LBL-35371
UC-414

Lawrence Berkeley Laboratory

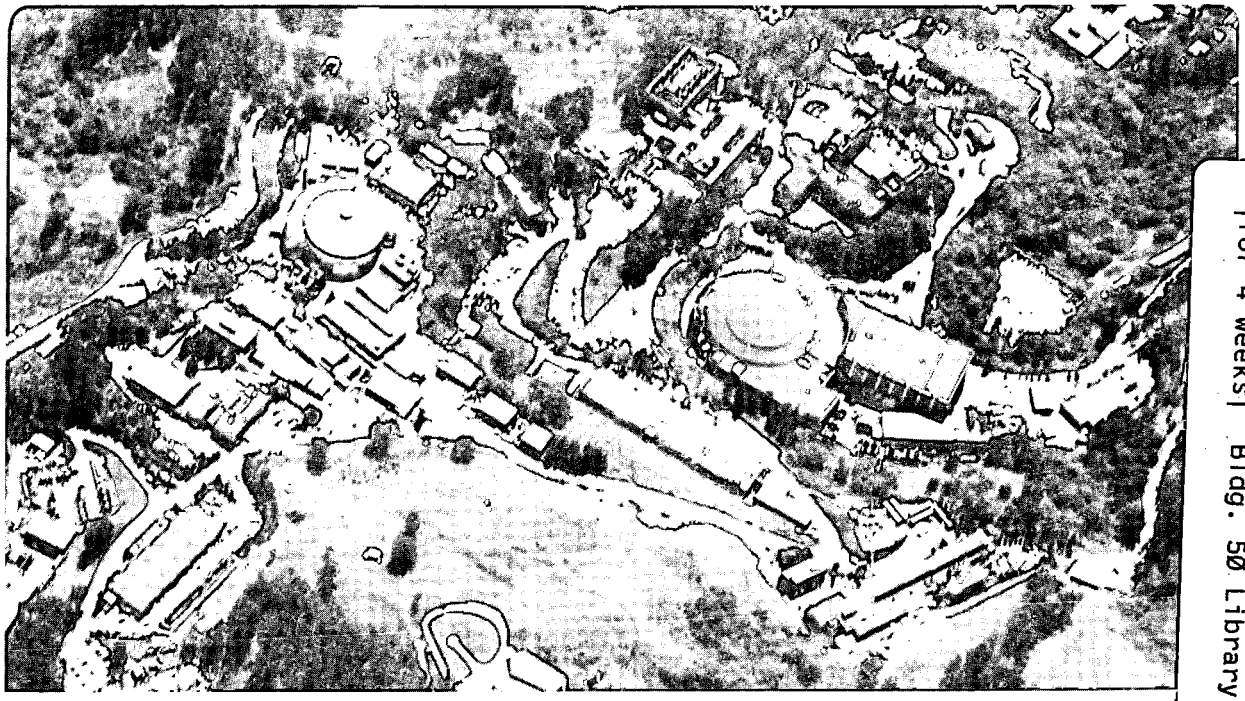
UNIVERSITY OF CALIFORNIA

Physics Division

Topics in the Structure of Hadronic Systems

R.F. Lebed
(Ph.D. Thesis)

April 1994



LOAN COPY
Circulates
for 4 weeks
Bldg. 50 Library.
LBL-35371
Copy 2

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF000

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBL-35371
UC-414

Topics in the Structure of Hadronic Systems

Richard Felix Lebed
Ph.D. Thesis

Physics Department
University of California

and

Physics Division
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

April 1994

This work was supported by the Director, Office of Energy Research, Office of High Energy and Nuclear Physics, Division of High Energy Physics of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Topics in the Structure of Hadronic Systems

Copyright © 1994

by

Richard Felix Lebed

The U.S. Department of Energy has the right to use this thesis for any purpose
whatsoever including the right to reproduce all or any part thereof

recycled paper

Abstract

Topics in the Structure of Hadronic Systems

by

Richard Felix Lebed

Doctor of Philosophy in Physics

University of California at Berkeley

Professor Mahiko Suzuki, Chair

In this dissertation we examine a variety of different problems in the physics of strongly-bound systems. Each is elucidated by a different standard method of analysis developed to probe the properties of such systems.

We begin with an examination of the properties and consequences of the current algebra of weak currents in the limit of heavy quark spin-flavor symmetry. In particular, we examine the assumptions in the proof of the Ademollo–Gatto theorem in general and for spin-flavor symmetry, and exhibit the constraints imposed upon matrix elements by this theorem.

Then we utilize the renormalization-group method to create composite fermions in a three-generation electroweak model. Such a model is found to reproduce the same low energy behavior as the top-condensate electroweak model, although in general it may have strong constraints upon its Higgs sector.

Next we uncover subtleties in the nonrelativistic quark model that drastically alter our picture of the physical origins of meson electromagnetic and hyperfine mass splittings; in particular, the explicit contributions due to $(m_d - m_u)$ and electrostatic potentials may be overwhelmed by other effects. Such novel effects are used to explain the anomalous pattern of mass splittings recently measured in bottom

mesons.

Finally, we consider the topic of baryon masses in heavy fermion chiral perturbation theory, including both tree-level and loop effects. We find that certain mass relations holding at second-order in symmetry breaking ($O(m_q^2)$ and $O(Q_q^2)$) have finite, computable, and numerically small loop corrections within the theory. The numerical values of these corrections are found to be in excellent agreement with experiment. We also find that, within chiral perturbation theory, the experimentally measured baryon masses alone are not enough to place stringent constraints upon the light quark masses.

Makoto Sugita 4/21/94

Committee Chair Date

To Allyson

Contents

List of Figures vi

List of Tables vii

Acknowledgements viii

List of Publications ix

Introduction 1

1 Current Algebra and $O(1/m)$ Corrections in Heavy Quark Spin-Flavor Symmetry 3

1.1 Introduction . . . 3
1.2 The Spin-Flavor Current Algebra . . . 6
1.3 The Algebra of Effective Weak Currents . . . 10
1.4 The Ademollo–Gatto Theorem . . . 13
1.5 The Ademollo–Gatto Theorem in Spin-Flavor Symmetry . . . 16
1.6 Scalar, Tensor, and Pseudoscalar Densities . . . 19
1.7 A Sample Application . . . 21

2 Composite Fermions in a Three-Generation Electroweak Model 23

2.1 Introduction . . . 23
2.2 Three Generations in the Top-Condensate Model . . . 26
2.3 The Composite-t_R Model . . . 29
2.4 Extending the Composite-t_R Model . . . 32
2.4.1 One Generation with One Higgs Doublet . . . 32
2.4.2 One Generation with Two Higgs Doublets . . . 34
2.4.3 More Than One Fermion Generation . . . 35
2.5 A Composite Left-Handed Fermion Model . . . 37
2.5.1 One Generation . . . 38
2.5.2 More Than One Fermion Generation . . . 39
2.6 Conclusions . . . 39

3 Meson Mass Splittings in Potential Models **41**
3.1 Introduction 41
3.2 Mass Computation in Field Theory 43
3.3 The Nonrelativistic Limit 50
3.4 Mass Splitting Formulas 53
3.5 Quantum-mechanical Theorems 55
3.6 Example: $V(r) = r/a^2 - \kappa/r$ 57
3.7 Numerical Results 60
3.8 Conclusions 65

4 Baryon Masses 1: Group Theory **68**
4.1 Introduction 68
4.2 The Structure of $SU(3)$ Breaking 69
4.3 Baryon Mass Relations 73
4.4 Quark Mass Parameters 77
4.5 Decuplet Mass Measurements 80

5 Baryon Masses 2: Chiral Dynamics **84**
5.1 Introduction 84
5.2 Heavy Baryon Theory and the Effective Lagrangian 86
5.3 Constructing the Lagrangian 88
5.3.1 Field Transformation Properties 88
5.3.2 Lagrangian Terms 91
5.4 Parameter Counting 95
5.5 Loop Corrections 99
5.6 Method of Calculation 102
5.7 Results and Predictions 105
5.7.1 Estimating Parameters 105
5.7.2 Decuplet Predictions: $\Delta_{1,2,3,4}$ 106
5.7.3 Octet Prediction: Δ_{CG} 108
5.7.4 Octet Prediction: Δ_{Σ} 108
5.8 Conclusions 109

Bibliography **111**

A Loop Corrections: Decuplet **118**

B Loop Corrections: Octet **122**

List of Figures

2.1 Fermion chain diagrams in the top-condensate model. 27
2.2 Chain diagrams in the composite-t_R model. 31

3.1 Diagrammatical representation of $\mathcal{M}_{fi}$. 45
3.2 Diagram for $\mathcal{M}_{fi}$ in the mesonic system. 46
3.3 Free quark Feynman amplitude $\mathcal{M}$. 47
3.4 Notation and conventions for the mesonic system. 48

5.1 "Keyhole" (quartic vertex) diagram contributing to baryon masses. 92
5.2 Trilinear vertex diagram contributing to baryon masses. 93

List of Tables

I Contributions to mass splittings of heavy mesons: Isospin pairs . . . 61
II Contributions to mass splittings of heavy mesons: $(1^-, 0^-)$ pairs . . 62
III Meson mass splittings compared to experiment 63

Acknowledgements

First and foremost, I wish to thank my advisor Mahiko Suzuki, without whose encouragement and wisdom this research would never have been possible. The gems of his insights from our discussions are to be found sprinkled liberally throughout the pages of this dissertation. Likewise, I wish to acknowledge the invaluable assistance of Dave Jackson, who taught me particle theory here at U. C. Berkeley, and who has helped me to answer more questions than I can count.

Were I to recognize individually every person in the LBL Theory Group who has aided me at one time or another during my graduate research, I would do as well to reproduce the whole group roster. But I would like to single out the following individuals: Lawrence Hall, who on more than one occasion pointed out to me some recent results which I subsequently incorporated into my research; Markus Luty, whose passion for discovery enhanced the quality of our work; and Paul Watts, whose incisive comments and painstaking exactitude helped to improve my understanding of numerous topics.

This work has also been enriched by my conversations or correspondences with Orlando Alvarez, Dick Arndt of Virginia Polytechnic Institute, John Donoghue of the University of Massachusetts, Amherst, Harris Kagan of Ohio State University, Joe Schechter of Syracuse University, and Charles Wohl of the LBL Particle Data Group.

Without the kind assistance of the Theory Group secretaries, Betty Moura and Luanne Neumann, I would forever have been lost in the maze of LBL bureaucracy. To them I convey my appreciation and gratitude.

I would also like to thank my earliest advocates, my parents Walter F. and Sally Lebed, for their encouragement which has sustained me from my earliest times to the present.

Finally, I would like to thank my beloved Allyson Ford, partly for her critical reading of this manuscript, but especially for her compassion and the infusion to me of her great emotional strength.

List of Publications

Publications in Particle Theory

1. "Meson Mass Splittings in the Nonrelativistic Model"
   Richard F. Lebed
   Physical Review D **47**, 1134 (1993).

2. "Making Electroweak Models of Composite Fermions Realistic"
   Richard F. Lebed and Mahiko Suzuki
   Physical Review D **45**, 1744 (1992).

3. "Current Algebra and the Ademollo–Gatto Theorem in
   Spin-Flavor Symmetry of Heavy Quarks"
   Richard F. Lebed and Mahiko Suzuki
   Physical Review D **44**, 829 (1991).

Preprints

4. "Baryon Decuplet Mass Relations in Chiral Perturbation Theory"
   Richard F. Lebed
   Lawrence Berkeley Laboratory preprint LBL-34704 and
   U.C. Berkeley preprint UCB-PTH-93/27.

5. "Baryon Masses Beyond Leading Order in Chiral Perturbation Theory"
   Richard F. Lebed and Markus A. Luty
   Lawrence Berkeley Laboratory preprint LBL-34779 and
   U.C. Berkeley preprint UCB-PTH-93/28.

Other Publications

6. "Charge State Distributions for Heavy Ions in Carbon Stripper Foils"
   M. A. McMahan, R. F. Lebed, B. Feinberg
   Presented at 1989 IEEE Particle Accelerator Conference,
   Chicago, IL, March 20–23, 1989
   IEEE Particle Accelerators 1989: pp. 536–538.

7. "High Temperature Radiator Materials for Applications in the Low Earth Orbital Environment"
   Sharon K. Rutledge, Bruce A. Banks, Michael J. Mirtich, Richard Lebed, Joyce Brady, Deborah Hotes and Michael Kussmaul
   Presented at the 1987 Spring Meeting of the Materials Research Society, Anaheim, CA, April 20–24, 1987
   NASA Technical Memorandum 100190.

Introduction

Every student of particle physics can attest to the brilliant successes of the Standard Model. We have seen the level of agreement between theory and experiment for the electron magnetic moment pushed to forty parts in a trillion; weak neutral currents and the weak gauge bosons themselves were predicted long before their discovery, based on the parameters of low-energy interactions; limits on the size of Standard Model loop corrections predicted a top quark with a mass less than 200 GeV, and it is believed that the announcement of evidence for the top is imminent, perhaps even before this dissertation is filed. At the time of this writing there is no serious discrepancy between the results of countless experiments and predictions of the theory using only eighteen input parameters in its most minimal form.

One path to further knowledge is to assume that nature is truly minimal, and that somehow these eighteen parameters must ultimately be related through some heretofore undiscovered symmetry. The practitioners of SUSY, GUTs, strings, and other theories clamber up the hierarchy of energy scales to propose new and beautiful symmetries, many of which could not be experimentally observed in our lifetimes, but in the spirit of pure science (or mathematics!) are well worth deeper investigation.

But there are still some serious gaps in our knowledge even at the lowest energy scales. What the Standard Model can predict, it can predict extremely well; when the assumption of perturbativity from free particle states is valid, perturbative calculations with the standard renormalization techniques lead to accurate and often astonishing predictions. But if systems are strongly bound, the standard methods

fail us, and we must resort to another approach. Exact field-theoretic methods for treating such systems, such as the Bethe–Salpeter equation, are cumbersome and have provided only limited results despite decades of effort. No one really knows how to build a proton with a mass of 1 GeV with nothing but three quarks of mass 5 MeV and their interactions.

In the Standard Model, we have managed to hide our ignorance by claiming that everything we do not understand about the strong interaction is somehow a consequence of QCD. We believe this gauge theory is correct because it explains some crucial phenomena, like the pattern of asymptotic freedom and the apparent threefold degeneracy of quarks, as well as providing some hints why we have only seen two kinds of hadron and why they appear in approximately degenerate multiplets. But our current understanding of QCD is not yet good enough to explain the *quantitative* details of the very interesting phenomena of confinement or chiral symmetry breaking.

