GRADUATE STUDENT SERIES IN PHYSICS

Series Editor:
Professor Douglas F Brewer, MA, DPhil
Emeritus Professor of Experimental Physics, University of Sussex

GAUGE THEORIES IN PARTICLE PHYSICS

A PRACTICAL INTRODUCTION
THIRD EDITION

Volume 2

Non-Abelian Gauge Theories: QCD and the Electroweak Theory

IAN J R AITCHISON
Department of Physics
University of Oxford

ANTHONY J G HEY
Department of Electronics and Computer Science
University of Southampton

IoP

INSTITUTE OF PHYSICS PUBLISHING
Bristol and Philadelphia

British Library Cataloguing-in-Publication Data

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

ISBN 0 7503 0950 4

Library of Congress Cataloging-in-Publication Data are available

Front cover image: Simulation by the ATLAS experiment of the decay of a Higgs boson into four muons (yellow tracks). © CERN Geneva.

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

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

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

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

Typeset in LaTeX 2_ε by Text 2 Text Limited, Torquay, Devon
Printed in the UK by MPG Books Ltd, Bodmin, Cornwall

To Jessie
and to
Jean, Katherine and Elizabeth

CONTENTS

PART 8
WEAK INTERACTIONS AND THE ELECTROWEAK THEORY

PREFACE TO VOLUME 2 OF THE THIRD EDITION

Volume 1 of our new two-volume third edition covers relativistic quantum mechanics, electromagnetism as a gauge theory, and introductory quantum field theory, and leads up to the formulation and application of quantum electrodynamics (QED), including renormalization. This second volume is devoted to the remaining two parts of the 'Standard Model' of particle physics, namely quantum chromodynamics (QCD) and the electroweak theory of Glashow, Salam and Weinberg.

It is remarkable that all three parts of the Standard Model are quantum gauge field theories: in fact, QCD and the electroweak theory are certain generalizations of QED. We shall therefore be able to build on the foundations of gauge theory, Feynman graphs and renormalization which were laid in Volume 1. However, QCD and the electroweak theory both require substantial extensions of the theoretical framework developed for QED. Most fundamentally, the discussion of global and local symmetries must be enlarged to include *non-Abelian symmetries*, and *spontaneous symmetry breaking*. At a somewhat more technical level, the *lattice (or path-integral) approach to quantum field theory*, and the *renormalization group* are both needed for access to modern work on QCD. For each of these theoretical elements, a self-contained introduction is provided in this volume. Together with their applications, this leads to a simple four-part structure (the numbering of parts, chapters and appendices continues on from Volume 1):

Part 5 Non-Abelian symmetries
Part 6 QCD and the renormalization group (including lattice field theory)
Part 7 Spontaneous symmetry breaking (including the spontaneous breaking of the approximate global chiral symmetry of QCD)
Part 8 The electroweak theory.

We have already mentioned several topics (path integrals, the renormalization group, and chiral symmetry breaking) which are normally found only in texts pitched at a more advanced level than this one—and which were indeed largely omitted from the preceding (second) edition. Nor, as we shall see, are these topics the only newcomers. With their inclusion in this volume, our book now becomes a comprehensive, practical and *accessible* introduction to the major theoretical and experimental aspects of the Standard Model. The emphasis is crucial: in once again substantially extending the scope of the book, we have tried hard not to compromise the title's fundamental aim—which is, as before, to make the chosen material accessible to the wide readership which the previous editions evidently attracted.

A glance at the contents will suggest that we have set ourselves a considerable challenge. On the other hand, not all of the topics are likely to be of equal interest to every reader. It may therefore be helpful to offer some more detailed guidance, while at the same time *highlighting* those items which are new to this edition.

First, then, non-Abelian symmetry. This refers to the fact that the symmetry transformations are matrices (acting on a set of fields), any two of which will generally not commute with each other, so that the order in which they are applied makes a difference. Much of the necessary mathematics already appears in the simpler case in which the symmetry is a global, rather than a local one. In chapter 12 we introduce global non-Abelian symmetries via the physical examples of the (approximate) SU(2) and SU(3) flavour symmetries of the strong interactions.

The underlying mathematics involved here is group theory. However, we take care to develop everything we need on a 'do-it-yourself' basis as we go along, so that no prior knowledge of group theory is necessary. Nevertheless, we have provided a new and fairly serious *appendix (M) on group theory*, which collects together the main relevant ideas, and shows how they apply to the groups we are dealing with (including the Lorentz group). We hope that this compact summary will be of use to those readers who want a sense of the mathematical unity behind the succession of specific calculations provided in the main text.

A further important global non-Abelian symmetry is also introduced in chapter 12—that of chiral symmetry, which is expected to be relevant if the quark masses are substantially less than typical hadronic scales, as is indeed the case. The apparent non-observation of this expected symmetry creates a puzzle, the resolution of which has to be deferred until part 7.

In chapter 13, the second in part 5, we move on to the local versions of SU(2) and SU(3) symmetry, arriving in section 13.5 at the corresponding non-Abelian gauge field theories which are the main focus of the book, being directly relevant to the electroweak theory and to QCD respectively. Crucial new physical phenomena appear, not present in QED—for example, the self-interactions among the gauge field quanta.

