^2} \right) \end{aligned} \quad (11.50)$$

Therefore, up to an irrelevant additive constant,

$$V(\tilde{a}) = \left(\frac{\chi}{2f_a^2} \right) \tilde{a}^2 \quad (11.51)$$

and so

$$m_a^2 = \frac{\chi}{f_a^2} \quad (11.52)$$

This shows that the axion mass is inversely proportional to f_a and proportional to the square root of the topological susceptibility.

The larger the value of f_a , the lighter axions would be, and the weaker would be their interactions with normal matter. Unfortunately, f_a is a parameter of the model which can *a priori* take on any value. However, it is subject to experimental bounds based on the failure to find evidence for axions despite numerous searches. Direct laboratory searches rule out the lowest values of f_a , and astrophysical bounds provide further limits because unless f_a is very large, axions would be produced in red-giant stars and supernovae more copiously than observations of those bodies would allow. Considering these systems leads to a limit of approximately $m_a < 10^{-3}$ eV, or $f_a > 10^{10}$ GeV.

Cosmology also provides an upper limit on f_a . It does so despite the fact that very light axion particles are very hard to produce, since their interactions scale inversely with f_a . However, at an early stage of the Big Bang, the axion field would generically *not* be at its minimum value, and the larger f_a , the more energy is thereby initially stored in the axion field. Since this energy density falls with the universal expansion only as fast as does non-relativistic matter (and so more slowly than does radiation), the mass density of axions remaining from early-universe production can easily exceed the mass density actually observed in the universe. This places a lower bound on f_a , corresponding to $m_a > 10^{-6}$ eV, which is somewhat dependent on the cosmological model. On the other hand, if axions are just now close to dominating the universal energy density, they make an ideal candidate for the universal *cold dark matter* which appears to make up about 25% of the energy density of the universe.

At the time of writing (2006), there is no evidence for the existence of axions. However, heroic experimental efforts should soon be sensitive to any relic axions left over from the Big Bang, relying for their detection on the axion's interaction with electromagnetic fields due to the k_1 and k_2 terms of the Lagrangian, Eq. (11.42). If these are nonzero, experiments should either discover or rule out the axion within the relatively near future.

11.4.4 The flavor problem

Another puzzle which involves the dimension-4 couplings asks: what is the origin of the observed pattern of quark and lepton masses and mixings? These are built into the standard model through the values taken by the dimensionless Yukawa couplings, but the model would have made equally good sense regardless of what values these couplings took (provided they

were small enough to justify the perturbative analysis). As such the standard model provides a description of, but not an explanation for, these masses and mixings. The puzzle of explaining these masses and mixings is called the *flavor problem*.

A considerable amount of thought has been put into coming up with possible explanations of the observed-mass patterns, typically through variations of the idea that physics at energies higher than the electroweak scale enjoys an exact or approximate symmetry which acts on the generation labels of the standard-model fermions (or the constituents from which they are made, if the standard model fermions are composites). The pattern of fermion masses then becomes related to the physics which breaks this “family” symmetry. What makes this kind of model-building challenging is the difficulty in ensuring the GIM-type suppression of flavor-changing interactions once this symmetry is broken. At present there is no compelling picture along these lines which explains the kinds of masses and mixing matrices which are observed.

Although a full theory of nature would hopefully explain the observed masses and mixings, we emphasize that the flavor problem is not a naturalness problem in the same sense as are the hierarchy or cosmological-constant problems. That is because once the pattern of Yukawa couplings is chosen in a microscopic theory, including very small values like the electron’s Yukawa coupling $y_e = m_e/v = 2 \times 10^{-6}$, these small couplings remain small as successive particles having very high energies are integrated out. Since these couplings remain small as high scales are integrated out, it is logically possible that we may not learn why the fermion masses and mixings are what they are until we understand the physics of extremely high energies, perhaps not even until we reach scales $\mu \sim M_p$. Contrast this with the situation for the hierarchy problem, which requires something new to happen not too far above the electroweak scale, or the cosmological constant problem, which seems to require something new already at energies as low as 10^{-3} eV.

The fermion Yukawa couplings remain small under renormalization because they renormalize multiplicatively. That is, the corrections to a coupling are proportional to the coupling itself. This can be seen directly from the one-loop running equations, which are given explicitly in Eq. (11.23). Evaluating these within a mass basis, for which the Yukawa matrices are diagonal and real, leads to the an evolution equation for any diagonal element, y_n , which has the form

$$16\pi^2\mu^2\frac{dy_n}{d\mu^2} = y_n \left[a_1g_1^2 + a_2g_2^2 + a_3g_3^2 + \sum_k c_k y_k^2 \right] \quad (11.53)$$

where the a_r s and c_k s are known dimensionless constants.

If y_n is initially small, then its cube can be neglected on the right-hand side of this equation, leading to the formal solution

$$y_n(\mu) = y_n(\mu_0) \exp \left[\int_{\ln \mu_0^2}^{\ln \mu^2} A(t) dt \right] \quad (11.54)$$

where

$$A = \frac{1}{16\pi^2} \left[a_1 g_1^2 + a_2 g_2^2 + a_3 g_3^2 + \sum_{k \neq n} c_k y_k^2 \right] \quad (11.55)$$

For instance, if the other couplings slowly vary, so A is approximately constant, then $y_n(\mu)/y_n(\mu_0) = (\mu/\mu_0)^{2A}$. We see from this that $[y_n(\mu) - y_n(\mu_0)]/y_n(\mu_0) \sim 2A \ln(\mu/\mu_0)$ for small A , showing that y_n doesn't change in order of magnitude even when run over many decades of scales.

There is a good reason why the diagonal Yukawa couplings in the mass basis renormalize multiplicatively in this way. They do so because in the limit where a diagonal element vanishes, $y_n = 0$, the corresponding fermion becomes massless. But in this limit the theory acquires a new symmetry, corresponding to axial rotations of this massless fermion. It is this symmetry which ensures that $y_n = 0$ is a fixed point of the running equations, and so ensures that every term on the right-hand side of Eq. (11.53) must involve at least one power of y_n .

11.5 Dimension six: baryon-number violation

The standard model represents the most general possible renormalizable Lagrangian consistent with the assumed field content and gauge symmetries. At the non-renormalizable level the same choice of symmetries and field content allows only one new type of interaction at dimension 5 – whose implications are examined in Subsection 10.4.2 for neutrino masses. After dimension 5 comes the deluge, however, with a host of possible effective interactions arising at dimension 6. These interactions describe the dominant contributions of very massive particles to a wide variety of observable deviations from standard model predictions, some of which are explored in Problem 11.2 and Problem 11.4.

Unlike the previous sections, which presented potential problems for the standard model, the dimension-6 operators represent a success. The renormalizable standard model predicts that they should all be absent, and indeed, no evidence currently exists for any of these operators, and the limits on some of them are quite restrictive. Instead, dimension-6 operators should

be seen as an opportunity; they are the most natural place to look if we expect the standard model (augmented with dimension-5 operators) to break down.

Rather than trying to survey all dimension-6 operators, in this section we focus instead only on those dimension-6 interactions which violate baryon-number conservation. The interactions we find are the lowest dimension possible which do so, and as such can be expected to capture robustly the low-energy implications of any very high energy baryon-number violating physics.

11.5.1 Baryon-number violation

There are six possible dimension-6 baryon-number violating interactions consistent with the standard model field content and gauge symmetries, $\mathcal{L}_B = \sum_{I=1}^6 c_{mnpq}^I \mathcal{O}_{mnpq}^I + \text{h.c.}$, where

$$\begin{aligned}
 \mathcal{O}_{mnpq}^1 &= \epsilon^{\alpha\beta\gamma} \epsilon_{ij} [\bar{Q}_{m\gamma}^i P_L L_n^j] [\bar{D}_{p\alpha} P_R U_{q\beta}] \\
 \mathcal{O}_{mnpq}^2 &= \epsilon^{\alpha\beta\gamma} \epsilon_{ij} [\bar{Q}_{m\alpha}^i P_L Q_{n\beta}^j] [\bar{U}_{p\gamma} P_R E_q] \\
 \mathcal{O}_{mnpq}^3 &= \epsilon^{\alpha\beta\gamma} \epsilon_{ij} \epsilon_{kl} [\bar{Q}_{m\alpha}^i P_L Q_{n\beta}^j] [\bar{Q}_{p\gamma}^k P_L L_q^l] \\
 \mathcal{O}_{mnpq}^4 &= \epsilon^{\alpha\beta\gamma} (\tau_a \epsilon)_{ij} (\tau_a \epsilon)_{kl} [\bar{Q}_{m\alpha}^i P_L Q_{n\beta}^j] [\bar{Q}_{p\gamma}^k P_L L_q^l] \\
 \mathcal{O}_{mnpq}^5 &= \epsilon^{\alpha\beta\gamma} [\bar{D}_{m\alpha} P_R U_{n\beta}] [\bar{U}_{p\gamma} P_R E_q] \\
 \mathcal{O}_{mnpq}^6 &= \epsilon^{\alpha\beta\gamma} [\bar{U}_{m\alpha} P_R U_{n\beta}] [\bar{D}_{p\gamma} P_R E_q]
 \end{aligned} \tag{11.56}$$

Here, as before, $\alpha, \beta, \gamma = 1, 2, 3$ denote color-triplet indices; $i, j, k, l = 1, 2$ are $SU_L(2)$ -doublet indices and $m, n, p, q = 1, 2, 3$ are generation labels. Any other $\Delta B \neq 0$ interaction (such as those involving γ^μ or $\gamma^{\mu\nu}$) can be rewritten as a linear combination of these six, possibly after the use of a Fiertz rearrangement of the fermions (see Problem 1.9).

On dimensional grounds, the effective couplings for these interactions are inversely proportional to two powers of a heavy mass: $c_{mnpq}^I = \tilde{c}_{mnpq}^I / M^2$, where M represents the mass scale at which baryon-violating interactions arise and the dimensionless couplings, \tilde{c}_{mnpq}^I , contain any relevant loop or coupling-constant factors. To the extent that processes which violate baryon number are rare we expect that M must be very large.

Since the dimension-6 interactions of Eq. (11.56) satisfy $\Delta B = \pm 1$, they can mediate proton decay, a process which is absolutely forbidden within the standard model. Experimental searches for proton decay are clearly of considerable interest, since there are no fundamental principles which preclude its occurrence (unlike, say, violations of electric-charge conservation, whose

presence would ruin $U_{\text{em}}(1)$ gauge invariance and so also destroy the unitarity of photon interactions), even though we also understand why they have not been observed to date (the accidental B conservation of the renormalizable interactions of the standard model). Since $\Delta B = \pm 1$ for all dimension-6 terms, $\Delta B = \pm 2$ processes – like $n\bar{n}$ oscillations – can only proceed suppressed by even more powers of $1/M$, and so should occur with negligibly small rates.

Despite diligent searching, proton decay has never been observed. As of this writing (2006), the present lower limit for proton decay into the πe^+ channel is $\tau > 1.6 \times 10^{33}$ years (90% CL). On dimensional grounds, the prediction is of order $\tau^{-1} \sim |c^I|^2 m_p^5$ (where m_p is the proton mass and c^I is the relevant effective coupling). This gives an upper bound $|c^I| \lesssim (10^{16} \text{ GeV})^{-2}$. This points to an enormous mass scale, which is interestingly close to the 10^{14} GeV mass scale indicated by the dimension-5 description of neutrino masses, as well as to the upper limit, $M_p \sim 10^{19}$ GeV, before which we know new gravitational physics must arise. As we shall see, within grand unified theories (see below) there is also a second line of evidence (coupling unification) which points to this same mass scale.

Identifying the dimension-6 operators allows an efficient identification of the selection rules which any $\Delta B \neq 0$ process like proton decay must satisfy, subject only to the assumption that only standard-model particles participate. Inspection of the operators of Eq. (11.56) reveals several of these.

- All of these interactions satisfy the selection rule $\Delta B = \Delta L$, which implies that protons must decay into anti-leptons, allowing $p \rightarrow \pi^0 \ell^+$ or $n \rightarrow \pi^- \ell^+$ but forbidding $p \rightarrow \pi^+ \pi^0$, $p \rightarrow \pi^+ \pi^+ \pi^-$, or $n \rightarrow \pi^+ \ell^-$. This shows that the standard-model field content ensures that B – L conservation is automatically satisfied at low energies to a better approximation than B conservation, regardless of whether or not this is true for the higher-energy particles whose virtual effects are ultimately responsible for baryon-number violation. Hence, even if proton decay is observed and it satisfies B – L conservation, this does not imply that B – L is exactly conserved at much higher energies.
- All of the dimension-6 interactions satisfy $\text{sign } \Delta S = -\text{sign } \Delta B$. For instance, any transition which lowers B by one unit can destroy zero, one or two strange quarks, but cannot create any strange quarks. This forbids decays like $p \rightarrow \bar{K}^0 \ell^+$ or $n \rightarrow \bar{K}^- \ell^+$.
- $\Delta S = 0$ processes satisfy the isospin selection rule $\Delta I = 1/2$, and imply

relations like

$$\begin{aligned} \Gamma(p \rightarrow \pi^0 \ell_R^+) &= \frac{1}{2} \Gamma(n \rightarrow \pi^- \ell_R^+) = \frac{1}{2} \Gamma(p \rightarrow \pi^+ \bar{\nu}) = \Gamma(n \rightarrow \pi^0 \bar{\nu}) \\ \text{and} \quad \Gamma(p \rightarrow \pi^0 \ell_L^+) &= \frac{1}{2} \Gamma(n \rightarrow \pi^- \ell_L^+) \end{aligned} \quad (11.57)$$

More detailed predictions are possible if it is known that only a subset of the dimension-6 operators contribute. For instance, in many grand unified models it is known that low-energy baryon-number violating interactions are generated by tree graphs involving the exchange of a super-heavy spin-one particle, and such an exchange can only generate a linear combination of interactions \mathcal{O}^1 and \mathcal{O}^2 .

11.5.1.1 *B-violation in the MSSM**

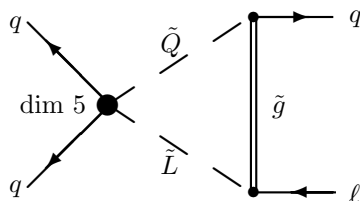
The key assumption on which the previous analysis relies is the presumption that the non-renormalizable interactions must be constructed only from standard model fields. This assumption would not be true if physics at the electroweak scale should prove to be supersymmetric, since it is then also possible to construct interactions using the many superpartners which then are also present.

We have already seen that once these fields are included it is possible to have baryon-number violating interactions already at the renormalizable level. In fact, this observation motivated the definition of R -parity, since the assumption of R -parity invariance was sufficient to preclude baryon-number non-conservation at dimension-4 or less. Owing to the strong motivations which supersymmetry has as a replacement for the standard model, we pause here to describe the lowest-dimension B - and L -violating interactions which preserve R -parity invariance.

It turns out that both B - and L -violating interactions are first possible at dimension-5 in a supersymmetric and R -parity invariant theory. These can be described by the following three types of quartic interactions in the superpotential of the low-energy supersymmetric theory:

$$(LLH_U H_U), \quad (QQQL) \quad \text{and} \quad (\bar{U}\bar{U}\bar{D}\bar{E}) \quad (11.58)$$

where all gauge and generation indices are suppressed. The first of these is the supersymmetric extension of the $\Delta L = \pm 2$ dimension-5 interaction which we saw can describe neutrino masses using only standard model

Fig. 11.1 $\Delta B \neq 0$ operator from a dimension-5 SUSY coupling.

fields. The others are new, and rely for their existence on the presence of superpartner fields, such as through dimension-5 quark–quark–squark–slepton interactions.

Such dimension-5 interactions lead to loop corrections which induce dimension-6 $\Delta B \neq 0$ interactions involving only quarks and leptons. For instance, a quark–quark–squark–slepton interaction can be combined with a renormalizable quark–squark–gaugino and lepton–slepton–gaugino interaction in the loop graph of Figure 11.1 to produce an effective 4-fermion 3-quark/1-lepton interaction which can mediate proton decay. The blob in the figure represents the dimension-5 coupling, the dashed line denotes a squark or slepton and the double line is a gaugino. This graph leads to the following estimate for the dimension-6 quark coupling,

$$c_6 \sim \left(\frac{g^2 m_g}{16\pi^2 M_s^2} \right) c_5 \quad (11.59)$$

where c_d is the effective coupling of the respective dimension- d interaction, g is the relevant gauge coupling, m_g is a gaugino mass, and M_s denotes a typical squark or slepton mass in the loop.

Recalling the bound $|c_6| \lesssim (10^{16} \text{ GeV})^{-2}$ and taking $g \sim 0.3$, $m_g \sim 100 \text{ GeV}$, and $M_s \sim 300 \text{ GeV}$ leads to the estimate $|c_5| \lesssim (10^{24} \text{ GeV})^{-1}$, which is potentially problematic. It turns out that such an operator is typically generated in high-energy models through the exchange of a superheavy boson, leading to the estimate $c_5 \sim y^2/M$, where y is of order the light-quark Yukawa couplings, $y \sim m_q/v \sim 10^{-4}$. With such a choice, the bound on c_5 implies $M \gtrsim 10^{16} \text{ GeV}$. We see that the observed proton decay rate is consistent with baryon-number violation as described by dimension-5 operators and new physics lighter than M_p only if the decay rate is suppressed by small dimensionless couplings like y .

Supersymmetric theories for proton decay can also imply different selection rules for nucleon decays. For instance, both of the baryon-number violating contributions to the superpotential, Eq. (11.58), must involve at least two generations of quarks, since color neutrality requires taking an antisymmetric combination of the three color indices. Because the fields involved in the superpotential are scalars and so obey Bose statistics, the resulting term vanishes if evaluated with all fields having the same generation index. Once expressed in terms of a basis of mass eigenstates this indicates that decays like $p \rightarrow \pi^0 e^+$ must be Cabbibo-suppressed compared to those like $p \rightarrow \pi^0 \mu^+$, say (although this suppression must be weighed against the reduced phase space associated with having second-generation fermions amongst the decay products).

11.5.2 Proposal: grand unified theories

Historically, baryon-number violating interactions were initially considered within the context of grand unified theories (GUTs). These theories were invented because of a dissatisfaction with the gauge structure of the standard model, for which the gauge group has three independent coupling constants corresponding to the three factors in the gauge group $SU_c(3) \times SU_L(2) \times U_Y(1)$. The fermion content of the theory is also complicated, containing a variety of fermions each of which transforms in a different way under the gauge group, and each of which has separate Yukawa couplings to the Higgs field.

On the other hand, if the gauge group were a *simple* group (see Appendix B), like $SU(N)$ or $O(N)$, consisting of a single factor, then there would be only one gauge coupling. Can the standard model be obtained from such a gauge group through a process of spontaneous symmetry breaking, in the same way that the gauge group $SU_c(3) \times U_{em}(1)$ emerges from the standard-model gauge group at the electroweak scale?

This turns out to be possible, and the framework for so doing is surprisingly simple and elegant. As a result, such unified models have captured the imagination of physicists, even though there is not yet any evidence that this actually happens in nature – with one tantalizing exception to which we return below.

The simplest group into which $SU_c(3) \times SU_L(2) \times U_Y(1)$ can be embedded in this way is $SU(5)$, the group of five-by-five unitary matrices having unit determinant. The gauge generators of $SU(5)$ are five-by-five traceless,

Hermitian matrices, having the schematic form

$$\begin{pmatrix} a_1 & a_4 & a_5 & b_1 & b_2 \\ a_4^* & a_2 & a_6 & b_3 & b_4 \\ a_5^* & a_6^* & a_3 & b_5 & b_6 \\ b_1^* & b_3^* & b_5^* & c_1 & c_3 \\ b_2^* & b_4^* & b_6^* & c_3^* & c_2 \end{pmatrix} \quad (11.60)$$

where the diagonal elements are real and tracelessness implies that their sum vanishes. The idea is to identify the upper-left three-by-three block (the a s) as the gauge group $SU_c(3)$, the lower-right two-by-two block (the c s) as $SU_L(2)$, and the diagonal generator which commutes with each of these as $U_Y(1)$:

$$Y = \begin{pmatrix} -\frac{1}{3} & & & & \\ & -\frac{1}{3} & & & \\ & & -\frac{1}{3} & & \\ & & & \frac{1}{2} & \\ & & & & \frac{1}{2} \end{pmatrix} \quad (11.61)$$

The normalization of Y is chosen in order to ensure that we obtain the proper charge assignments for quarks and leptons.

Can the known quarks and leptons be assembled into representations of $SU(5)$? With the above identification of the $SU_c(3) \times SU_L(2) \times U_Y(1)$ gauge group, the 5-dimensional column vector (the defining representation of $SU(5)$) contains the following $SU_c(3) \times SU_L(2)$ representations: $\mathbf{5} = (\mathbf{3}, \mathbf{1}) \oplus (\mathbf{1}, \mathbf{2})$. That is, writing the $\mathbf{5}$ as a column vector, $(f_1, f_2, f_3, h_1, h_2)^T$, the f s transform as $(\mathbf{3}, \mathbf{1})$ and the h s as $(\mathbf{1}, \mathbf{2})$. We see that the top three components (the f s) transform in the same way as do right-handed quarks, while the bottom two components (the h s) transform like the lepton doublet. But since all $SU(5)$ generators are traceless, the hypercharges for these two must satisfy $3y_f + 2y_h = 0$, and so the hypercharge assignments can only work if we take $y_f = -1/3$ and $y_h = 1/2$, indicating that f must be the right-handed down-type singlet, D_R , while h is the right-handed doublet, L_R . (This is also what tells us how to normalize the hypercharge generator, Y .)

It remains to find an $SU(5)$ representation (or representations) which can contain Q , U , and E . Since Q carries both an $SU_L(2)$ index and an $SU_c(3)$ index, this suggests we try a representation built from the tensor product of two 5-dimensional representations, and one of these – the antisymmetric product, $\Phi_{IJ} = -\Phi_{JI} \in \mathbf{10}$ – does the job. Explicitly, the standard-model content of the $\Psi_I \in \mathbf{5}$ and $\Phi_{IJ} \in \mathbf{10}$ representations can be written in the

following way,

$$\Psi_I = \begin{pmatrix} D_\alpha \\ L_i \end{pmatrix}_R \quad \Phi_{IJ} = \begin{pmatrix} \epsilon_{\alpha\beta\gamma} U^\gamma & -Q_{\alpha j} \\ Q_{i\beta} & \epsilon_{ij} E \end{pmatrix}_L \quad (11.62)$$

which also correctly gives the hypercharge assignments. This shows that one generation of left-handed standard-model fermions precisely fits into the $\bar{\mathbf{5}} \oplus \mathbf{10}$ representation of $SU(5)$, with no states left over and no additional states needed.