But the particular case of QCD is not the end of the story. Even if some clever person uncovered a method of completely solving the theory of QCD tomorrow, we would probably still find the same situation to be repeated every time a new strong interaction were discovered at higher energy scales. The lesson is that techniques that lead to a predictive understanding of strongly-bound systems in general may have implications far beyond the problems to which they are currently being applied.

The purpose of my own doctoral research gradually developed into a study of approaches designed to obtain insight about these systems that do not willingly yield their secrets. The specific problems described in this dissertation are, in themselves, unrelated. But the methods—heavy quark spin-flavor symmetry, renormalization-group methods, potential models, chiral Lagrangians, and others—are connected by the philosophy described above. In the future I hope to add to this list of novel and insightful approaches.

Chapter 1

Current Algebra and $O(1/m)$ Corrections in Heavy Quark Spin-Flavor Symmetry

1.1 Introduction

The idea of spin-flavor symmetry of a flavor multiplet of N_f quarks, namely, enlarging the separate symmetries of spin $SU(2)$ and flavor $SU(N_f)$ to the symmetry $SU(2N_f)$ in order to increase the predictive power of the theory, is quite old. It was applied to light quarks [1] almost immediately after the development of the flavor $SU(3)$ model of strong interactions from which it originated. However, the various *no-go* theorems of the late 1960's, culminating in the Coleman–Mandula theorem [2], showed that the light-flavor $SU(6)$ cannot be an exact symmetry because of problems arising from mixing internal (flavor) symmetry with the Poincaré symmetry induced by the inclusion of spin angular momentum.

Nevertheless, spin-flavor symmetry made a comeback in the late 1980's when it was applied to heavy quarks [3]–[8]. The physical picture of this symmetry is actually quite simple: A hadron containing a heavy quark Q is viewed as a pointlike static source of color weakly coupled to a cloud of light degrees of freedom consisting of the light valence quarks, light sea quark-antiquark pairs, and gluons. "Weak" in

this sense means that the interactions of the cloud with itself and with the heavy quark have typical energies of $O(\Lambda_{QCD}) \ll m_Q$. In this limit, the properties of the heavy quark, namely its flavor and spin, are decoupled from the light degrees of freedom. Unlike the old $SU(6)$ symmetry, this spin-flavor symmetry has a natural Lorentz-covariant formulation within QCD for all flavors of quark that may be considered infinitely heavy. In physical terms, decay of one heavy quark flavor to another occurs without changing the heavy quark's velocity or the makeup of the light cloud, and moreover, the process is independent of the initial spin state of the heavy quark.

Now let us consider what this means in the real world. The only quarks that may be reliably considered heavy compared to Λ_{QCD} are c, b, and t. Since the experimental lower bound on the top quark mass is currently $m_t \geq 131$ GeV $> M_W$ [9], the top quark should decay into a *real* W in a time $O(1/\alpha_2 m_t)$, and so we would never see a top hadron. Thus the usefulness of heavy quark spin-flavor symmetry (HQS) is limited to c and b hadrons. Extraction of information is straightforward in semileptonic weak decays of heavy hadrons [10]–[15]. We may use HQS to relate the properties of, for example, D and B mesons, or D with D^* and B with B^*, as in this work. From this, one can gain insight into the elusive Standard Model weak mixing angles V_{ub} and V_{cb} [12]. One can also use HQS to relate charm and bottom baryons [15].

Of course, m_c and m_b are not infinite, and therefore HQS is not an exact symmetry in the real world. However, as with any approximate symmetry, we approach the problem by expanding physical quantities in a symmetry-breaking parameter that must be small for the symmetry to make sense. Since the requirement of HQS is $\Lambda_{QCD}/m_{c,b} \ll 1$, these ratios are exactly the expansion parameters we need. Let us henceforth describe them with the shorthand $1/m_{c,b}$.

Unfortunately, HQS as described above cannot work if the heavy quark interacts with hard gluons (*i.e.*, with energies $O(m_Q)$ or larger), because then the coupling between the heavy quark and the cloud is no longer weak, and the description of the heavy quark as having an approximately conserved velocity is invalid. Fortunately, there is a standard method to eliminate this problem: One simply

"integrates out" via the renormalization group the hard gluon degrees of freedom to the mass of the heaviest quark (here m_b), and then successively integrates out the gluons to the next heaviest quark, until one reaches the lightest of the heavy quarks. This method is valid because QCD is perturbative at these energy scales. In general, this procedure induces new expansion parameters of the ratios of heavy quark masses. Thus, in the real world, implementation of HQS requires an expansion not only in $1/m_{c,b}$, but m_c/m_b as well. This program of computing Wilson coefficients and matching was carried out in Refs. [16]–[21].

Once we have computed the quantities of interest in an operator product expansion, we may use HQS to relate matrix elements of these operators between various heavy hadron states. Essentially, HQS provides us with a Wigner–Eckart theorem: There is a geometric (Clebsch–Gordan) coefficient due to the spin-flavor symmetry, and a model-dependent (reduced matrix element) portion, which cannot be evaluated with further assumptions. However, this does not imply that nothing can be said about the latter; there is a renormalization-free theorem (*i.e.*, a theorem that forbids the presence of some subleading terms) due to Luke [18] constraining matrix elements of certain weak currents. It was subsequently shown by Boyd and Brahm [22] that this theorem is a consequence of an extension of the Ademollo–Gatto (AG) theorem [23]. This Chapter, based on Ref. [24], focuses upon issues surrounding the proof of the theorem and its consequences.

The first emphasis involves the generality of the statement that the theorem is "renormalization-free." In order to verify this claim, we must first carry out the renormalization-group evolution of the weak currents. The current algebra itself is defined by a set of canonical commutation relations at equal times established by quantization of the fields. Given the Lagrangian, one must first compute the Noether currents, quantize, evolve the currents down to the lower energy scale, and then examine precisely how much the commutation relations have changed. What we find is that, apart from an overall factor due to short-distance QCD corrections, the Noether currents and the effective currents differ only at $O(1/m^2)$, not $O(1/m)$. This fact turns out to be crucial to the validity of the AG theorem in HQS, and we show that this result is most easily seen in the $\mathbf{v} = 0$ heavy quark rest frame.

Moreover, we find that canonical quantization in the $\mathbf{v} = 0$ frame allows us to circumvent quantization difficulties present in a general frame.

Once we have established the algebra of effective currents, we show from the AG theorem that the matrix elements of certain weak effective currents deviate from their symmetry limits only at $O(1/m^2)$. The usefulness of the $\mathbf{v} = 0$ frame is again demonstrated: In an arbitrary frame, we would find states of different spin mixed due to the contribution of orbital angular momentum allowed by $\mathbf{v} \neq 0$. It turns out that, in this case, the AG theorem would not prevent $O(1/m)$ corrections to current matrix elements; in the $\mathbf{v} = 0$ frame, however, the $O(3)$ symmetry of space forbids such mixing and permits us to ignore this potential problem. As a result, the physical consequences are most transparent in this frame.

In Sec. 1.2 we study the current algebra of the spin-flavor group, with our focus upon the cancellations of $O(1/m)$ terms that occur in the $\mathbf{v} = 0$ frame. In Sec. 1.3 we examine the algebra of the effective currents after integrating out hard gluons, and see that this algebra satisfies the $SU(4)$ algebra (for b and c) up to $O(1/m_c^2)$. The AG theorem modified for HQS is presented in Sec. 1.4. Luke's results are rederived using the AG theorem in Sec. 1.5, and additional consequences of the AG theorem in HQS for the scalar, pseudoscalar, and tensor densities are presented in Sec. 1.6. Finally, in Sec. 1.7, we demonstrate some of the consequences of the AG theorem in a particular nonrelativistic quark model.

1.2 The Spin-Flavor Current Algebra

A particularly convenient method for exhibiting spin symmetry in the infinite-mass limit is provided by Georgi [8]. The usual Dirac spinor $\psi(x)$ in the Lagrangian

$$L = \overline{\psi}\,(i\,\partial\!\!\!/ - m)\,\psi \tag{1.1}$$

is transformed into the spinor

$$h_v(x) \equiv \exp\left[im\not{v}\,(v_\mu x^\mu)\right]\psi(x), \tag{1.2}$$

with the *massless* Lagrangian

$$L_v = \overline{h}_v \, i \, \partial\!\!\!/ h_v, \tag{1.3}$$

where $v^\mu \equiv p^\mu/m$ is the four-velocity of the fermion. In the rest frame, $\frac{1}{2}(1 \pm \gamma^0)$ projects out the upper (lower) two components of h_v; thus, in the limit of infinite mass, these project out the quark (antiquark) spinors in the rest frame. We generalize to an arbitrary frame via

$$P^\pm \equiv \frac{(1 \pm v\!\!\!/)}{2} \Longrightarrow h_v^\pm \equiv P^\pm h_v. \tag{1.4}$$

Then the Lagrangian Eq. 1.3 may be expressed in terms of $h_v^\pm$. In the infinite-mass limit, terms creating fermion-antifermion pairs are suppressed, leading to

$$L_v = \overline{h}_v \, i \, v\!\!\!/ \, (v_\mu \partial^\mu) \, h_v. \tag{1.5}$$

We now render the derivative covariant by $D^\mu \equiv \partial^\mu + igA^\mu$, where A^μ is the gluon field. When the mass is instead large but finite, the expansion reads [7],[17]–[19]

$$\begin{aligned} L_v &= \overline{h}_v \, i \, v\!\!\!/ \, (vD) \, h_v \\ &\quad + \frac{1}{2m} \overline{h}_v \, (vD)^2 \, h_v - \frac{1}{2m} \overline{h}_v D^2 h_v - \frac{g}{4m} \overline{h}_v \sigma_{\mu\nu} G^{\mu\nu} h_v + O\left(\frac{1}{m^2}\right), \end{aligned} \tag{1.6}$$

where $G^{\mu\nu}$ is the gluon field strength. The second term is formally $O(1/m)$, but because the equation of motion from this Lagrangian is $(vD)h_v = O(1/m)$, we see that it can only produce matrix elements that are $O(1/m^2)$. The rest of the expansion follows from the operator identity

$$D\!\!\!/ D\!\!\!/ = D^2 + \frac{g}{2} \sigma_{\mu\nu} G^{\mu\nu}. \tag{1.7}$$

If we now expand to include N_f heavy quark flavors with their antiquarks, the Lagrangian Eq. 1.5 with a covariant derivative possesses $SU(2N_f)_+ \times SU(2N_f)_-$ symmetry for each v_μ. In the physically relevant case, we embed the b and c quarks in an $SU(4)$ symmetry, using the spinor notation

$$h_v = \begin{pmatrix} c_v \\ b_v \end{pmatrix}, \qquad M = \begin{pmatrix} m_c & \\ & m_b \end{pmatrix}, \tag{1.8}$$

and the Lagrangian

$$\begin{aligned} L_v &= \overline{h}_v\, i\, \not{v}(v\vec{D})h_v \\ &\quad -\frac{1}{2}\overline{h}_v M^{-1}(v\overleftarrow{D})(v\vec{D})h_v + \frac{1}{2}\overline{h}_v M^{-1}\overleftarrow{\not{D}}\vec{\not{D}}h_v + O(1/m^2), \end{aligned} \tag{1.9}$$

where the left-handed covariant derivative, defined by $\overline{h}_v\overleftarrow{D}^\mu \equiv \partial^\mu\overline{h}_v - ig\overline{h}_v A^\mu$, is introduced for ease of obtaining currents and equations of motion. Also, the m in $O(1/m^n)$ here and henceforth refers to the *lightest* heavy quark mass. We introduce the fifteen $SU(4)$ generators [8]

$$\Gamma_\alpha = \tau_a^\pm/2, \quad S_i^\pm/2, \quad S_i^\pm\tau_a^\pm/2 \quad (a = 1,2,3;\ i = 1,2,3), \tag{1.10}$$

where

$$\tau_a^\pm \equiv \tau_a P^\pm, \qquad S_i^\pm \equiv \frac{i}{2}\epsilon_{ijk}\not{e}_j\not{e}_k P^\pm, \tag{1.11}$$

and e_i^μ form an orthonormal set with v^μ:

$$e_{i\mu}e_j{}^\mu = -\delta_{jk}, \qquad v_\mu e_j{}^\mu = 0. \tag{1.12}$$

Then the $\Gamma_\alpha^\pm$ satisfy the $SU(4)$ algebra

$$\begin{aligned} [\tau_a^\pm/2, \tau_b^\pm/2] &= i\epsilon_{abc}\tau_c^\pm/2, \\ [\tau_a^\pm/2, S_i^\pm/2] &= 0, \\ [\tau_a^\pm/2, S_i^\pm\tau_b^\pm/2] &= i\epsilon_{abc}S_i^\pm\tau_c^\pm/2, \\ [S_i^\pm/2, S_j^\pm/2] &= i\epsilon_{ijk}S_k^\pm/2, \\ [S_i^\pm/2, S_j^\pm\tau_a^\pm/2] &= i\epsilon_{ijk}S_k^\pm\tau_a^\pm/2, \\ [S_i^\pm\tau_a^\pm/2, S_j^\pm\tau_b^\pm/2] &= i\delta_{ab}\,\epsilon_{ijk}S_k^\pm/2 + i\delta_{ij}\,\epsilon_{abc}\tau_c^\pm/2. \end{aligned} \tag{1.13}$$

The $SU(4)$ symmetry induces

$$\delta_\alpha h_v^\pm = i\Gamma_\alpha^\pm h_v^\pm, \tag{1.14}$$

and using these in the Lagrangian Eq. 1.9, we obtain the Noether currents

$$\begin{aligned} (J_\alpha^\pm)^\mu &= \overline{h}_v\, \not{v}v^\mu\Gamma_\alpha^\pm h_v \\ &\quad -\frac{i}{2}\overline{h}_v\left[\Gamma_\alpha^\pm M^{-1}v^\mu(v\vec{D}) - (v\overleftarrow{D})v^\mu M^{-1}\Gamma_\alpha^\pm\right]h_v \\ &\quad +\frac{i}{2}\overline{h}_v\left[\Gamma_\alpha^\pm M^{-1}\gamma^\mu\vec{\not{D}} - \overleftarrow{\not{D}}\gamma^\mu M^{-1}\Gamma_\alpha^\pm\right]h_v + O(1/m^2). \end{aligned} \tag{1.15}$$