On the mathematical side, the algebraic (or group-theoretic) aspects developed in chapter 12 carry through unchanged into chapter 13, but the 'gauging' of the symmetry brings in some new geometrical concepts, such as *'covariant derivative', 'parallel transport', 'connection', and 'curvature'*. We decided against banishing these matters to an appendix, since they are such a significant part of the conceptual structure of all gauge theories, and moreover their inclusion allows instructive reference to be made to a theory otherwise excluded from mention, namely general relativity. All the same, practically-minded readers may want to pass quickly over sections 13.2 and 13.3, and also section 13.5.3, which explains why obtaining the correct Feynman rules for loops in a non-Abelian gauge theory is such a difficult problem, within the 'canonical' approach to quantum field theory as developed in volume 1.

Immediate application of the formalism can now be made to QCD, and this occupies most of the next three chapters, which form part 6. Chapter 14 introduces 'colour' as a dynamical degree of freedom, and leads on to the QCD Lagrangian. Some simple tree-graph applications are then described, using the techniques learned for QED. These provide a good first orientation to data, following 'parton model' and 'scaling' ideas.

But of course a fundamental question immediately arises: how can such an approach, based on perturbation theory, possibly apply to QCD which, after all, describes the strong interactions between quarks? The answer lies in the profound property (possessed only by non-Abelian gauge theories) called 'asymptotic freedom'—that is, the decrease of the effective interaction strength at high energies or short distances. Crucially, this property cannot be understood in terms of tree graphs: loops must be studied, and this immediately involves renormalization. In fact, perturbation theory becomes useful at high energies only after an infinite series of loop contributions has been effectively re-summed. The technique required to do this goes by the name of *the renormalization group* (RG), and it is described in chapter 15, along with applications to asymptotic freedom, and to the calculation of scaling violations in deep inelastic scattering.

We do not expect the majority of our readers to find chapter 15 easy going. But there is no denying the central importance of the RG in modern field theory, nor its direct relevance to experiment. In section 15.2 we have tried to provide an elementary introduction to the RG, by considering in detail the much simpler case of QED, using no more theory than is contained in chapter 11 of volume 1. Sections 15.4 and 15.5 are less central to the main argument, as is an *appendix (N) on dimensional regularization*.

In chapter 16, the third of part 6, we turn to the problem of how to extract predictions from a quantum field theory (in particular, QCD) in the non-perturbative regime. The available technique is computational, based on the discretized (lattice) version of *Feynman's path-integral formulation of quantum field theory*, to which we provide a simple introduction in section 16.4. A substantial bonus of this formulation is that it allows fruitful analogies to be drawn with the statistical mechanics of spin systems on a lattice. In particular, we hope that readers who may have struggled with the formal manipulations of chapter 15 will be refreshed by seeing RG ideas in action from a different and more physical point of view—that of 'integrating out' short distance degrees of freedom, leading to an *effective theory* valid at longer distances. The chapter ends with some illustrative results from lattice QCD calculations, in section 16.7. An *appendix (O) on Grassmann variables* is provided for those interested in seeing how the path-integral formalism can be made to work for fermions.

At this half-way stage, QCD has been established as the theory of strong interactions, by the success of both RG-improved perturbation theory and non-perturbative numerical computations.

Further progress requires one more fundamental idea—the subtle concept of spontaneous symmetry breaking, which forms the subject of part 7. Chapter 17 sets out *the basic theory of spontaneously broken global symmetries*, and also considers two physical examples in considerable detail, namely *the Bogoliubov superfluid* in section 17.3, and *the BCS superconductor* in section 17.7. It is of course true that these systems are not part of the standard model of particle physics. However, the characteristic methods and concepts developed for such systems provide valuable background for the particle physics applications of the idea, which follow in the next two chapters. In particular, our presentation of *chiral symmetry breaking* in chapter 18 follows Nambu's remarkable original analogy between fermion mass generation and the appearance of an energy gap in a superconductor. Section 18.3, on *linear and nonlinear sigma models*, is rather more optional, as is our brief introduction to *chiral anomalies* in section 18.4. In chapter 19, the third in part 7, we consider the spontaneous breaking of local (gauge) symmetries. Here the fundamental point is that it is possible for gauge quanta to acquire mass, while still preserving the local gauge symmetry of the Lagrangian. We consider applications both to the Abelian case of a superconductor (sections 19.2 and 19.4—once again, a valuable working model of the physics), and to the non-Abelian case required for the electroweak theory.

The way is now clear to develop the electroweak theory, in part 8. Chapter 20 is a self-contained review of weak interaction phenomenology, based on Fermi's 'current–current' model. New material here includes discussion of *the discrete symmetries* **C** *and* **P**, and of lepton number conservation taking into account the possibility that *neutrinos may be Majorana particles*, in support of which we provide *an appendix (P) on Majorana fermions*. Chapter 21 describes what goes wrong with the current–current model, and with theories in which the W and Z bosons are given a 'naive' mass, and suggests why a gauge theory is needed to avoid these difficulties. Finally, in chapter 22, all the pieces are put together in the presentation of the electroweak theory. New additions here include *three-family mixing via the CKM matrix*, together with more detail on *higher order (one-loop) corrections*, the *top quark*, and aspects of *Higgs phenomenology*. The remarkably precise agreement—thus far—between theory and experiment, which depends upon the inclusion of one-loop effects, makes it hard to deny that, when interacting weakly, Nature has indeed made use of the subtle intricacies of a renormalizable, spontaneously broken, non-Abelian chiral gauge theory.