We remark in passing that all of the known fermions can fit into an even simpler representation if we are prepared to enlarge the gauge group even further, to $SO(10)$. The main observation in this case is that an entire generation of fermions can be fit into a single spinor representation – the $\mathbf{16}$ – of $SO(10)$, *provided* it is supplemented by a standard-model singlet fermion, N . (This observation is most succinctly summarized by giving the transformation properties of this representation in terms of $SU(5) \subset SO(10)$: $\mathbf{16} = \bar{\mathbf{5}} \oplus \mathbf{10} \oplus \mathbf{1}$.) As we have seen, such a singlet fermion has a natural interpretation as a sterile neutrino, allowing $SO(10)$ models to incorporate a seesaw pattern of neutrino masses within a very appealing framework.

Since there are a total of 24 $SU(5)$ gauge generators, but only $8 + 3 + 1 = 12$ standard-model gauge generators, $SU(5)$ contains 12 new gauge fields which are not in the standard model. These fields correspond to the b s of Eq. (11.60), and so transform under the $SU_c(3) \times SU_L(2)$ gauge group as a $(\mathbf{3}, \mathbf{2})$, with hypercharge $-1/2$. We write these as a doublet of color-triplet fields: $X_{i\alpha}$. In the same way that the W and Z acquire mass due to electroweak symmetry breaking, we expect the X bosons to acquire large masses, $M \gg M_W$, due to the symmetry breaking from $SU(5)$ down to $SU_c(3) \times SU_L(2) \times U_Y(1)$ (which requires that the theory contain scalar fields which develop *v.e.v.s*, in analogy with the Higgs field of the standard model).

Baryon-number violation necessarily enters into such a theory because quarks and leptons coexist within a single gauge representation. This ensures the existence of gauge interactions for which leptons convert to quarks (or back) together with the emission or absorption of an X boson. Virtual X boson exchange therefore gives rise at low energies to an effective dimension-6 baryon-number violating interaction, with coefficient $c^I \sim g_5^2/M^2$ where g_5 is the $SU(5)$ gauge coupling constant. The constraints on c^I from the proton lifetime then show that the X boson mass must be very large, $M \gtrsim 10^{15}$ GeV if $g_5 \sim 0.1$.

Such a large X boson mass also fits well with another observational issue,

which we now describe. This issue asks: given that the gauge couplings of the three factors of the standard model gauge group are different, how can they all be described by a single $SU(5)$ coupling? This can be possible in principle because the gauge couplings run with energy, according to

$$\frac{1}{\alpha(\mu)} = \frac{1}{\alpha(\mu_0)} - b \ln \left(\frac{\mu^2}{\mu_0^2} \right) \quad (11.63)$$

where, for generic spin-zero, spin-half, and spin-one matter the coefficient b is given by (generalizing Eq. (7.40), Eq. (7.42), and Eq. (11.22))

$$b = \frac{1}{12\pi} [T(R_0) + 2T(R_{1/2}) - 11T(A)] \quad (11.64)$$

Here the Dynkin index, $T(R)$, for a representation R is defined in terms of the trace of the group generators by $\text{tr}[t_a t_b] = T(R)\delta_{ab}$. R_0 here denotes the gauge representation of the complex scalars, $R_{1/2}$ is the same for the left-handed Weyl fermions and A represents the adjoint representation. With these conventions we have $T(F) = 1/2$ and $T(A) = N$ for the fundamental (F) and adjoint (A) representations of $SU(N)$.

At energies above the scale M where $SU(5)$ breaks, the running is due to loops involving complete $SU(5)$ multiplets of particles, and so the couplings of all $SU(5)$ gauge bosons run in the same way, with $b = b_5$ computed using the appropriate $SU(5)$ representation content at these energies (i.e. $\mathbf{\bar{5}}$ and $\mathbf{10}$ for the fermions, the adjoint $\mathbf{24}$ for the gauge bosons plus whatever Higgs representations are required to accomplish the desired pattern of spontaneous symmetry breaking).

However, below M $SU(5)$ is broken, and so it need not be possible to assemble the available particles into $SU(5)$ multiplets. (Although we have seen that this is possible for the quarks and leptons, it is not for the gauge bosons because the heavy X particles are too heavy to be present in the low-energy theory below M .) Consequently, the three standard-model gauge couplings can run differently below M , with

$$\begin{aligned} \frac{1}{\alpha_1(\mu)} &= \frac{1}{\alpha_1(\mu_0)} - b_1 \log \left(\frac{\mu^2}{\mu_0^2} \right) \\ \frac{1}{\alpha_2(\mu)} &= \frac{1}{\alpha_2(\mu_0)} - b_2 \log \left(\frac{\mu^2}{\mu_0^2} \right) \\ \frac{1}{\alpha_3(\mu)} &= \frac{1}{\alpha_3(\mu_0)} - b_3 \log \left(\frac{\mu^2}{\mu_0^2} \right) \end{aligned} \quad (11.65)$$

Because M is so large, the couplings can become very different when run to the electroweak scale.

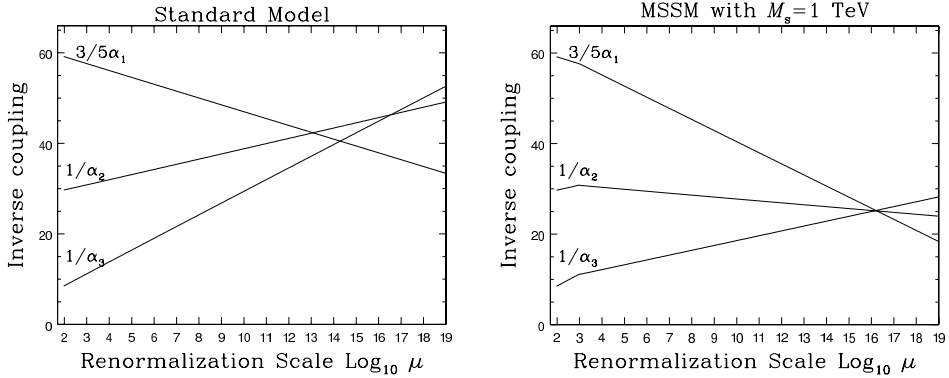


Fig. 11.2 Coupling unification in the MSSM.

Can the couplings change by enough to give the values observed in the standard model? To find out, it is useful to start with the known experimental values for the couplings at $\mu = M_Z$ and run these to higher energies to see if there is a scale $\mu = M$ where they all unify. When doing so, it is important to identify the standard-model gauge couplings with the hypercharge generator normalized using the same conventions as those for the generators of the other gauge groups. Since for these we use $\text{tr}[t_a t_b] = \frac{1}{2} \delta_{ab}$ in the fundamental representation, inspection of Eq. (11.61) shows that it is $T = \sqrt{3/5} Y$ whose coupling should equal the couplings g_2 and g_3 at some scale.

We therefore plot the inverse couplings, $1/\alpha_3$, $1/\alpha_2$, and $3/5 \alpha_1$, vs. $\ln \mu^2$, starting with their measured values at $\mu = M_Z$. The result in the standard model is shown on the left in Figure 11.2. Eq. (11.65) implies the resulting curves are straight lines whose slopes are given by $-b_i$. (The plot actually uses 2-loop evolution equations, but the difference is insignificant.) The three lines approach one another as μ increases, because $-b_1 < 0 < -b_2 < -b_3$. Any two converging straight lines must cross; the question is whether the third curve also crosses the others at this same scale (within errors due to our imperfect knowledge of the α s, particularly α_3). We see that in the standard model, they quite clearly fail to cross.

How the couplings vary with scale depends on the field content with masses below the scale of interest. The failure of the three couplings to unify (become equal) at a single scale, which we just found, was conditional on the standard-model field content being correct all the way up to the scale

M where the unification might have occurred. If there are additional fields with masses between M_Z and M , and particularly if they do not form complete multiplets of $SU(5)$, then the story will be different. (New fields which *do* form $SU(5)$ multiplets change the value of α where lines cross, but not the μ where the crossing occurs, at least in the 1-loop approximation.)

An interesting thing happens if we assume instead that it is the MSSM particle content which governs the running at energies above, say, 1 TeV. In this case, because the elements of a supermultiplet transform in the same way under gauge transformations, we may group their contributions to Eq. (11.64) together, giving

$$b = \frac{1}{4\pi} [T(R_m) - 3T(A)] \quad (11.66)$$

where R_m is the representation of the matter supermultiplets and A is the (adjoint) representation of the gauge supermultiplets. Using the MSSM field content in this way, we find that the gauge couplings *do* cross at a common scale, as shown on the right in Figure 11.2. Even more interesting, the scale at which they cross is about $M = 1.7 \times 10^{16}$ GeV, very close to the scale required for the X bosons if they are not to produce too rapid proton decay.

What does this tell us? It could be telling us two very exciting things: that supersymmetry is a symmetry of nature which is broken at energies just out of reach; and that the standard-model gauge group really is unified at very high energies. On the other hand it could instead just be a cruel coincidence. At this point we do not yet know how much importance to give to this tantalizing high-energy convergence of the standard-model couplings.

11.6 Problems

[11.1] Quadratic contributions to the Higgs mass

Suppose the standard model is supplemented by an additional real scalar field, s , subject to a discrete symmetry, $s \rightarrow -s$. The new terms in the Lagrangian are

$$\mathcal{L}_{\text{new}} = -\frac{1}{2}\partial_\mu s \partial^\mu s - \frac{1}{2}m_s^2 s^2 - \frac{\lambda_s}{24}s^4 - \frac{\lambda'}{2}s^2\phi^\dagger\phi \quad (11.67)$$

Compute the loop correction to the Higgs mass by evaluating the following loop graph in the limit where no momentum flows through the external lines. In the graph it is the s particle which circulates about the loop and

the external lines denote ϕ fields.

$$\begin{aligned}
 \text{Diagram: a circle with a dot on its bottom edge, connected to a horizontal line passing through the dot.} &\Rightarrow \delta m_{\text{H}}^2 = -\frac{i\lambda'}{2} \int \frac{d^4 q}{(2\pi)^4} \left[\frac{1}{q^2 + m_s^2 - i\epsilon} \right] \\
 &= \frac{\lambda'}{32\pi^2} \left[\Lambda^2 - m_s^2 \ln \left(\frac{\Lambda^2}{m_s^2} \right) + O \left(\frac{m_s^2}{\Lambda^2} \right) \right]
 \end{aligned}$$

where Λ represents an ultraviolet cut-off. Hint: going from the first to the second line requires Wick rotation, see Subsection 7.4.5.

[11.2] Fermion dipole moments

Identify the gauge-invariant dimension-6 operators which, in unitary gauge, have the Lorentz structure $[\bar{f}_1 \gamma^{\mu\nu} P_L f_2] F_{\mu\nu}$, where f_i denote spinor fields and $F_{\mu\nu}$ is the electromagnetic or gluon field. The matrix elements of such operators contribute to the anomalous magnetic (or chromomagnetic) dipole moments (and transition moments) of the quarks and leptons. The corresponding electric moments are obtained from these using the replacement $F_{\mu\nu} \rightarrow \epsilon_{\mu\nu\lambda\rho} F^{\lambda\rho}$.

[11.3] Redundant effective interactions

Consider the expansion of an arbitrary action in powers of a small parameter, ϵ (which could represent powers of $1/M$ in a low-energy effective theory):

$$S[\varphi] = S_0[\varphi] + \epsilon S_1[\varphi] + \cdots \quad (11.68)$$

where φ^i denotes the relevant field content. In such an expansion any term in S_1 which vanishes on the use of the zeroth-order equations, such as

$$\hat{S}_1[\varphi] = \int d^4 x \, c^i(x) \left[\frac{\delta S_0}{\delta \varphi^i(x)} \right] \quad (11.69)$$

is called a *redundant interaction* because it can be removed by performing an appropriate field redefinition, $\varphi^i \rightarrow \varphi^i + \epsilon \delta \varphi^i$. Neglecting terms which are $O(\epsilon^2)$, find the required redefinition, $\delta \varphi^i$, which removes the interaction, Eq. (11.69), given above. What is the generalization of this result to the $O(\epsilon^n)$ term in the action?

The ability to remove such interactions by a field redefinition shows that they cannot contribute to physical observables. This justifies the use of the lower-order equations of motion to simplify the kinds of interactions which must be entertained at any given order in a low-energy effective theory. Notice that this statement is *not* restricted to the use of such interactions in tree-level graphs in the effective theory.

[11.4] **Anomalous gauge-boson self-couplings**

Identify the most general $SU_L(2) \times U_Y(1)$ invariant dimension-6 interactions which contribute to gauge-boson self-couplings. How many independent new kinds of couplings are allowed in this case? Classify these according to their transformation properties with respect to C, P, and T.

[11.5] **Neutralinos in the MSSM**

The MSSM contains four new neutral Majorana fermionic fields: the Bino \tilde{B} , the neutral Wino \tilde{W}^0 , and the two neutral components of the Higgsinos, \tilde{H}_U^0 and \tilde{H}_D^0 . In the absence of electroweak symmetry breaking, \tilde{B} and \tilde{W}^0 have Majorana masses, while the Higgsinos have a Dirac mass arising from the μ term:

$$\mathcal{L}_{\text{neutral}} = -\frac{m_{\tilde{B}}}{2}\tilde{B}\tilde{B} - \frac{m_{\tilde{W}}}{2}\tilde{W}\tilde{W} - \left(\mu\tilde{H}_U^0 P_L \tilde{H}_D^0 + \text{h.c.}\right)$$

However, electroweak symmetry breaking adds several terms when the Higgs fields are replaced by their expectation values.

Use Eq. (11.9) to show that, if the four fields are written as a column vector $\chi \equiv \left(\tilde{B}\tilde{W}\tilde{H}_U^0\tilde{H}_D^0\right)^T$, then the mass term including contributions from the Higgs *v.e.v.s* becomes

$$\mathcal{L}_{\text{neutral}} = -\frac{1}{2}(\bar{\chi} M P_L \chi + \text{h.c.}), \quad M = \begin{pmatrix} m_{\tilde{B}} & 0 & \frac{g_1 v_u}{2} & -\frac{g_1 v_d}{2} \\ 0 & m_{\tilde{W}} & \frac{g_2 v_u}{2} & -\frac{g_2 v_d}{2} \\ \frac{g_1 v_u}{2} & \frac{g_2 v_u}{2} & 0 & \mu \\ -\frac{g_1 v_d}{2} & -\frac{g_2 v_d}{2} & \mu & 0 \end{pmatrix}$$

Argue that the eigenvectors and the absolute values of the eigenvalues of the matrix M are the masses of the four chargeless Majorana states, and that in general none of the states coincide precisely with any of the original fields.

Appendix A

Experimental values for the parameters

We present here a summary of the values for the various parameters of the standard model as has been determined by experiment, together with a brief description of the physical processes from which the comparison between theory experiment has been used to obtain these parameters. The accuracy with which we know these parameters is improving with time, so the reader should consult the recent literature for the most up to date values of parameters. The Particle Data Group maintains exhaustive and current lists at a public website, <http://pdg.lbl.gov/> and also issues booklets and reviews every two years. Most of the values quoted below are taken from the 2004 review by the Particle Data Group.

Most of the masses listed are determined either in particle-physics experiments, or for the very long lived or stable particles, in trap experiments. The masses of quarks are difficult to determine, since the quarks themselves do not appear as particles; in fact the masses of the light quarks are even difficult to define unambiguously. The charm and bottom masses are defined as the “running mass” at the scale set by the “running mass,” in the $\overline{\text{MS}}$ renormalization scheme. For the top quark, such ambiguities are subdominant to experimental error. The photon mass is known to be less than 10^{-24} GeV from geomagnetic observations, and is believed theoretically to be exactly zero. The mass of the gluon is theoretically zero, but since the gluon exists only within bound states, this refers to its Lagrangian mass, rather than the mass of particles containing a gluon. At the time of writing, the best limits on neutrino masses come from cosmology; specifically, larger neutrino masses than 0.2 eV would change the formation of large-scale structure in the universe, contrary to observation. An arguably less model-dependent particle physics limit on the electron mass, $m_{\nu_e} < 3$ eV, has been obtained from the tritium beta decay endpoint. A limit on one neutrino type can be interpreted as a limit on all three, because neutrino oscillation

Table A.1 *Particle masses*

Particle name	symbol	mass (GeV)	source
electron	e^-	0.000 510 998 918(44)	trap experiments
muon	μ^-	0.105 658 369(9)	NMR
tau	τ^-	1.776 99(29)	BES ($e^+e^- \rightarrow \tau^+\tau^-$)
electron neutrino	ν_e	$< 2 \times 10^{-10}$	cosmology
muon neutrino	ν_μ	$< 2 \times 10^{-10}$	cosmology
tau neutrino	ν_τ	$< 2 \times 10^{-10}$	cosmology
photon	γ	0	classical E&M
gluon	g	"0"	theoretical value
W boson	W^\pm	80.425(38)	LEP II, Tevatron
Z boson	Z^0	91.187 6(21)	LEP I
Higgs boson	H	> 114	LEP II
up quark	u	0.0019(3)($\overline{\text{MS}}, \mu = 2 \text{ GeV}$)	Lattice
down quark	d	0.0044(3)($\overline{\text{MS}}, \mu = 2 \text{ GeV}$)	Lattice
strange quark	s	0.087(4)($\overline{\text{MS}}, \mu = 2 \text{ GeV}$)	Lattice
charm quark	c	1.31(7)($\overline{\text{MS}}, \mu = m$)	$c\bar{c}$ meson masses
bottom quark	b	4.24(11)($\overline{\text{MS}}, \mu = m$)	$b\bar{b}$ meson masses
top quark	t	172.7(29) (pole mass)	Tevatron
charged pion	π^\pm	0.139 570 18(35)	pionic atoms
neutral pion	π^0	0.134 976 6(6)	π^\pm decays
proton	p	0.938 272 029(80)	trap experiments
neutron	n	0.939 565 36(8)	trap etc.

Table A.2 *Numerical values of coupling constants*

Coupling	symbol	renorm. point	value
Electromagnetic	α	$\ll m_e$	1/137.035 999 11(46)
Electromagnetic	α	$\mu_{\overline{\text{MS}}} = M_Z$	1/127.934(27)
weak	g_2, α_2	$\mu_{\overline{\text{MS}}} = M_Z$	0.652, 1/29.6
hypercharge	g_1, α_1	$\mu_{\overline{\text{MS}}} = M_Z$	0.357, 1/98.4
Weinberg angle	$\sin^2 \theta_w$	$\mu_{\overline{\text{MS}}} = M_Z$	0.231 21(15)
Fermi	G_F	$\ll M_W$	$1.166\,37(1) \times 10^{-5} \text{ GeV}^{-2}$
strong	α_3	$\mu_{\overline{\text{MS}}} = M_Z$	0.1187(20)
strong Θ	Θ_3	$\ll M_Z$	$ \Theta_3 < 10^{-9}$

experiments show that the differences in the neutrino masses squared are less than $5 \times 10^{-3} \text{ eV}^2$.

Most of the numerical values for coupling constants presented are in the $\overline{\text{MS}}$ renormalization scheme, evaluated at the scale of the Z mass. We also

quote the electromagnetic coupling evaluated at low scales – it has no scale dependence below the electron-mass scale. The electromagnetic interaction strength is determined from laboratory experiments such as the Josephson effect and the quantum Hall effect, and (most accurately) by comparing the measured magnetic moment of the electron to a very high precision calculation within the theory of electromagnetism, QED. The high energy value is determined partly from data (necessary to account for the strong interactions of light, charged hadrons) and partly perturbatively. The other couplings are determined largely from the Z resonance studies of LEP I.

The best values for the CKM matrix elements, in magnitude, are currently

$$|V_{mn}| = \begin{bmatrix} 0.9741 - 0.9756 & 0.219 - 0.226 & 0.0025 - 0.0048 \\ 0.219 - 0.226 & 0.9732 - 0.9748 & 0.038 - 0.044 \\ 0.004 - 0.014 & 0.037 - 0.044 & 0.9990 - 0.9993 \end{bmatrix} \quad (\text{A.1})$$

Unitarity has been used to constrain elements which are poorly constrained experimentally. In particular, direct constraints on the bottom row, involving the top quark, are extremely weak.

Appendix B

Symmetries and group theory review

We give here a very short review of symmetries, and of the theory of Lie groups needed in their study. The treatment is not intended to be complete or rigorous, just to give a brief introduction for readers unfamiliar with the material.

B.1 Symmetry transformations as a group

A symmetry transformation is a transformation on the states of a theory $|\psi\rangle$ and the operators \mathcal{O} ,

$$|\psi\rangle \rightarrow U|\psi\rangle, \quad \mathcal{O} \rightarrow U\mathcal{O}U^* \quad (\text{B.1})$$

which preserves “all physics.” In particular, amplitudes must be preserved (up to a phase, a complication we ignore here and in the following)

$$\langle\psi_1|\psi_2\rangle \rightarrow \langle\psi_1|U^*U|\psi_2\rangle = \langle\psi_1|\psi_2\rangle \quad (\text{B.2})$$

which shows that U must either be unitary or anti-unitary. We will only consider unitary symmetries here. (Here U^* is the Hermitian conjugate of U . We write U^\dagger only if U is a matrix.)

A symmetry in which a local operator $\mathcal{O}(x)$ (an operator built out of fields at point x only) is transformed into a local operator at the same point, is called an *internal symmetry*. Such symmetries can be considered separately from *spacetime symmetries*, such as translations, rotations, and boosts. In fact, it is a theorem that the full group of symmetries is always a product of the internal symmetries and the spacetime symmetries. In this appendix we concentrate on internal symmetries; spacetime symmetries are discussed in the next appendix.

Because the states and operators under discussion may appear at different times, the symmetry operator must commute with time evolution,

$[H, U] = 0$. Similarly, it must commute with the momentum operators. This is summarized by saying that it must commute with the Lagrangian density, $[\mathcal{L}, U] = 0$. Therefore the symmetries of the theory can usually be identified by looking at the symmetries of the Lagrangian.

In a renormalizable theory, an operator \mathcal{O} and its symmetry transform $U \mathcal{O} U^* = \mathcal{O}'$ must be of the same dimension. Since the fields are the operators of the smallest dimension, this means the fields transform linearly among themselves,

$$\varphi_a \rightarrow U \varphi_a U^* = M_{ab}^{-1} \varphi_b \quad (\text{B.3})$$

where φ collectively symbolizes the fields of the theory and a, b are indices on the set of fields. Classically, the symmetries of a theory are the set of transformations on the fields of this form, under which \mathcal{L} is unchanged. The relation between φ' and φ involves a matrix inverse essentially because φ annihilates a particle $|\varphi\rangle$, and must therefore have the inverse transformation properties of the particle.

Acting with two symmetry transformations successively, $|\psi\rangle \rightarrow U_1 |\psi\rangle \rightarrow U_2 U_1 |\psi\rangle$, yields another symmetry transformation, namely, the one induced by the operator $(U_2 U_1)$. This defines a multiplication rule under which symmetry transformations form a group. In general the symmetry group can be factorized as a product of nonfactorizable subgroups, and it is sufficient to examine the behavior of the subgroups individually. In particle physics these subgroups are usually small discrete groups (which we will not discuss) and continuous (Lie) groups. The latter can generally be factorized into a product of simple Lie groups and $U(1)$ groups.