Note especially the asymmetry of derivatives of h_v and $\overline{h}_v$ in the first term of Eq. 1.9. The natural impulse is to symmetrize this term as $\frac{i}{2}\overline{h}_v \not{v}(v\overleftrightarrow{D})h_v$ and to take both h_v and $h_v^\dagger$ as coordinate fields in the canonical quantization. However, we show that this leads to an inconsistency. For then the fields h_v and $\overline{h}_v$ have the conjugate momenta

$$\begin{aligned}
\pi_v &\equiv \partial L/\partial(\partial^0 h_v) \\
&= \tfrac{1}{2}\overline{h}_v\left[i\not{v}v_0 - M^{-1}(v\overleftarrow{D})v_0 + M^{-1}\overleftarrow{\not{D}}\gamma_0\right] + O(1/m^2), \\
\pi_v^\dagger &\equiv \partial L/\partial(\partial^0 h_v^\dagger) \\
&= \tfrac{1}{2}\left[-i\gamma_0\not{v}v_0 - \gamma_0 M^{-1}v_0(v\overleftarrow{D}) + M^{-1}\overrightarrow{\not{D}}\right]h_v + O(1/m^2),
\end{aligned} \tag{1.16}$$

whereas the charge densities are

$$(J_\alpha^\pm)^0 = -i\pi_v(\Gamma_\alpha^\pm)h_v + ih_v^\dagger(\Gamma_\alpha^\pm)^\dagger\pi_v^\dagger + O(1/m^2). \tag{1.17}$$

Imposing the usual canonical anticommutation relations

$$\left\{\pi_v(\mathbf{x},t)_i\,, h_v(\mathbf{y},t)_j\right\} = i\delta_{ij}\delta(\mathbf{x}-\mathbf{y})\,,\ \text{etc.} \tag{1.18}$$

leads to the $SU(4)$ charge algebra

$$\left[Q_\alpha^\pm, Q_\beta^\pm\right] = if_{\alpha\beta\gamma}Q_\gamma^\pm, \tag{1.19}$$

with structure constants $f_{\alpha\beta\gamma}$ and charges $Q_\alpha^\pm \equiv \int d^3x(J_\alpha^\pm)^0$. Now if we take the limit $m \to \infty$, the coordinate–momentum anticommutation relation in the $\mathbf{v} = 0$ frame collapses to

$$\left\{h_v^\dagger(\mathbf{x},t)_i\,, h_v(\mathbf{y},t)_j\right\} = 2\delta_{ij}\delta(\mathbf{x}-\mathbf{y})\,. \tag{1.20}$$

The offending factor of 2 is a well-known inconsistency in canonical quantization due to a haphazard handling of time derivatives of fields; however, it is not merely an error in our initial approach of symmetrizing the free kinetic term and treating the field $h_v^\dagger$ as an independent coordinate field, for factors of $\partial^0 h_v^\dagger$ already appear in the Lagrangian. There exist methods [25, 26] of treating such problems in a consistent fashion, but we offer a much simpler prescription: Do not symmetrize, and work in the frame $\mathbf{v} = 0$.

This frame has a very important cancellation in the Lagrangian Eq. 1.6 (generalized to N_f flavors), namely that the D_0^2 terms cancel between the second and third terms, leaving us with

$$L_{\mathbf{v}=0} = h^\dagger i D_0 h - \frac{1}{2}(\overline{D_i h})M^{-1}(D_i h) - \frac{g}{4}\bar{h}M^{-1}\sigma_{\mu\nu}G^{\mu\nu}h + O(1/m^2), \tag{1.21}$$

where we suppress the $\mathbf{v} = 0$ subscript on the spinor h. We no longer have the difficulty of higher time derivatives, and $h^\dagger$ is no longer an independent coordinate field. The momentum conjugate to h is

$$\pi = \partial L_{\mathbf{v}=0}/\partial(\partial^0 h) = ih^\dagger + O(1/m^2), \tag{1.22}$$

and the Noether charge densities are

$$(J_\alpha^\pm)^0 = h^\dagger \Gamma_\alpha^\pm h + O(1/m^2). \tag{1.23}$$

Now we have no difficulties in defining a charge algebra with consistent canonical commutation relations.

Notice that the Noether charges assume the same form, up to $O(1/m^2)$, as they would had we worked explicitly in the $m \to \infty$ limit; the cancellation of the $O(1/m)$ time derivatives in the frame $\mathbf{v} = 0$ makes this possible. Therefore, in physical applications, one may use the symmetric charges instead of the physical charges in the $\mathbf{v} = 0$ frame without fear of inducing $O(1/m)$ errors.

1.3 The Algebra of Effective Weak Currents

The next step in our program is to evolve the weak currents associated with the $b \to c$ transition via the renormalization group. For now, let us neglect $O(m_c/m_b)$ terms in the evolution. Upon integrating out the gluons down to $-q^2 = m_c^2$, the $V - A$ current $\bar{c}\gamma^\mu(1-\gamma_5)b$ becomes [17, 18]

$$\begin{aligned} J_{cb}^\mu &= [\alpha_s(m_b)/\alpha_s(m_c)]^{a_I}\, \bar{c}_{v'}\left[\gamma^\mu(1-\gamma_5) - \frac{i}{2m_c}\overleftarrow{\not{D}}\gamma^\mu(1-\gamma_5)\right] b_v \\ &\quad + O(1/m_c^2), \end{aligned} \tag{1.24}$$

where $a_I = -\frac{6}{25}$. The spinors $c_{v'}$ and b_v now satisfy the equations of motion derived from the Lagrangian Eq. 1.6. Note also that, for the first time, we consider $v' \neq v$, namely a breaking of the exact recoil-free limit for the heavy quarks. We must now integrate out the gluons in the momentum range $\mu^2 < -q^2 < m_c^2$ in order to evaluate matrix elements at the low momentum scale μ, where we believe that the residual gluon interactions represented by the spin-flip interaction $(g/4m)\overline{h}_v\sigma_{\mu\nu}G^{\mu\nu}h_v$ are a small symmetry-breaking perturbation. Then the matrix elements of the resulting Lagrangian,

$$\begin{aligned} L_v &= L_{m=\infty} \\ &\quad + \frac{1}{2m_c}(\overline{D_\mu h_v})(D^\mu h_v) - \frac{g}{4m_c}[\alpha_s(\mu)/\alpha_s(m_c)]^{-\frac{9}{25}}\,\overline{h}_v\sigma_{\mu\nu}G^{\mu\nu}h_v, \end{aligned} \tag{1.25}$$

may be evaluated in the approximate $SU(4)$ symmetry. The $(vD)^2$ term has been suppressed, as mentioned above, because it gives rise to $O(1/m^2)$ matrix elements. The weak current of Eq. 1.24 is modified by this further evolution to [17]–[19]

$$\begin{aligned} J^\mu_{cb} &= [\alpha_s(m_b)/\alpha_s(m_c)]^{a_I}\left\{[\alpha_s(m_c)/\alpha_s(\mu)]^{-\frac{8}{27}a_0}\,\overline{c}_{v'}\gamma^\mu(1-\gamma_5)b_v \right. \\ &\quad \left. -[\alpha_s(m_c)/\alpha_s(\mu)]^{-\frac{8}{27}a_1}\left(\frac{i}{2m_c}\right)\overline{c}_{v'}\overleftarrow{\not{D}}\gamma^\mu(1-\gamma_5)b_v\right\}, \end{aligned} \tag{1.26}$$

where now a_0 and a_1 are functions of vv' that vanish in the symmetry limit $vv' = 1$. It is important to realize that the spinors in this expression satisfy equations of motion derived from the *renormalized* effective Lagrangian Eq. 1.25; with this caveat, we may use the currents (1.26) to evaluate matrix elements using the approximate $SU(4)$ spin-flavor symmetry.

The kinematic point of interest, of course, is the symmetry point $\mathbf{v} = \mathbf{v}' = 0$, or $vv' = 1$, for which the heavy quark has zero recoil, and the two spinors $\overline{c}$ and b are labeled by the same four-velocity. Thus their bilinears give rise to $SU(4)$ generators. We then define the effective current

$$\overline{J}^\mu_{cb} \equiv [\alpha_s(m_b)/\alpha_s(m_c)]^{-a_I}J^\mu_{cb}, \tag{1.27}$$

and compute its matrix elements in the $vv' = 1$ limit. The second term of Eq. 1.26 is quickly seen to be $O(1/m^2)$ for the time part of the vector current and the space

part of the axial current, because a_0 and a_1 vanish in this limit, and also

$$(1/m)\overline{h}\,\overleftarrow{D}_0 = O(1/m^2) \tag{1.28}$$

from the $\mathbf{v} = 0$ equation of motion, and

$$\begin{aligned} (1/m)\overline{h}\gamma_i\gamma_0 h &= O(1/m^2), \\ (1/m)\overline{h}\gamma_i\gamma_j\gamma_5 h &= O(1/m^2) \quad (i,j = 1,2,3), \end{aligned} \tag{1.29}$$

which are results of the standard nonrelativistic Pauli expansion of bilinears. The surviving components of the effective $V - A$ current $\overline{J}$ in the symmetry limit, up to $O(1/m^2)$, are therefore

$$\begin{aligned} V_0 &= \overline{c}\,\gamma_0 b &= c^\dagger b, \\ A_i &= \overline{c}\,\gamma_i\gamma_5 b &= \epsilon_{ijk} c^\dagger(\sigma_{jk}/2)b. \end{aligned} \tag{1.30}$$

Thus, up to $O(1/m^2)$, V_0 and A_i are just the Noether charge densities of $\Gamma_\alpha = (\tau_1 + i\tau_2)/2$ and $S_i(\tau_1 + i\tau_2)/2$. Defining the set of charges by

$$\begin{aligned} Q^a &\equiv \int d^3x\, h^\dagger(\tau_a/2)h, \\ Q_{5i} &\equiv \int d^3x\, \epsilon_{ijk} h^\dagger(\sigma_{jk}/4)h, \\ Q^a_{5i} &\equiv \int d^3x\, \epsilon_{ijk} h^\dagger(\sigma_{jk}\tau_a/4)h, \end{aligned} \tag{1.31}$$

and using the set of canonical commutation relations including $\{h(\mathbf{x},t)_i\,, h^\dagger(\mathbf{y},t)_j\}$ $= \delta_{ij}\delta(\mathbf{x}-\mathbf{y})$, we generate the $SU(4)$ charge algebra, which holds to $O(1/m^2)$:

$$\begin{aligned} [Q^a, Q^b] &= i\epsilon^{abc} Q^c, \\ [Q^a, Q_{5i}] &= 0, \\ [Q^a, Q^b_{5i}] &= i\epsilon^{abc} Q^c_{5i}, \\ [Q_{5i}, Q_{5j}] &= i\epsilon_{ijk} Q_{5k}, \\ [Q_{5i}, Q^a_{5j}] &= i\epsilon_{ijk} Q^a_{5k}, \\ [Q^a_{5i}, Q^b_{5j}] &= i\epsilon^{abc}\,\delta_{ij} Q^c + i\epsilon_{ijk}\,\delta^{ab} Q_{5k}. \end{aligned} \tag{1.32}$$

Some components of the weak currents are thus related to generators of the $SU(4)$ algebra up to $O(1/m^2)$; therefore, the form of the AG theorem introduced in the next two Sections applies to matrix elements of these currents.

1.4 The Ademollo–Gatto Theorem

Originally, the AG theorem [23] was used to obtain model-independent information about matrix elements in the flavor-$SU(3)$ symmetry of light quarks, for example, to provide an accurate determination of the Cabibbo angle. The initial statement of the theorem is just this: *Matrix elements of a charge operator in a broken symmetry deviate from their symmetry values only at second order in the symmetry breaking.* This statement actually hides some assumptions, which we elucidate in this Section; then the extension of the theorem to HQS is simple.

The generators of the $SU(3)$ algebra in the symmetric limit (denoted by $SU(3)_0$) are given by the Noether charges $Q_a^0 \equiv \int d^3x\, q_0^\dagger(\lambda_a/2)q_0$. The algebra itself is defined by the set of structure constants f_{abc} and the commutation relations $[Q_a^0, Q_b^0] = if_{abc}Q_c^0$. But of course, the symmetry is broken in the real world. For the moment, let us assume that the u and d quarks are degenerate in mass, and the dominant breaking is given by $L_{brk} = -\Delta m\bar{q}_0(\lambda_8/2)q_0$. Does this affect the algebra? No, because the symmetry breaking contains no time derivatives of the quark fields, which means that the expressions for the zero-components of the Noether currents are unaltered. However, the charges themselves are different, because the quark fields now satisfy the equation of motion from $L_0 + L_{brk}$; they are denoted by q instead of q_0. To sum up, the broken charges $Q_a \equiv \int d^3x\, q^\dagger(\lambda_a/2)q$ satisfy an $SU(3)$ algebra formally equivalent to the symmetric algebra $SU(3)_0$ defined above. It is $SU(3)$, not $SU(3)_0$, which was used in the original proof of the theorem. If we attempted to work in $SU(3)_0$ with L_{brk} as an interaction-picture perturbation, the states used in computing physical matrix elements would be unphysical symmetric limits, and the corresponding operators would explicitly contain the first-order symmetry-breaking parameter Δm. Then the proof of the AG theorem would be unnecessarily complicated. It is far better to work in the Heisenberg picture, where we use the asymptotic states of the Lagrangian including L_{brk}, and the charges of $SU(3)$.

We therefore consider the physical matrix element

$$\langle\tilde{\beta}(\mathbf{p})|Q_a|\tilde{\alpha}(\mathbf{p})\rangle, \tag{1.33}$$

where the states $|\tilde{\alpha}, \tilde{\beta}(\mathbf{p})\rangle$ are physical states not only in being described by $SU(3)$ instead of $SU(3)_0$, but also in consisting of a superposition of group-theoretically pure states of $SU(3)$. Of course, $SU(3)$ is useful to us because the physical states approximately equal eigenstates of the symmetry. We denote these as $|\alpha, \beta(\mathbf{p})\rangle$, and other states induced by the mixing by $|n(\mathbf{p})\rangle$. We then write

$$\begin{aligned} |\tilde{\alpha}(\mathbf{p})\rangle &= c_{\alpha\alpha}|\alpha(\mathbf{p})\rangle + \textstyle\sum_n c_{\alpha n}|n(\mathbf{p})\rangle, \\ |\tilde{\beta}(\mathbf{p})\rangle &= c_{\beta\beta}|\beta(\mathbf{p})\rangle + \textstyle\sum_n c_{\beta n}|n(\mathbf{p})\rangle, \end{aligned} \tag{1.34}$$

with the normalization conditions

$$\langle n(\mathbf{p})|\alpha(\mathbf{p})\rangle = \langle n(\mathbf{p})|\beta(\mathbf{p})\rangle = 0, \tag{1.35}$$

$$\langle n_1(\mathbf{p})|n_2(\mathbf{p})\rangle = \delta_{n_1 n_2}. \tag{1.36}$$

Let us denote the dimensionless symmetry-breaking parameter by ε; in the case at hand, it is proportional to Δm. In general, the coefficients $c_{(\alpha,\beta)n}$ vanish with ε. We thus write

$$c_{\alpha n} = O(\varepsilon) \tag{1.37}$$

$$c_{\alpha\alpha} = \left[1 - \sum_n |c_{\alpha n}|^2\right]^{\frac{1}{2}} = 1 + O(\varepsilon^2), \tag{1.38}$$

and similarly for β, where the latter relation is merely the normalization condition on the physical state $|\tilde{\alpha}(\mathbf{p})\rangle$. This trivial fact actually turns out to be the essential content of the AG theorem. It should be pointed out that there are ways to break symmetries (*e.g.*, spontaneous chiral symmetry breaking in effective Lagrangians) such that vanishing matrix elements do not approach zero analytically (typically with logarithmic singularities); in such cases, what we call $O(\varepsilon^n)$ is actually $O(\varepsilon^n, \varepsilon^n \ln \varepsilon)$.