But the story of the Standard Model is not yet quite complete. One vital part—the Higgs sector—remains virtual, and phenomenological. Further progress in understanding the mechanism of electroweak symmetry breaking, and of mass generation, requires input from the next generation of experiments, primarily at the LHC. We hope that we leave our readers with a sound grasp of what is at stake in these experiments, and a lively interest in their outcome.

Acknowledgments

Our expression of thanks to friends and colleagues, made in the first volume, applies equally to this one. More particularly: Keith Hamilton was a willing 'guinea pig' for chapters 14–17, and we thank him for his careful reading and encouraging comments; we are grateful to Chris Allton for advice on lattice gauge theory results, for section 16.7; and Nikki Fathers again provided essential help in making the electronic version. Above all, the constant and unstinting help of our good friend George Emmons throughout the genesis and production of the book has, once again, been invaluable.

In addition, we are delighted to thank two new correspondents, whom we look forward to greeting physically, as well as electronically. Paolo Strolin and Peter Williams both worked very carefully through Volume 1, and between them found a good many misprints and infelicities. An up-to-date list will be posted on the book's website:
http://bookmarkphysics.iop.org/bookpge.htm?&book=1130p.
Paolo also read drafts of chapters 12 and 14, and made many useful comments. We wish it had been possible to send him more chapters: however, he made numerous excellent suggestions for improving our treatment of weak interactions in the second edition (parts 4 and 5), and—where feasible—many of them have been incorporated into the present part 8. Errors, old and new, there still will be, of course: we hope readers will draw them to our attention.

Ian J R Aitchison and Anthony J G Hey
October 2003

PART 5

NON-ABELIAN SYMMETRIES

12

GLOBAL NON-ABELIAN SYMMETRIES

In the preceding volume, a very successful dynamical theory—QED—has been introduced, based on the remarkably simple *gauge principle*: namely that the theory should be invariant under local phase transformations on the wavefunctions (chapter 3) or field operators (chapter 7) of charged particles. Such transformations were characterized as *Abelian* in section 3.6, since the phase factors commuted. The second volume of this book will be largely concerned with the formulation and elementary application of the remaining two dynamical theories within the Standard Model—that is, QCD and the electroweak theory. They are built on a generalization of the gauge principle, in which the transformations involve more than one state, or field, at a time. In that case, the 'phase factors' become matrices, which generally do not commute with each other, and the associated symmetry is called a '*non-Abelian*' one. When the phase factors are independent of the spacetime coordinate x, the symmetry is a 'global non-Abelian' one; when they are allowed to depend on x, one is led to a non-Abelian gauge theory. Both QCD and the electroweak theory are of the latter type, providing generalizations of the Abelian U(1) gauge theory which is QED. It is a striking fact that all three dynamical theories in the Standard Model are based on a gauge principle of local phase invariance.

In this chapter we shall be mainly concerned with two global non-Abelian symmetries, which lead to useful conservation laws but not to any specific dynamical theory. We begin in section 12.1 with the first non-Abelian symmetry to be used in particle physics, the *hadronic isospin* 'SU(2) symmetry' proposed by Heisenberg (1932) in the context of nuclear physics, and now seen as following from the near equality of the u and d quark masses (on typical hadronic scales), and the flavour independence of the QCD interquark forces. In section 12.2 we extend this to $SU(3)_f$ flavour symmetry, as was first done by Gell-Mann (1961) and Ne'eman (1961)—an extension seen, in its turn, as reflecting the rough equality of the u, d and s quark masses, together with flavour independence of QCD. The 'wavefunction' approach of sections 12.1 and 12.2 is then reformulated in field-theoretic language in section 12.3.

In the last section of this chapter, we shall introduce the idea of a global *chiral* symmetry, which is a symmetry of theories with massless fermions. This may be expected to be a good approximate symmetry for the u and d quarks. But the anticipated observable consequences of this symmetry (for example, nucleon parity doublets) appear to be absent. This puzzle will be resolved in part 7, via the profoundly important concept of 'spontaneous symmetry breaking'.

The formalism introduced in this chapter for SU(2) and SU(3) will be required again in the following one, when we consider the local versions of these non-Abelian symmetries and the associated dynamical gauge theories. The whole modern development of non-Abelian gauge theories began with the attempt by Yang and Mills (1954) (see also Shaw 1955) to make hadronic isospin into a local symmetry. However, the beautiful formalism developed by these authors turned out *not* to describe interactions between hadrons. Instead, it describes the interactions between the *constituents* of the hadrons, namely quarks—and this in two respects. First, a local SU(3) symmetry (called $SU(3)_c$) governs the strong interactions of quarks, binding them into hadrons (see part 6). Second, a local SU(2) symmetry (called *weak isospin*) governs the weak interactions of quarks (and leptons); together with QED, this constitutes the electroweak theory (see part 8). It is important to realize that, despite the fact that each of these two local symmetries is based on the same group as one of the earlier global (flavour) symmetries, the physics involved is completely different. In the case of the strong quark interactions, the $SU(3)_c$ group refers to a new degree of freedom ('colour') which is quite distinct from flavour u, d, s (see chapter 14). In the weak interaction case, since the group is an SU(2), it is natural to use 'isospin language' in talking about it, particularly since flavour degrees of freedom are involved. But we must always remember that it is *weak* isospin, which (as we shall see in chapter 20) is an attribute of leptons as well as of quarks and, hence, physically quite distinct from hadronic spin. Furthermore, it is a parity-violating chiral gauge theory.