B.2 Lie groups and Lie algebras

A Lie group is a group which is also a manifold. In particular, there is a small neighborhood around the identity $\mathbf{1}$ which looks like a piece of \mathbf{R}^n , with n the dimension of the group. One can always choose a coordinate basis for this region; the coordinate unit vectors t_a are called the *Lie algebra* and an arbitrary element g of the group which is close to the identity can always be expanded in the coordinates,

$$g = \mathbf{1} + i\omega^\alpha t_\alpha \quad (\text{B.4})$$

with ω_α (infinitesimal) parameters. (The i is customary so that for groups of unitary matrices, the t_α are Hermitian.)

Now consider two elements of the group which are each close to the identity, say, $g_1 = \mathbf{1} + i\omega_1^\alpha t_\alpha$ and $g_2 = \mathbf{1} + i\omega_2^\alpha t_\alpha$. The multiplication rule to first

order in these parameters is given by

$$g_1 g_2 = \mathbf{1} + i(\omega_1^\alpha + \omega_2^\alpha) t_\alpha + O(\omega^2) \quad (\text{B.5})$$

which is addition of the departures from the identity. At the next order, $g_1 g_2$ and $g_2 g_1$ can differ:

$$\begin{aligned} g_1 g_2 (g_2 g_1)^{-1} &= (\mathbf{1} + i\omega_1^\alpha t_\alpha) (\mathbf{1} + i\omega_2^\beta t_\beta) (\mathbf{1} - i\omega_1^\gamma t_\gamma) (\mathbf{1} - i\omega_2^\sigma t_\sigma) \\ &= \mathbf{1} - \omega_1^\alpha \omega_2^\beta [t_\alpha, t_\beta] \end{aligned} \quad (\text{B.6})$$

Therefore, to determine the multiplication rule to second order we need to know the commutators of the Lie algebra elements. Since $g_1 g_2 g_1^{-1} g_2^{-1}$ is still close to the identity, it can still be expressed in terms of coefficients multiplying Lie algebra elements, so the commutator must also be a sum of elements of the Lie algebra:

$$[t_\alpha, t_\beta] = i f_{\alpha\beta}^\gamma t_\gamma \quad (\text{B.7})$$

The *structure constants*, $f_{\alpha\beta}^\gamma$ are real valued and explicitly antisymmetric in the last two indices, and are antisymmetric in all indices if the t_α are chosen orthonormal.

The Lie algebra elements and the structure constants together constitute the *Lie algebra* of the group. They turn out to be sufficient to determine the group almost uniquely.[†]

The groups of interest in particle physics are compact Lie groups. These can all be thought of as groups of matrices. Of particular interest is the group of $N \times N$ special (unit determinant) unitary matrices, $SU(N)$, which we describe in more detail in the next section.

B.3 Group representations

We saw in Eq. (B.3) that a symmetry transformation acts on a field operator like a matrix multiplication. Successive symmetry transformations act like a series of matrix multiplications, which gives us a condition on the matrices which can appear in Eq. (B.3). Namely, under successive transformations by two group elements,

$$U(g_2 g_1) \varphi_a U^*(g_2 g_1) = U(g_2) U(g_1) \varphi_a U^*(g_1) U^*(g_2)$$

[†] A Lie group can have several disconnected pieces; the Lie algebra specifies only the connected piece containing the identity. For simple compact Lie groups, the Lie algebra gives a unique simply connected group, and any other connected group with the same Lie algebra must be a quotient of this group over a discrete identification map. For instance, the Lie algebra of rotations gives the group $SU(2)$. The group $SO(3)$ has the same Lie algebra, but differs in that a rotation by 360° , represented in $SU(2)$ by $\text{diag}[-1, -1]$, is identified with the identity in $SO(3)$.

$$\begin{aligned}
M_{ab}^{-1}(g_2 g_1) \varphi_b &= U(g_2) M_{ab}^{-1}(g_1) \varphi_b U^*(g_2) \\
M_{ac}^{-1}(g_2 g_1) \varphi_c &= M_{ab}^{-1}(g_1) M_{bc}^{-1}(g_2) \varphi_c
\end{aligned} \tag{B.8}$$

Since this must hold for any field φ , the matrices themselves must be equal,

$$M^{-1}(g_2 g_1) = M^{-1}(g_1) M^{-1}(g_2) \quad \text{or} \quad M(g_2 g_1) = M(g_2) M(g_1) \tag{B.9}$$

Matrix multiplication must respect group element multiplication. A set of matrices associated with elements of a group which satisfy this condition are called a *representation* of the group. Since the matrices can be thought of as operating on column vectors, physicists often refer to column vectors (or fields) which are multiplied by such matrices as representations. More properly, one should say that such column vectors or fields are “acted on” or “transform under” the representation. To understand the ways symmetries can act on fields and field products we must understand representations and their tensor products.

In any representation, the identity element of the group must be mapped into the identity matrix $\mathbf{1}$. An element close to the identity must map into an element close to the identity matrix, so

$$M(\mathbf{1} + i\omega^\alpha t_\alpha) = \mathbf{1} + i\omega^\alpha T_\alpha \tag{B.10}$$

with T_α some matrices particular to the representation. (We use Greek letters to index the Lie algebra and Roman letters for matrix indices.) It then follows by considering products of such matrices and using Eq. (B.9) that the matrices T_α must satisfy a Lie algebra with the same structure constants as the t_α :

$$[T_\alpha, T_\beta] = if_{\alpha\beta}^\gamma T_\gamma \tag{B.11}$$

Furthermore, any set of matrices T_α which satisfy this identity can be exponentiated to give a representation. Frequently a basis of fields can be found under which the T_α are all block diagonal, in which case the representation is said to *reduce* into the blocks. A representation which cannot be block diagonalized by any basis change is called *irreducible*. The problem of classifying representations of a group G is the problem of finding all sets of matrices T_α which obey the same commutation relations as the Lie algebra of the group.

Every group has a representation, called the *singlet* or *trivial* representation, in which $M(g)$ is the 1×1 identity matrix for each g , and $T_\alpha = 0$ for all α . (Equation (B.9) is satisfied because $1 \times 1 = 1$, and Eq. (B.11) is automatically satisfied since both sides are zero.) Invariance of the Lagrangian under a symmetry is equivalent to the requirement that the Lagrangian

transform under the singlet representation. Therefore it will be important to see how other representations can be combined together to give the singlet representation.

Every group also contains a representation called the *adjoint* representation, made up of $n \times n$ real matrices, with n the number of elements in the Lie algebra, and with T_α given by $(T_\alpha)_{bc} = -if_{bc}^\alpha$ (with b, c the matrix indices). For the case of an abelian group (a group where the f vanish) the adjoint representation is the same as the singlet representation. For the group of rotations, $SU(2)$, it is the spin-one representation.

Physicists generally refer to the several fields which transform together in an irreducible representation of the symmetry group as “a” field transforming under that representation. If two such fields ϕ, χ transform under two different representations with representation matrices M, N which are respectively $m \times m$ and $n \times n$, then the operator $\phi_a \chi_b$ transforms as

$$U(g^{-1})\phi_a \chi_b U^*(g) = M_{ac}(g)N_{bd}(g)\phi_c \chi_d \quad (\text{B.12})$$

The object $M_{ac}N_{bd}$ can be considered an $(mn) \times (mn)$ matrix obtained as the tensor product of the matrices M and N . So the product of two operators transforms under the tensor product of the representation matrices. In general such tensor products are reducible – for instance, in the familiar example of angular momentum $SU(2)$, two spin-half operators can combine into a spin-one or a spin-zero operator, because the tensor product $\frac{1}{2} \otimes \frac{1}{2}$ is reducible, $\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0$.

Representations and the rules for their tensor products are quite group dependent. We will quickly outline what happens for $U(1)$ and $SU(N)$, since these arise the most often in physics and in particular are the only groups needed in the standard model.

B.3.1 Representations of $U(1)$

The group $U(1)$ is the group of phase rotations. A generic element is $e^{i\theta}$ and the group is parametrized by θ . Any irreducible representation can be written as a 1×1 complex matrix (a complex number) and the representation is determined by a charge q , with the group element $g = e^{i\theta}$ being represented by $e^{iq\theta}$. The tensor product of two representations is just a representation with the sum of the charges. Therefore, the charge of a product of operators is the sum of their charges. For the Lagrangian to have a $U(1)$ symmetry, each term in the Lagrangian must have the charges of the fields add up to zero.

B.3.2 Representations of $SU(N)$

The group $SU(N)$ consists of complex $N \times N$ matrices U which are unitary, $U^\dagger = U^{-1}$, and satisfy $\det U = 1$. (The U in $SU(N)$ stands for unitary, the S for “special,” meaning determinant 1.)

A generic element of $SU(N)$ can be written $U(\omega) = \exp(i\omega_\alpha t_\alpha)$, with t_α a standard set of $N \times N$ complex matrices and ω_α parameters. Before imposing unitarity and determinant 1, there are $2N^2$ independent t_α . However, unitarity requires each t_α be Hermitian, eliminating half, and the unit determinant condition requires each t_α be traceless, eliminating one more possibility. Therefore there are $N^2 - 1$ independent elements t_α of the Lie algebra, which should be chosen to be orthogonal and to satisfy the same normalization condition. For $SU(2)$ they can be chosen to be half the Pauli matrices, Eq. (2.12). In this case the structure functions are

$$\left[\frac{\tau_i}{2}, \frac{\tau_j}{2} \right] = i f_{kij} \frac{\tau_k}{2}, \quad f_{kij} = \epsilon_{kij} \quad (\text{B.13})$$

the totally antisymmetric tensor. For $SU(3)$ the Lie algebra elements can be chosen to be half the Gell-Mann matrices of Eq. (1.186). There is no simple expression for the resulting structure functions.

Besides the singlet representation, the smallest representation for $SU(N)$ is the $SU(N)$ matrices themselves, $M(U) = U$. This is called the *fundamental representation*.

$$\text{Fundamental representation: } T_\alpha = t_\alpha \quad (\text{B.14})$$

This is, for instance, the representation quarks transform under in QCD. It is customary to refer to the representations of $SU(N)$ according to the rank of the representation matrices, so the singlet representation is called the **1** representation and the fundamental representation is called the **N** representation.

Equally important is the *antifundamental representation*, $M(U) = U^*$, given by complex conjugating (but not transposing) the $SU(N)$ matrices.

$$\text{Antifundamental representation: } T_\alpha = -t_\alpha^* \quad (\text{B.15})$$

This is the representation which antiquarks transform under. To see that it is a valid representation, note that

$$\left[-t_\alpha^*, -t_\beta^* \right] = \left(\left[t_\alpha, t_\beta \right] \right)^* = (i f_{\alpha\beta}^\gamma t_\gamma)^* = i f_{\alpha\beta}^\gamma (-t_\gamma^*) \quad (\text{B.16})$$

Therefore the $-t_\alpha^*$ obey the same commutation relations as the t_α . These matrices have rank N , but since the symbol **N** is taken, the representation is called the $\overline{\mathbf{N}}$ representation.

A field which transforms under the fundamental representation is left-multiplied by U , and can be thought of as a column vector. For the antifundamental representation, the column vector is left-multiplied by U^* , and is more conveniently thought of as a row vector right-multiplied by $U^\dagger = U^{-1}$. In general, if M is a representation of a group, M^* is also, and is called the *conjugate representation* to M . The contraction of a fundamental and an antifundamental field (or any operators in conjugate representations to each other) forms a singlet,

$$\text{for } \chi_a \bar{\mathbf{N}}, \quad \phi_a \mathbf{N}, \quad \chi^T \phi \equiv \chi_a \phi_a \text{ is singlet } \mathbf{1} \quad (\text{B.17})$$

At the same time, inserting the generators of the representation between them,

$$\text{for } \chi_a \bar{\mathbf{N}}, \quad \phi_a \mathbf{N}, \quad \chi^T T_\alpha \phi \equiv \chi_a (T_\alpha)_{ab} \phi_b \text{ is adjoint} \quad (\text{B.18})$$

Since there are $N^2 - 1$ elements in the adjoint representation and one in the singlet, this uses up the $N \times N = N^2$ objects in the tensor product of fundamental and antifundamental representations;

$$\bar{\mathbf{N}} \otimes \mathbf{N} = (\mathbf{N}^2 - \mathbf{1}) \oplus \mathbf{1} \quad (\text{B.19})$$

It is also important to know how multiple fundamental representations tensor together. Here it is important to know that the totally antisymmetric object $\epsilon_{ab\dots}$, which contracts N fundamental indices, is an invariant. Using it to contract $N - 1$ objects transforming in the fundamental representation gives an object transforming in the antifundamental representation. For $SU(2)$, this means that a single fundamental representation object can be “flipped” into an antifundamental representation object by ϵ , as we did with the Higgs field in Eq. (2.24). In $SU(2)$ the fundamental and antifundamental representations are equivalent and generally not distinguished from each other. In $SU(3)$, contracting two fundamental fields with the antisymmetric tensor, $\epsilon_{abc} \phi_b \psi_c$, produces an antifundamental object, and three gives a singlet. The other (symmetrized) linear combination of two fundamental fields has six components and is called the **6** representation:

$$\mathbf{3} \otimes \mathbf{3} = \bar{\mathbf{3}} \oplus \mathbf{6} \quad \text{in } SU(3) \quad (\text{B.20})$$

More generally one gets representations containing $N(N - 1)/2$ and $N(N + 1)/2$ elements. In general, $SU(N)$ groups have a large number of representations, all of which can be found by taking antisymmetrized and symmetrized combinations of fundamental representation objects. Further enumerating them is beyond the scope of this appendix.

Appendix C

Lorentz group and the Dirac algebra

This appendix provides a review and summary of the Lorentz group, its properties, and the properties of its infinitesimal generators. It then reviews representations of the Lorentz group and the Dirac algebra. This material is intended to supplement Chapter 1, for those students who are not as familiar with the Lorentz group and Dirac equation as they find they need to be.

C.1 Lorentz group

According to special relativity, physical laws are unchanged by a linear change of coordinates,

$$x'^{\mu} = \Lambda^{\mu}_{\nu} x^{\nu} + \xi^{\mu} \quad (\text{C.1})$$

with Λ and ξ real, provided it leave unchanged the invariant separation between points,

$$(x - y)_{\mu} (x - y)^{\mu} = \eta_{\mu\nu} (x - y)^{\mu} (x - y)^{\nu} = -[(x - y)^0]^2 + [\vec{x} - \vec{y}]^2$$

This condition does not constrain ξ , since it cancels in the difference, but it imposes a constraint on Λ ,

$$x_{\mu} x^{\mu} = x'_{\mu} x'^{\mu} = \eta_{\mu\nu} \Lambda^{\mu}_{\alpha} x^{\alpha} \Lambda^{\nu}_{\beta} x^{\beta} \quad (\text{C.2})$$

for all x^{μ} . A transformation of the form shown in Eq. (C.1) which satisfies Eq. (C.2) is called a Poincaré transformation. These transformations close and form a group, called the Poincaré group. The subset where Λ is the identity matrix and ξ is arbitrary is a subgroup called the group of translations. We assume that this group and its implications, such as conservation of energy and momentum, are familiar to the reader. Instead we concentrate on the subgroup in which $\xi = 0$, which is called the Lorentz group.

It is convenient to think of an element of the Lorentz group as a matrix

which is operating on the coordinate x^μ . This is possible if we always write Λ with its first index raised and second index lowered, so it carries x^μ to x'^μ , both with raised index. Writing in this way, repeated Lorentz transformations are implemented via matrix multiplications of the respective Λ matrices:

$$x'^\mu = \Lambda^\mu_\nu x^\nu \text{ and } x''^\mu = \Lambda'^\mu_\nu x'^\nu \Rightarrow x''^\mu = \Lambda'^\mu_\nu \Lambda^\nu_\alpha x^\alpha \equiv \left[\Lambda' \Lambda \right]^\mu_\alpha x^\alpha \quad (\text{C.3})$$

We see from Eq. (C.2) that the condition on Λ^μ_ν to be a Lorentz transformation is

$$\eta_{\mu\nu} x^\mu x^\nu = \eta_{\alpha\beta} \Lambda^\alpha_\mu \Lambda^\beta_\nu x^\mu x^\nu \quad (\text{C.4})$$

for all x^μ . Since this must hold for all x^μ , we have

$$\eta_{\mu\nu} = \Lambda^\alpha_\mu \eta_{\alpha\beta} \Lambda^\beta_\nu \quad (\text{C.5})$$

or (writing $\eta_{\mu\nu}$ as η when using matrix notation)

$$\eta = \Lambda^T \Lambda \quad (\text{C.6})$$

The group of matrices Λ satisfying Eq. (C.6) is called $O(3,1)$, and is a Lie group. Therefore the same technology of Lie algebra generation may be applied to it as to the groups of the previous appendix.

As we will discuss momentarily, not all elements of $O(3,1)$ can be built infinitesimally from the identity. Those elements which can, form a subgroup written $SO(3,1)$, which we will now analyze. A Lorentz transformation Λ^μ_ν which is infinitesimally close to the identity must be of the form,

$$\Lambda^\mu_\nu = \delta^\mu_\nu + \omega^\mu_\nu \quad (\text{C.7})$$

with ω^μ_ν a matrix of infinitesimal coefficients. The condition on ω^μ_ν for Λ^μ_ν to be a valid Lorentz transformation is found by inserting Eq. (C.7) into Eq. (C.6) and expanding to linear order in ω :

$$\begin{aligned} \eta_{\mu\nu} &= (\delta^\alpha_\mu + \omega^\alpha_\mu) \eta_{\alpha\beta} (\delta^\beta_\nu + \omega^\beta_\nu) \\ &= \eta_{\mu\nu} + (\omega_{\nu\mu} + \omega_{\mu\nu}) + O(\omega^2) \\ 0 &= \omega_{\nu\mu} + \omega_{\mu\nu} \end{aligned} \quad (\text{C.8})$$

That is, the condition on ω^μ_ν is that $\omega_{\mu\nu}$ be antisymmetric on its indices. The space of real antisymmetric 4×4 matrices is six-dimensional, so the Lorentz group is six-dimensional.

Now ω^μ_ν is related to $\omega_{\mu\nu}$ as $\omega^\mu_\nu = \eta^{\mu\alpha} \omega_{\alpha\nu}$. Since $\eta^{00} = -1$ and $\eta^{ii} = 1$ for $i = 1, 2, 3$, the sign of the spacetime component of ω^μ_ν must be the same as

the sign of the time-space component, while the space-space components must be antisymmetric. Thus, the most general form of ω_ν^μ is

$$\omega_\nu^\mu = \begin{pmatrix} 0 & b_1 & b_2 & b_3 \\ b_1 & 0 & -r_3 & r_2 \\ b_2 & r_3 & 0 & -r_1 \\ b_3 & -r_2 & r_1 & 0 \end{pmatrix} \quad (\text{C.9})$$

symmetric in the space-time entries and antisymmetric in the space-space entries. The b_1, b_2, b_3 entries respectively cause infinitesimal boosts in the 1, 2, 3 directions; the r_1, r_2, r_3 entries cause infinitesimal rotations about the 1, 2, 3 axes. A general element of $SO(3, 1)$ can be written as an exponential of a *finite* ω_ν^μ ,

$$\Lambda^\mu{}_\nu = (\exp \omega)_\nu^\mu = \delta_\nu^\mu + \omega_\nu^\mu + \frac{1}{2}\omega_\alpha^\mu \omega_\nu^\alpha + \frac{1}{6}\omega_\alpha^\mu \omega_\beta^\alpha \omega_\nu^\beta + \dots \quad (\text{C.10})$$

When only the r_i are nonzero, this gives a rotation by angle $|\vec{r}|$ about the \hat{r} axis. When only the b_i are nonzero, this gives a boost by velocity $\tanh |\vec{b}|$ along the \hat{b} axis. When both \vec{r} and \vec{b} are nonzero the Lorentz transformation cannot be described either solely as a rotation or as a boost. Note that, while a rotation by angle $|\vec{r}| = 2\pi$ gives the identity Λ , no nonzero magnitude of boost $|\vec{b}|$ returns the identity. Hence the group $SO(3, 1)$ is *noncompact*.

Now we argue that $O(3, 1)$ has four disconnected pieces, one of which is $SO(3, 1)$. To see this, take the determinant of Eq. (C.6):

$$\det \eta = \det \Lambda_\eta^T \Lambda = \det \eta \times (\det \Lambda)^2 \quad (\text{C.11})$$

Since η is nonsingular, we can divide by $\det \eta$:

$$(\det \Lambda)^2 = 1 \quad \Rightarrow \quad \det \Lambda = \pm 1 \quad (\text{C.12})$$

The determinant must vary continuously within a path connected region of $O(3, 1)$, but you cannot go continuously from 1 to -1 , so any elements of $O(3, 1)$ with $\det \Lambda = -1$ cannot be elements of the connected group $SO(3, 1)$. An element of $O(3, 1)$ with $\det \Lambda = 1$ is called *proper*, and an element with $\det \Lambda = -1$ is called *improper*.

Furthermore, if we write out the $\mu = 0, \nu = 0$ element of Eq. (C.6), it is

$$\begin{aligned} \eta_{00} &= \Lambda_0^\mu \eta_{\mu\nu} \Lambda_0^\nu \\ -1 &= -\Lambda_0^0 \Lambda_0^0 + \sum_{i=1,2,3} \Lambda_0^i \Lambda_0^i \\ (\Lambda_0^0)^2 &= 1 + \sum_{i=1,2,3} (\Lambda_0^i)^2 \geq 1 \end{aligned} \quad (\text{C.13})$$

so the square of the time-time component of any Λ must always be at least

1, and Λ_0^0 must be either ≥ 1 or ≤ -1 . Again, you cannot go continuously from ≥ 1 to ≤ -1 , so no elements of $SO(3,1)$ have $\Lambda_0^0 < 0$. An element of $O(3,1)$ with $\Lambda_0^0 \geq 1$ is called *orthochronous*, and an element with $\Lambda_0^0 \leq -1$ is called *nonorthochronous*.

The canonical example of an improper (but orthochronous) element of $O(3,1)$ is the *parity transformation*,

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (\text{C.14})$$

which satisfies Eq. (C.6) but has determinant -1 . The canonical example of a nonorthochronous (and also improper) transformation is the *time reversal transformation*,

$$T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (\text{C.15})$$

which also satisfies Eq. (C.6) but has $T^0_0 = -1$. It turns out that any element of $O(3,1)$ must be an element of $SO(3,1)$, times either the identity (proper orthochronous), P (improper orthochronous), T (improper nonorthochronous), or PT (proper nonorthochronous). The improper or nonorthochronous Lorentz transformations need not be symmetries of nature – in fact, in the standard model, they are not – but it is an axiom of field theory that the elements of $SO(3,1)$ must be symmetries.