Suppose for the moment that the states $|n(\mathbf{p})\rangle$ are all in irreducible representations of $SU(3)$ different from that of $|\alpha, \beta(\mathbf{p})\rangle$. For example, it is easy to check that λ_8 symmetry breaking, as in our example, does not connect any two distinct members of the pseudoscalar meson octet. Then the charges Q_a, being generators

of the group, connect $|\alpha(\mathbf{p})\rangle$ and $|\beta(\mathbf{p})\rangle$ to each other, but not to $|n(\mathbf{p})\rangle$. The $SU(3)$ structure constants alone determine the matrix element

$$(U^a)_{\beta\alpha} \equiv \langle\beta(\mathbf{p})|Q_a|\alpha(\mathbf{p})\rangle, \tag{1.39}$$

and analogously for $(U^a)_{mn}$, so that the physical matrix element is given by

$$\begin{aligned}\langle\tilde{\beta}(\mathbf{p})|Q_a|\tilde{\alpha}(\mathbf{p})\rangle &= c^*_{\beta\beta}c_{\alpha\alpha}(U^a)_{\beta\alpha} + \sum_{m,n} c^*_{\beta m}c_{\alpha n}(U^a)_{mn} \\ &= c^*_{\beta\beta}c_{\alpha\alpha}(U^a)_{\beta\alpha} + O(\varepsilon^2) \\ &= (U^a)_{\beta\alpha} + O(\varepsilon^2),\end{aligned} \tag{1.40}$$

where we have used Eqs. 1.37 and 1.38. Thus, in this case, the AG theorem is proved as stated.

Now suppose that one of the $|n(\mathbf{p})\rangle$ in the expansion of $|\tilde{\beta}(\mathbf{p})\rangle$ is in the same irreducible representation as $|\beta(\mathbf{p})\rangle$. In the example of flavor $SU(3)$, this is provided by explicit isospin symmetry breaking $L_{iso} = (m_u - m_d)(\overline{u}u - \overline{d}d)/2 = (m_u - m_d)\overline{q}\,(\lambda_3/2)q$. Such a breaking permits the mixing of the π^0 and the η eigenstates of $SU(3)$. In this case, the physical π^0 is expanded as

$$|\tilde{\pi}^0(\mathbf{p})\rangle = c_{\pi^0\pi^0}|\pi^0(\mathbf{p})\rangle + c_{\pi^0\eta}|\eta(\mathbf{p})\rangle + \sum_n c_{\pi^0 n}|n(\mathbf{p})\rangle, \tag{1.41}$$

where the states $|n(\mathbf{p})\rangle$ are not members of the pseudoscalar octet. The key here is that although $c_{\pi^0\pi^0}$ is again $1 + O(\varepsilon^2)$, we must contend with the additional coefficient $c_{\pi^0\eta} = O(\varepsilon)$. Now consider the matrix element of the weak current transition from K^+ to π^0. The K^+ does not suffer from a similar mixing when we expand the physical state $|\tilde{K}^+(\mathbf{p})\rangle$, and we find

$$\begin{aligned}\langle\tilde{\pi}^0(\mathbf{p})|Q_4 - iQ_5|\tilde{K}^+(\mathbf{p})\rangle &= -\sqrt{1/2}c^*_{\pi^0\pi^0}c_{K^+K^+} - \sqrt{3/2}c^*_{\pi^0\eta}c_{K^+K^+} \\ &\quad + \sum_{m,n} c^*_{\pi^0 m}c_{K^+n}(U^{4-i5})_{mn} \\ &= -\sqrt{1/2} + O(\varepsilon),\end{aligned} \tag{1.42}$$

where the $O(\varepsilon)$ term originates from $-\sqrt{3/2}\,c^*_{\pi^0\eta}c_{K^+K^+}$. Correspondingly, we must qualify the statement of the AG theorem: *For any two states in an irreducible representation of some symmetry group, and symmetry breaking that neither contains*

time derivatives of fields nor mixes these states with other states in the same irreducible representation, the matrix elements of a charge operator with these states deviate from their symmetry values only at second order in the symmetry breaking (up to logarithms in the symmetry-breaking parameter).

We may worry about this mixing in HQS; in fact, it does occur for frames with arbitrary v^μ, but we see in the next Section that working in the frame $\mathbf{v} = 0$ eliminates these mixings, and the AG theorem holds in its simpler form.

1.5 The Ademollo–Gatto Theorem in Spin-Flavor Symmetry

In the previous Section the AG theorem was proved in a general context, although we used flavor-$SU(3)$ symmetry broken by the inequality of quark masses as a concrete example. Now the symmetry is the spin-flavor symmetry $SU(4)_+$ of b and c quarks (or $SU(4)_-$ for their antiparticles), and the symmetry-breaking parameter is $(1/m)$.

In Section 1.2 we observed that we can eliminate $O(1/m)$ terms containing time derivatives of fields by working in the $\mathbf{v} = 0$ frame, which permits us to use the symmetry charges in place of the physical charges up to $O(1/m^2)$. Obviously this is important for the usefulness of the AG theorem in HQS. In flavor-$SU(3)$, choosing a frame was never an issue because the symmetry group commutes with the Lorentz group, and thus all frames are trivially equivalent. HQS, however, is rendered Lorentz-covariant, and therefore the use of the $\mathbf{v} = 0$ frame has substantial implications.

As $\mathbf{q} \equiv m(\mathbf{v}-\mathbf{v}') \to 0$, the chief symmetry-breaking term is the spin-flip interaction of gluon absorption and emission, which assumes the limiting form

$$\delta L = -\frac{ig}{2m}\chi^\dagger \boldsymbol{\sigma} \cdot (\mathbf{q} \times \mathbf{A})\chi, \tag{1.43}$$

where χ is the Pauli spinor, and $\mathbf{q}$ is interpreted as the soft-gluon momentum. The symmetry-breaking parameter ε is $\sqrt{-q^2/m^2}$, where "soft" means that $-q^2 < \mu^2$, μ

being the scale of the softest gluons that have been integrated out. Since spin-flavor rotations do not act upon color indices, the interaction in Eq. 1.43 transforms under an adjoint representation of $SU(4)$.

Exactly as before, we use a basis of states created by the generators of the *broken* symmetry ($SU(4)$, not $SU(4)_0$)—remember that, up to $O(1/m^2)$, the generators of the unbroken symmetry satisfy the same algebra. The symmetry breaking induces a mixture of pure states of $SU(4)$ to form the physical states:

$$|\tilde{\alpha}\rangle = c_{\alpha\alpha}|\alpha\rangle + \sum_n c_{\alpha n}|n\rangle, \tag{1.44}$$

where the states contain a heavy quark and (spectator) light degrees of freedom. Note that the interaction (1.43) is diagonal in flavor, but it can change the spin of the heavy quarks. The induced states $|n\rangle$ can therefore have spins different from the orthogonal state $|\alpha\rangle$ but still be in the same irreducible representation. A concrete example should make this clear: If $|\tilde{\alpha}\rangle$ is the B^- meson, $|\alpha\rangle$ consists of the b quark and a superposition of the $\overline{u}$ valence quark plus sea quark-antiquark pairs and gluons (denoted as "$\overline{u}$") with the quantum numbers appropriate to build an eigenstate of $SU(4)$:

$$|\alpha\rangle = \frac{1}{\sqrt{2}}[|b(\uparrow)\text{ ``}\overline{u}(\uparrow)\text{''}\rangle + |b(\downarrow)\text{ ``}\overline{u}(\downarrow)\text{''}\rangle], \tag{1.45}$$

where the arrows indicate helicities. On the other hand, the B^{*-} state in $SU(4)$ is obtained from the state $|\alpha\rangle$ by rotating the spin of $|b\rangle$ with the generator Q_{5i}, and thus belongs to the same irreducible representation as B^-. In general, the *physical* state $|\tilde{B}^-\rangle$ consists primarily of the $SU(4)$-*symmetric* state $|B^-\rangle$, but can also have components from other representations:

$$|\tilde{B}^-\rangle = c_{BB}|B^-\rangle + \sum_n c_{Bn}|n\rangle. \tag{1.46}$$

The $|B^{*-}\rangle$ component, although in the same irreducible representation of $SU(4)$ as $|B^-\rangle$, never enters into the $|\tilde{B}^-\rangle$ expansion in any Lorentz frame. This result follows because the interaction (1.43) is an $O(3)$ scalar in the heavy quark limit $|\mathbf{v} - \mathbf{v}'| \ll 1$, and therefore cannot mix vector and pseudoscalar states. Transitions between B and B^* (or D and D^*) states are thus forbidden if no other particles are

produced. Of course, since particle spin and parity are defined in the $\mathbf{v} = 0$ (*i.e.*, heavy quark or hadron at rest) frame, the use of the frame $\mathbf{v}, \mathbf{v}'$ both approximately zero allows us very conveniently to distinguish spin states.

Consequently, the AG theorem in its simpler form holds for the frame $\mathbf{v} = 0$, both because of the cancellation of field time-derivatives and the separation of spin states in this frame. Anticipating the relations of weak form factors in the $SU(4)$-symmetry limit [6, 16], we parametrize the matrix elements as

$$\begin{aligned}
\langle D(p)|V_\mu^{1+i2}|\overline{B}(p')\rangle &= \sqrt{m_D m_B}\left[(v'_\mu + v_\mu) f_+(v'v)\right. \\
&\quad \left. +(v'_\mu - v_\mu) f_-(v'v)\right], \qquad (1.47)\\
\langle D^*(p,\epsilon)|A_\mu^{1+i2}|\overline{B}(p')\rangle &= \sqrt{m_{D^*} m_B}\left\{\left[\epsilon^*_\mu(1+v'v) - v_\mu(v'\epsilon^*)\right] g_1(v'v)\right. \\
&\quad \left. +v_\mu(v'\epsilon^*) g_2(v'v) + v'_\mu(v'\epsilon^*) g_3(v'v)\right\}, \qquad (1.48)\\
\langle D^*(p,\epsilon)|V_\mu^{1+i2}|\overline{B}(p')\rangle &= \sqrt{m_{D^*} m_B}\left[i\epsilon_{\mu\nu\kappa\lambda}\epsilon^{*\nu} v^\kappa v'^\lambda f_3(v'v)\right], \qquad (1.49)
\end{aligned}$$

where V^μ and A^μ are the polar and axial vector components of the short-distance corrected effective weak current $\overline{J}^\mu_{cb}$ introduced in Section 1.3. The form factors in the symmetry limit assume the relations

$$\begin{aligned}
f_+(v'v) &= f_3(v'v) = g_1(v'v), \\
f_-(v'v) &= g_2(v'v) = g_3(v'v) = 0,
\end{aligned} \qquad (1.50)$$

up to phases in the normalization of the meson states. We now follow Ref. [16] and define the symmetry limit of $f_+(v'v)$ to be $\xi(v'v)$ (called the Isgur–Wise function). Because V^0 is a generator of $SU(4)$ in the limit $\mathbf{v} = \mathbf{v}' = 0$, the normalization condition of this function is

$$\xi(1) = 1. \qquad (1.51)$$

Once we turn on the $O(1/m)$ symmetry breaking, the AG theorem comes into force. We again use $\varepsilon \equiv \sqrt{-q^2/m^2}$, where q^μ is the typical soft-gluon momentum. We have seen in Section 1.3 that V^0 and A^i are components of the charge algebra up to $O(1/m^2)$, and thus we find

$$\begin{aligned}
f_+(1) &= 1 + O(\varepsilon^2), \\
g_1(1) &= 1 + O(\varepsilon^2).
\end{aligned} \qquad (1.52)$$

On the other hand, for all form factors whose kinematical coefficient vanishes in the symmetry limit, there is simply no way to apply the AG theorem. One then finds

$$\begin{aligned} f_3(1) &= 1 + O(\varepsilon), \\ f_1(1), g_2(1), g_3(1) &= O(\varepsilon). \end{aligned} \tag{1.53}$$

Equations 1.52–1.53 are Luke's conclusions [18], for which the relevance of the AG theorem was first demonstrated by Boyd and Brahm [22]. One can check these predictions in a particular model; in Section 1.7 we apply them to a calculation by Isgur *et al.* [14].