Despite the attractive conceptual unity associated with the gauge principle, the way in which each of QCD and the electroweak theory 'works' is actually quite different from QED and from each other. Indeed it is worth emphasizing very strongly that it is, *a priori*, far from obvious why either the strong interactions between quarks or the weak interactions should have anything to do with gauge theories at all. Just as in the U(1) (electromagnetic) case, gauge invariance forbids a mass term in the Lagrangian for non-Abelian gauge fields, as we shall see in chapter 13. Thus it would seem that gauge field quanta are necessarily massless. But this, in turn, would imply that the associated forces must have a long-range (Coulombic) part, due to exchange of these massless quanta—and of course in neither the strong nor the weak interaction case is that what is observed.[1] As regards the former, the gluon quanta are indeed massless but the contradiction is resolved by *non-perturbative* effects which lead to *confinement*, as we indicated in chapter 2. We shall discuss this further in chapter 16. In weak interactions, a third realization appears: the gauge quanta acquire mass via (it is believed) a second instance of *spontaneous symmetry breaking*, as will be explained in part 7. In fact, a further application of this idea is required in the electroweak theory because of

[1] Pauli had independently developed the theory of non-Abelian gauge fields during 1953 but did not publish any of this work because of the seeming physical irrelevancy associated with the masslessness problem (Enz 2002, pp 474-82; Pais 2002, pp 242-5).

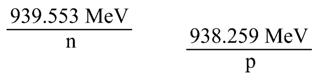

Figure 12.1. Early evidence for isospin symmetry.

the chiral nature of the gauge symmetry in this case: the quark and lepton masses also must be 'spontaneously generated'.

12.1 The flavour symmetry $\mathrm{SU(2)_f}$

12.1.1 The nucleon isospin doublet and the group SU(2)

The transformations initially considered in connection with the gauge principle in section 3.5 were just global phase transformations on a single wavefunction

$$\psi' = \mathrm{e}^{\mathrm{i}\alpha}\psi. \tag{12.1}$$

The generalization to non-Abelian invariances comes when we take the simple step—but one with many ramifications—of considering more than one wavefunction, or state, at a time. Quite generally in quantum mechanics, we know that whenever we have a set of states which are *degenerate* in energy (or mass) there is no unique way of specifying the states: any linear combination of some initially chosen set of states will do just as well, provided the normalization conditions on the states are still satisfied. Consider, for example, the simplest case of just two such states—to be specific, the neutron and proton (figure 12.1). This single near coincidence of the masses was enough to suggest to Heisenberg (1932) that, as far as the strong nuclear forces were concerned (electromagnetism being negligible by comparison), the two states could be regarded as truly degenerate, so that any arbitrary linear combination of neutron and proton wavefunctions would be entirely equivalent, as far as this force was concerned, for a single 'neutron' or single 'proton' wavefunction. This hypothesis became known as the 'charge independence of nuclear forces'. Thus redefinitions of neutron and proton wavefunctions could be allowed, of the form

$$\psi_\mathrm{p} \to \psi'_\mathrm{p} = \alpha\psi_\mathrm{p} + \beta\psi_\mathrm{n} \tag{12.2}$$
$$\psi_\mathrm{n} \to \psi'_\mathrm{n} = \gamma\psi_\mathrm{p} + \delta\psi_\mathrm{n} \tag{12.3}$$

for complex coefficients α, β, γ and δ. In particular, since ψ_p and ψ_n are degenerate, we have

$$H\psi_\mathrm{p} = E\psi_\mathrm{p} \qquad H\psi_\mathrm{n} = E\psi_\mathrm{n} \tag{12.4}$$

from which it follows that

$$H\psi'_\mathrm{p} = H(\alpha\psi_\mathrm{p} + \beta\psi_\mathrm{n}) = \alpha H\psi_\mathrm{p} + \beta H\psi_\mathrm{n} \tag{12.5}$$
$$= E(\alpha\psi_\mathrm{p} + \beta\psi_\mathrm{n}) = E\psi'_\mathrm{p} \tag{12.6}$$

and, similarly,

$$H\psi_{\rm n}' = E\psi_{\rm n}' \tag{12.7}$$

showing that the redefined wavefunctions still describe two states with the same energy degeneracy.

The two-fold degeneracy seen in figure 12.1 is suggestive of that found in spin-$\frac{1}{2}$ systems in the absence of any magnetic field: the $s_z = \pm\frac{1}{2}$ components are degenerate. The analogy can be brought out by introducing the *two-component nucleon isospinor*

$$\psi^{(1/2)} \equiv \begin{pmatrix} \psi_{\rm p} \\ \psi_{\rm n} \end{pmatrix} \equiv \psi_{\rm p}\chi_{\rm p} + \psi_{\rm n}\chi_{\rm n} \tag{12.8}$$

where

$$\chi_{\rm p} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \chi_{\rm n} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{12.9}$$

In $\psi^{(1/2)}$, $\psi_{\rm p}$ is the amplitude for the nucleon to have 'isospin up' and $\psi_{\rm n}$ is that for it to have 'isospin down'.