C.2 Generators of the Lorentz group

As discussed in Section B.1, each element $\Lambda \in SO(3,1)$ must have associated with it a unitary operator $U(\Lambda)$ which implements it on the Hilbert space, and which represents the group operation

$$U(\Lambda_1)U(\Lambda_2) = U(\Lambda_1\Lambda_2) \quad (\text{C.16})$$

For an element infinitesimally close to the identity, it must be possible to expand these operators in a Lie algebra of generators,

$$U(\omega) = \mathbf{1} + \frac{i}{2}\omega_{\mu\nu}\hat{J}^{\mu\nu} + O(\omega^2) \quad (\text{C.17})$$

for some operators $\hat{J}^{\mu\nu}$, antisymmetric in μ, ν . Similarly, there are generators for translations,

$$U(\xi) = \mathbf{1} - i\xi_\mu \hat{P}^\mu \quad (\text{C.18})$$

The \hat{P}^i are also called momentum operators, and the \hat{J}^{ij} are called angular momentum operators.

The commutation relations between the operators \hat{P}^μ , $\hat{J}^{\mu\nu}$ can be worked out by using Eq. (C.16). For instance, consider a translation by a small distance ξ^μ , either preceded or followed by a Lorentz transformation involving ω_α^ν . We will evaluate the difference between the two orders of operation, in two ways.

First, if we translate first and then rotate, then the coordinate is transformed according to

$$\begin{aligned} x'^\mu &= x^\mu + \xi^\mu \\ x''^\mu &= (\delta_\nu^\mu + \omega_\nu^\mu)(x^\nu + \xi^\nu) \\ &= x^\mu + (\omega_\nu^\mu x^\nu) + (\xi^\mu + \omega_\nu^\mu \xi^\nu) \end{aligned} \quad (\text{C.19})$$

where the first and second parenthesis represent a rotation and a translation. The result is the same rotation, and a translation by ξ^μ plus an extra piece involving ω and ξ . If the rotation is performed first, we get

$$\begin{aligned} x'^\mu &= x^\mu + \omega_\nu^\mu x^\nu \\ x''^\mu &= x^\mu + (\omega_\nu^\mu x^\nu) + (\xi^\mu) \end{aligned} \quad (\text{C.20})$$

which is the rotation and the translation just by ξ . The unitary operators for these transformations are

$$\begin{aligned} U(\omega\xi) &= \mathbf{1} + \frac{i}{2}\omega_{\mu\nu}\hat{J}^{\mu\nu} - i\eta_{\mu\nu}(\xi^\mu + \omega_\alpha^\mu \xi^\alpha)\hat{P}^\nu \\ U(\xi\omega) &= \mathbf{1} + \frac{i}{2}\omega_{\mu\nu}\hat{J}^{\mu\nu} - i\eta_{\mu\nu}(\xi^\mu)\hat{P}^\nu \end{aligned} \quad (\text{C.21})$$

The difference of the operators, to second order in the infinitesimals, is

$$U(\omega\xi) - U(\xi\omega) = -i\eta_{\mu\nu}\omega_\alpha^\mu \xi^\alpha \hat{P}^\nu \quad (\text{C.22})$$

(There is actually also a second order in ω piece, but it is the same for the two U s and therefore cancels in this difference.)

Alternately, we can say using Eq. (C.16) that

$$\begin{aligned} U(\omega\xi) = U(\omega)U(\xi) &= \left(\mathbf{1} + \frac{i}{2}\omega_{\mu\nu}\hat{J}^{\mu\nu}\right) \left(\mathbf{1} - i\eta_{\alpha\beta}\xi^\alpha \hat{P}^\beta\right) \\ U(\xi\omega) = U(\xi)U(\omega) &= \left(\mathbf{1} - i\eta_{\alpha\beta}\xi^\alpha \hat{P}^\beta\right) \left(\mathbf{1} + \frac{i}{2}\omega_{\mu\nu}\hat{J}^{\mu\nu}\right) \end{aligned}$$

$$U(\omega\xi) - U(\xi\omega) = -\frac{i^2}{2}\omega_{\mu\nu}\eta_{\alpha\beta}\xi^\alpha [\hat{J}^{\mu\nu}, \hat{P}^\beta] \quad (\text{C.23})$$

Now equating Eq. (C.22) and Eq. (C.23), we learn what the commutator of \hat{P} with \hat{J} must be

$$-\frac{i^2}{2}\omega_{\mu\nu}\xi_\alpha [\hat{J}^{\mu\nu}, \hat{P}^\alpha] = -i\omega_{\mu\nu}\xi_\alpha \eta^{\nu\alpha} \hat{P}^\mu \quad (\text{C.24})$$

This must hold for any antisymmetric $\omega_{\mu\nu}$ and any ξ_α , so (antisymmetrizing over the indices on ω) the operators must satisfy

$$[\hat{J}^{\mu\nu}, \hat{P}^\alpha] = i \left(\eta^{\mu\alpha} \hat{P}^\nu - \eta^{\nu\alpha} \hat{P}^\mu \right) \quad (\text{C.25})$$

By a completely analogous but more involved procedure one can also show

$$[\hat{J}^{\mu\nu}, \hat{J}^{\alpha\beta}] = i \left(\eta^{\nu\beta} \hat{J}^{\mu\alpha} + \eta^{\mu\alpha} \hat{J}^{\nu\beta} - \eta^{\mu\beta} \hat{J}^{\nu\alpha} - \eta^{\nu\alpha} \hat{J}^{\mu\beta} \right) \quad (\text{C.26})$$

and (this is simpler)

$$[\hat{P}^\mu, \hat{P}^\nu] = 0 \quad (\text{C.27})$$

These commutation relations are called the Poincaré algebra.

To make contact with the more familiar generators of rotations and boosts, it is convenient to define

$$\hat{J}_i \equiv \frac{\epsilon_{ijk}}{2} \hat{J}_{jk} \quad (\text{C.28})$$

$$\hat{K}_i \equiv \hat{J}_i^0 \quad (\text{C.29})$$

which are respectively the generator of rotations about the i axis and of boosts along the i axis, so a rotation by $\vec{\theta}$ is $\exp(-iJ_i\theta_i)$ and a boost by \vec{v} is $\exp(-iK_i v_i)$. They satisfy the commutation relations, following from Eq. (C.26),

$$[\hat{J}_i, \hat{J}_j] = i\epsilon_{ijk} \hat{J}_k \quad (\text{C.30})$$

$$[\hat{J}_i, \hat{K}_j] = i\epsilon_{ijk} \hat{K}_k \quad (\text{C.31})$$

$$[\hat{K}_i, \hat{K}_j] = -i\epsilon_{ijk} \hat{J}_k \quad (\text{C.32})$$

The first expression is the familiar commutator between rotations. The second means that, if a rotation is performed before a boost, the boost will be in a different direction than before the rotation is performed, which is intuitively clear. The third result is more surprising; the commutator of two boosts is a rotation. More importantly, the sign is opposite on the last result than on the previous two.

C.3 Representations of the Lorentz group

Just as for an internal symmetry, an $SO(3, 1)$ transformation will carry a field to a linear combination of fields, so the fields must transform under representations of the group. The difference is that the transformed field will be at the Lorentz transformed point:

$$U(\omega)\varphi_a(x)U^*(\omega) = D_{ab}^{-1}(\omega)\varphi_b(\Lambda x) \quad (\text{C.33})$$

with $D^{-1}(\omega) = D(-\omega)$ an ω dependent matrix in the space of fields. The fields can be chosen to block-diagonalize the matrix D into irreducible representations of the Lorentz group. For instance, in QED, the components of the gauge potential A^μ mix with each other under Lorentz transformations, but they never mix with the different spin components of the electron e_i ; so there is one “block” of D which mixes the A^μ and an independent block mixing the e_i . Our goal now is to find the possible structures D can take.

Just as for internal symmetries, there are two very simple irreducible representations, which are also physically important. One is the trivial (singlet) representation,

$$U(\omega)\phi(x)U^*(\omega) = \phi(\Lambda x) \quad (\text{C.34})$$

which for $SO(3, 1)$ is called the *scalar representation*. Lorentz symmetry demands that the Lagrangian be a Lorentz scalar. The other is the *vector representation*, for which the field index is a four-vector index and the representation matrix is Λ itself:

$$U(\omega)A^\mu(x)U^*(\omega) = (\Lambda^{-1})^\mu_\nu A^\nu(\Lambda x) \quad (\text{C.35})$$

A representation is determined by a set of matrices $\mathcal{J}^{\mu\nu}$ with the same commutation relations as the $\hat{J}^{\mu\nu}$. That is, the matrix D_{ab} must be of the form

$$D_{ab}(\omega) = \exp\left(\frac{i}{2}\omega_{\mu\nu}\mathcal{J}_{ab}^{\mu\nu}\right) \quad (\text{C.36})$$

with the exponentiation interpreted as matrix exponentiation with a, b the matrix indices, and $\mathcal{J}^{\mu\nu}$ satisfying

$$[\mathcal{J}^{\mu\nu}, \mathcal{J}^{\alpha\beta}] = i\left(\eta^{\nu\beta}\mathcal{J}^{\mu\alpha} + \eta^{\mu\alpha}\mathcal{J}^{\nu\beta} - \eta^{\mu\beta}\mathcal{J}^{\nu\alpha} - \eta^{\nu\alpha}\mathcal{J}^{\mu\beta}\right) \quad (\text{C.37})$$

The problem of finding representations is the problem of finding all sets of matrices with this algebra.

It is believed that only field theories containing finite numbers of fields are well defined. Therefore we need only look for finite-dimensional representations of $SO(3, 1)$. The classification of the representations is made easier by

the following convenient property of the group. Using Eq. (C.30)–Eq. (C.32), one can show that the operators

$$\hat{\mathcal{L}}_i \equiv \frac{\hat{J}_i + i\hat{K}_i}{2}, \quad \hat{\mathcal{R}}_i \equiv \frac{\hat{J}_i - i\hat{K}_i}{2} \quad (\text{C.38})$$

satisfy the commutation relations

$$[\hat{\mathcal{L}}_i, \hat{\mathcal{L}}_j] = i\epsilon_{ijk}\hat{\mathcal{L}}_k \quad (\text{C.39})$$

$$[\hat{\mathcal{R}}_i, \hat{\mathcal{R}}_j] = i\epsilon_{ijk}\hat{\mathcal{R}}_k \quad (\text{C.40})$$

$$[\hat{\mathcal{L}}_i, \hat{\mathcal{R}}_j] = 0 \quad (\text{C.41})$$

Therefore the generators of $SO(3,1)$ can be split into two subsets which commute with each other, and each satisfy the same commutation relations as the group $SU(2)$. This group is familiar as the group of rotations and its representations are well known; they are the spin-zero representation, the spin-half representation, the spin-one representation, and so forth. A general irreducible representation can be described by its transformation properties under $\hat{\mathcal{L}}$ and under $\hat{\mathcal{R}}$, e.g., spin- $m/2$ under $\hat{\mathcal{L}}$ and spin- $n/2$ under $\hat{\mathcal{R}}$.

Only four representations will be of any interest to us in studying the standard model, because it turns out that only four representations can participate in renormalizable interactions in a theory satisfying the basic principles laid out in Section 1.2.

The first of these is the scalar representation already introduced. The scalar representation transforms as $(0,0)$, that is, as spin-zero under $\hat{\mathcal{L}}$ and spin-zero under $\hat{\mathcal{R}}$. The Lie algebra representations are $\hat{\mathcal{J}}^{\mu\nu} = 0$ and the transformation matrix $D = 1$ is the identity.

The second common representation is the vector representation, which transforms as $(\frac{1}{2}, \frac{1}{2})$. The Lie algebra is represented as $\mathcal{J}_{\alpha\beta}^{\mu\nu} = -i(\eta^{\mu\alpha}\eta_{\beta}^{\nu} - \eta_{\beta}^{\mu}\eta^{\nu\alpha})$, and $D = \Lambda$, as already discussed. Note that for both of these representations, the matrix D is always real; therefore it is consistent to consider real-valued scalar or vector fields.

The other two interesting representations are called *spinor representations*, and consist of two fields which mix with each other under Lorentz transformations. Since these are probably less familiar to the reader and are in some ways more complicated than the scalar and vector representations, we will discuss them at length in the next section.

C.4 Spinors and the Dirac algebra

We now introduce the other two physically important representations of the Lorentz group, the left- and right-handed spinor representations. A field transforming in one of these can be rewritten in terms of the other, and it is convenient to combine them together using *Majorana notation*, which we will also introduce and which we use throughout this book.

C.4.1 Spinor representations

The simplest nontrivial matrices which satisfy Eq. (C.39) are the Pauli matrices,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (\text{C.42})$$

which satisfy the commutation relation

$$\left[\frac{\sigma_i}{2}, \frac{\sigma_j}{2} \right] = i\epsilon_{ijk} \frac{\sigma_k}{2} \quad (\text{C.43})$$

Therefore, if the matrices representing $\hat{\mathcal{L}}_i$ and $\hat{\mathcal{R}}_i$ are $\sigma_i/2$ and 0 respectively, we get a representation of the Lorentz algebra. Inverting Eq. (C.38), rotations and boosts are implemented by the matrices,

$$\mathcal{J}_i = \frac{\sigma_i}{2}, \quad \mathcal{K}_i = -i\frac{\sigma_i}{2}, \quad (\text{left-handed spinor}) \quad (\text{C.44})$$

which it is easy to show satisfy Eq. (C.30) through Eq. (C.32).

Therefore, a pair of fields $\psi_a, a = 1, 2$ can transform under Lorentz transformations according to

$$U(-\omega)\psi_a U^*(-\omega) = D_{ab}(\omega)\psi_b, \quad D(\omega) = [\exp(-i(r_i - ib_i)\sigma_i/2)] \quad (\text{C.45})$$

where r_i, b_i are the amount of rotation and boost performed, as introduced in Eq. (C.9). The two fields ψ_a are generally referred to as the components of a single *spinor field* with a the *spinor index*, which is almost always suppressed by writing ψ and D in matrix notation (ψ as a column vector, D as a matrix). Such a spinor field is called a *left-handed Weyl spinor* ψ_L .

Alternately, \mathcal{R}_i could be represented by the Pauli matrices and \mathcal{L}_i by 0s,

$$\mathcal{J}_i = \frac{\sigma_i}{2}, \quad \mathcal{K}_i = i\frac{\sigma_i}{2}, \quad (\text{right-handed spinor}) \quad (\text{C.46})$$

in which case a Lorentz transform acts on ψ via

$$U(-\omega)\psi_a U^*(-\omega) = D_{ab}(\omega)\psi_b, \quad D(\omega) = [\exp(-i(r_i + ib_i)\sigma_i/2)] \quad (\text{C.47})$$

A field transforming this way is called a *right-handed Weyl spinor* ψ_R .

Since the matrices D we just constructed are in general complex, a spinor ψ_L or ψ_R must be a pair of complex fields. We can ask how the complex conjugate of ψ_L transforms. Because complex conjugation flips the i in front of \mathcal{K} in Eq. (C.44), the answer is that it transforms as a right-handed Weyl spinor. More properly, defining the matrix

$$\epsilon \equiv i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{satisfying} \quad \epsilon \sigma_i^* = -\sigma_i \epsilon \quad (\text{C.48})$$

we see that ϵ times the conjugate of ψ_L transforms according to

$$\begin{aligned} U(-\omega)\epsilon\psi_L^*U^*(-\omega) &= \epsilon \left(\exp \left[-i(r_i - ib_i)\frac{\sigma_i}{2} \right] \psi_L \right)^* \\ &= \epsilon \exp \left[+i(r_i + ib_i)\frac{\sigma_i^*}{2} \right] \psi_L^* \\ &= \exp \left[-i(r_i + ib_i)\frac{\sigma_i}{2} \right] \epsilon\psi_L^* \end{aligned} \quad (\text{C.49})$$

which is precisely the transformation rule for a right-handed Weyl spinor. Similarly, $-\epsilon\psi_R^*$ transforms as a left-handed Weyl spinor, and $-\epsilon(\epsilon\psi_L^*)^* = \psi_L$ transforms as a left-handed Weyl spinor again. Both the field and its complex conjugate will typically appear in the Lagrangian so it is important to have a notation which can deal with each. Whether we consider the left- or right-handed version as the field rather than the conjugated object is a matter of convention.

C.4.2 Weyl, Majorana, Dirac

There are two common notational ways of dealing with the fact that a field can be written either as a left- or a right-handed spinor.

One, called *Weyl notation*, expresses the fields as two component objects, and then specifies whether one is referring to ψ_L or to its right-handed conjugate $\epsilon\psi_L^*$ by using either an undotted or a dotted index: $\psi_\alpha = \psi_L$ and $\psi^{\dot{\alpha}} = \epsilon\psi_L^*$. (Indices are raised and lowered using ϵ and dotted and undotted according to whether they are conjugated.) This notation is common in the supersymmetry and string theory literature.

An alternative which we will use, *Majorana notation*, writes a single *four* component field ψ_M , defined as

$$\psi_M = \begin{pmatrix} \psi_L \\ \epsilon\psi_L^* \end{pmatrix} \quad (\text{C.50})$$

that is, ψ_M redundantly records both the left-handed and the right-handed

ways of writing the field. The individual pieces can be accessed separately by using the projection operators

$$P_L \equiv \begin{pmatrix} \mathbf{1} & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad P_R \equiv \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{1} \end{pmatrix} \quad (\text{C.51})$$

The action of rotations and boosts on ψ_M are respectively,

$$\mathcal{J}_i = \begin{pmatrix} \frac{\sigma_i}{2} & 0 \\ 0 & \frac{\sigma_i}{2} \end{pmatrix}, \quad \mathcal{K}_i = \begin{pmatrix} \frac{-i\sigma_i}{2} & 0 \\ 0 & \frac{i\sigma_i}{2} \end{pmatrix} \quad (\text{C.52})$$

If a left-handed spinor transforms nontrivially under an internal symmetry group, then since the right-handed version involves complex conjugation, the right-handed version $\epsilon\psi_L^*$ transforms under the conjugate representation. In particular, if ψ_L has charge q under a $U(1)$ symmetry and is in the fundamental representation of an $SU(N)$ symmetry, then $\epsilon\psi_L^*$ has charge $-q$ and transforms under the antifundamental representation of $SU(N)$. One must keep this in mind when constructing Lagrangians out of Majorana spinors.

In QED and QCD, if we write the spinor fields as left-handed objects, the fields form pairs with conjugate symmetry transformation properties. For instance, in QED, there is a field \mathcal{E}_L which is charge -1 under $U_{\text{em}}(1)$, called the left-handed electron, and a field $-\epsilon E_R^*$ which is charge 1 under $U_{\text{em}}(1)$, called the left-handed positron. In this case it is most convenient to think of the latter as the conjugate of a right-handed field with charge -1 , E_R , called the right-handed electron, and to combine them together in a single 4-component object called a *Dirac spinor*, $e = [\mathcal{E}_L E_R]^T$.

The Lorentz transformation properties of Majorana and Dirac spinors are the same. The two distinctions are that the upper and lower components of a Dirac spinor generally have the same transformation properties under internal symmetries, while for Majorana spinors they have conjugate transformation properties; and the upper and lower components of a Dirac spinor are independent, while for a Majorana spinor they are redundant notations for the same field.

C.4.3 Tensor products of spinors

Since the Lagrangian must be a Lorentz scalar, it must be a sum of terms even in spinorial fields. Therefore we need to know how products of two spinor fields transform. We will only consider the combination of a spinor field ψ_1 with the complex conjugate of another, ψ_2^\dagger . This is sufficient for Majorana spinors because $\psi_2^T \psi_1$ can be re-expressed in terms of $\psi_2^\dagger \psi_1$, and

it suffices for Dirac spinors with internal symmetries because only such combinations are invariant under the internal symmetries.

The Hermitian conjugate of a spinor field ψ transforms as

$$U(-\omega)\psi^\dagger U^*(-\omega) = (D(\omega)\psi)^\dagger = \psi^\dagger D^\dagger(\omega) \quad (\text{C.53})$$

The \mathcal{J}_i are Hermitian, but the \mathcal{K}_i are anti-Hermitian, so $D(\omega)$ is not in general unitary. Therefore ψ^\dagger does not have the inverse transformation property of ψ . However, there is a Hermitian, unit determinant matrix β ,

$$\beta \equiv \begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix}, \quad \beta \mathcal{J}_i = \mathcal{J}_i \beta, \quad \beta \mathcal{K}_i = -\mathcal{K}_i \beta \quad (\text{C.54})$$

which flips the sign of \mathcal{K} but not \mathcal{J} when commuted across D^\dagger , so $D^\dagger \beta = \beta D^{-1}$. Therefore, defining $\bar{\psi} = \psi^\dagger \beta$, called the *Dirac conjugate* of ψ ,

$$U(-\omega)\bar{\psi}\beta U^*(-\omega) = \psi^\dagger D^\dagger(\omega)\beta = \psi^\dagger \beta D^{-1}(\omega) = \bar{\psi} D^{-1}(\omega) \quad (\text{C.55})$$

so $\bar{\psi}$ has the inverse transformation property of ψ .

Since ψ has four components, there are sixteen independent 4×4 matrices Γ which can be used to combine spinors, $\bar{\psi}_2 \Gamma \psi_1$. These can all be gotten from four such matrices, called the *gamma matrices* γ^μ , given in Eq. (1.87). These satisfy anticommutation relations called the *Clifford algebra*,

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \mathbf{1} \quad (\text{C.56})$$

The matrices $\mathcal{J}^{\mu\nu}$ can be expressed in terms of the gamma matrices:

$$\mathcal{J}^{\mu\nu} = \frac{-i}{4} [\gamma^\mu, \gamma^\nu] \quad (\text{C.57})$$

which together with Eq. (C.56) is enough to prove that $\mathcal{J}^{\mu\nu}$ satisfies the Lorentz algebra, Eq. (C.37). Further, these relations ensure that

$$[\mathcal{J}^{\mu\nu}, \gamma^\alpha] = i(\eta^{\mu\alpha}\gamma^\nu - \eta^{\nu\alpha}\gamma^\mu) \quad (\text{C.58})$$

from which it follows that

$$D^{-1}(\omega)\gamma^\mu D(\omega) = \Lambda^\mu_\nu \gamma^\nu \quad (\text{C.59})$$

Therefore, while the combination $\bar{\psi}_2 \psi_1$ is a scalar

$$U(-\omega)\bar{\psi}_2 \psi_1 U^{-1}(-\omega) = \bar{\psi}_2 D^{-1}(\omega) D(\omega) \psi_1 = \bar{\psi}_2 \psi_1 \text{ is a scalar} \quad (\text{C.60})$$

the combination $\bar{\psi}_2 \gamma^\mu \psi_1$ is a vector,

$$U(-\omega)\bar{\psi}_2 \gamma^\mu \psi_1 U^{-1}(-\omega) = \bar{\psi}_2 D^{-1}(\omega) \gamma^\mu D(\omega) \psi_1 = \Lambda^\mu_\nu \bar{\psi}_2 \gamma^\nu \psi_1 \text{ is a vector} \quad (\text{C.61})$$

Similarly, defining $\sigma^{\mu\nu} = 2i\mathcal{J}^{\mu\nu}$, the combination

$$U(-\omega)\bar{\psi}_2\sigma^{\mu\nu}\psi_1U^{-1}(-\omega) = \Lambda_\alpha^\mu\Lambda_\beta^\nu\bar{\psi}_2\sigma^{\alpha\beta}\psi_1 \quad (\text{C.62})$$

is a rank-2 antisymmetric tensor.