1.6 Scalar, Tensor, and Pseudoscalar Densities

We have thus far only considered the implications of the AG theorem for the matrix elements of the polar and axial vector currents because these are, of course, the dominant effects in the $V-A$ weak interaction. When we include terms in the renormalization group expansion of $O(m_c/m_b)$, the scalar and pseudoscalar densities also appear [17]. It is also interesting to consider scalar, pseudoscalar, and tensor densities, not only for their possible usefulness in future quark-model calculations, but also logically to complete the analysis. The densities in a general v^μ frame are

$$\overline{h}_v(\tau_a/2)h_{v'}, \quad i\overline{h}_v\gamma_5(\tau_a/2)h_{v'}, \quad \overline{h}_v\sigma_{\mu\nu}h_{v'}, \quad \overline{h}_v\sigma_{\mu\nu}(\tau_a/2)h_{v'}. \tag{1.54}$$

We find that all of these densities can be expressed in terms of a single scalar function $\eta(v'v)$ in the symmetry limit, exactly as for the V and A currents. This results from the fact that proper Lorentz transformations relate the space-space and space-time components of the tensor densities, and the spin symmetry relates these to the scalar and pseudoscalar densities, respectively. We parametrize these form factors as

$$\langle D(p)|\overline{c}_v b_{v'}|\overline{B}(p')\rangle = \sqrt{m_D m_B}(1+v'v)s(v'v), \tag{1.55}$$

$$\langle D^*(p,\epsilon)|i\overline{c}_v\gamma_5 b_{v'}|\overline{B}(p')\rangle = -i\sqrt{m_{D^*}m_B}(v'\epsilon^*)p(v'v), \tag{1.56}$$

$$\langle D(p)|\overline{c}_v\sigma_{\mu\nu}b_{v'}|\overline{B}(p')\rangle = i\sqrt{m_D m_B}(v_\mu v'_\nu - v_\nu v'_\mu)t(v'v), \tag{1.57}$$

$$\begin{aligned}\langle D^*(p,\epsilon)|\overline{c}_v\sigma_{\mu\nu}b_{v'}|\overline{B}(p')\rangle &= -\sqrt{m_{D^*}m_B}\epsilon_{\mu\nu\kappa\lambda}[\epsilon^{*\kappa}(v^\lambda + v'^\lambda)t_+(v'v) \\ &\quad +\epsilon^{*\kappa}(v^\lambda - v'^\lambda)t_-(v'v) + v^\kappa v'^\lambda(v'\epsilon^*)t'(v'v)].\end{aligned} \tag{1.58}$$

In the symmetry limit, we have

$$\begin{aligned} s(v'v) &= p(v'v) = t(v'v) = t_+(v'v) = \eta(v'v), \\ t_-(v'v) &= t'(v'v) = 0, \end{aligned} \tag{1.59}$$

with

$$\eta(1) = 1. \tag{1.60}$$

In fact we can improve upon the results of Ref. [24]—the function $\eta(v'v)$ is none other than the Isgur–Wise function $\xi(v'v)$. There are several ways to see this result. The original assumption of a fully covariant formulation (Ref. [16]) renders this result trivial. Furthermore, one may use an analysis similar to that in Ref. [6] to relate the V, A to the S, T, P matrix elements by including *improper* Lorentz transformations via a parity flip. This point was described in a general setting by Politzer [27], who argued that only one function exists for the lowest-lying mesons because of angular momentum conservation and parity conservation of the light degrees of freedom.

To apply the AG theorem, we need only find which bilinear components have vanishing $O(1/m)$ corrections in the Pauli expansion, for then they are generators of the $SU(4)$ algebra up to $O(1/m^2)$. These turn out to be the scalar current and space-space components of the tensor current; they are related to the V^0 and A^i currents, respectively, through a parity flip of one heavy quark field, which induces an additional γ_0. As an example, the scalar current in the $\mathbf{v} = 0$ frame is

$$\overline{c}b = c^\dagger\gamma_0 b = (c^+)^\dagger(b^+) - (c^-)^\dagger(b^-), \tag{1.61}$$

whereas the time component of the vector current is

$$\overline{c}\gamma_0 b = c^\dagger b = (c^+)^\dagger(b^+) + (c^-)^\dagger(b^-). \tag{1.62}$$

The parity flip changes the sign of the *small* components of the spinors ($h^\mp$ for particles (antiparticles)), which are $O(1/m)$; therefore, the scalar current and the

time component of the vector current differ only at $O(1/m^2)$ in this frame. Similar arguments relate A^i to T^{jk}. Consequently, the AG theorem predicts the form factors to have the limiting behavior

$$\begin{aligned} s(1) &= 1+O(\varepsilon^2), \\ p(1) &= 1+O(\varepsilon), \\ t(1) &= 1+O(\varepsilon), \\ t_+(1) &= 1+O(\varepsilon), \\ t_-(1), t'(1) &= O(\varepsilon). \end{aligned} \tag{1.63}$$

1.7 A Sample Application

We conclude this Chapter with an example of the application of the AG theorem in a particular nonrelativistic quark model calculation. From the Appendix of Isgur *et al.* [14], we find the following expressions for our form factors in terms of their notation:

$$f_+(1) = g_1(1) = [2\beta_D\beta_B/(\beta_D^2+\beta_B^2)]^{\frac{3}{2}}, \tag{1.64}$$

$$\begin{aligned} f_3(1) &= [2\beta_D\beta_B/(\beta_D^2+\beta_B^2)]^{\frac{3}{2}}\{1+(m_D-m_c)/m_c-[(m_b-m_c)/2m_cm_b]m_{u,d} \\ &\quad -[(m_b-m_c)/2m_cm_b][(\beta_B^2-\beta_D^2)/(\beta_B^2+\beta_d^2)]m_{u,d}\}, \end{aligned} \tag{1.65}$$

where $m_{u,d}$ are the constituent light quark masses, and $\beta_{B,D}$ are the characteristic widths of the B and $D(D^*)$ bound-state wavefunctions in the nonrelativistic quark model. In obtaining the expression for $f_+(1)$ we have used the relation

$$m_b - m_c = m_B - m_D + O\left([m_{u,d}(m_b-m_c)/m_bm_c]\Delta B\right), \tag{1.66}$$

where ΔB is the universal binding energy of heavy mesons in the limit of $m_{b,c} \to \infty$. The correction term originates from reduced mass effects on binding; it is negligible because it contributes only $O(m_{u,d}\Delta B/m_c^2) = O(\varepsilon^2)$ to the form factors (ε is the same as in Secs. 1.5–1.6).

The widths of the wavefunctions may be expected to deviate from their

symmetry values at first order in symmetry breaking, up to logarithms:

$$\begin{aligned} \beta_B(m_B) &= \beta(\infty) + O(\sqrt{-q^2}/m_B), \\ \beta_D(m_B) &= \beta(\infty) + O(\sqrt{-q^2}/m_D). \end{aligned} \tag{1.67}$$

From Eq. 1.67 we find

$$\begin{aligned} 2\beta_D\beta_B/(\beta_D^2+\beta_B^2) &= 1 - [(\beta_B-\beta_D)^2/(\beta_B^2+\beta_D^2)] \\ &= 1 + O(\varepsilon^2), \end{aligned} \tag{1.68}$$

$$(\beta_B^2-\beta_D^2)/(\beta_B^2+\beta_D^2) = O(\varepsilon). \tag{1.69}$$

Substituting these relations into Eqs. 1.64,1.65, we can confirm the results of the AG theorem in Section 1.5 for the form factors $f_+(1)$, $f_3(1)$, and $g_1(1)$.

Chapter 2

Composite Fermions in a Three-Generation Electroweak Model

2.1 Introduction

The top quark is extraordinarily heavy. Strictly speaking, the current lower bound on its mass ($m_t \geq 131$ GeV) [9] is large—far larger than the masses of all the known quarks, larger even than the masses of the weak gauge bosons. Indeed, its mass is of the order of the electroweak symmetry-breaking scale, which naturally leads one to wonder whether the top quark itself might somehow be responsible for the symmetry breakdown.

A model that realizes this suggestion was proposed by Nambu [28], and by Miransky, Tanabashi, and Yamawaki [29]. The usual vacuum expectation value (VEV) arises through a condensation of $t\bar{t}$-pairs via a Nambu–Jona-Lasinio mechanism [30], in analogy to the formation of Cooper pairs in superconductivity. This condensation is the consequence of a new interaction whose low-energy form is non-renormalizable within the Standard Model, *i.e.*, an interaction whose scale Λ is large compared to the electroweak scale v. Such a $t\bar{t}$-condensate carries exactly the symmetry-breaking properties and quantum numbers of what we call the Higgs

boson; we thus conclude that the Higgs is not an elementary particle in this model, but a *dynamically-generated* composite of t and $\bar{t}$. Henceforth we refer to this model as the top-condensate or composite Higgs model.

Because the Higgs field is not fundamental in this model, its kinetic term in the Lagrangian is generated only through the momentum dependence of loop diagrams. Therefore, at the compositeness scale, where the $t\bar{t}$-condensates break apart, the Higgs must cease to exist as a propagating field, and the dynamically-generated kinetic term vanishes. However, in the low-energy Lagrangian, we normalize kinetic terms to unity at all scales. As a result, *renormalized* Yukawa and quartic Higgs couplings in this model rise to infinity at the scale Λ; we call these the *compositeness conditions* of the model. The running of these couplings give rise to renormalization group equations (RGE's), which we use to predict a relation between the top and Higgs masses ($m_H = 2m_t$ in Refs. [28, 29]) through the presence of infrared fixed points.

Unfortunately, only the toy form of this model, in which the effective couplings are determined by summing simple chain diagrams alone (Fig. 2.1, p. 27), is exactly solvable. Without additional particle lines connecting the "links," such diagrams may be summed in a simple geometric series. This special case is equivalent to assuming the dominance of color loops of fermions, which is the limit of large N_C with gauge couplings turned off. However, the top-condensate model was extended to the physical case including finite N_C and gauge couplings by Bardeen, Hill, and Lindner [31] by using the following principle: The RGE's for two models with the same effective interactions at a given energy scale are the same, whether or not any of the particles are composite; the compositeness conditions serve merely as initial values for the RGE's. Consequently, the *full* RGE's are the same as in the Standard Model, and therefore the RGE's derived in the top-condensate model must be exactly the large-N_C limits of the usual RGE's without gauge couplings. We therefore have a prescription for computing the running of couplings in this model, despite the fact that we are not smart enough to compute and add up all of the complicated diagrams involved.

In the top-condensate model, the nonrenormalizable interaction is a dimen-

sion six four-fermion operator. In fact it turns out that the most general form of the model possesses, through higher-dimensional operators, exactly the same number of parameters as the minimal standard model, and so one cannot use it to predict a *unique* relation between the top and Higgs couplings [32, 33]. However, if the dimensionless coefficients of these higher-dimension operators are natural ($O(1)$ or smaller), and if the running length from the compositeness to electroweak scales is large ($\ln(\Lambda/v) \gg 1$), then the couplings may be expected to approach their infrared fixed points quite closely, undisturbed by the precise form of the interaction. (In fact, Marciano [34] made numerical predictions from an RGE analysis without reference to a particular model.) The numerical predictions of low-energy parameters in the model are thus altered only minimally by these uncertainties [35].

This is, however, not the end of the story. As pointed out by Suzuki [36], it is precisely the inability of the RGE to discern compositeness that permits a number of distinct models with the same compositeness scale Λ and the same compositeness condition and hence the same low-energy predictions. One simply observes that the divergence at some large scale Λ of the Yukawa coupling connecting the Higgs and t_L, t_R fields would occur if *any one* of these three fields were a composite of the other two. In the case of the composite-t_R model constructed in Ref. [36], the toy-model chain diagrams (Fig. 2.2, p. 31) include a sum over the number N_F of particles in the left-handed weak multiplets, and so the model is solvable in the large-N_F limit. The prescription of Ref. [31] then generates the full RGE. Because this is the same as the RGE in the top-condensate model, the two models are equivalent at low energies.

Of course we may argue that the composite-t_R model is unaesthetic, because it treats the two chiralities of top quark very differently, and we still have an elementary scalar in the theory. Physically, however, the theories are superficially indistinguishable at low energies, and so we must look deeper for distinctions. It is when we attempt to include additional fermions in the composite-t_R model [37] that we begin to see differences from the top-condensate model. These differences come about because of the quantum numbers of the composite particle: Unlike the Higgs, the t_R carries spinor and color indices, and thus we find the matching of the

indices to be trickier in the composite-t_R model.

We consequently find that the Higgs sector in this model is subject to strong constraints, with the least unnatural model of composite right-handed fermions consisting of all third-generation fermions (t_R, b_R, τ_R) being formed from the condensation of two Higgs doublets with the corresponding left-handed fermion doublets, and all other fermions being elementary. We also repeat this analysis for a proposed composite-t_L model.

This Chapter is organized as follows: We begin in Sec. 2.2 by seeing how one may naturally generalize the top-condensate model to three fermion generations. In Sec. 2.3, we briefly review the content of the composite-t_R model of Ref. [36]. The composite-t_R model is enlarged in Sec. 2.4, first by the addition of the b quark alone, and then the τ lepton and other generations. In Sec. 2.5, we develop an analogous composite-t_L model and consider its extension to the b_L and other generations. We summarize and comment upon the results in Sec. 2.6.

2.2 Three Generations in the Top-Condensate Model

As mentioned in the Introduction, arranging for three generations of massive fermions in the top-condensate model [29, 38, 39] is straightforward because the quantum numbers of the composite Higgs are easy to accommodate. When only one Higgs doublet is desired, one forms the composite Higgs field in a linear superposition channel of the fields $\overline{u_{Rj}}\,\psi_{Li}$, $\overline{d^c_{Rj}}\,\psi^c_{Li}$, and $\overline{e^c_{Rj}}\,\ell^c_{Li}$, where the charge-conjugate fields have been introduced to supply the appropriate quantum numbers, and $i, j = 1, 2, 3$ are generation indices. Also,

$$\psi^c_{Li} \equiv \begin{pmatrix} -d^c_{Li} \\ u^c_{Li} \end{pmatrix}, \quad \ell^c_{Li} \equiv \begin{pmatrix} -e^c_{Li} \\ \nu^c_{Li} \end{pmatrix}. \tag{2.1}$$

The binding interaction responsible for fermion condensation may be written

$$L_{int} = GJ^\dagger J, \tag{2.2}$$

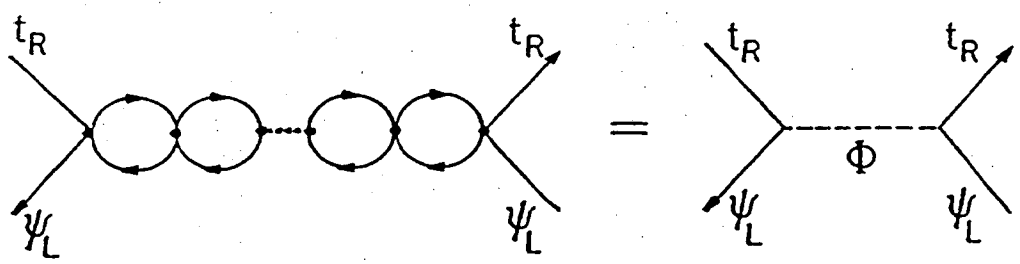


Figure 2.1: Fermion chain diagrams in the top-condensate model that dominate the calculation of effective Yukawa couplings for large N_C. The ellipsis indicates iteration of the chain.

where the scalar density J with the quantum numbers of the Higgs is defined by

$$J = \sum_{ij} (k_{ij}\, \overline{u_{Rj}}\, \psi_{Li} + k_{i+3,j+3}\, \overline{d^c_{Rj}}\, \psi^c_{Li} + k_{i+6,j+6}\, \overline{e^c_{Rj}}\, \ell^c_{Li}). \tag{2.3}$$