As far as the states are concerned, this terminology arises, of course, from the formal identity between the 'isospinors' of (12.9) and the two-component eigenvectors (4.59) corresponding to eigenvalues $\pm\frac{1}{2}\hbar$ of (true) spin: compare also (4.60) and (12.8). It is important to be clear, however, that the degrees of freedom involved in the two cases are quite distinct; in particular, even though both the proton and the neutron have (true) spin-$\frac{1}{2}$, the transformations (12.2) and (12.3) leave the (true) spin part of their wavefunctions completely untouched. Indeed, we are suppressing the spinor part of both wavefunctions altogether (they are of course 4-component Dirac spinors). As we proceed, the precise mathematical nature of this 'spin-$\frac{1}{2}$' analogy will become clear.

Equations (12.2) and (12.3) can be compactly written in terms of $\psi^{(1/2)}$ as

$$\psi^{(1/2)} \rightarrow \psi^{(1/2)'} = \mathbf{V}\psi^{(1/2)} \qquad \mathbf{V} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \tag{12.10}$$

where $\mathbf{V}$ is the indicated complex 2×2 matrix. Heisenberg's proposal, then, was that the physics of strong interactions between nucleons remained the same under the transformation (12.10): in other words, a symmetry was involved. We must emphasize that such a symmetry can *only* be exact in the *absence* of electromagnetic interactions: it is, therefore, an intrinsically approximate symmetry, though presumably quite a useful one in view of the relative weakness of electromagnetic interactions as compared to hadronic ones.

We now consider the general form of the matrix $\mathbf{V}$, as constrained by various relevant restrictions: quite remarkably, we shall discover that (after extracting an overall phase) $\mathbf{V}$ has essentially the same mathematical form as the matrix $\mathbf{U}$ of (4.81), which we encountered in the discussion of the transformation of (real) spin wavefunctions under rotations of the (real) space axes. It will be instructive to see how the present discussion leads to the same form (4.81).

We first note that $\mathbf{V}$ of (12.10) depends on four arbitrary complex numbers or, alternatively, on eight real parameters. By contrast, the matrix $\mathbf{U}$ of (4.81) depends on only three real parameters: two to describe the axis of rotation represented by the unit vector $\widehat{\boldsymbol{n}}$, together with a third for the angle of rotation θ. However, $\mathbf{V}$ is subject to certain restrictions and these reduce the number of free parameters in $\mathbf{V}$ to three, as we now discuss. First, in order to preserve the normalization of $\psi^{(1/2)}$, we require

$$\psi^{(1/2)'\dagger}\psi^{(1/2)'} = \psi^{(1/2)\dagger}\mathbf{V}^{\dagger}\mathbf{V}\psi^{(1/2)} = \psi^{(1/2)\dagger}\psi^{(1/2)} \tag{12.11}$$

which implies that $\mathbf{V}$ has to be *unitary*:

$$\mathbf{V}^{\dagger}\mathbf{V} = \mathbf{1}_2 \tag{12.12}$$

where $\mathbf{1}_2$ is the unit 2×2 matrix. Clearly this unitarity property is in no way restricted to the case of two states—the transformation coefficients for n degenerate states will form the entries of an $n \times n$ unitary matrix. A trivialization is the case $n = 1$, for which, as we noted in section 3.6, $\mathbf{V}$ reduces to a single phase factor as in (12.1), indicating how all the previous work is going to be contained as a special case of these more general transformations. Indeed, from the elementary properties of determinants, we have

$$\det\mathbf{V}^{\dagger}\mathbf{V} = \det\mathbf{V}^{\dagger}\cdot\det\mathbf{V} = \det\mathbf{V}^{*}\cdot\det\mathbf{V} = |\det\mathbf{V}|^2 = 1 \tag{12.13}$$

so that

$$\det\mathbf{V} = \exp(\mathrm{i}\theta) \tag{12.14}$$

where θ is a real number. We can separate off such an overall phase factor from the transformations mixing 'p' and 'n', because it corresponds to a rotation of the phase of both p and n wavefunctions by the *same* amount:

$$\psi'_{\mathrm{p}} = \mathrm{e}^{\mathrm{i}\alpha}\psi_{\mathrm{p}} \qquad \psi'_{\mathrm{n}} = \mathrm{e}^{\mathrm{i}\alpha}\psi_{\mathrm{n}}. \tag{12.15}$$

The $\mathbf{V}$ corresponding to (12.15) is $\mathbf{V} = \mathrm{e}^{\mathrm{i}\alpha}\mathbf{1}_2$, which has determinant $\exp(2\mathrm{i}\alpha)$ and is, therefore, of the form (12.1) with $\theta = 2\alpha$. In the field-theoretic formalism of section 7.2, such a symmetry can be shown to lead to the conservation of baryon number $N_{\mathrm{u}} + N_{\mathrm{d}} - N_{\bar{\mathrm{u}}} - N_{\bar{\mathrm{d}}}$, where bar denotes the anti-particle.