Next, define

$$\gamma^5 = \gamma_5 \equiv \frac{i}{24}\epsilon_{\mu\nu\alpha\beta}\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta = -i\gamma^0\gamma^1\gamma^2\gamma^3 \quad (\text{C.63})$$

where the latter follows from the anticommutation of the distinct gamma matrices. We have that

$$\begin{aligned} U(-\omega)\bar{\psi}_2\gamma_5\psi_1U^{-1}(-\omega) &= \frac{i}{24}\epsilon_{\mu\nu\alpha\beta}\bar{\psi}_2D^{-1}\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta D\psi_1 \\ &= \frac{i}{24}\epsilon_{\mu\nu\alpha\beta}\Lambda_\sigma^\mu\Lambda_\rho^\nu\Lambda_\kappa^\alpha\Lambda_\zeta^\beta\bar{\psi}_2\gamma^\sigma\gamma^\rho\gamma^\kappa\gamma^\zeta\psi_1 \\ &= (\det\Lambda)\frac{i}{24}\epsilon_{\sigma\rho\kappa\zeta}\bar{\psi}_2\gamma^\sigma\gamma^\rho\gamma^\kappa\gamma^\zeta\psi_1 \\ &= (\det\Lambda)\bar{\psi}_2\gamma_5\psi_1 \end{aligned} \quad (\text{C.64})$$

Therefore $\bar{\psi}_2\gamma_5\psi_1$ is a pseudoscalar, a scalar under $SO(3,1)$ which flips sign under parity transformations. Finally, the quantity $\bar{\psi}_2\gamma^\mu\gamma_5\psi$ transforms as a pseudovector,

$$U(-\omega)\bar{\psi}_2\gamma^\mu\gamma_5\psi_1U^{-1}(-\omega) = \bar{\psi}_2D^{-1}\gamma^\mu\gamma_5D\psi_1 = (\det\Lambda)\Lambda^\mu{}_\nu\bar{\psi}_2\gamma^\nu\gamma_5\psi_1 \quad (\text{C.65})$$

Since this gives $1 + 4 + 6 + 4 + 1 = 16$ independent contractions, the above are exhaustive; any other matrix sandwiched between $\bar{\psi}_2$ and ψ_1 must be a linear combination of $\mathbf{1}$, γ^μ , $\sigma^{\mu\nu}$, $\gamma^\mu\gamma_5$, and γ_5 .

The choice of matrices made above is called the chiral basis and is convenient because the right- and left-handed components of ψ factorize. However, multiplying ψ by an arbitrary unitary matrix S and all matrices by $\Gamma \rightarrow S\Gamma S^{-1}$ leaves the theory unchanged. While the explicit expressions for the matrices are obviously changed, certain relations are not, and are therefore particularly valuable. In particular, the Clifford algebra, Eq. (C.56), the relations Eq. (C.57), Eq. (C.58), Eq. (C.59), the definition Eq. (C.63) of γ_5 in terms of the other γ matrices, and the relations between the projection operators and γ_5 ,

$$P_L = \frac{1+\gamma_5}{2}, \quad P_R = \frac{1-\gamma_5}{2} \quad (\text{C.66})$$

are basis independent and should therefore be sufficient to evaluate any invariant quantities.

Appendix D

ξ -gauge Feynman rules

Unitary gauge, as described in Chapter 5, has several disadvantages that make it inappropriate for most calculations that go beyond tree level in the perturbative expansion. One of these difficulties is that the spin-one propagator does not fall to zero for large momenta, $p \rightarrow \infty$, thereby making the ultraviolet behavior of the theory appear to be worse than it really is. It is therefore usually more convenient to use in these computations a gauge in which the proper ultraviolet behavior is more manifest. A one-parameter family of such gauges is given by the Lorentz-covariant ξ -gauges.

Although no loop graphs are attempted in this book, the modification of the Feynman rules appropriate for ξ -gauges are included here for the sake of completeness. There are three new types of propagator that arise in ξ -gauge. The first of these is a modified spin-one boson propagator that was given in Chapter 5.

The particular cases $\xi = 1$ and $\xi = 0$ are respectively known as *Feynman–’t Hooft Gauge* and *Landau Gauge*. There are also two other types of *unphysical* particles that arise in the ξ -gauge graphs. These are the *unphysical scalars* and the *Fadeev–Popov–DeWitt ghosts*. Their role is to cancel the contributions within loops of the various unphysical components of the vector boson propagator. Neither the unphysical scalars nor the ghosts *ever* appear in external lines in scattering amplitudes.

The unphysical scalars are just the three remaining real scalars of the original Higgs doublet that were explicitly “eaten” by the massive spin-one particles in unitary gauge: $\phi = \left(\frac{1}{\sqrt{2}} \begin{pmatrix} w^+ \\ (v+H+iz) \end{pmatrix} \right)$. Two of these scalars may be combined into an electrically-charged complex scalar field, w^+ (together with its complex conjugate) which is absorbed in unitary gauge by the W boson. The third, z , is the real spinless particle that was eaten in unitary gauge by the Z^0 boson. Unitary gauge corresponds to the $\xi \rightarrow \infty$ limit of the ξ -gauge,

in which the unphysical particles become infinitely heavy. When computing loop corrections, this limit becomes singular, so *strict* unitary gauge should not be used in loop calculations.

D.1 Internal lines

The charged unphysical scalar, w^\pm , has a propagator that is given by

$$w^\pm : \quad \bullet \text{---} \text{---} \text{---} \bullet \quad G(p) = -i \int \frac{d^4 p}{(2\pi)^4} \frac{1}{p^2 + \xi M_W^2 - i\epsilon} \quad (\text{D.1})$$

The propagator for the other unphysical scalar, z , is

$$z : \quad \bullet \text{---} \text{---} \text{---} \bullet \quad G(p) = -i \int \frac{d^4 p}{(2\pi)^4} \frac{1}{p^2 + \xi M_Z^2 - i\epsilon} \quad (\text{D.2})$$

The ordinary gauge boson propagator differs from the one we found in unitary gauge; it is given in Eq. (5.55), but for completeness we repeat it here:

$$\mu \text{---} \text{---} \text{---} \nu \quad G_{\mu\nu}(p) = -i \int \frac{d^4 p}{(2\pi)^4} \frac{\eta_{\mu\nu} + (\xi-1) \frac{p_\mu p_\nu}{p^2 + \xi m^2}}{p^2 + m^2 - i\epsilon} \quad (\text{D.3})$$

Note, however, that the contraction of external state polarizations still satisfies Eq. (1.119), $\sum_\lambda \epsilon_\mu(\lambda) \epsilon_\nu^*(\lambda) = \eta_{\mu\nu} + p_\mu p_\nu / m^2$, the m^2 is not modified by ξ .

Finally there are the ghost particles. These are complex scalar particles which are required by consistency to carry the the statistics of a *fermion*. That is, the operators which create and annihilate them, satisfy anticommutation relations; when using them in diagrams, there is a factor of (-1) from each closed loop of ghosts, just as there is a factor of (-1) from a closed fermionic loop. There is one such (complex) ghost particle for each generator in the gauge group: we will denote them ω_g^α , ω_W^\pm , ω_Z , and ω_γ . Each ghost has a mass that is given by ξ times the mass of the corresponding gauge boson, so the eight color ghosts, ω_g^α , and the electromagnetic ghost, ω_γ , are massless while the ghosts corresponding to the massive W (or Z) bosons have squared-masses ξM_W^2 (respectively: ξM_Z^2). Their propagators are therefore,

$$\omega_g : \quad \bullet \cdots \blacktriangleright \cdots \bullet \quad G(p) = -i \int \frac{d^4 p}{(2\pi)^4} \frac{1}{p^2 - i\epsilon} \quad (\text{D.4})$$

$$\omega_\gamma : \quad \bullet \cdots \blacktriangleright \cdots \bullet \quad G(p) = -i \int \frac{d^4 p}{(2\pi)^4} \frac{1}{p^2 - i\epsilon} \quad (\text{D.5})$$

$$\omega_Z : \quad \bullet \cdots \blacktriangleright \cdots \bullet \quad G(p) = -i \int \frac{d^4 p}{(2\pi)^4} \frac{1}{p^2 + \xi M_Z^2 - i\epsilon} \quad (\text{D.6})$$

$$\omega_W^\pm : \quad \bullet \cdots \blacktriangleright \cdots \bullet \quad G(p) = -i \int \frac{d^4 p}{(2\pi)^4} \frac{1}{p^2 + \xi M_W^2 - i\epsilon} \quad (\text{D.7})$$

Note that, since ghosts are complex and have fermionic statistics, we must specify the sense of propagation.

It should be clear from the above propagators that the limit $\xi \rightarrow \infty$ naively returns unitary gauge; the unphysical particles' masses are all proportional to ξ , and the large ξ limit of Eq. (D.3) returns Eq. (5.54).

D.2 Vertices

We next turn to the Feynman rules for the interactions of the various particles in ξ -gauge. The interaction terms for all physical particles are the same as those given in Chapter 5, so we do not repeat them here. Instead we list only the “new” couplings involving the unphysical scalars and ghosts.

D.2.1 Unphysical scalar couplings

The unphysical scalars, w^\pm and z , have all of the couplings of the Higgs doublet – that is to say, Yukawa couplings and gauge-boson couplings. Their Yukawa couplings are

$zf\bar{f}$ Coupling

$$j \longrightarrow \bullet \longrightarrow i \quad \begin{array}{c} | \\ | \end{array} \quad \left(\eta_f \frac{m_f}{v} \right) (\gamma_5)_{ij} (2\pi)^4 \delta^4(k+l+p) \quad (\text{D.8})$$

in which $\eta_f = +1$ if f is an “up-type” fermion (u , c , or t) and $\eta_f = -1$ if f is “down-type” (d , s , b , or e , μ , and τ).

Also

$w^+ d \bar{u}$, $w^+ e \bar{\nu}$ Coupling

$$j \longrightarrow \bullet \longrightarrow i \quad \begin{array}{c} | \\ | \end{array} \quad i\sqrt{2}V_{ud} \left(\frac{m_u}{v} P_L - \frac{m_d}{v} P_R \right)_{ij} (2\pi)^4 \delta^4(k+l+p) \quad (\text{D.9})$$

and

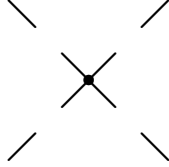
$w^- u \bar{d}, w^- \nu \bar{e}$ Coupling

$$\begin{array}{c} | \\ j \longrightarrow \bullet \longrightarrow i \end{array} \quad i\sqrt{2}V_{ud}^* \left(\frac{m_u}{v} P_R - \frac{m_d}{v} P_L \right)_{ij} (2\pi)^4 \delta^4(k+l+p) \quad (\text{D.10})$$

Here m_u, m_d are understood to mean 0, m_e in the case of the leptonic coupling.

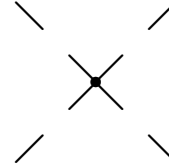
Next, there are interactions between unphysical scalars and either other unphysical scalars or the Higgs scalar, generated from the Higgs potential term

$zzzz$ Coupling



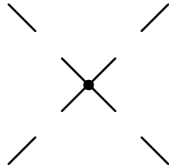
$$\frac{-3im_{\text{H}}^2}{24v^2} (2\pi)^4 \delta^4(k+l+p+q) \quad (\text{D.11})$$

zzw^+w^- Coupling



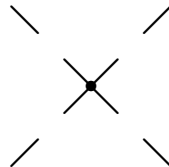
$$\frac{-im_{\text{H}}^2}{2v^2} (2\pi)^4 \delta^4(k+l+p+q) \quad (\text{D.12})$$

$w^+w^+w^-w^-$ Coupling



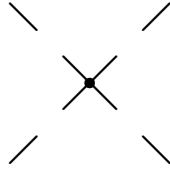
$$\frac{-2im_{\text{H}}^2}{4v^2} (2\pi)^4 \delta^4(k+l+p+q) \quad (\text{D.13})$$

$zzHH$ Coupling



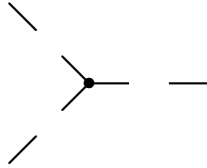
$$\frac{-im_{\text{H}}^2}{4v^2} (2\pi)^4 \delta^4(k+l+p+q) \quad (\text{D.14})$$

w^+w^-HH Coupling



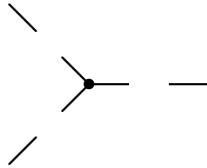
$$\frac{-im_{\text{H}}^2}{2v^2}(2\pi)^4\delta^4(k+l+p+q) \quad (\text{D.15})$$

$z^2 H$ Coupling



$$\frac{-im_H^2}{2v}(2\pi)^4\delta^4(k+l+p) \quad (D.16)$$

w^+w^-H Coupling



$$\frac{-im_H^2}{v}(2\pi)^4\delta^4(k+l+p) \quad (D.17)$$

Finally, there are cubic and quartic scalar-gauge field interactions, arising from the Higgs field kinetic term. Those terms containing only H and gauge fields have already been listed in Section 5.4; here we give the terms containing z and w^\pm fields.

w^+w^-A Coupling

$$\begin{array}{c} A: \mu; q \\ | \\ w^+: p - \text{---} \bullet \text{---} w^-: k \end{array} \quad i e(p-k)^\mu (2\pi)^4 \delta^4(p+k+q) \quad (\text{D.18})$$

w^+w^-Z Coupling

$$\begin{array}{c}
Z: \mu; q \\
\text{---} \bullet \text{---} \\
w^+: p \text{ ---} \quad w^-: k
\end{array}
\frac{ie(1 - 2 \sin^2 \theta_w)}{2 \sin \theta_w \cos \theta_w} (p - k)^\mu (2\pi)^4 \delta^4(p + k + q)
\quad (D.19)$$

zHZ Coupling

$$\begin{array}{c}
 Z: \mu; q \\
 \text{---} \bullet \text{---} \\
 z: p \quad \quad H: k
 \end{array}
 \quad \frac{e}{2 \sin \theta_W \cos \theta_W} (p-k)^\mu (2\pi)^4 \delta^4(p+k+q)
 \quad (\text{D.20})$$

 $w^\pm z W^\mp$ Coupling

$$\begin{array}{c}
 W^\mp: \mu; q \\
 \text{---} \bullet \text{---} \\
 w^\pm: p \quad \quad z: k
 \end{array}
 \quad \frac{g_2}{2} (p-k)^\mu (2\pi)^4 \delta^4(p+k+q)
 \quad (\text{D.21})$$

 $w^\pm H W^\mp$ Coupling

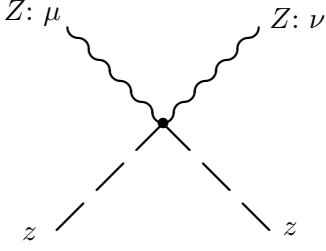
$$\begin{array}{c}
 W^\mp: \mu; q \\
 \text{---} \bullet \text{---} \\
 w^\pm: p \quad \quad H: k
 \end{array}
 \quad \frac{\pm i g_2}{2} (p-k)^\mu (2\pi)^4 \delta^4(p+k+q)
 \quad (\text{D.22})$$

 $w^\pm W^\mp A$ Coupling

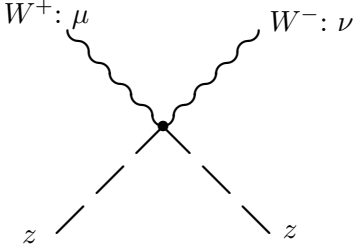
$$\begin{array}{c}
 w^\pm \\
 | \\
 W^\mp: \mu \quad \quad A: \nu
 \end{array}
 \quad -ie M_W \eta_{\mu\nu} (2\pi)^4 \delta^4(p+k+q)
 \quad (\text{D.23})$$

 $w^\pm W^\mp Z$ Coupling

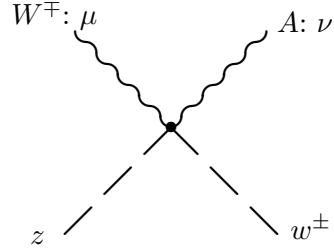
$$\begin{array}{c}
 w^\pm \\
 | \\
 W^\mp: \mu \quad \quad Z: \nu
 \end{array}
 \quad i g_2 \sin^2 \theta_W M_Z \eta_{\mu\nu} (2\pi)^4 \delta^4(p+k+q)
 \quad (\text{D.24})$$

zzZZ Coupling

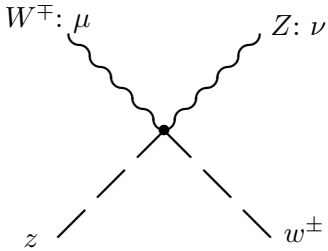
$$\frac{-ie_z^2}{2 \times 4} \eta_{\mu\nu} (2\pi)^4 \delta^4(p+k+l+q) \quad (\text{D.25})$$

zzW⁺W⁻ Coupling

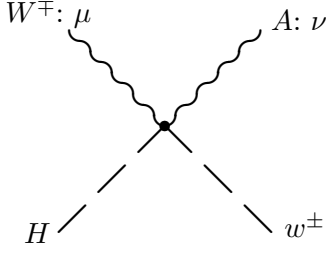
$$\frac{-ig_2^2}{2 \times 2} \eta_{\mu\nu} (2\pi)^4 \delta^4(p+k+l+q) \quad (\text{D.26})$$

zw[±]W[∓]A Coupling

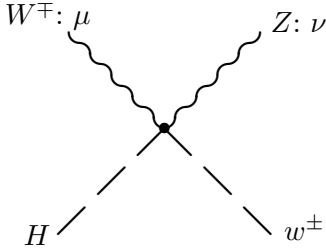
$$\frac{\mp eg}{2} \eta_{\mu\nu} (2\pi)^4 \delta^4(p+k+l+q) \quad (\text{D.27})$$

zw[±]W[∓]Z Coupling

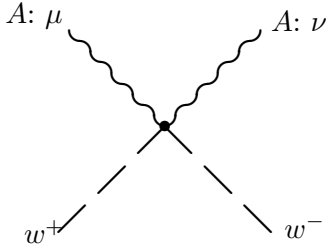
$$\frac{\pm e_z g \sin^2 \theta_w}{2} \eta_{\mu\nu} (2\pi)^4 \delta^4(p+k+l+q) \quad (\text{D.28})$$

$Hw^\pm W^\mp A$ Coupling

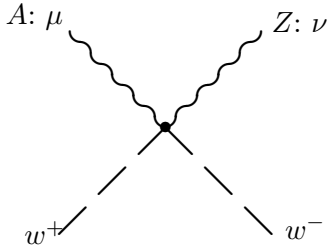
$$\frac{-ie g}{2} \eta_{\mu\nu} (2\pi)^4 \delta^4(p+k+l+q) \quad (\text{D.29})$$

 $Hw^\pm W^\mp Z$ Coupling

$$\frac{ie_z g \sin^2 \theta_W}{2} \eta_{\mu\nu} (2\pi)^4 \delta^4(p+k+l+q) \quad (\text{D.30})$$

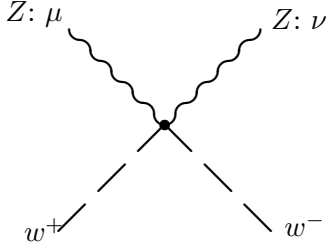
 $w^+ w^- A A$ Coupling

$$\frac{-2ie^2}{2} \eta_{\mu\nu} (2\pi)^4 \delta^4(p+k+l+q) \quad (\text{D.31})$$

 $w^+ w^- A Z$ Coupling

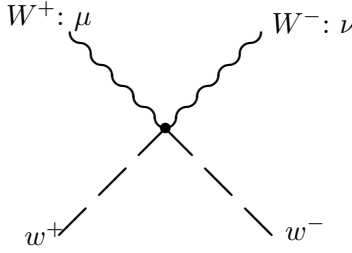
$$-ie e_z (1 - 2 \sin^2 \theta_W) \eta_{\mu\nu} (2\pi)^4 \delta^4(p+k+l+q) \quad (\text{D.32})$$

w^+w^-ZZ Coupling



$$\frac{-ie_Z^2(1-2\sin^2\theta_w)^2}{2 \times 2} \eta_{\mu\nu} (2\pi)^4 \delta^4(p+k+l+q) \quad (\text{D.33})$$

 $w^+w^-W^+W^-$ Coupling



$$\frac{-ig_2^2}{2} \eta_{\mu\nu} (2\pi)^4 \delta^4(p+k+l+q) \quad (\text{D.34})$$

D.2.2 Ghost couplings

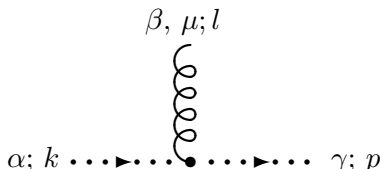
The ghosts couple principally to the gauge bosons. The ghosts that correspond to the W and Z generators of the gauge group have additional couplings to the scalars of the theory. In a general gauge theory the ghost Lagrangian takes the form

$$-\partial_\mu \omega_\alpha^* \partial^\mu \omega^\alpha - (\partial_\mu \omega_\alpha^*) \omega^\gamma A_\mu^\beta c_{\beta\gamma}^\alpha - \xi \omega_\alpha^* \omega^\beta \left(\langle \phi \rangle^T t_\alpha t_\beta \phi + \phi^\dagger t_\beta t_\alpha \langle \phi \rangle \right) \quad (\text{D.35})$$

Here ϕ is the full Higgs field and $\langle \phi \rangle$ is just the vacuum expectation value.