As in the Standard Model, we may perform separate unitary rotations upon left- and right-handed fermions to diagonalize the interaction into fermion mass eigenstates:

$$J = \sum_{i} (k_i\, \overline{u_{Ri}}\, \psi_{Li} + k_{i+3}\, \overline{d^c_{Ri}}\, \psi^c_{Li} + k_{i+6}\, \overline{e^c_{Ri}}\, \ell^c_{Li}), \tag{2.4}$$

where now

$$\psi_{Li} \equiv \begin{pmatrix} u_{Li} \\ (V d_L)_i \end{pmatrix}, \quad \psi^c_{Li} \equiv \begin{pmatrix} -d^c_{Li} \\ (V^T u^c_L)_i \end{pmatrix}, \tag{2.5}$$

with V being the Kobayashi–Maskawa matrix, and the coefficients k_i being the eigenvalues of the diagonalized couplings k_{ij} normalized to unity:

$$\sum_{i} (|k_i|^2 + |k_{i+3}|^2 + |k_{i+6}|^2) = 1. \tag{2.6}$$

It is clear that the coefficients k_i are proportional to the fermion masses in the limit of the top-condensate toy (large-N_C) model, for one may compute the effective Yukawa couplings from the chain diagrams of Fig. 2.1:

$$\begin{aligned} f_{u_i}(\mu)^2 &= 16\pi^2|k_i|^2/[N_C \ln(\Lambda^2/\mu^2)], \\ f_{d_i}(\mu)^2 &= 16\pi^2|k_{i+3}|^2/[N_C \ln(\Lambda^2/\mu^2)], \\ f_{\ell_i}(\mu)^2 &= 16\pi^2|k_{i+6}|^2/[N_C \ln(\Lambda^2/\mu^2)]. \end{aligned} \tag{2.7}$$

The compositeness condition of the Higgs doublet then reads

$$\sum_i (f_{u_i}(\mu)^2 + f_{d_i}(\mu)^2 + f_{\ell_i}(\mu)^2) = 16\pi^2/[N_c \ln(\Lambda^2/\mu^2)] \to \infty \text{ as } \mu \to \Lambda. \tag{2.8}$$

Note that this is a condition only on the *sum* of the squared Yukawa couplings; this constraint assures that the given combination of the couplings satisfies the large-N_C limit of its corresponding RGE. However, the individual couplings are not analogously constrained. Thus in the large-N_C limit one may trivially choose the ratios of parameters k_i to recover the experimentally observed fermion mass spectrum.

In the real world of finite N_C, we may use the prescription of Ref. [31] to abstract the full (Standard Model) RGE's including gauge contributions and nonleading effects in N_C, which in turn modifies the running of the Yukawa couplings from the compositeness scale. Defining the differential operator $D \equiv 16\pi^2\mu\partial/\partial\mu$, and ignoring quark mixing, one finds the one-loop single-generation RGE's

$$\begin{aligned} D(f_{d_i}/f_{u_i}) &= 3f_{u_i}f_{d_i}(f_{d_i}^2/f_{u_i}^2 - 1) + g_1^2 f_{d_i}/f_{u_i}, \\ D(f_{\ell_i}/f_{\ell_j}) &= \tfrac{3}{2}f_{\ell_i}f_{\ell_j}(f_{\ell_i}^2/f_{\ell_j}^2 - 1), \end{aligned} \tag{2.9}$$

where g_1, the $U(1)$ gauge coupling, is small at all scales and is hence negligible in this analysis. It is particularly notable that the $SU(3)_C \times SU(2)_F$ gauge loops cancel in these expressions, in the first expression because the two members of the weak doublet differ physically only by hypercharge, and in the second because of family replication symmetry. Then we find that the full RGE's in the realistic model possess infrared fixed points at $f_{u_i}/f_{d_i} = 1$ and $f_{\ell_i}/f_{\ell_j} = 1$. If the running distance $\ln(\Lambda/v)$ were sufficiently large, the Yukawa couplings would approach their infrared fixed points regardless of initial values, and with only one VEV in this one-doublet model, we would obtain various unfortunate predictions such as $m_t = m_b$ and $m_\tau = m_e$. But $\ln(\Lambda/v)$ is happily large yet finite, and thus sufficiently asymmetric

boundary conditions (*e.g.*, $|k_6| \ll |k_3|$ for b and t quarks) evade the fixed point expressions like $f_b(v)/f_t(v) = 1$, allowing $m_b \ll m_t$ and other realistic ratios of fermion masses.

If we extend this model to include two composite Higgs doublets, then there are two *distinct* compositeness conditions. Consider, for example, the case in which one doublet couples only to up-type quarks, and the other couples to down-type quarks and leptons [38]. Then the Yukawa couplings in the toy model still behave as in Eq. 2.7, but now the normalization conditions on k_i read $\sum_i |k_i|^2 = 1$ and $\sum_i(|k_{i+3}|^2 + |k_{i+6}|^2) = 1$. Now we are compelled to choose comparable values for the coefficients $|k_6|$ and $|k_3|$, and the Yukawa ratio f_b/f_t is driven to near unity at low energies. This running behavior was confirmed numerically by Luty [40], whose two-doublet model generates a t-quark mass essentially the same as in the one-doublet model. But the masses of the b and t need not be equal in the two-doublet model, for we have an additional VEV at our disposal, which can be used to adjust the mass ratio of m_b/m_t even though the Yukawa couplings are essentially equal.

2.3 The Composite-t_R Model

Models with a composite t_R, in contrast with the top-condensate model, turn out to require nontrivial particle content at high energies, owing to the additional balancing that must be done with spinor and color indices in the former model. Here we briefly review the minimal model proposed by Suzuki [36]. We begin with the gauge group $SU(3)_C \times SU(N_F) \times U(1)$, with left-handed fermion and Higgs multiplets transforming under this group as

$$\begin{aligned} \psi_L &= (3, N_F, Y/2), \\ \Phi &= (1, \overline{N}_F, Q_t - Y/2), \end{aligned} \tag{2.10}$$

where Q_t is the electric charge of the t quark. A composite field with the correct chirality and gauge quantum numbers to be a composite t_R is

$$\xi_R \sim (\partial\!\!\!/ \Phi_a)\psi_L^a = (3, 1, Q_t), \tag{2.11}$$

where a indicates the $SU(N_F)$ index. However, a composite fermion must appear with both chiralities, or else its Dirac mass is zero. Because there exist many radiative corrections that could spoil such a coincidence, the only natural way to generate this scenario is for ξ_R to be a Goldstino of broken supersymmetry. Barring this possibility, there exists in this model the composite state

$$\xi_L \sim \Phi_a \psi_L^a = (3, 1, Q_t). \tag{2.12}$$

The field ξ_L represents a particle with quantum numbers not observed in the low-energy Standard Model; the simplest way to remove it is to postulate an elementary right-handed "quark"

$$\eta_R = (3, 1, Q_t), \tag{2.13}$$

such that ξ_L and η_R together have a Dirac mass at the compositeness scale Λ. One may also observe that a particle with the quantum numbers of η_R is actually required to cancel the electroweak anomaly induced by ξ_L.

Now we have the situation that ξ_L couples to both ξ_R and η_R. We must diagonalize the couplings to obtain the physical mass eigenstates. The combination $(\overline{\xi_L}\xi_R + \text{H.c.})$ obtains a mass through both its bare value $(O(\Lambda))$ and through loop diagrams with intermediate Φ and ψ_L fields, which in sum can be fine-tuned to a value of $O(v)$; this is equivalent to the fine-tuning of the Higgs mass in the usual Standard Model. The combination $(\overline{\xi_L}\eta_R + \text{H.c.})$ obtains its mass through the VEV of a supermassive Higgs and so remains at $O(\Lambda)$. Thus under diagonalization, the light eigenstate consists of nearly all ξ_R and a small admixture of η_R, with coefficient $O(v/\Lambda) \ll 1$; we identify this eigenstate as the physical t_R. The heavy eigenstate mass remains at $O(\Lambda)$ and therefore has no effect on low energy physics.

The equivalent nonrenormalizable binding interaction that forms the composite states $\xi_{R,L}$, as given in Eqs. 2.11, 2.12, is of the form

$$L_{int} = -i\frac{g}{\Lambda^2}\overline{\psi_{L_a}}\Phi^{\dagger a}(\partial\!\!\!/\Phi_b)\psi_L^b + \text{ H.c.} \tag{2.14}$$

We may now use this interaction to calculate the effective Yukawa coupling using the chain diagrams of Fig. 2.2. Each loop contains a sum over the $SU(N_F)$ multiplets,

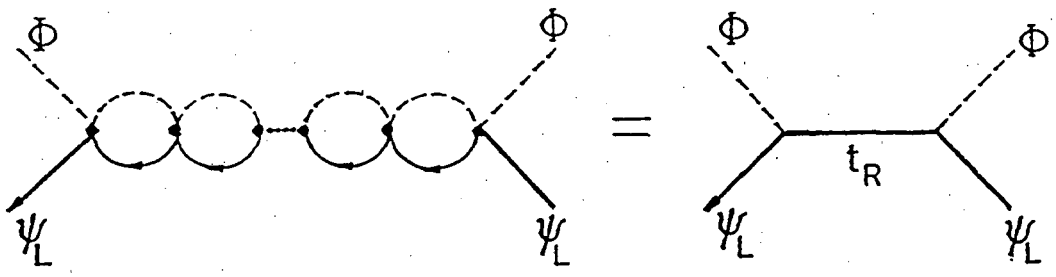


Figure 2.2: Chain diagrams in the composite-t_R model that dominate the calculation of effective Yukawa couplings for large N_F.

and thus these chain diagrams dominate the calculation in the large-N_F limit. The result is

$$f_t(\mu)^2 = 32\pi^2/[N_F \ln(\Lambda^2/\mu^2)] \Rightarrow Df_t = \frac{N_F}{2} f_t^3. \tag{2.15}$$

We compare this to the result from the minimal top-composite model (as can be extracted from Eq. 2.7):

$$f_t(\mu)^2 = 16\pi^2/[N_C \ln(\Lambda^2/\mu^2)] \Rightarrow Df_t = N_C f_t^3. \tag{2.16}$$

The full one-loop RGE for the top Yukawa coupling with arbitrary $N_{C,F}$, including all gauge bosons but neglecting other fermions, reads

$$\begin{aligned} Df_t = &\left(N_C + \frac{N_F}{2} + \frac{1}{2}\right) f_t^3 \\ &-3\left\{[(N_C^2-1)/N_C]g_{N_C}^2 + [(N_F^2-1)/2N_F]g_{N_F}^2 + (Y^2/4 + Q_t^2)g_1^2\right\} f_t, \end{aligned} \tag{2.17}$$

from which we can easily see how the prescription of Ref. [31] applies to both models.

There is a superficial difference between the top-condensate and composite-t_R *toy* models in that, although the Yukawa couplings diverge in both, the Higgs is elementary in the latter; we therefore naively expect that the quartic Higgs coupling $\lambda(\mu)$ does not diverge as $\mu \to \Lambda$. Thus the composite-t_R model does not appear to have as many compositeness conditions as the top-condensate model. However,

this is only true for the toy models; when we leave the large-N_F limit, quark box diagrams renormalize the quartic coupling with a logarithmic divergence:

$$\lambda(\mu) \approx \lambda_0 + f(\mu)^4 \ln(\Lambda^2/\mu^2) \approx 1/\ln(\Lambda^2/\mu^2) \to \infty \text{ as } \mu \to \Lambda, \tag{2.18}$$

because $f(\mu)^2 \approx 1/\ln(\Lambda^2/\mu^2)$ in both models.

2.4 Extending the Composite-t_R Model

2.4.1 One Generation with One Higgs Doublet

The most basic extension to the model is to generate t_R and b_R simultaneously as composite, massive particles. The physical case of $N_F = 2$ is actually special, because all representations of $SU(2)$ are real, whereas N_F and $\overline{N}_F$ are inequivalent representations for $N_F > 2$. In the Standard Model, this fact provides us with two distinct but group-theoretically equivalent forms for the Higgs doublet, $\Phi \equiv (\phi^0, \phi^+)$ and $\tilde{\Phi}^\dagger \equiv (-i\tau_2\Phi^\dagger)^T = (-\phi^-, \overline{\phi}^0)$. Since no similar relation preserving the transformation structure of the Higgs multiplet exists for $N_F > 2$, one cannot form two $SU(N_F)$ singlets t_R and b_R. Formally, we evade this restriction by pretending for the moment that Φ and $\tilde{\Phi}^\dagger$ are *distinct* multiplets transforming under the $\overline{N}_F$ representation of $SU(N_F)$, solving the toy model in the large-N_F limit, and then setting $N_F = 2$ so that the two Higgs multiplets may be related.

The composite t_R and b_R are assumed to form in the channels

$$t_R \sim \Phi\psi_L \quad \text{and} \quad b_R \sim \tilde{\Phi}^\dagger\psi_L, \tag{2.19}$$

and the Yukawa couplings are defined through

$$L_{int} = -f_t\,\overline{t_R}(\Phi\psi_L) - f_b\,\overline{b_R}(\tilde{\Phi}^\dagger\psi_L) + \text{H.c.}, \tag{2.20}$$

where the parentheses indicate a sum over $SU(N_F)$ indices. Now we may solve the toy model of Ref. [36] by summing chain diagrams as in Fig. 2.2 in the large-N_F limit to obtain

$$f_t(\mu)^2 = f_b(\mu)^2 = 32\pi^2/[N_F \ln(\Lambda^2/\mu^2)]. \tag{2.21}$$

Note that, in contrast to the top-condensate model, we have two *distinct* compositeness conditions, for t_R and b_R separately; we remind the reader that this results from the comparative difficulty of forming composite states with the quantum numbers of chiral fermions rather than scalars.