The new physics will lie in the remaining transformations which satisfy

$$\det\mathbf{V} = +1. \tag{12.16}$$

Such a matrix is said to be a *special* unitary matrix—which simply means it has unit determinant. Thus, finally, the $\mathbf{V}$'s we are dealing with are *special, unitary*, 2×2 *matrices*. The set of all such matrices form a *group*. The general defining properties of a group are given in appendix M. In the present case, the elements of the group are all such 2×2 matrices and the 'law of combination' is just ordinary matrix multiplication. It is straightforward to verify (problem 12.1) that all the defining properties are satisfied here; the group is called ‘SU(2)’, the ‘S’ standing for ‘special’, the ‘U’ for ‘unitary’ and the ‘2’ for ‘2×2’.

SU(2) is actually an example of a *Lie group* (see appendix M). Such groups have the important property that their physical consequences may be found by considering ‘infinitesimal’ transformations, that is—in this case—matrices $\mathbf{V}$ which differ only slightly from the ‘no-change’ situation corresponding to $\mathbf{V} = \mathbf{1}_2$. For such an infinitesimal SU(2) matrix $\mathbf{V}_{\text{infl}}$, we may therefore write

$$\mathbf{V}_{\text{infl}} = \mathbf{1}_2 + \mathrm{i}\xi \tag{12.17}$$

where ξ is a 2×2 matrix whose entries are all first-order small quantities. The condition $\det \mathbf{V}_{\text{infl}} = 1$ now reduces, on neglect of second-order terms $O(\xi^2)$, to the condition (see problem 12.2)

$$\operatorname{Tr} \xi = 0. \tag{12.18}$$

The condition that $\mathbf{V}_{\text{infl}}$ be unitary, i.e.

$$(\mathbf{1}_2 + \mathrm{i}\xi)(\mathbf{1}_2 - \mathrm{i}\xi^\dagger) = \mathbf{1}_2 \tag{12.19}$$

similarly reduces (in first order) to the condition

$$\xi = \xi^\dagger. \tag{12.20}$$

Thus ξ is a 2×2 traceless Hermitian matrix, which means it must have the form

$$\xi = \begin{pmatrix} a & b - \mathrm{i}c \\ b + \mathrm{i}c & -a \end{pmatrix} \tag{12.21}$$

where a, b, c are infinitesimal parameters. Writing

$$a = \epsilon_3/2 \qquad b = \epsilon_1/2 \qquad c = \epsilon_2/2 \tag{12.22}$$

(12.21) can be put in the more suggestive form

$$\xi = \boldsymbol{\epsilon} \cdot \boldsymbol{\tau}/2 \tag{12.23}$$

where $\boldsymbol{\epsilon}$ stands for the three quantities

$$\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \epsilon_3) \tag{12.24}$$

which are all first-order small. The three matrices $\boldsymbol{\tau}$ are just the familiar Hermitian Pauli matrices

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \tau_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \qquad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{12.25}$$

here called 'tau' precisely in order to distinguish them from the mathematically identical 'sigma' matrices which are associated with the real spin degree of freedom. Hence, a general infinitesimal SU(2) matrix takes the form

$$\mathbf{V}_{\mathrm{infl}} = (\mathbf{1}_2 + \mathrm{i}\boldsymbol{\epsilon} \cdot \boldsymbol{\tau}/2) \tag{12.26}$$

and an infinitesimal SU(2) transformation of the p–n doublet is specified by

$$\begin{pmatrix} \psi_{\mathrm{p}}' \\ \psi_{\mathrm{n}}' \end{pmatrix} = (\mathbf{1}_2 + \mathrm{i}\boldsymbol{\epsilon} \cdot \boldsymbol{\tau}/2) \begin{pmatrix} \psi_{\mathrm{p}} \\ \psi_{\mathrm{n}} \end{pmatrix}. \tag{12.27}$$

The $\boldsymbol{\tau}$-matrices clearly play an important role, since they determine the forms of the three independent infinitesimal SU(2) transformations. They are called the *generators* of infinitesimal SU(2) transformations; more precisely, the matrices $\boldsymbol{\tau}/2$ provide a particular *matrix representation* of the generators, namely the two-dimensional or 'fundamental' one (see appendix M). We note that they do not commute amongst themselves: rather, introducing $\mathbf{T}^{(1/2)} \equiv \boldsymbol{\tau}/2$, we find (see problem 12.3)

$$[T_i^{(1/2)}, T_j^{(1/2)}] = \mathrm{i}\epsilon_{ijk} T_k^{(1/2)}, \tag{12.28}$$

where i, j and k run from 1 to 3 and a sum on the repeated index k is understood as usual. The reader will recognize the commutation relations (12.28) as being precisely the same as those of angular momentum operators in quantum mechanics:

$$[J_i, J_j] = \mathrm{i}\epsilon_{ijk} J_k. \tag{12.29}$$

In that case, the choice $J_i = \sigma_i/2 \equiv J_i^{(1/2)}$ would correspond to a (real) spin-$\frac{1}{2}$ system. Here the identity between the tau's and the sigma's gives us a good reason to regard our 'p–n' system as formally analogous to a 'spin-$\frac{1}{2}$' one. Of course, the 'analogy' was made into a mathematical identity by the judicious way in which ξ was parametrized in (12.23).