In the following, the ghost field on the left is the ω and the field on the right is the ω^* ; the indices on the structure constant have been rearranged, leading to an extra $-$ sign. The gauge couplings are:

Gluon-ghost-ghost coupling



$$-g_3 p_\mu f^{\alpha\beta\gamma} (2\pi)^4 \delta^4(k+l+p) \quad (\text{D.36})$$

$A\omega_W^\pm\omega_W^\pm$ coupling

$$\begin{array}{c}
 \gamma: \mu; l \\
 \text{wavy line} \\
 \omega_W^\pm: k \dots \blacktriangleright \bullet \dots \blacktriangleright \omega_W^\pm: p
 \end{array}
 \quad \mp ie p_\mu (2\pi)^4 \delta^4(k+l+p) \quad (\text{D.37})$$

 $Z\omega_W^\pm\omega_W^\pm$ coupling

$$\begin{array}{c}
 Z: \mu; l \\
 \text{wavy line} \\
 \omega_W^\pm: k \dots \blacktriangleright \bullet \dots \blacktriangleright \omega_W^\pm: p
 \end{array}
 \quad \mp ie \cot \Theta_W p_\mu (2\pi)^4 \delta^4(k+l+p) \quad (\text{D.38})$$

 $W^\pm\omega_W^\mp\omega_\gamma$ coupling

$$\begin{array}{c}
 W^\pm: \mu; l \\
 \text{wavy line} \\
 \omega_W^\mp: k \dots \blacktriangleright \bullet \dots \blacktriangleright \omega_\gamma: p
 \end{array}
 \quad \mp ie p_\mu (2\pi)^4 \delta^4(k+l+p) \quad (\text{D.39})$$

 $W^\pm\omega_W^\mp\omega_Z$ coupling

$$\begin{array}{c}
 W^\pm: \mu; l \\
 \text{wavy line} \\
 \omega_W^\mp: k \dots \blacktriangleright \bullet \dots \blacktriangleright \omega_Z: p
 \end{array}
 \quad \mp ie \cot \Theta_W p_\mu (2\pi)^4 \delta^4(k+l+p) \quad (\text{D.40})$$

 $W^\pm\omega_\gamma\omega_W^\pm$ coupling

$$\begin{array}{c}
 W^\pm: \mu; l \\
 \text{wavy line} \\
 \omega_\gamma: k \dots \blacktriangleright \bullet \dots \blacktriangleright \omega_W^\pm: p
 \end{array}
 \quad \pm ie p_\mu (2\pi)^4 \delta^4(k+l+p) \quad (\text{D.41})$$

 $W^\pm\omega_Z\omega_W^\pm$ coupling

$$\begin{array}{c}
 W^\pm: \mu; l \\
 \text{wavy line} \\
 \omega_Z: k \dots \blacktriangleright \bullet \dots \blacktriangleright \omega_W^\pm: p
 \end{array}
 \quad \pm ie \cot \Theta_W p_\mu (2\pi)^4 \delta^4(k+l+p) \quad (\text{D.42})$$

The scalar couplings are:

$H\omega_Z\omega_Z$ Coupling

$$\begin{array}{c}
 H \\
 | \\
 \omega_Z \cdots \blacktriangleright \cdots \bullet \cdots \blacktriangleright \cdots \omega_Z
 \end{array}
 \quad
 \frac{-i\xi M_Z^2}{v} (2\pi)^4 \delta^4(k+l+p) \quad (D.43)$$

$H\omega_W^\pm\omega_W^\pm$ Coupling

$$\begin{array}{c}
 H \\
 | \\
 \omega_W^\pm \cdots \blacktriangleright \cdots \bullet \cdots \blacktriangleright \cdots \omega_W^\pm
 \end{array}
 \quad
 \frac{-i\xi M_W^2}{v} (2\pi)^4 \delta^4(k+l+p) \quad (D.44)$$

which, note, are both proportional to the particle mass, like all

Higgs couplings;

$z\omega_W^\pm\omega_W^\pm$ Coupling

$$\begin{array}{c}
 z \\
 | \\
 \omega_W^\pm \cdots \blacktriangleright \cdots \bullet \cdots \blacktriangleright \cdots \omega_W^\pm
 \end{array}
 \quad
 \frac{\pm\xi M_W^2}{v} (2\pi)^4 \delta^4(k+l+p) \quad (D.45)$$

$w^\pm\omega_W^\mp\omega_Z$ Coupling

$$\begin{array}{c}
 w^\pm \\
 | \\
 \omega_W^\mp \cdots \blacktriangleright \cdots \bullet \cdots \blacktriangleright \cdots \omega_Z
 \end{array}
 \quad
 \frac{i\xi M_W M_Z}{v} (2\pi)^4 \delta^4(k+l+p) \quad (D.46)$$

The $w^\pm\omega_W^\mp\omega_A$ coupling vanishes.

$w^\pm\omega_Z\omega_W^\pm$ Coupling

$$\begin{array}{c}
 w^\pm \\
 | \\
 \omega_Z \cdots \blacktriangleright \cdots \bullet \cdots \blacktriangleright \cdots \omega_W^\pm
 \end{array}
 \quad
 \frac{-i\xi M_W M_Z (1-2\sin^2\theta_W)}{v} (2\pi)^4 \delta^4(k+l+p) \quad (D.47)$$

$w^\pm\omega_A\omega_W^\pm$ Coupling

$$\begin{array}{c}
 w^\pm \\
 | \\
 \omega_A \cdots \blacktriangleright \cdots \bullet \cdots \blacktriangleright \cdots \omega_W^\pm
 \end{array}
 \quad
 \frac{-2i\xi M_W^2 \sin\theta_W}{v} (2\pi)^4 \delta^4(k+l+p) \quad (D.48)$$

Appendix E

Metric convention conversion table

In this book we have systematically used the metric convention, $\eta_{\mu\nu} = \text{Diag}[-1, +1, +1, +1]$, the “Pauli,” “East Coast,” or “mostly plus” metric. The other common convention, the “Bjorken and Drell,” “West Coast,” or “mostly minus” convention, takes $\eta_{\mu\nu} = \text{diag}[+1, -1, -1, -1]$. The “mostly minus” metric convention is currently in more common use in the field of phenomenology. The “mostly plus” convention predominates in the general relativity, string theory, supersymmetry, and formal field-theory communities.

To make this book more useful to its intended audience, who primarily use the opposite metric convention, we describe in this appendix the differences between these conventions, culminating in a “translation table” between the conventions, which should ease the difficulty in hopping between our conventions and the conventions appearing in most of the relevant literature. There are other conventions besides the metric which must be decided on, and these are not uniform in either community; since it would be too complicated to discuss every possible set of conventions, we will focus only on the most common coherent set of “mostly minus” conventions, taken to be those of Peskin and Schroeder, *“An Introduction to Quantum Field Theory,”* Westview, 1995.

Finally, we will end this section with an explanation of why we prefer the “mostly plus” metric. We postpone that discussion to the end because some physicists approach this issue with almost religious conviction, and it is important to us that you not slam this book shut before reading the rest of this appendix.

E.1 Propagation of the differences

In going between metric conventions, it is necessary to decide what is kept the *same*. Generally $x^\mu = (t, \vec{x})$ in both conventions, and similarly $p^\mu = (E, \vec{p})$. Therefore $\partial_\mu = (\partial_t, \partial_{\vec{x}})$ is also the same. This means that x_μ , p_μ , and ∂^μ are opposite. The relations for spinorial objects, gauge fields, and generators of transformations are more complicated and are discussed in turn.

E.1.1 Dirac algebra

The Dirac matrices γ^μ are required to obey the anticommutation relations (the Clifford algebra)

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\eta^{\mu\nu} \quad (\text{E.1})$$

The difference in sign in $\eta^{\mu\nu}$ then requires a difference in normalization for the Dirac matrices. We will only discuss the chiral basis for the Dirac matrices. In this case, the relationship is

$$\gamma_{\text{mostly plus}}^\mu = -i\gamma_{\text{mostly minus}}^\mu \quad \text{or} \quad i\gamma_{\text{mostly plus}}^\mu = \gamma_{\text{mostly minus}}^\mu \quad (\text{E.2})$$

In the mostly plus metric, all gamma matrices are Hermitian except γ^0 , which is anti-Hermitian. In the mostly minus metric, the reverse is true; γ^0 is Hermitian but the others are anti-Hermitian. Because of this change, it is necessary in the mostly plus case to introduce a new matrix $\beta = i\gamma_{\text{mostly plus}}^0 = \gamma_{\text{mostly minus}}^0$, which is Hermitian and which serves the role of relating $\bar{\psi}$ and ψ^\dagger . In the mostly minus case, this matrix is the same as γ^0 and is generally not given an independent symbol.

The factor of i changes the appearance of the fermionic kinetic term in the Lagrangian;

$$\mathcal{L}_{\text{fermion}} = -\bar{\psi}(\not{D} + m)\psi, \text{ “mostly plus,”} \quad \bar{\psi}(i\not{D} - m)\psi, \text{ “mostly minus”} \quad (\text{E.3})$$

This makes a corresponding change in the Dirac equation,

$$\text{“mostly plus”}: [\not{p} + m]u(p) = 0; \quad \text{“mostly minus”}: [\not{p} - m]u(p) = 0 \quad (\text{E.4})$$

We reiterate that these expressions look different, but they have exactly the same content; the different appearance is because the definition of the symbol γ^μ has changed, and because $\gamma^\mu p_\mu$ in one case is $-\gamma^0 p^0 + \vec{\gamma} \cdot \vec{p}$, while in the other it is $\gamma^0 p^0 - \vec{\gamma} \cdot \vec{p}$.

The sign of the matrix γ^5 is also convention dependent. The eigenspinors of γ^5 are the spinors of definite chirality. The old convention for γ^5 was

that right-handed spinors should have eigenvalue 1 and left-handed spinors should have eigenvalue -1 . The “mostly minus” literature has generally maintained this convention. However, this convention was established before the discovery of the weak interactions, which couple exclusively to left-handed particles; so it is in some ways more convenient to adopt the opposite sign convention. Weinberg does this and we have followed his convention. Therefore,

$$\begin{aligned}\gamma_{\text{mostly plus}}^5 &= -i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \\ \gamma_{\text{mostly minus}}^5 &= i\gamma^0\gamma^1\gamma^2\gamma^3 = -\gamma_{\text{mostly plus}}^5\end{aligned}\quad (\text{E.5})$$

We emphasize that this is *not* a consequence of our choice of metric, it is a separate choice to modernize the notation which we choose to make at the same time. Because of this choice, the ubiquitous $2P_L = (1 + \gamma^5)$ projectors which appear in this book would be $2P_L = (1 - \gamma^5)$ in the notation of almost all “mostly minus” authors.

In addition, there are differing conventions in the definition of the charge conjugation matrix C (which should be distinguished from the charge conjugation operator \mathcal{C}). We have followed the most common practice of defining C such that, for Majorana fermions, Eq. (1.97) holds: $\psi_M = C\bar{\psi}_M^T$. We cannot compare with Peskin and Schroeder because C is not defined there. Therefore we take the mostly minus convention also to give $\psi_M = C\bar{\psi}_M^T$, which requires $C = -i\gamma^2\gamma^0$ (mostly minus).

Finally, because of the different factor of i in γ^μ , the behavior of fermion bilinears under Hermitian conjugation differs in the two conventions. In Problem 1.1, the first set of relations all hold unchanged in mostly minus, but in the second and third sets, the equations involving γ^μ and $\gamma^\mu\gamma^5$ have the opposite sign.

These differences are summarized in Table E.1.

E.1.2 Gauge fields and Poincaré generators

We follow the convention, $D_\mu = \partial_\mu - iT^a A_\mu^a$ for gauge fields (with $T^a = q$ the charge for a $U(1)$ field). Peskin and Schroeder choose the opposite sign for $U(1)$ fields but the same sign for non-abelian fields. Therefore,

$$A_{\text{mostly plus}}^\mu = A_{\text{mostly minus}}^\mu, \quad \text{but} \quad G_{\text{mostly plus}}^\mu = -G_{\text{mostly minus}}^\mu \quad (\text{E.6})$$

With this sign convention, $A^\mu = (\Phi, \vec{A})$, with Φ and \vec{A} the conventional scalar and vector potentials, in both conventions. With this choice, the

Table E.1 *Metric-convention conversion table*

Equation	“mostly plus”	“mostly minus”
Clifford Algebra	$[\gamma^\mu, \gamma^\nu] = 2\eta^{\mu\nu}$	$[\gamma^\mu, \gamma^\nu] = 2\eta^{\mu\nu}$
Dirac Lagrangian	$-\bar{\psi}(\not{D} + m)\psi$	$\bar{\psi}(i\not{D} - m)\psi$
Dirac Equation	$[i\not{p} + m]u(p) = 0$	$[\not{p} - m]u(p) = 0$
Spinor bilinears	$\sum_\sigma u\bar{u}/v\bar{v}(p, \sigma) = -i\not{p} \pm m$	$\sum_\sigma u\bar{u}/v\bar{v}(p, \sigma) = \not{p} \pm m$
γ Hermiticity	$\beta\gamma_\mu^\dagger = -\gamma_\mu\beta, \quad \beta \equiv i\gamma^0$	$\beta\gamma_\mu^\dagger = +\gamma_\mu\beta, \quad \beta \equiv \gamma^0$
Charge conjugation	$\psi_M = C\bar{\psi}_M^T, \quad C^T = -C$	same
Gamma transposes	$\gamma_\mu^T C = -C\gamma_\mu$	same

relation between the field strength and the non-covariant field strength is opposite for “mostly plus” as for “mostly minus”: the field strengths are

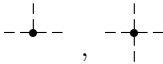

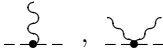
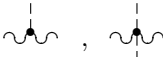


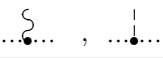
$$\begin{aligned}
 \text{“mostly plus” : } \vec{E}_i &= F_{i0} = F^{0i}, \quad \vec{B}_i = \frac{1}{2}\epsilon_{ijk} \times (F_{jk} = F^{jk}) \\
 \text{“mostly minus” : } \vec{E}_i &= F_{0i} = F^{i0}, \quad \vec{B}_i = \frac{1}{2}\epsilon_{ijk} \times (F_{kj} = F^{kj}) \quad (\text{E.7})
 \end{aligned}$$

As for the generators of the Lorentz group, we adopt the convention that \hat{P}^μ acting on a state return its four-momentum p^μ . Therefore $P^\mu_{\text{mostly plus}} = P^\mu_{\text{mostly minus}}$. The lowered components are opposite, $P_\mu = -i\partial_\mu$ in “mostly plus” whereas it is $i\partial_\mu$ in “mostly minus.” The phase $e^{-i\omega t + i\vec{p}\cdot\vec{x}}$ becomes $e^{ip_\mu x^\mu}$ in “mostly plus,” while it is $e^{-ip_\mu x^\mu}$ in “mostly minus.” An active translation by ξ^μ is therefore implemented by the operator $e^{-i\hat{P}^\mu x_\mu}$. (To remember this, recall that an active transformation is one which changes the position or time of a particle. If a particle with wave function $e^{i\vec{p}\cdot\vec{x}}$ is shifted so what was its origin is now at $\vec{\xi}$, then the phase is 0 at $\vec{\xi}$ and must be $e^{-i\vec{p}\cdot\vec{\xi}}$ at the origin.)

In both conventions, $J^{\mu\nu} = x^\mu p^\nu - p^\mu x^\nu$ when acting on a scalar field. That is, $J^{\mu\nu}_{\text{mostly plus}} = J^{\mu\nu}_{\text{mostly minus}}$. The object with both lowered indices is also the same in the two conventions, but when one index is raised and one is lowered, the conventions differ. With these conventions, $\vec{J}_i = \frac{1}{2}\epsilon_{ijk} J_{jk}$ is the conventional angular momentum operator. Because of the opposite sign

on the metric, the signs in Appendix C Eq. (C.25) and Eq. (C.26) are flipped between the two conventions.

Table E.2 *Metric-convention conversion table for Feynman rules*

Propagators Spin	symbol	“mostly plus”	“mostly minus”
0	-----	$\frac{-i}{p^2 + M^2 - i\epsilon}$	$\frac{i}{p^2 - M^2 + i\epsilon}$
$\frac{1}{2}$	→	$\frac{-i(-i\not{p} + m)}{p^2 + m^2 - i\epsilon}$	$\frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon}$
1 Massless	~~~~~	$\frac{-i\eta^{\mu\nu}}{p^2 - i\epsilon}$	$\frac{-i\eta^{\mu\nu}}{p^2 + i\epsilon}$
1 Unitary	~~~~~	$\frac{-i\left(\eta^{\mu\nu} + \frac{p^\mu p^\nu}{M^2}\right)}{p^2 + M^2 - i\epsilon}$	$\frac{-i\left(\eta^{\mu\nu} - \frac{p^\mu p^\nu}{M^2}\right)}{p^2 - M^2 + i\epsilon}$
1 R_ξ	~~~~~	$\frac{-i\left(\eta^{\mu\nu} + \frac{(\xi-1)p^\mu p^\nu}{p^2 + \xi M^2}\right)}{p^2 + M^2 - i\epsilon}$	$\frac{-i\left(\eta^{\mu\nu} + \frac{(\xi-1)p^\mu p^\nu}{p^2 - \xi M^2}\right)}{p^2 - M^2 + i\epsilon}$
Vertices type	symbol	“mostly plus”	“mostly minus”
Scalar		$-i\lambda\nu, -i\lambda$	$-i\lambda\nu, -i\lambda$
Yukawa		$-if_{mn}$	$-if_{mn}$
Gauge-scalar		$ie(p-k)_\mu, -ie^2\eta_{\mu\nu}$	$ie(p-k)_\mu, +ie^2\eta_{\mu\nu}$
Gauge-Higgs		$-ie^2\nu\eta_{\mu\nu}, -ie^2\eta_{\mu\nu}$	$ie^2\nu\eta_{\mu\nu}, ie^2\eta_{\mu\nu}$
$A\bar{\psi}\psi$		$-e\gamma_\mu$	$ie\gamma_\mu [-i, \text{abelian}]$
A^3, A^4		$+gf_{abc}\dots, -ig^2\dots$	$+gf_{abc}\dots, -ig^2\dots$
ghost		$-gp_\mu, -i\xi g^2 v$	$+gp_\mu, -i\xi g^2 v$

We have taken these differences, the differences due to the Dirac algebra described above, and the differences due to the metric, and used them to find the changes to Feynman rules, summarized in Table E.2. Vertices involving gauge bosons are for non-abelian interactions; for QED, the sign of e in “mostly minus” must be switched.

E.2 Why we use “mostly plus”

There are two key advantages of the “mostly plus” metric.

First, it is much easier to go back and forth between covariant and non-covariant notation when using “mostly plus.” This is because the non-covariant notation also involves a metric, δ_{ij} , the metric on the three spatial indices. It would be awkward to make this metric purely negative, and no one does. In the “mostly plus” convention, in passing to the non-covariant notation, the part involving a metric stays unchanged, only the temporal part, which is already being split off and treated differently, has its sign flipped. In the “mostly minus” convention, to pass to non-covariant notation, it is the part which is otherwise being treated as special which retains its sign, and the term which has a metric in it must have its sign flipped.

Second and even more important, most complicated calculations in quantum field theory involve Wick rotation, that is, continuation to an imaginary value of the time or momentum variable. Indeed, field theories are probably only formally well defined after such continuation. When using the “mostly plus” convention, this continuation is very simple; the negative term in the metric is switched to being positive. When using the “mostly minus” convention, if one merely continues the time or frequency coordinate to imaginary values, one is left with a totally negative metric. One either has to work with a totally negative metric, or flip the sign convention of the metric at the same time as analytically continuing. Either approach introduces extra opportunities for confusion and error, and neither one is very appealing.

In addition to these major advantages, there are a number of minor advantages to the “mostly plus” metric.

- (i) Photon polarization vectors ϵ_μ have positive squares. Similarly, most components of the gauge field A^μ have positive norm. The unphysical, negative norm gauge field states which can arise in certain quantization procedures (Gupta–Bleuler) arise from the negative piece of the metric, rather than from the positive piece.
- (ii) The sign on scalar and vector propagators is the same.

(iii) Most of the Dirac matrices are Hermitian.

To be fair, there are also advantages to the “mostly minus” metric. Indeed, if there were not, then everyone would agree by now on a metric convention. In our view, these advantages are,

- (i) most four-momenta encountered in particle physics are timelike, and this convention gives them positive squares, $p^2 = m^2 > 0$;
- (ii) the matrix β and the matrix $\gamma^0 = \gamma_0$ are the same;
- (iii) the Dirac propagator involves $\not{p} + m$, which does not contain a relative factor of i between the two terms. In addition, because the γ matrices satisfy $\gamma_\mu^\dagger \beta = \beta \gamma_\mu$ without a minus sign, there are no “surprise minus signs” emerging from the complex conjugation of a matrix element.

Besides these technical reasons is the practical one: “mostly minus” users are in the majority in the phenomenology community and communication within the community is easiest if people converge on a set of conventions. The importance of this final reason should not be underestimated.

Based on our experience we prefer the “mostly plus” metric, and have written this book in that convention. We hope that this appendix, and in particular the two lookup tables it contains, eases the translation between the conventions in practical calculations and renders this book usable to both communities.

Select bibliography

Here are some references the reader may find useful to get background material before reading this book, to get complementary field theory material to better understand the more advanced topics treated superficially here, or to go beyond where this book stops. We will not make any effort to give a proper bibliography of the primary literature.

First and foremost, the reader should be aware of the excellent resources provided by the *Particle Data Group*, which makes extensive listings of particles and their properties available on their website, <http://pdg.lbl.gov>, which also provides reviews on a number of subjects relevant to particle physics. They also provide (free!) a nifty pocket-sized “Particle physics booklet,” which contains summary versions of particle listings and reviews. A new edition is printed every two years. Generally, experimental values for measurables quoted in this book are taken from the PDG. Their most recent edition is W. M. Yao *et al.*, *Journal of Physics* **G33**, 1 (2006).

Complementary books

Now we list some books which may be useful to our readers. We start with some introductory books which may be useful complementary references:

- “An Introduction to the Standard Model,” by W. Cottingham and D. Greenwood, Cambridge Press, 1998. Reviews the standard model at a more introductory level than this book (roughly, senior undergraduate level), and may be useful for setting the stage.
- “Quarks and Leptons: An Introductory Course in Particle Physics,” by F. Halzen and A. Martin, Wiley, 1984. Their book is to QED and QCD as this book attempts to be for the standard model; an elementary introduction aimed at developing the computational tools and getting people calculating, with a minimum of formal baggage. An excellent introduction to utilitarian field theory.

Textbooks on quantum field theory

Some of the classic textbooks on quantum field theory are:

- “An Introduction to Quantum Field Theory,” by M. Peskin and D. Schroeder, Westview, 1995. An excellent and readable book on field theory. The first 4 chapters of their book gives all required background preparation for this book, and the remaining chapters give the tools needed to take the material presented in this book to the next level (renormalization, higher order effects, and a more solid theoretical foundation).
- “The Quantum Theory of Fields, I–III”, by S. Weinberg, Cambridge Press, 2000. An original and encyclopedic presentation of quantum field theory from one of the masters. Most of the field theoretical arguments alluded to in this book are laid out in detail here. You will learn something new every time you read it, but it is probably not as good for a first text as Peskin and Schroeder’s book.
- “Quantum Field Theory,” by L. Ryder, Cambridge Press, 1996 (2’nd ed). A somewhat more introductory text on field theory, providing more than enough background material to understand this book.
- “Quantum Field Theory,” by L. Brown, Cambridge Press, 1994. An idiosyncratic but clear introduction to quantum field theory, which gives an extremely solid underpinning but does not cover nonabelian gauge theories.
- “Quantum Field Theory”, by G. Sterman, Cambridge Press, 1993. A clear and systematic exposition of modern field theoretic techniques which includes a number of topics (like infrared divergences and factorization) which are not covered in other texts.
- “Renormalization”, by J. Collins, Cambridge Press 1984. A focussed and thorough treatise on renormalization theory.