Now we employ the prescription of Ref. [31] and proceed directly to $N_F = 2$. The full one-loop RGE's with one Higgs doublet for the top and bottom Yukawa couplings (but ignoring other generations) read (*cf.* Eq. 2.17)

$$\begin{aligned} Df_t &= \tfrac{9}{2}f_t^3 + \tfrac{3}{2}f_b^2 f_t \\ &\quad -(8g_3^2 + \tfrac{9}{4}g_2^2 + \tfrac{17}{12}g_1^2)f_t, \\ Df_b &= \tfrac{9}{2}f_b^3 + \tfrac{3}{2}f_t^2 f_b \\ &\quad -(8g_3^2 + \tfrac{9}{4}g_2^2 + \tfrac{5}{12}g_1^2)f_b. \end{aligned} \tag{2.22}$$

From Eq. 2.21 we see that the values of $f_t(\mu)$ and $f_b(\mu)$ are large and comparable near $\mu = \Lambda$; on the other hand, the ratio of Yukawa couplings obeys the running given in Eq. 2.9:

$$D(f_b/f_t) = 3f_t f_b(f_t^2 f_b^2 - 1) + O(g_1^2). \tag{2.23}$$

Thus, because $f_b(\mu)/f_t(\mu) \approx 1$ for $\mu \to \Lambda$, this ratio must very nearly approach its infrared fixed point of $f_t(\mu) = f_b(\mu)$ as $\mu \to v$. In the one-doublet scenario, Φ and $\tilde{\Phi}^\dagger$ have identical VEV's:

$$\langle\Phi\rangle = \left(\frac{v}{\sqrt{2}}, 0\right), \quad \langle\tilde{\Phi}^\dagger\rangle = \left(0, \frac{v}{\sqrt{2}}\right). \tag{2.24}$$

Thus the quark masses $m_{t,b} = f_{t,b}(v)v/\sqrt{2}$ are equal, in sharp contradiction to reality. We conclude that one Higgs doublet in the composite-t_R model is just not enough to give the appropriate masses to composite t and b quarks.

One might be troubled by the outright neglect of the $U(1)$ coupling g_1 in this analysis. It has been suggested [29, 41] that this coupling may upset the RGE's to the extent that the ratio $f_t(v)/f_b(v)$ is driven far from unity; such an effect is called *critical instability*. It is possible, however [42], to absorb such instabilities into the VEV's of the theory. In the usual Standard Model, this has the effect of fine-tuning the Higgs masses and thus has no physical manifestation once these

masses are fixed. Similarly, in the composite-t_R model, the instability is absorbed in the fine-tuning of the Dirac mass of the field ξ from Section 2.2, which has negligible effect when we diagonalize to obtain mass eigenstates: At low energies, it represents corrections to an $O(v/\Lambda)$ effect.

Therefore, with a single Higgs doublet, we conclude that the only option is to introduce b_R as an elementary particle and choose its Yukawa coupling at the scale Λ to be tiny so that we avoid the infrared fixed point upon evolving down to v. This is possible because $Df_b = O(f_b)$ for small f_b, so it is possible to keep Yukawa couplings small at all scales; thus one can obtain an acceptable value for m_b. Such a problem persists with more than one generation of fermions present: Elementary fermions with small Yukawa couplings at Λ do not substantially alter the RGE flow for the diverging composite fermion Yukawa couplings. On the other hand, each composite fermion is forced to have the same Yukawa coupling at scale Λ, and with only one VEV in the theory, all composite fermions necessarily have nearly equal masses.

2.4.2 One Generation with Two Higgs Doublets

The natural solution to the problem is simply to produce another VEV, which requires a second Higgs doublet. It is easy to generalize the model. We define

$$\Phi_t \equiv (\phi_t^0, \phi_t^+), \quad \tilde{\Phi}_b^\dagger \equiv (-\phi_b^-, \overline{\phi}_b^0), \tag{2.25}$$

to produce the Yukawa coupling

$$L_{int} = -f_t \overline{t_R}(\Phi_t \psi_L) - f_b \overline{b_R}(\tilde{\Phi}_b^\dagger \psi_L) + \text{H.c.}, \tag{2.26}$$

so that t_R and b_R are composite of $\Phi_t \psi_L$ and $\tilde{\Phi}_b^\dagger$, respectively. Apart from having two distinct VEV's in this case ($v_t \equiv \sqrt{2}\langle\phi_t^0\rangle \neq v_b \equiv \sqrt{2}\langle\overline{\phi}_b^0\rangle$ in general), the only real difference from the one-doublet case is that the RGE (*cf.* Eq. 2.23) has a different coefficient [43]:

$$D(f_b/f_t) = 4 f_t f_b (f_b^2/f_t^2 - 1) + O(g_1^2). \tag{2.27}$$

Nevertheless, the infrared fixed point is still $f_b/f_t = 1$, and so the previous methods still apply. Because in the two-doublet model we are using the same RGE's and boundary conditions as in Ref. [40], the same numerical analysis applies, confirming that $f_b/f_t = 1$ is indeed a true fixed point of the RGE. It is then a simple matter to obtain any value of m_t/m_b by choosing the equivalent ratio v_t/v_b; thus the two-doublet model produces a composite b_R naturally.

We can even improve upon this result. Although the g_2^2 and g_3^2 terms vanish in the RGE for the ratio f_b/f_t, this is not true if we compute the ratio of either quark Yukawa coupling with f_τ. Neglecting gauge couplings, the fixed point of such a ratio is still unity, but now the *strong* coupling g_3^2 (unlike g_1^2) can easily drive the ratio away from one. That this actually happens has been demonstrated by calculations in grand unified models [44], in which one may, for example, couple the same Higgs (and hence the same Yukawa coupling and VEV) to both the b quark and the τ and still find $m_\tau/m_b \approx \frac{1}{3}$.

2.4.3 More Than One Fermion Generation

The most natural way to include additional quark and lepton generations is to assume that only the third-generation quarks (and possibly the τ) are composite, and all of the other fermions are elementary fields. Then the RGE's as in Eq. 2.27 are modified by additional tiny Yukawa couplings. The numerical results must then be essentially identical to those in the previous Subsection. Such a model is not unnatural within the context of our original motivation that the third generation quarks, particularly the t, are rather different from the lighter quarks.

Neither is quark mixing difficult in this model. One simply assumes that t_R and b_R are actually formed in the combined channels

$$t_R \sim \Phi_t(\varepsilon_1\psi_{L1} + \varepsilon_2\psi_{L2} + \psi_{L3}), \quad b_R \sim \tilde{\Phi}_b^\dagger(\varepsilon_1'\psi_{L1} + \varepsilon_2'\psi_{L2} + \psi_{L3}), \tag{2.28}$$

where $|\varepsilon_1| \ll |\varepsilon_2| \ll 1$ and $|\varepsilon_1'| \ll |\varepsilon_2'| \ll 1$. Only in the very special case of $\varepsilon_i = \varepsilon_i'$ can one remove this mixing by redefining the combined channel to be ψ_{L3}; in general, diagonalization of this mass matrix gives rise to a nontrivial Kobayashi–Maskawa mixing.

It is also true that this model with only two Higgs doublets cannot generate more than one generation of massive quarks. In short, this follows because the RGE's force composite fermions to have the same Yukawa couplings at low energies, but there are only two independent VEV's. However, the mixings of Yukawa couplings in the full RGE's makes this conclusion somewhat less transparent, and so let us consider the particular example of only two generations of quarks, (t, b) and (c, s). The compositeness conditions then read

$$f_t(\mu)^2 = f_b(\mu)^2 = f_c(\mu)^2 = f_s(\mu)^2 = 32\pi^2/[N_F \ln(\Lambda^2/\mu^2)], \tag{2.29}$$

and the one-loop RGE's for the Yukawa coupling ratios read [43]

$$\begin{aligned} D\ln(f_b/f_t) &= 4(f_b^2 - f_t^2) + 3(f_s^2 - f_c^2) + O(g_1^2), \\ D\ln(f_s/f_c) &= 4(f_s^2 - f_c^2) + 3(f_b^2 - f_t^2) + O(g_1^2), \\ D\ln(f_c/f_t) &= \tfrac{3}{2}(f_c^2 - f_t^2) + \tfrac{1}{2}(f_s^2 - f_b^2), \\ D\ln(f_s/f_b) &= \tfrac{3}{2}(f_s^2 - f_b^2) + \tfrac{1}{2}(f_c^2 - f_t^2), \end{aligned} \tag{2.30}$$

where these equations are actually linearly dependent. To solve them, we choose large and comparable boundary values for the Yukawa couplings near Λ as suggested by the conditions Eq. 2.29, and evolve down to low energy. We now show that the deviations of these ratios from unity are damped exponentially as the energy scale decreases. Define the variables

$$\begin{aligned} \xi &\equiv f_b/f_t - 1, \quad & \eta &\equiv f_s/f_c - 1, \\ \zeta &\equiv f_c/f_t - 1, \quad & \omega &\equiv f_s/f_b - 1, \end{aligned} \tag{2.31}$$

and expand Eqs. 2.30 about $\xi = \eta = \zeta = \omega = 0$; such an expansion is reasonable since these variables are approximately zero at the scale Λ, and remains reasonable if (as we now show) their values do not increase as the energy scale decreases. After diagonalizing the equations, we find

$$\begin{aligned} D(\xi + \eta) &= 14 f_t^2(\xi + \eta), \\ D(\xi - \eta) &= 2 f_t^2(\xi - \eta), \\ D(\zeta + \omega) &= 4 f_t^2(\zeta + \omega), \\ D(\zeta - \omega) &= 2 f_t^2(\zeta - \omega). \end{aligned} \tag{2.32}$$

Thus the Yukawa coupling ratios are driven to unity very quickly as μ decreases. We conclude that

$$f_t(v) = f_b(v) = f_c(v) = f_s(v), \tag{2.33}$$

where v is a shorthand for v_1, v_2; considering the rapidity of convergence, the numerical distinction between them in this equation is irrelevant. Therefore, we have only one distinct Yukawa coupling and two VEV's, which means that only two distinct masses are possible. This conclusion holds even if we include realistic quark mixings and three generations.

The natural means by which one might hope to make any number of right-handed fermions composite is to introduce more Higgs doublets so that we have enough VEV's to adjust all of the fermion masses. Apart from the unappealing proliferation of elementary scalars in the theory, we encounter the very serious problem of flavor-changing neutral currents from mixing of scalars in the Higgs potential. Indeed, such currents can only be naturally suppressed [45] by coupling all of the up-type quarks to one doublet and all of the down-type quarks to another. (These two doublets may be the same, as in the minimal Standard Model.) Furthermore, the RGE analysis for only two doublets [40] suggests that at least one of the neutral Higgs bosons should have mass of $O(m_t)$; with more doublets, one might expect still lighter scalars and achieve a contradiction with experimental lower bounds on scalar masses. We conclude that it is impossible in this model to make any massive fermions composite except for t_R, b_R, and τ_R.

2.5 A Composite Left-Handed Fermion Model

In Ref. [36] it was pointed out that a composite-t_L model may be developed in analogy to the top-condensate and composite-t_R models. This case, however, possesses no natural large-N expansion with which to justify the dominance of simple chain diagrams near the compositeness scale. Nevertheless, if we are willing to take this dominance as an *ad hoc* assumption (recall that N_C and N_F are in reality not large integers), then the model may be constructed in parallel with the other

two cases. However, as has been emphasized elsewhere in this Chapter, it is the balancing of quantum numbers of the composite fields that accounts for the unique restrictions on the Higgs sector of each model. We focus upon the restrictions on the composite-t_L model in the following.

2.5.1 One Generation

Because t_L and b_L, unlike t_R and b_R, are related by $SU(2)_L$, we expect some differences from the composite-t_R model. In the present case, the doublet $\psi_L = (t_L, b_L)^T$ is a composite of t_R, b_R, and Φ; in general, it may be formed in the channel

$$\psi_L \sim \Phi^\dagger t_R \cos\alpha + \tilde{\Phi} b_R \sin\alpha, \tag{2.34}$$

where α is a mixing constant. Note that such a combination is not possible in the composite-t_R model. The Yukawa couplings are computed by iteration of chain diagrams identical to those in Fig. 2.2 except that the component fields are now $\Phi^\dagger t_R$ and $\tilde{\Phi} b_R$. In analogy to the other models, one finds

$$\begin{aligned} f_t(\mu)^2 &= 32\pi^2 \cos^2\alpha \,/\, \ln(\Lambda^2/\mu^2), \\ f_b(\mu)^2 &= 32\pi^2 \sin^2\alpha \,/\, \ln(\Lambda^2/\mu^2), \end{aligned} \tag{2.35}$$

with the *single* compositeness constraint $f_t(\mu)^2 + f_b(\mu)^2 \to \infty$ as $\mu \to \Lambda$. Thus, even with only one VEV, one can tune $f_b/f_t = \tan\alpha$ near the compositeness scale to a sufficiently small value that the ratio m_b/m_t is correctly reproduced. The composite-t_L,b_L model, in contrast with the composite-t_R,b_R model, requires only *one* Higgs doublet.

However, if we now attempt to add the τ lepton as a composite particle, as in the composite-t_R model, the single compositeness condition works against us: Then the RGE evolution driven by the coupling g_3^2 produces the result $f_\tau/\sqrt{f_t^2 + f_b^2} \approx 1/3$, or $m_\tau \approx \frac{1}{3}\sqrt{m_t^2 + m_b^2}$. One requires two doublets to make the τ composite in this model.

2.5.2 More Than One Fermion Generation

The problems of the previous Section appear also in this model when one attempts to include more than one generation of composite fermions. If we stick with only one doublet, we have only one VEV, and the Yukawa coupling combinations $f_{u_i}^2 + f_{d_i}^2$ are attracted to the same fixed point for all generations i, leading to the incorrect relations

$$m_t^2 + m_b^2 = m_c^2 + m_s^2 = m_u^2 + m_d^2. \tag{2.36}$$

We conclude that it is not possible to generate more than two distinct composite fermions in this model with only one doublet. On the other hand, if we introduce too many Higgs doublets (for example, just one more doublet to give mass to composite (c_L, s_L)), we again encounter the problem of flavor-changing neutral currents. The only natural extension of this model is to introduce a second doublet to give mass only to a composite τ, an extension that seems ill-motivated and needlessly complicated.

2.6 Conclusions

We have seen that it is possible to build realistic models of composite left- and right-handed fermions (as well as top-condensate models) with more than one generation present, although it appears to be impossible in this context to construct models in which particles from more than one generation are composite. Furthermore, we have found such models are distinguishable from the top-condensate model and each other at low energies only through their Higgs spectra. The least unnatural composite right-handed model consists of a composite t_R obtaining mass from one Higgs doublet, and composite b_R and τ_R obtaining (distinct) masses from a second doublet. The least unnatural composite left-handed model requires only one Higgs doublet for t_L and b_L, but another is needed if we wish to construct a composite τ. Obviously the discovery of only one Higgs doublet would invalidate the former model, but the discovery of two doublets would mean that we cannot distinguish between the three models at low energies.