The form for a *finite* SU(2) transformation $\mathbf{V}$ may then be obtained from the infinitesimal form using the result

$$\mathrm{e}^A = \lim_{n\to\infty} (1 + A/n)^n \tag{12.30}$$

generalized to matrices. Let $\boldsymbol{\epsilon} = \boldsymbol{\alpha}/n$, where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ are three real finite (not infinitesimal) parameters, apply the infinitesimal transformation n times and let n tend to infinity. We obtain

$$\mathbf{V} = \exp(\mathrm{i}\boldsymbol{\alpha} \cdot \boldsymbol{\tau}/2) \tag{12.31}$$

so that

$$\psi^{(1/2)\prime} \equiv \begin{pmatrix} \psi_{\mathrm{p}}' \\ \psi_{\mathrm{n}}' \end{pmatrix} = \exp(\mathrm{i}\boldsymbol{\alpha} \cdot \boldsymbol{\tau}/2) \begin{pmatrix} \psi_{\mathrm{p}} \\ \psi_{\mathrm{n}} \end{pmatrix} = \exp(\mathrm{i}\boldsymbol{\alpha} \cdot \boldsymbol{\tau}/2)\psi^{(1/2)}. \tag{12.32}$$

Note that in the finite transformation, the generators appear in the exponent. Indeed, (12.31) has the form

$$\mathbf{V} = \exp(\mathrm{i}G) \tag{12.33}$$

where $G = \boldsymbol{\alpha} \cdot \boldsymbol{\tau}/2$, from which the unitary property of $\mathbf{V}$ easily follows:

$$\mathbf{V}^\dagger = \exp(-\mathrm{i}G^\dagger) = \exp(-\mathrm{i}G) = \mathbf{V}^{-1}, \tag{12.34}$$

where we used the Hermiticity of the tau's. Equation (12.33) has the general form

$$\text{unitary matrix} = \exp(\mathrm{i}\ \text{Hermitian matrix}), \tag{12.35}$$

where the 'Hermitian matrix' is composed of the generators and the transformation parameters. We shall meet generalizations of this structure in the following section for SU(2), again in section 12.2 for SU(3), and a field-theoretic version of it in section 12.3.

As promised, (12.32) is of essentially the same mathematical form as (4.81). In each case, three real parameters appear: in (4.81) there are three parameters to describe the axis $\hat{\boldsymbol{n}}$ and angle θ of rotation; in (12.32) there are just the three components of $\boldsymbol{\alpha}$. We can always[2] write $\boldsymbol{\alpha} = |\boldsymbol{\alpha}|\hat{\boldsymbol{\alpha}}$ and identify $|\boldsymbol{\alpha}|$ with θ and $\hat{\boldsymbol{\alpha}}$ with $\hat{\boldsymbol{n}}$.

In the form (12.32), it is clear that our 2×2 isospin transformation is a generalization of the global phase transformation of (12.1), except that

(a) there are now *three* 'phase angles' $\boldsymbol{\alpha}$; and
(b) there are non-commuting matrix operators (the $\boldsymbol{\tau}$'s) appearing in the exponent.

The last fact is the reason for the description 'non-Abelian' phase invariance. As the commutation relations for the $\boldsymbol{\tau}$ matrices show, SU(2) is a non-Abelian group in that two SU(2) transformations do not, in general, commute. By contrast, in the case of electric charge or particle number, successive transformations clearly commute: this corresponds to an Abelian phase invariance and, as noted in section 3.6, to an Abelian U(1) group.

We may now put our initial 'spin-$\frac{1}{2}$' analogy on a more precise mathematical footing. In quantum mechanics, states within a degenerate multiplet may conveniently be characterized by the eigenvalues of a complete set of Hermitian operators which commute with the Hamiltonian and with each other. In the case of the p–n doublet, it is easy to see what these operators are. We may write (12.4), (12.6) and (12.7) as

$$H_2\psi^{(1/2)} = E\psi^{(1/2)} \tag{12.36}$$

and

$$H_2\psi^{(1/2)'} = E\psi^{(1/2)'} \tag{12.37}$$

[2] It is not completely obvious that the general SU(2) matrix *can* be parametrized by an angle θ with $0 \le \theta \le 2\pi$, and $\hat{\boldsymbol{n}}$: for further discussion, see appendix M, section M.7.

where H_2 is the 2×2 matrix

$$H_2 = \begin{pmatrix} H & 0 \\ 0 & H \end{pmatrix}. \tag{12.38}$$

Hence H_2 is proportional to the unit matrix in this two-dimensional space and it therefore commutes with the tau's:

$$[H_2, \boldsymbol{\tau}] = 0. \tag{12.39}$$

It then also follows that H_2 commutes with $\mathbf{V}$ or, equivalently,

$$\mathbf{V} H_2 \mathbf{V}^{-1} = H_2 \tag{12.40}$$

which is the statement that H_2 is invariant under the transformation (12.32). Now the tau's are Hermitian and, hence, correspond to possible observables. Equation (12.39) implies that their eigenvalues are constants of the motion (i.e. conserved quantities), associated with the invariance (12.40). But the tau's do not commute amongst themselves and so, according to the general principles of quantum mechanics, we cannot give definite values to more than one of them at a time. The problem of finding a classification of the states which makes the maximum use of (12.39), given the commutation relations (12.28), is easily solved by making use of the formal identity between the operators $\tau_i/2$ and angular momentum operators J_i (cf (12.29)). The answer is[3] that the total squared 'spin'