Group theoretical methods

Two very useful introductions to the theory of Lie groups for physicists, and a very useful reference with extensive tables, are

- “Lie Algebras in Particle Physics,” by H. Georgi, Perseus, 1999 (2’nd ed). A thorough treatment of group theory for particle physics, essential for those who find appendix B either too telegraphic or far too incomplete.
- “Semi-Simple Lie Algebras and Their Representations”, R. N. Cahn, Benjamin/Cummings, 1984. A book which picks up where the previous book leaves off, presenting more of the properties of groups and their representations.

- “Group Theory For Unified Model Building”, R. Slansky, *Physics Reports* **79**, (1981) 1–128. This provides very useful summary of the properties of the representations of Lie groups, including detailed tables showing how representations decompose in terms of representations of subgroups.

More advanced texts

Other more specialized textbooks on particle physics are

- “Gauge Theory of Elementary Particle Physics,” by T. Cheng and L. Li, Oxford Press, 1988. An exhaustive study of nonabelian gauge theory at an advanced level, especially QCD, electroweak theory, and $SU(5)$ GUT’s.
- “Weak Interactions and Modern Particle Theory”, by H. Georgi, Benjamin Cummings, 1984. A very physical discussion of much of the standard model and some of the techniques used to compute with it, which survives well despite its age.
- “Particle Physics and Introduction to Field Theory”, by T. D. Lee, Harwood, 1981. This very physical and insightful introduction to field theoretical techniques in particle physics also survives very well.
- “Effective Lagrangians for the Standard Model,” by A. Dobado, A. Gomez-Nicola, A. Maroto, and J. Pelaez, Springer, 1997. An advanced treatment of effective field theory and its application to the standard model at energies below the weak scale. It gives a detailed discussion of chiral perturbation theory and nonlinearly realized electroweak symmetry. The assumed level of field theory background is fairly high.
- “Dynamics of the Standard Model,” by J. Donoghue, E. Golowich, and B. Holstein, Cambridge Press, 1992. An advanced discussion of the standard model, with particular emphasis on bound states in QCD, chiral symmetry, and radiative corrections.
- “Journeys Beyond the Standard Model,” by P. Ramond, Westview, 2003. This book provides a nice discussion of the field theory of the standard model, the physics it contains, and the physics of its proposed extensions, assuming a much broader field theory background than is assumed in this book.
- “The Standard Model in the Making: Precision Study of the Electroweak Interactions,” by D. Bardin and G. Passarino, Clarendon Press, 1999. An advanced text on the standard model, for someone who has already mastered everything in this book and in Peskin and Schroeder. Their book explains in detail how to carry out loop level calculations within electroweak theory and provides a compendium of high order results in precision electroweak theory and in perturbative QCD.

- “The Higgs Hunter’s Guide,” by H. Haber, G. Kane, S. Dawson, J. Gunion, Perseus, 2000. An extensive review of the physics of the Higgs boson, including complete discussions of extensions of the model such as multi-Higgs theories and the Higgs sector in supersymmetric theories.
- “Neutrino Physics,” by K. Zuber, Taylor and Francis, 2004. A comprehensive and up-to-date presentation of neutrino physics, including a more complete discussion of the issues in neutrino phenomenology discussed in this book, and careful discussion of the experimental situation and of the details of the experiments.
- “Neutrino Astrophysics,” by J. N. Bahcall, Cambridge Press, 1989. The classic text, which contains everything you ever wanted to know about the solar neutrino problem.
- “Stars as Laboratories for Fundamental Physics,” by G. G. Raffelt, University of Chicago Press, 1996. An extensive description of what can be learned about particle physics from the study of astrophysical environments.

That other interaction

More and more, particle physicists are expected to be knowledgeable about gravitational physics, and the geometrical techniques which are used in its study. Some useful texts for these purposes are:

- “Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity”, S. Weinberg, Wiley, 1972. An oldie but a goodie: a very physical introduction to general relativity and its applications in astrophysics, the solar system and cosmology.
- “Gravitation”, C. W. Misner, K. S. Thorne and J. A. Wheeler, Freeman, 1973. The classic book with the quirky style, which sets the standard for its comprehensive and modern treatment of geometrical techniques.
- “General Relativity”, R. M. Wald, University of Chicago Press, 1984. A modern update of the two previous classics, containing more of the modern mathematical techniques.
- “Introduction to Supersymmetry and Supergravity”, by P. West, World Scientific, 1986. A well-written introduction to a difficult subject, using notation which is not too different from that used here.
- “Supersymmetry and Supergravity”, by J. Wess and J. Bagger, Princeton Press, 1983. The classic textbook on supersymmetry from some of those who invented it.

Chapter-specific references

These more specific references are for those interested in greater detail or more rigor on some of the subjects covered in specific chapters. Many papers are available online; when we write, say, [arXiv:hep-ph/9708416] it is shorthand for <http://arxiv.org/abs/hep-ph/9708416>.

Chapter 1:

- Our discussion of the foundations of field theory is here modelled on the treatment in “The Quantum Theory of Fields I”, by S. Weinberg, *op. cit.*;
- For a detailed (but advanced) discussion of C, P, and T symmetries, the spin-statistics theorem, and related topics, try “PCT, Spin and Statistics, and All That,” by A. Wightman and R. Streater, Princeton University Press, 2000.

Chapter 2:

- Since the book is about the standard model, we make our sole excursion into history to record some of the seminal steps in the development of the model. The standard model did not spring from the ground whole, but was constructed in a number of steps. The proposal that the electroweak interactions amongst leptons could be described using an approximate $SU(2) \times U(1)$ symmetry was made in S. L. Glashow, *Nucl. Phys.* **22**, 579 (1961); A. Salam and J. Ward, *Phys. Lett.* **13**, 168 (1964).
- The recognition that the electroweak interactions could be described by a spontaneously broken gauge symmetry was by S. Weinberg, *Phys. Rev. Lett.* **19**, 1264 (1967); A. Salam, in *Elementary Particle Physics*, ed. by N. Svartholm (Almqvist and Wiksells, Stockholm, 1968, p. 367).
- The inclusion of quarks and the recognition that a c quark would suppress flavor-changing charged-current loops, S. L. Glashow, J. Iliopoulos and L. Maiani, *Phys. Rev.* **D2**, 1285 (1970), as well as flavor-changing neutral currents, S. Weinberg, *Phys. Rev. Lett.* **27**, 1688 (1971); *Phys. Rev.* **D5**, 1413 (1972), came later.
- The incorporation of the quark ‘mixing angles’ starts with the $d-s$ mixing, N. Cabibbo, *Phys. Rev. Lett.* **10**, 531 (1963), and its full 3-generation formulation comes with M. Kobayashi and K. Maskawa, *Prog. Theor. Phys.* **49**, 282 (1972).
- A more complete discussion of anomalies may be found in the field theory textbooks listed above, such as “The Quantum Theory of Fields, II” *op. cit.*.

Chapter 3:

- Scattering theory is described by most advanced texts on quantum mechanics. A classic is “Collision Theory”, by M. L. Goldberger and K. M. Watson, Wiley, 1964 (Dover edition, 2004), although our treatment follows more that of “The Quantum Theory of Fields I” *op. cit.*.

Chapter 4:

- For an early argument for the ability to perturbatively predict hadronic jet distributions from QCD see G. Sterman and S. Weinberg, *Phys. Rev. Lett.* **39**, 1436 (1977).

Chapter 5:

- The discussion of the cancellations between propagators and non-covariant terms in H , leading to the ‘miracle’ of Lorentz invariance, follows that of “The Quantum Theory of Fields I”, *op. cit.*.
- The Feynman rules for general gauge theories were first given in unitary gauge by S. Weinberg, *Phys. Rev.* **D7**, 1068 (1973).
- The forms more useful for loops were developed by G. ’t Hooft, *Nucl. Phys.* **B35**, 167 (1971); B. W. Lee, *Phys. Rev.* **D5**, 823 (1972); K. Fujikawa, B. W. Lee and A. Sanda, *Phys. Rev.* **D6**, 2923 (1972).

Chapter 6:

- As is frequently the case, our treatment of resonances is a cartoon motivated by the discussion in “The Quantum Theory of Fields I”, *op. cit.*.
- The cancellation of infrared divergences for photons goes back to F. Bloch and A. Nordsieck, *Phys. Rev.* **37**, 54 (1937); D. R. Yennie, S. C. Frautschi and H. Suura, *Ann. Phys. (NY)* **13**, 379 (1955).
- A clear and detailed exposition of soft radiation within field theory was presented by D. R. Yennie, S. C. Frautschi and H. Suura, *Annals Phys.* **13**, 379 (1961).
- The implications for soft-photon emission are given for photons in S. Weinberg, *Phys. Lett.* **9**, 357 (1964); *Phys. Rev.* **135**, B1049 (1964), and for gravitons in S. Weinberg, *Phys. Rev.* **140**, **B515** (1965).
- Initial state radiation to high order and the Z -pole is treated in F. A. Berends, W. L. van Neerven and G. J. H. Burgers, *Nucl. Phys.* **B297**, 429 (1988) [Erratum-ibid. **B304**, 921 (1988)].

Chapter 7:

- For an informal introduction to the renormalization and the origin of large logarithms see *Why The Renormalization Group Is A Good Thing*, S. Weinberg, in ‘Cambridge 1981, Proceedings, Asymptotic Realms Of Physics,’ 1–19.
- The 2-loop beta functions for the Standard Model were worked out by M. Machacek and M. Vaughn, *Nucl. Phys.* **B222**, 83 (1983); *Nucl. Phys.* **B236**, 221 (1984); *Nucl. Phys.* **B249**, 70 (1985).
- The 4-loop beta function for QCD was found in T. van Ritbergen, J. A. M. Vermaseren and S. A. Larin, *Phys. Lett.* **B400**, 379 (1997).
- The logic of effective field theories is clearly laid out in *Phenomenological Lagrangians*, S. Weinberg, *Physica* **A96**, 327 (1979).
- The theory of nonlinear realizations of symmetries is given by S. Coleman, C. Callan, J. Wess and B. Zumino, *Phys. Rev.* **177**, 2239 (1969); *Phys. Rev.* **177**, 2247 (1969).
- Our discussion of the equivalence between nonlinearly realized gauge symmetries and no gauge symmetry at all is based on that of C. P. Burgess and D. London, *Phys. Rev. Lett.* **69**, 3428 (1992); *Phys. Rev.* **D48** (1993) 4337 [[arXiv:hep-ph/9203216](#)].
- There is an extensive literature which treats a heavy Higgs sector using nonlinearly realized gauge symmetries, starting with A. C. Longhitano, *Phys. Rev.* **D22**, 1166 (1980); *Nucl. Phys.* **B188**, 118 (1981). The discussion here follows C. P. Burgess, S. Godfrey, H. Konig, D. London and I. Maksymyk, *Phys. Rev.* **D49**, 6115 (1994) [[arXiv:hep-ph/9312291](#)].
- The S , T , U ‘oblique’ electroweak parameters were introduced in M. E. Peskin and T. Takeuchi, *Phys. Rev. Lett.* **65**, 964 (1990), and their ϵ_i counterparts by G. Altarelli, R. Barbieri and S. Jadach, *Nucl. Phys.* **B369**, 3 (1992), Erratum-*ibid.* **B376**, 444 (1992).

Chapter 8:

- For a readable introduction to lattice gauge theory, “Lattice Gauge Theories: An Introduction,” by H. K. Rothe, World Scientific, 2005.
- A recent lattice treatment of the string tension, glueballs, *etc.*, is found in C. R. Allton *et al.* [UKQCD Collaboration], *Phys. Rev.* **D65**, 054502 (2002).
- The effective field theory approach to chiral perturbation theory were first developed in S. Weinberg *Phys. Rev.* **166** 1568 (1968). The loop corrections to the theory were first worked out by J. Gasser and H. Leutwyler,

Annals Phys. (NY) **158**, 142 (1984). Our presentation follows the review: C. P. Burgess, *Phys. Rept.* **330**, 193 (2000) [arXiv:hep-th/9808176].

- A recent review of heavy quarkonium physics is given by N. Brambilla *et al.*, “Heavy quarkonium physics,” arXiv:hep-ph/0412158.

Chapter 9:

- The original papers developing heavy quark effective methods are N. Isgur and M. B. Wise, *Phys. Lett. B* **237**, 527 (1990); *Phys. Lett. B* **232**, 113 (1989).
- Several groups make their fits to parton distribution functions available online, for instance at <http://www.phys.psu.edu/~cteq> (CTEQ group) and <http://durpdg.dur.ac.uk/hepdata/> (MRST group)
- The parton idea was proposed in R. P. Feynman, *Phys. Rev. Lett.* **23**, 1415 (1969). A recent thorough overview of partons and perturbative QCD is R. Brock *et al.* [CTEQ Collaboration], *Rev. Mod. Phys.* **67**, 157 (1995).
- A thorough discussion of factorization by its developers is J. C. Collins, D. E. Soper and G. Sterman, *Adv. Ser. Direct. High Energy Phys.* **5**, 1 (1988) [arXiv:hep-ph/0409313].
- The Ademollo-Gatto theorem was proven in M. Ademollo and R. Gatto, *Phys. Rev. Lett.* **13**, 264 (1964).
- The influence of anomalies on low-energy meson interactions is first examined in J. Wess and B. Zumino, *Phys. Lett.* **B37**, 95 (1971).
- Up to date results on the experimental constraints on the CKM parameters and the unitarity triangle are maintained by the CKMfitter group, <http://www.slac.stanford.edu/xorg/ckmfitter/ckm.welcome.html> whose most recent publication is J. Charles *et al.*, *Euro. J. Phys.* **C41**, 1–131 (2005), which also contains an extensive review of the theoretical and experimental status of neutral meson mixing.
- An extensive and advanced review on neutral meson mixing, including higher order QCD corrections to box diagrams, is given in G. Buchalla, A. J. Buras and M. E. Lautenbacher, *Rev. Mod. Phys.* **68**, 1125 (1996) [arXiv:hep-ph/9512380].
- The original derivation of the box diagram contributions to neutral meson mixing appears in T. Inami and C. S. Lim, *Prog. Theor. Phys.* **65**, 297 (1981) [Erratum-ibid. **65**, 1772 (1981)].

Chapter 10:

- For recent results on cosmological neutrino mass bounds, see D. N. Spergel *et al.* [WMAP collaboration], arXiv:astro-ph/0603449; U. Seljak, A. Slosar and P. McDonald, arXiv:astro-ph/0604335.

- A detailed treatment of the standard solar model and its neutrinos appears in J. N. Bahcall, M. H. Pinsonneault and S. Basu, *Astrophys. J.* **555**, 990 (2001).
- For a recent fit to world data on neutrino oscillations, see for instance M. Maltoni, T. Schwetz, M. A. Tortola and J. W. F. Valle, *New J. Phys.* **6**, 122 (2004); [arXiv:hep-ph/0405172](#) is updated as of 2006.
- Bounds on sterile neutrinos: J. N. Bahcall, M. C. Gonzalez-Garcia and C. Pena-Garay, *Phys. Rev.* **C66**, 035802 (2002); B. C. Chauhan and J. Pulido, *JHEP* **0412**, 040 (2004); J. Pulido and B. C. Chauhan, *Nucl. Phys. Proc. Suppl.* **145**, 37 (2005).
- Bounds on sterile neutrinos from supernovae are discussed by J. Fetter *et.al.*, *Astropart. Phys.* **18**, 433 (2003) [[arXiv:hep-ph/0205029](#)], and G. Cacciapaglia *et.al.*, *Phys. Rev.* **D68**, 033013 (2003).
- Seesaw neutrino masses were independently proposed by P. Minkowski, *Phys. Lett.* **B67** 421 (1977); M. Gell-Mann, P. Ramond and R. Slansky, in *Supergravity*, ed. by D.Z. Freedman *et.al.*, North Holland, 1979; T. Yanagida, in “the Proceedings of the Workshop on the Baryon Number of the Universe and Unified Theories,” ed. by O. Sawada and A. Sugamoto, 1979.

Chapter 11:

- The technicolor proposal was made in S. Weinberg, *Phys. Rev.* **D19**, 1277 (1979); L. Susskind, *Phys. Rev.* **D19**, 2619 (1979). For a review of technicolor: K. Lane, “Two lectures on technicolor,” [arXiv:hep-ph/0202255](#).
- Interest in large extra dimensions as a solution of the hierarchy problem was ignited by N. Arkani-Hamed, S. Dimopoulos and G. R. Dvali, *Phys. Lett.* **B 429**, 263 (1998) [[arXiv:hep-ph/9803315](#)] and by L. Randall and R. Sundrum, *Phys. Rev. Lett.* **83**, 3370 (1999) [[arXiv:hep-ph/9905221](#)].
- Early discussions of the MSSM are P. Fayet, *Phys. Lett.* **B69**, 489 (1977); *Phys. Lett.* **B70**, 461 (1977); S. Dimopoulos and H. Georgi, *Nucl. Phys.* **B193**, 150 (1981); S. Weinberg, *Phys. Rev.* **D26**, 287 (1982).
- For an exhaustive review of SUSY and the MSSM, see H. E. Haber and G. L. Kane, *Phys. Rept.* **117**, 75 (1985).
- An up to date review of supersymmetry breaking and experimental constraints on the MSSM can be found in D. J. H. Chung *et al.*, *Phys. Rept.* **407**, 1 (2005).
- Vacuum stability bounds are discussed in M. Sher, *Phys. Lett.* **B317**, 159 (1993); T. Hambye and K. Riesselmann, [arXiv:hep-ph/9708416](#).
- A clear introduction to instanton methods may be found in “The Uses

Of Instantons”, by S.R. Coleman, in “the proceedings of the 1977 Erice International School on Subnuclear Physics.”

- A symmetry proposal to solve the strong CP problem was made by R. D. Peccei and H. R. Quinn, *Phys. Rev. Lett.* **38** 1440 (1977); *Phys. Rev.* **D16** 1791 (1977), after which it was realized that it implied the existence of a very light axion; S. Weinberg, *Phys. Rev. Lett.* **40** 223 (1978). For a review of strong CP and axions, see H. Y. Cheng, *Phys. Rept.* **158**, 1 (1988).
- The axion in astrophysics and cosmology is reviewed in M. S. Turner, *Phys. Rept.* **197**, 67 (1990).
- We follow here the dimension-6 analysis of refs. S. Weinberg, *Phys. Rev. Lett.* **43**, 1566 (1979); *Phys. Rev.* **D22**, 1694 (1980); F. Wilczek and A. Zee, *Phys. Rev. Lett.* **43**, 1571 (1979).
- The Grand Unification proposal is made by J. Pati and A. Salam, *Phys. Rev. Lett.* **31**, 275 (1973); the $SU(5)$ proposal in H. Georgi and G. Glashow, *Phys. Rev. Lett.* **32**, 438 (1974); and $SO(10)$ by H. Georgi, in *Particles and Fields – 1974*, ed. by C. Carlson (American Institute of Physics, 1975).
- Gauge coupling unification was described by H. Georgi, H. R. Quinn and S. Weinberg, *Phys. Rev. Lett.* **33**, 451 (1974), and for supersymmetric models in S. Dimopoulos, S. Raby and F. Wilczek, *Phys. Rev.* **D24**, 1681 (1981).

Index

- 2π counting, 156
- $\Delta I = \frac{1}{2}$ Rule, 356
- Δ baryon, 302
- Λ baryon, 302
- Ω baryon, 302
- Σ baryon, 302
- Σ^* baryon, 302
- Θ_3 , and electric dipole moments, 465
- Ξ baryon, 302
- Ξ gauges, 507
- Ξ^* baryon, 302
- γ^5 , convention for, 519
- Λ_{QCD} , 278
- ρ parameter, 236
- $SU_c(3) \times SU_L(2) \times U_Y(1)$, 54

- A terms, supersymmetry breaking, 450
- abelian Higgs model, 40
- accidental conservation laws, in the standard model, 95
- accidental symmetries, 239, 241
 - and the Fermi theory, 241
- accidental symmetry, 68
- action, 12
- Ademollo–Gatto theorem, 325, 362
- adjoint Higgs models, 107
- adjoint representation, 491
- Altarelli–Parisi (DGLAP) equations, 348
- angular momentum operators, 498
- annihilation operators, 4, 13, 21
- annihilation to photons, 217
- anomalies, 96
 - and π^0 decay, 297, 364
 - and $Q_e + Q_p$, 101
 - and Q_ν , 101
 - and baryon number, 102
 - and lepton number, 102
 - and neutrino stability, 154
 - between electromagnetism and chiral symmetries, 297
 - gauge, 99
 - involving the Lorentz group, 102
- anomalous magnetic moment of a proton, 323

- anomaly matching, 366
- antifundamental representation, 492
- antiparticles, 14
- asymmetries, left-right, 200
- antisymmetry of the baryonic wave function, 301
- asymmetry, forward–backward, 201
- asymptotic freedom, 455
- atmospheric neutrino oscillations, 404
- axioms of field theory, 7
- axion, 460
- axion potential, 468
- axions, as cold dark matter, 469

- B–L conservation, 104
- B–L symmetry, gauged, 107
- BaBar experiment, 228
- barn, 198
- baryon masses, 300
 - and chiral symmetry breaking, 317
- baryon number, 95, 286
 - anomalies, 102
 - of QCD bound states, 287
- baryons, 71, 281
 - and flavor symmetry, 300
 - defined, 281
 - octet and decuplet, 302
 - valence quark content of, 302
 - wave function antisymmetry, 301
- basic principles, 7
- Belle experiment, 228
- beta decay, double, 418
- Bhabha scattering, 223
- Bjorken x , 331
- Bloch–Nordsieck resummation, 221
- boosts, generator of, 499
- boson decays, 127
- bottomonium, 285
- bound states, 113
 - $b\bar{b}$, 200
 - $c\bar{c}$, 200
 - of heavy quarks, 284
 - strongly interacting, 75