$$(\mathbf{T}^{(1/2)})^2 = (\tfrac{1}{2}\boldsymbol{\tau})^2 = \tfrac{1}{4}(\tau_1^2 + \tau_2^2 + \tau_3^2) = \tfrac{3}{4}\mathbf{1}_2 \tag{12.41}$$

and one component of spin, say $T_3^{(1/2)} = \frac{1}{2}\tau_3$, can be given definite values simultaneously. The corresponding eigenfunctions are just the χ_{p}'s and χ_{n}'s of (12.9), which satisfy

$$\tfrac{1}{4}\boldsymbol{\tau}^2 \chi_{\mathrm{p}} = \tfrac{3}{4}\chi_{\mathrm{p}} \qquad \tfrac{1}{2}\boldsymbol{\tau}_3 \chi_{\mathrm{p}} = \tfrac{1}{2}\chi_{\mathrm{p}} \tag{12.42}$$

$$\tfrac{1}{4}\boldsymbol{\tau}^2 \chi_{\mathrm{n}} = \tfrac{3}{4}\chi_{\mathrm{n}} \qquad \tfrac{1}{2}\boldsymbol{\tau}_3 \chi_{\mathrm{n}} = -\tfrac{1}{2}\chi_{\mathrm{n}}. \tag{12.43}$$

The reason for the 'spin' part of the name 'isospin' should by now be clear: the term is actually a shortened version of the historical one 'isotopic spin'.

In concluding this section we remark that, in this two-dimensional p–n space, the electromagnetic charge operator is represented by the matrix

$$\mathbf{Q}_{\mathrm{em}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}(\mathbf{1}_2 + \tau_3). \tag{12.44}$$

It is clear that although $\mathbf{Q}_{\mathrm{em}}$ commutes with τ_3, it does not commute with either τ_1 or τ_2. Thus, as we would expect, electromagnetic corrections to the strong interaction Hamiltonian will violate SU(2) symmetry.

[3] See, for example, Mandl (1992).

12.1.2 Larger (higher-dimensional) multiplets of SU(2) in nuclear physics

For the single nucleon states considered so far, the foregoing is really nothing more than the general quantum mechanics of a two-state system, phrased in 'spin-$\frac{1}{2}$' language. The real power of the isospin (SU(2)) symmetry concept becomes more apparent when we consider states of *several* nucleons. For A nucleons in the nucleus, we introduce three 'total isospin operators' $\mathbf{T} = (T_1, T_2, T_3)$ via

$$\mathbf{T} = \tfrac{1}{2}\boldsymbol{\tau}_{(1)} + \tfrac{1}{2}\boldsymbol{\tau}_{(2)} + \cdots + \tfrac{1}{2}\boldsymbol{\tau}_{(A)} \tag{12.45}$$

which are Hermitian. The Hamiltonian H describing the strong interactions of this system is presumed to be invariant under the transformation (12.40) for all the nucleons independently. It then follows that

$$[H, \mathbf{T}] = 0. \tag{12.46}$$

Thus, the eigenvalues of the $\mathbf{T}$ operators are constants of the motion. Further, since the isospin operators for different nucleons commute with each other (they are quite independent), the commutation relations (12.28) for each of the individual $\boldsymbol{\tau}$'s imply (see problem 12.4) that the components of $\mathbf{T}$ defined by (12.45) satisfy the commutation relations

$$[T_i, T_j] = \mathrm{i}\epsilon_{ijk} T_k \tag{12.47}$$

for $i, j, k = 1, 2, 3$, which are simply the standard angular momentum commutation relations, once more. Thus the energy levels of nuclei ought to be characterized—after allowance for electromagnetic effects, and correcting for the slight neutron–proton mass difference—by the eigenvalues of $\mathbf{T}^2$ and T_3, say, which can be simultaneously diagonalized along with H. These eigenvalues should then be, to a good approximation, 'good quantum numbers' for nuclei, if the assumed isospin invariance is true.

What are the possible eigenvalues? We know that the $\mathbf{T}$'s are Hermitian and satisfy exactly the same commutation relations (12.47) as the angular momentum operators. These conditions are all that are needed to show that the eigenvalues of $\mathbf{T}^2$ are of the form $T(T+1)$, where $T = 0, \frac{1}{2}, 1, \ldots$, and that for a given T the eigenvalues of T_3 are $-T, -T+1, \ldots, T-1, T$; that is, there are $2T+1$ *degenerate states* for a given T. These states all have the same A value, and since T_3 counts $+\frac{1}{2}$ for every proton and $-\frac{1}{2}$ for every neutron, it is clear that successive values of T_3 correspond physically to changing one neutron into a proton or *vice versa*. Thus we expect to see 'charge multiplets' of levels in neighbouring nuclear isobars. These are precisely the multiplets of which we have already introduced examples in chapter 1 (see figure 1.8) which we reproduce here as figure 12.2 for convenience. These level schemes (which have been adjusted for Coulomb energy differences, and for the neutron–proton mass difference) provide clear evidence of $T = \frac{1}{2}$ (doublet), $T = 1$ (triplet) and $T = \frac{3}{2}$ (quartet) multiplets. It is important to note that states in the same T-multiplet must have the same J^P quantum numbers