- branching fraction, 137
- branching fractions, for Z^0 bosons, 138
 - in W^\pm decays, 145
- branching ratios of the τ lepton, 159
- breaking of $U_V(3)$ symmetry by masses, 295
- Breit frame, 332
- Breit–Wigner lineshape, 206
- Breit–Wigner resonance, 208
- C symmetry, 74, 75, 77, 80–1, 85
 - and QCD bound states, 287
 - broken by Z and W couplings, 89
- C, definition, 86
- Cabbibo angle, 81
- Callan–Gross sum rule, 337
- canonical form: fermions, 23
 - scalars, 17
- Casimir, in $SU_c(3)$, 346
- causality, 13
- center-of-mass frame, 191
- CERN, 202
- charge conjugation symmetry, 86
- charge radius of the proton, 324
- charged-current interactions, 78, 154, 235
 - and W^\pm boson decay, 143
 - and Fermi theory, 235
- Charmonium, 285
- Cherenkov radiation, 224
- Chern–Simons number, 461
- chiral Lagrangian, 310
- chiral perturbation theory, 279
- chiral symmetry, 103, 290, 292, 321
 - breaking, and baryon masses, 317
 - and spectral degeneracies, 297
 - broken by electromagnetism, 296
 - explicit breaking, 294
 - spontaneous breaking, 304
- circular polarization, 27
- CKM matrix, 79
- Clifford algebra, 505, 519
- cluster decomposition, 7
- cold dark matter, 469
- collinear photon emission, 218
- color, 54
- color confinement, 282
- commutation relations, 38
- Compton scattering, 214
- confinement, 282
- confinement hypothesis, 282
- confinement scale, Λ_c , 283
- confinement, linear, 283
- conservation laws, and stable particles, 153
 - approximate, 239
- conservation of baryon number, 95
 - lepton number, 95
- conserved vector current, 325
- conventions, and γ^5 , 519
- corrections near a resonance, 203
- cosmological constant problem, 437–8
- Coulomb barrier, 326
- coupling constants, 485
 - couplings, fermion- Z^0 boson, 137
 - gluon–fermion, 74
 - covariant derivative, 38
 - covariant derivatives of standard model fields, 58
 - CP symmetry, 460
 - broken by Kobayashi–Maskawa matrix phases, 89
 - CP violation, 401
 - in neutrino physics, 401
 - creation operators, 4, 13, 21
 - cross section, 112
 - crossed graphs, 212
 - crossing symmetry, 14, 211, 217
 - CTEQ collaboration, 349
 - cubic electroweak interactions, 77
 - current conservation, 215
 - custodial $SU(2)$ symmetry, 66
- D term, supersymmetric, 446
- decay rate, of a polarized Z^0 boson, 136
 - of the Z^0 boson, 135
- decay rates, 121
- decay, of the μ^- lepton, 155
 - of the τ^- lepton, 158
- decays, of elementary bosons, 127
 - of leptons, 153
- decoupling, 436
- deep inelastic scattering, 331
 - kinematics, 331
- delta baryon, 302
- density matrix, 129
- DGLAP (Altarelli–Parisi) equations, 348
- dilepton, 339
- dimension-5 operators in the standard model, 426
- Dirac conjugate, 22, 505
 - field, 69
 - neutrino, 397
 - neutrino masses, 415
 - neutrinos, 411
 - spinor, 69, 504
 - traces, 131
 - traces, evaluation, 131
- discrete symmetries, 85
- divergences, soft, 347
- divergent soft photon corrections, 220
- double beta decay, neutrinoless, 418
- Drell–Yan process, 339
 - differential cross section, 341
- dyadics, of fermion spinors, 131
- Dynkin index, 479
- effective Lagrangian, 231
- effective Lagrangians, 231
- effective theory, of QED with hadrons, 321
- electric charge, 39
 - and anomalies, 101
- electric dipole moment, for the neutron, 465
- electric dipole moments, and Θ_3 , 465
- electromagnetic breaking of chiral symmetry, 296

- electromagnetic interactions, 83
- electromagnetism, as a symmetry of the vacuum, 64
- electron charge, 83
- electron number, 95
- electroweak interactions, 76
- elementary boson decays, 127
- endpoint, 169
- eta' meson, mass, 313
- evolution equations for PDFs, 348
- expansion, in large M_W , 163
- experimental values, of masses, 485
- external lines, Feynman rules, 172
- F term, supersymmetric, 446
- factorization, 207, 330
 - on resonance, 207
- Fadeev–Popov ghosts, 507
- families, 54
- Fermi coupling constant, 164
- Fermi effective theory, 225
- Fermi Lagrangian, 235
- Fermi statistics—and signs in interference terms, 213
- Fermi theory, 231, 233
 - of neutral currents, 236
 - of the weak interactions, 425
 - and accidental symmetries, 241
- fermion electroweak couplings, Feynman rules, 178
- fermion loops, minus sign associated with, 180
- fermion masses, 69
- fermionic content of the standard model, 56
- fermionic fields: canonical form, 23
- fermions, charges in the standard model, 56
 - Dirac, 69
- Feynman gauge, 174, 507
- Feynman graph, symmetry factor, 180
- Feynman rules, 153, 172, 508
 - and metric convention, 522
 - combinatorial factors, 175
 - for external lines, 172
 - for incoming lines, 172
 - for internal lines, 173
 - for outgoing lines, 173
 - ghosts, 515
 - ghosts and unphysical scalars, 509
 - ordering of spinorial indices, 173
 - vertices for the standard model, 175
- field strength, 27
- fields, Dirac, 69
- fields, transformation under a symmetry, 488
- Fierz identity, 52, 225
- Fierz rearrangements, 472
- flavor change, absent in Z couplings, 84
- flavor changing, 81
 - neutral currents, 73, 84
 - absent in strong interactions, 75
- flavor octet mesons, 298
- flavor problem, 470
- flavor singlet mesons, 298
- Fock space, 4
- forward–backward asymmetry, 201
 - on resonance, 207
- fragmentation, 344
- fragmentation function, 344
- free quarks, non-observation, 282
- fundamental representation, 492
- G_F , the Fermi coupling constant, 164
- gamma matrices, 21, 505
 - Hermiticity of, 519
- gauge anomalies, 99
 - bosons, 37
 - choice, and Feynman rules, 174
 - unitary, 42
 - coupling unification, 479
 - fixing, 62
 - group, of the standard model, 54
 - invariance, 29, 215
 - renormalizable ξ , 507
 - unitary, 62, 507
- gaugino, 444
- gaugino Yukawa couplings, 446
- Gell-Mann–Okubo mass relation, 317
- Gell-Mann matrices, 44
- generators of Lorentz group, 498
- ghosts, 507
 - Feynman rules, 515
- GIM mechanism, 376, 378, 400
 - leptons, 412
- global symmetries, and Yukawa interactions, 94
- glueballs, 71
- gluon, 44
- gluon emission, and gauge invariance, 217
- gluon self-couplings, Feynman rules, 176
- gluon–fermion coupling, Feynman rules, 177
- gluon–fermion couplings, 74
- gluons, 54, 74
 - self-couplings, 74
- Goldberger–Treiman relation, 364
- Goldstone boson, 305
- Goldstone's Theorem, 306
- grand unified theories, 473, 476
- group theory, 487
- gyroscopic ratio, of protons and neutrons, 323
- hadron–hadron collisions, 338
- hadronic heavy quark pair production, 341
- hadronic scattering, 320
- hadronization, 344
- hadrons, 54, 70, 75, 140
 - light, 290
- Hamiltonian, 3
 - of a free system, 6
 - physical constraints on, 7
- Heaviside function, 116, 162
- heavy particles, and non-renormalizable interactions, 233
- heavy-quark effective theory, 326

- heavy-quark pair production, hadronic, 341
- heavy quarks, bound states, 284
- helicity conservation, 136
 - suppression, 413
 - and angular distributions of final states, 136
- hierarchy problem, 440
- Higgs-boson decays, and longitudinal gauge boson polarizations, 149
- Higgs-boson mass, 63, 74
- Higgs boson, decays, 146
 - decays into bosons, 148
 - decays into fermions, 147
- Higgs field, 60
- Higgs mechanism, 42
- Higgs self-couplings, 72
 - Feynman rules, 175
- Higgs-fermion couplings, 73
 - Feynman rules, 176
- Higgs-gauge boson couplings, 72
 - Feynman rules, 175
- Higgsino, 445
- Higgsstrahlung, 228, 431
- higher-order corrections, 141
 - parametric estimates, 141
- Hilbert space, 3, 112
 - and symmetries, 487
- hypercharge, 54
- inclusive reactions, 275
- incoming lines, Feynman rules, 172
- infinite momentum frame, 332
- initial-state radiation, 222, 345
 - and PDF evolution, 345
- instanton, 462
- interaction Hamiltonian, 128
 - and temporal gauge-field components, 128
- interactions, 34
 - W -fermion, 78
 - Z -fermion, 82–3
 - Z -fermion coupling strengths, 84
 - and gauge invariance, 36
 - and vector particles, 36
 - between W and Z bosons, 76
 - charged current, 78
 - electromagnetic, 83
 - electroweak, 76
 - electroweak cubic, 77
 - electroweak quartic, 77
 - gluon self, 74
 - gluon-fermion, 74
 - Higgs self, 72
 - Higgs-fermion, 73
 - Higgs-gauge boson, 72
 - neutral current, 82–3
 - non-renormalizable, 233
 - strong, 75
 - scalar, 35
 - scalar-spinor, 35
- interference, 212
 - and minus signs, 180
- internal lines, Feynman rules, 173
- internal symmetry, 32
- inverted hierarchy of neutrino masses, 405
- irreducible representation, 490
- isospin, 295
- J/ψ particle, 284
- Jets, 330
 - and collinear radiation, 344
- Josephson effect, 251
- K_L decays, 73
- kinematics, 189
 - in deep inelastic scattering, 331
- Klein-Gordon equation, 16
- Kobayashi-Maskawa matrix, 79
- Lab frame, 190
- Lagrangian, 12
- Lagrangian density, 12
- Lagrangian terms, involving the Higgs boson, 60
- Lagrangian, of a spin-zero field, 17
 - of a spin-one field, 26
 - of a spin-half field, 24
 - of the standard model, 58
- lambda baryon, 302
- Landau gauge, 174, 507
- Landau pole, 456
- large extra dimensions, 442
- large logarithms, 218
- large mass expansion, 163, 231
- lattice QCD, 278
- left-handed spinor, 502
- left-right asymmetry, 200
 - on resonance, 207
- LEP collider, 183, 198, 209, 228
- LEP I experiment, 202
- LEP II experiment, 431
- lepton number, 95
 - anomalies, 102
- Leptons, 54, 75
- Levi-Civita tensor, 132
- Lie algebra, 38, 489
 - for the Lorentz group, 499
- Lie groups, 488
- lifetime, of the μ^- lepton, 157
 - of the τ^- lepton, 159
- linear confinement, 283
- linear potential, and confinement, 283
- lineshape, 206
- LMA oscillations, 404
- local symmetry, 487
- locality, 7, 8, 12
- loop corrections, 242
- loops, 242
- Lorentz covariance, and gauge boson propagators, 162

- Lorentz group, 494
 - and anomalies, 102
 - generators, 498
 - Lie algebra, 499
 - representations, 13, 19, 500
 - reducible representations, 19
- Lorentz invariance, 13, 162
 - of time ordering operation, 117
- Lorentz symmetry, 8
- Lorentz transformation, orthochronous, 497
 - proper, 497
- low-energy expansion, 231
- LSP (lightest supersymmetric particle), 448
 - as dark matter candidate, 448
- luminosity, 198

- magnetic moment, of a proton, 323
- Majorana mass, 399, 422
- Majorana neutrino, 399
- Majorana notation, 502, 503
- Majorana spinor, 20
- Majorana spinors, identities, 46
- Mandelstam variables, 189, 210
 - kinematically allowed ranges, 192
- mass formula, for mesons, 313
- mass relations, for baryons, 317
 - Gell-Mann–Okubo, 317
- mass shell, 202
- mass, A boson, 65
 - W boson, 64
 - Z boson, 65
 - fermions, 69
 - Higgs boson, 63, 74
 - Majorana, 399
 - neutrino, 70
- masses of the baryons, 300
 - of particles, 485
- matrix element, 117
 - for Z^0 decay, 129
- meson decays, 352
 - nonleptonic, 354
- meson mass formula, 313
- meson masses, and $SU(3)$ flavor symmetry, 298
 - related by chiral symmetry breaking, 313
- mesons, 71, 281, 298
 - defined, 281
- metric convention, 9, 21
 - and Feynman rules, 522
- metric conventions, 518
- microcausality, 7
- minimal supersymmetric standard model (MSSM), 444
- minus signs, in interference between diagrams, 180
- mixing, between ϕ and ω mesons, 316
- MNS matrix, 401
- MRST collaboration, 349
- MSW effect, 406
- MSW oscillations, adiabatic, 409
- MSW resonance, 409

- mu decays, 155
- muon decay rate, computed value, 168
- muon lifetime, 157
- muon number, 95
- muon-hadron production ratio R , 199
- Møller scattering, 212
- Møller wave operators, 114

- nanobarn, 198
- neutral K system, 73
- neutral-current interactions, 82–3, 128
 - and Z^0 decay, 128
- neutral currents, 236
 - and Fermi theory, 236
- neutral-pion decay, 364
 - and broken axial symmetry, 297
- neutrino-beam oscillations, 404
- neutrino CP violation, 82
- neutrino mass, 70, 82
 - absent in the standard model, 70
- neutrino masses, 106, 427
 - Dirac, 415
 - experimental values, 404
 - upper and lower bounds, 406
- neutrino oscillations, 395, 402
- neutrino, right-handed, 57, 70
 - sterile, 421
- neutrinoless double-beta decay, 418
- neutrinos propagating through solar medium, 406
- neutrinos, pseudo-Dirac, 415
- neutron, 302
 - decay, 362
- neutron electric-dipole moment, 465
- Noether current, 31
 - and gauge couplings, 37
- Noether's theorem, 13, 31
- non-abelian symmetries, 37
- non-leptonic meson decays, 354
- non-renormalizable interactions, 233, 424
- non-renormalizable operators, and the standard model, 426
- normal hierarchy of neutrino masses, 405
- nucleon–pion scattering, 364
- numerical values, of couplings, 485

- $O(3,1)$, 495
- oblique electroweak corrections, 266
- omega baryon, 302
- order parameter, 306
- orthochronous, 497
- oscillation length, 403
- outgoing lines, Feynman rules, 173

- P symmetry, 74, 75, 77, 80, 81, 85, 460
 - and QCD bound states, 287
 - broken by Z and W couplings, 89
- parameters of the standard model, 61
- parametric estimates, 139, 156
- parity, 497
 - definition, 85

- Particle Data Book*, 284
- Particle Data Group, 79
 - parameterization of CKM matrix, 79
- particle masses, 485
- parton distribution functions (PDFs), 335
- partons, 329
- Pascos–Wolfenstein ratio, 388
- Pauli matrices, 19, 58
- PDFs, 335
 - approximate scale independence, 345
 - scale dependence, 345
- pentaquark, 281
- pentaquarks, 303
- perturbative expansion, breakdown for low
 - energy QCD, 278
- phase space, 138
- phase-space integrals, three body, 165
- photon, 54
- photon emission, collinear, 218
- photon exchange, 197
- photons, as external states, 214
- pion-decay constant, from decay rates, 353
- pion decay, neutral, 364
- pion scattering, 357
- pions, as pseudo-Goldstone bosons, 308
- PMNS matrix, 262, 401
- Poincaré algebra, 499
- Poincaré invariance, 8
- polarization vectors, 27
- polarizations, summed over, 215
- polarized Z^0 boson decay rate, 136
- precision electroweak measurements, 266
- propagator, 161
 - Feynman rules, 508
- propagators, near resonance, 203
- proper Lorentz transformation, 497
- proton, 302
 - charge radius, 324
 - magnetic moment, 323
- pseudo-Dirac neutrinos, 415
- pseudo-Goldstone boson, 305, 308
- pseudoscalar, 506
- pseudoscalar mesons, 299
- pseudovector mesons, 299

- QCD, 43, 276
- QCD bound states, quantum numbers, 285
- QCD Lagrangian, 276
- QCD vacuum, and topology, 461
 - spontaneously breaks chiral symmetry, 304
- QED, 276
- QED Lagrangian, 276
- QED, including hadrons, 321
- quadratic Casimir C_F , 346
- quantum chromodynamics (QCD), 275
- quantum electrodynamics, 39
- quantum numbers, of QCD bound states, 285
- quark, 43
- quarkonia, 284
- quarkonium, 285

- quarks, 54
 - valence and sea, 336
- quartic electroweak interactions, 77

- R -parity, 448
- R_ξ gauge, 174
- radiative corrections, 218
 - and the Z line shape, 222
- rapidity, 340
- reactor-neutrino oscillations, 404
- redundant effective interactions, 482
- reflection positivity, 467
- Regge theory, 321
- renormalizability, 9, 15, 424
- representation theory, 489
- representation, adjoint, 491
 - irreducible, 490
 - singlet or trivial, 490
- representations, and Lie algebra, 490
 - of the Lorentz group, 500
- resonance, Z , 206
- resonances, 203
 - hadronic, 321
- rho parameter, 236, 242
- right-handed neutrinos, 57, 105
- right-handed spinor, 502
- rotations, generator of, 499

- S -matrix, in terms of interaction Hamiltonian, 116
- s -channel process, 195, 211
- S -matrix, 111, 114
- scalar fields: canonical form, 17
 - vacuum energy, 18
- scalar particles, 16
- scalar representation, 500
- scalar self-interactions, 35
- scalar–spinor interactions, 35
- scattering process, s -channel, 195
- scattering states, 112
- sea quarks, 336
- seesaw mechanism, 416, 428
- seesaw neutrino masses, 478
- self-energy corrections, 204
- semi-leptonic decays, 352
- sigma baryon, 302
- sigma* baryon, 302
- singlet representation, 490
- SLAC, 202
- SLC experiment, 202
- slepton, 445
- SO(10) grand unified models, 478
- SO(3,1), 495
- soft-photon emission, 220
- soft singularities, 347
- soft-pion scattering, 350
- solar-neutrino oscillations, 404
- solar neutrinos, 396
- spectral degeneracies, and chiral symmetry, 297
- spectral relations, 33

- spin average, in initial state, 129
- spin dependent structure functions, 334
- spin-half particles, 19
- spin-one particles, 25
 - massive, 25
 - massless, 27
- spin-statistics theorem, 14
- spinor, 501
 - fields, 19
 - Dirac, 69
 - Majorana, 20
 - Weyl, 20
- spinorial indices, ordering in Feynman rules, 173
- spinors, Majorana versus Dirac, 55
- splitting functions, 348
- spontaneous breaking of chiral symmetry, 304
- spontaneous symmetry breaking, 40, 42
- spontaneously broken symmetry, 34
- squark, 445
- stability, 9, 14
- stability of neutrinos, 153
- stability of the electron, 153
- standard-model fields, covariant derivatives, 58
 - gauge group, 54
 - Lagrangian, 58, 60
 - continuous symmetries, 90
 - fermionic content, 56
 - parameters, 61
- step function, 116, 162
 - integral representation, 162
- sterile neutrino, 421
 - bounds, 416
- sterile neutrinos, 411
- strangeness, 286, 296
- string tension, in QCD, 283
- strong CP problem, 460, 461
- strong interactions, 54, 74–5
 - and bound states, 75
- structure constants, 489
- structure functions, 333
 - in terms of PDFs, 337
 - spin dependent, 334
- $SU(2)$ symmetry, custodial, 66
- $SU(3)$ flavor, 298
- $SU(3) \times SU(2) \times U(1)$, 54
- $SU(5)$ grand unified model, 476
- $SU(N)$, representation theory, 492
- Sum rule, Callan–Gross, 337
- Supernova 1987A, 224
- superpotential, 446
- supersymmetric D term, 446
- supersymmetric F term, 446
- supersymmetry, 439, 443
- supersymmetry breaking, 449
- supersymmetry-breaking A terms, 450
 - and gauge coupling unification, 481
- symmetries, 487
 - accidental, 85, 241
 - and the standard model, 85
 - as a group, 488
 - C, P, and T, 74, 75, 77, 80, 81, 85
 - continuous global, 90
 - discrete, 85
 - of the standard model, 90
 - symmetry-breaking vacuum, 41
 - symmetry factor, of a Feynman diagram, 180
 - symmetry transformation, 487
 - and fields, 488
 - symmetry, accidental, 68
 - broken by the ground state, 33
 - CP, 460
 - local, 487
 - P, 460
 - spontaneously broken, 34
- T -matrix, 117
- T -matrix element, 117
- T symmetry, 74, 75, 77, 80–1, 85
- t-channel process, 211
- tachyons, 40
 - absent about correct vacuum, 41
- tau decay rate, computed value, 168
- tau decay, branching ratios, 159
- tau decays, 158
- tau lifetime, 159
- tau number, 95
- technicolor models, 257, 442
- tensor product, 491
 - of spinor fields, 504
- tensor, totally antisymmetric, 132
- theta vacua, 462
- Thompson scattering, 386
- three-body phase-space integrations, 165
- time-dependent perturbation theory, 115
- time-ordering operator, 116
- time reversal, 497
 - definition, 85
- top-quark decay, 151
- top quark, decay width, 151
 - virtual, 238
- topological susceptibility, 468
- topology, and QCD vacuum, 461
- total derivative, in the Lagrangian, 461
- trace theorems, 132
- translation invariance, 13
- trivial representation, 490
- triviality, 455
- triviality bounds, on Higgs mass, 74
- triviality problem, 456
- twist, 338
- two Higgs doublet models, 106
- $U_A(1)$ anomaly, 293
- $U_L(3) \times U_R(3)$ approximate symmetry, 292
- $U_V(3)$ breaking, 295
- $U(1)$, representation theory, 491
- $U(3)$ chiral symmetry, 104
- uncrossed graphs, 212
- unitarity, 7, 12
- unitarity bounds, on Higgs mass, 74

- unitary gauge, 42, 62, 507
 - and Feynman rules, 174
- unphysical scalars, Feynman rules, 509
- upsilon particle, 284
- vacuum alignment, in the MSSM, 452
- vacuum energy: scalars, 18
- vacuum expectation value, 41
- vacuum, symmetry breaking, 41
- valence quarks, 336
- vector particles, 25
 - massive, 25
 - massless, 27
- vector representation, 500
- vertex, Feynman rules involving ghosts, 515
- vertices, for unphysical scalars, 509
- virtual Higgs exchange, 237
- virtual particles, 153
- virtual top-quark effects, 238
- W-boson decay, 143
 - interaction Hamiltonian, 143
- W-boson mass, 64
 - and renormalization, 251
- W bosons, 54
- W decay width, 143
- W decay, branching fractions, 145
- W lifetime, 145
- W-fermion interactions, 78
- wave packets, 112, 118
- weak interactions, 54
 - weak mixing angle, 65
 - weak universality, 81
 - Weinberg angle, 65
 - Wess–Zumino term, 366
 - Weyl notation, 503
 - Weyl spinor, 20, 502
 - Wick rotation, 523
 - Wolfenstein parameterization, 80
- X boson, of grand unified theories, 478
- x , Bjorken, 331
- Yukawa couplings, 35
 - and global symmetries, 94
 - from a superpotential, 446
- Z-boson decay, 127
- Z-boson pole, imaginary part, 204
- Z boson, couplings to fermions, 137
- Z bosons, 54
- Z charges of the fermions, 84
- Z decay rate, 135
- Z decay, and hadronic final states, 140
- Z lifetime, 138
- Z-pole line shape, 222
- Z resonance, 206
- Z width, 138
 - and higher order corrections, 141
 - parametric estimate, 139
- Z-fermion coupling strengths, 84
- Z-fermion interactions, 82–3