All three models, however, seem to suffer from the same problem of an m_t prediction [31, 40] large enough to be in conflict with recent analyses of electroweak precision measurements. One obvious escape from this difficulty in the top-condensate model is to postulate a composite fourth fermion generation [31, 46]. Unfortunately, this skirts the fundamental issue by pushing the problem to a higher, unobserved regime. As before, we must either choose only the fourth-generation fermions to be composite, which denies the initial motivation of the approach, or make both the third and fourth generations composite, and then worry about whether we need to suppress flavor-changing currents between them. A third, and much more optimistic, possibility is to assume that someone will develop a natural way to decrease m_t in such models.

Chapter 3

Meson Mass Splittings in Potential Models

3.1 Introduction

The mass splittings of hadrons in an isospin doublet are often called "electromagnetic splittings" because of the traditional belief that their physical origin lies primarily in the differences of (nonrelativistic) Coulombic expectation values, which are distinguished by the charges of the valence quarks ($Q_u \neq Q_d$). Such an approach may of course be augmented by including hyperfine, spin-orbit, and other well-known interactions. When extensive efforts from the 1950's through the early 1970's [47] failed to produce a reliable model for explaining the $(p-n)$ mass difference in purely electromagnetic terms, the assumption of corrections due to unequal intrinsic masses of the quarks, $m_d - m_u = O(5 \text{ MeV})$, resolved the problem. A model including only this effect and the standard electromagnetic interactions serves to explain the observed splittings (as appearing in the 1992 *Review of Particle Properties* [48]) $(K^0 - K^+) = 4.024 \pm 0.032$ MeV and $(D^+ - D^0) = 4.77 \pm 0.27$ MeV (and even that of the very tightly-bound pions, $(\pi^+ - \pi^0) = 4.5936 \pm 0.0005$ MeV), but has failed in light of the surprisingly small $(B^0 - B^+) = 0.1 \pm 0.8$ MeV.

It is precisely the last mass difference that has led to a variety of calculations. Some of these [49, 50, 51, 52] are based on the nonrelativistic model of

hadron masses put forth by De Rújula, Georgi, and Glashow [53] soon after the development of QCD; such models have the unfortunate tendency to predict numbers no smaller than $B^0 - B^+ \approx 2$ MeV, well outside the current experimental limits. Using more phenomenological models [54, 55], one can obtain a smaller splitting in closer agreement with experiment. Nevertheless, it may seem odd that the usual nonrelativistic model, which works well for the D and even the K mesons, should fail in the case of the B, which boasts an even heavier quark.

In fact, a careful analysis [56] of masses starting from field theory yields the usual Breit–Fermi interaction used in Ref. [53]; however, we show that a number of novel effects appear, owing to the dependence of quantum-mechanical expectation values upon mass, relativistic kinetic energies of the constituent particles, and running of the gauge couplings. We find that, using 1992 numbers, it is possible to explain the mass splittings of heavy mesons (D and B, but not K) in such an ordinary nonrelativistic model with a linear-plus-Coulomb potential, as long as we take into account *all* of these corrections to consistent orders of magnitude. However, as we discuss, more recent measurements from CLEO [57] tend to somewhat alter this numerical conclusion.

In this spirit, the Chapter is organized as follows: In Sec. 3.2 we consider the problem of computing mesonic mass contributions in field theory. Then, in Sec. 3.3, we confirm that the nonrelativistic limit of the field-theoretic result leads to kinematic terms and the Breit–Fermi interaction, exactly as stated in De Rújula *et al.* This is followed in Sec. 3.4 by an exhibition of the full mass splitting relations for isodoublet 0^- and 1^- meson pairs, as well as (0^-,1^-) pairs with the same valence quarks. Sec. 3.5 discusses the application of quantum-mechanical theorems, including a very useful generalized virial theorem, to the problem of reducing the number of independent expectation values in the splitting formulas. These theorems are applied to the popular choice of a linear-plus-Coulomb potential in Sec. 3.6, with numerical results presented in Sec. 3.7. We conclude and comment upon the effects of more recent experimental results in Sec. 3.8.

3.2 Mass Computation in Field Theory

Typically, the computation of hadronic mass splittings in a nonrelativistic model is accomplished by starting with the Breit–Fermi interaction [58, Secs. 38-42]

$$
\begin{aligned}
H_{BF} = \sum_{i>j} & (\alpha Q_i Q_j + k\alpha_s) \\
& \left\{ \frac{1}{|\mathbf{r}_{ij}|} - \frac{1}{2m_i m_j |\mathbf{r}_{ij}|} (\mathbf{p}_i \cdot \mathbf{p}_j + \hat{\mathbf{r}}_{ij} \cdot (\hat{\mathbf{r}}_{ij} \cdot \mathbf{p}_i)\mathbf{p}_j) \right. \\
& -\frac{\pi}{2}\delta^3(\mathbf{r}_{ij}) \left\{ \frac{1}{m_i^2} + \frac{1}{m_j^2} + \frac{4}{m_i m_j} \left[\frac{4}{3}\mathbf{s}_i \cdot \mathbf{s}_j + \left(\frac{3}{4} + \mathbf{s}_i \cdot \mathbf{s}_j \right) \delta_{q_i \overline{q}_j} \right] \right\} \\
& -\frac{1}{2|\mathbf{r}_{ij}|^3} \left[\frac{1}{m_i^2} \mathbf{r}_{ij} \times \mathbf{p}_i \cdot \mathbf{s}_i - \frac{1}{m_j^2} \mathbf{r}_{ij} \times \mathbf{p}_j \cdot \mathbf{s}_j \right. \\
& \qquad + \frac{2}{m_i m_j} (\mathbf{r}_{ij} \times \mathbf{p}_i \cdot \mathbf{s}_j - \mathbf{r}_{ij} \times \mathbf{p}_j \cdot \mathbf{s}_i \\
& \qquad\qquad \left. \left. + 3(\mathbf{s}_i \cdot \hat{\mathbf{r}}_{ij})(\mathbf{s}_j \cdot \hat{\mathbf{r}}_{ij}) - \mathbf{s}_i \cdot \mathbf{s}_j) \right] \right\} ,
\end{aligned}
\tag{3.1}
$$

where $\mathbf{r}_i$, $\mathbf{p}_i$, m_i, $\mathbf{s}_i$, and Q_i denote the coordinate, momentum, (constituent) mass, spin, and charge (in units of the protonic charge) of the ith quark, respectively; $\mathbf{r}_{ij} \equiv \mathbf{r}_i - \mathbf{r}_j$; α and α_s are the (running) QED and QCD coupling constants; and $k = -\frac{4}{3}(-\frac{2}{3})$ is a color binding factor for mesons (baryons). This expression includes an annihilation term if $q_i = \overline{q}_j$ are in a relative $j = 1$ state. From this expression, one chooses the terms that are considered significant and then calculates the appropriate quantum-mechanical expectation values. We pursue this course of action in the next Section; however, to be confident that no novel interactions arise when obtaining this expression from the more fundamental field theories of QED and QCD, we perform a detailed derivation for the meson case.

We first consider the question of the mass of a composite system from the point of view of the S-matrix and interaction-picture perturbation theory. The mass of a system, defined as the expectation value of the total Hamiltonian in the center-of-momentum frame of the constituents, receives contributions from both the noninteracting and interacting pieces of the Hamiltonian; the former gives rise to the masses and kinetic energies of the constituents, and the latter produces the interaction energy. Technically, this division is not exact in the interaction

picture because the noninteracting Hamiltonian is not necessarily the same as the noninteracting Hamiltonian in Schrödinger or Heisenberg pictures; they are different by perturbations in the interactions that serve partially to "dress" the states in the interaction picture, a topic to which we return momentarily.

Let us follow the method of Gupta [59] to derive the form of the interaction from the field-theoretical interaction Hamiltonian. We begin by writing the S matrix in the Cayley form

$$S = \frac{1 - \frac{1}{2}iK}{1 + \frac{1}{2}iK}, \tag{3.2}$$

and expand the Hermitian operator K in a perturbative series:

$$K = \sum_n K_n. \tag{3.3}$$

The purpose of this expansion, rather than the usual expansion of S, is to preserve unitarity in each partial sum of S. The physical effect of this parametrization is to eliminate diagrams with real intermediate states from the S-matrix expansion.

Computing the terms K_n, one finds

$$K_1 = \int_{-\infty}^{+\infty} dt\, H^I_{\text{int}}(t), \tag{3.4}$$

where I indicates the interaction picture. Now observe that we may invent an effective Hamiltonian, H^I_{eff}, such that its first-order contribution is equivalent to the contribution from H^I_{int} *to all orders.* Thus,

$$K = \int_{-\infty}^{+\infty} dt\, H^I_{\text{eff}}(t). \tag{3.5}$$

The interaction energy is then

$$\Delta E = \frac{\left\langle f^I \middle| H^I_{\text{eff}}(0) \middle| i^I \right\rangle}{\langle f^I | i^I \rangle}, \tag{3.6}$$

where $\left|i^I\right\rangle$ and $\left|f^I\right\rangle$ are actually the same state because the system is stable.

For the case of quark-antiquark interactions via QED and QCD, the lowest-order contribution to K comes through two interaction vertices, *i.e.*, the exchange

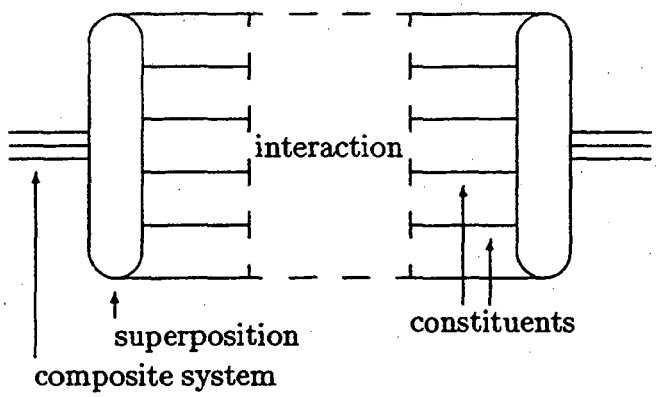


Fig. 3.1

Figure 3.1: Diagrammatical representation of $\mathcal{M}_{fi}$.

of one vector boson. The equivalent portions of H^I_{eff} and K are thus labeled with a 2, and in terms of the usual invariant amplitude $\mathcal{M}$, we have

$$K_2 = iS_2 = (2\pi)^4\delta^4(P_f - P_i)\mathcal{M}^{(2)}_{fi} = \left\langle f^I \middle| \int dt\, H^{I(2)}_{\text{eff}}(t) \middle| i^I \right\rangle . \tag{3.7}$$

Eliminating the delta functions that arise in the rightmost expressions, we find

$$\Delta E^{(2)} = \frac{\left\langle f^I \middle| H^{I(2)}_{\text{eff}}(0) \middle| i^I \right\rangle}{\langle f^I | i^I \rangle} = \mathcal{M}^{(2)}_{fi}. \tag{3.8}$$

Beyond second order the relation between interaction energy and the invariant amplitude becomes less trivial, but nevertheless Gupta has shown that it can be found. However, we do not continue to fourth-order in this work, and henceforth suppress the (2) in the following.

In general, $\mathcal{M}_{fi}$ at any given order is represented by diagrams of the form indicated in Fig. 3.1. The composite state is formed by superposition of the constituent particle wavefunctions in such a manner that one obtains the desired overall quantum numbers. For the mesonic system, $\mathcal{M}_{fi}$ is represented by the diagram in Fig. 3.2, where the lowest-order interaction is the exchange of a single gauge boson. This class of diagrams allows for only the valence quark and antiquark (no sea $q\bar{q}$ pairs or glue), and thus would induce a poor model if we chose them to be current quarks. Instead, the quarks in our diagrams must be constituent

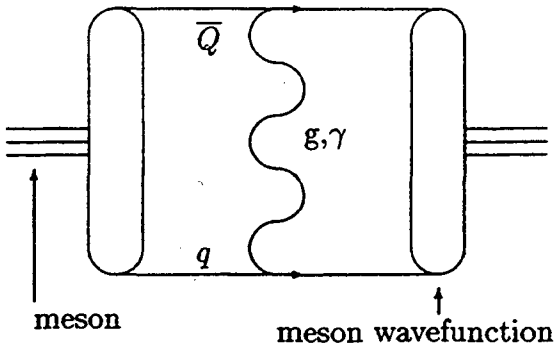


Fig. 3.2

Figure 3.2: Diagram for $\mathcal{M}_{fi}$ in the mesonic system.

quarks, whereas the loop effects involving the vector boson serve to renormalize the gauge couplings to their running values. In this way we can model the hadronic cloud, as well as renormalizations of the lines and vertices of our diagram, so that its particles are "dressed" in two senses. There is also an annihilation diagram if the quark and antiquark are of the same flavor; in this work we consider only the exchange diagram, since the mesons of greatest interest to us are those with one heavy and one light quark.

Is it legitimate to factorize the meson wavefunction into two distinct and weakly-interacting constituent quark clouds? Similar issues were discussed in the Introduction to Chapter 1; there we argued that every quark that is heavy compared to the QCD scale approximately decouples from the light degrees of freedom, so the assumption of a perturbative coupling to the heavy quark seems to be reasonable. The scheme could be greatly improved, in principle, by a better understanding of the interaction with the light constituent degrees of freedom.

The next step is to obtain the amplitude $\mathcal{M}_{fi}$, in which the constituent legs are bound in the composite system, from the Feynman amplitude $\mathcal{M}$ (Fig. 3.3) for the same interaction with *free* external constituent legs. To do this, we need only constrain the free external legs in a manner that reflects the wavefunction and rotational properties of the meson state. In general, if the variables z_n are the

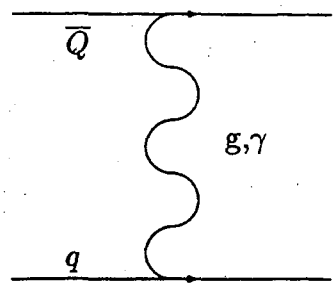


Fig. 3.3

Figure 3.3: Free quark Feynman amplitude $\mathcal{M}$.

degrees of freedom of the meson state $|\Phi\rangle$, then we may write

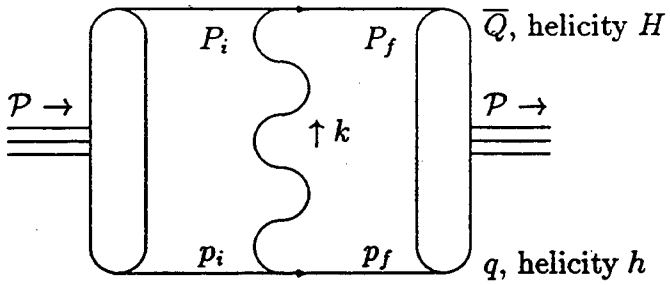
$$|\Phi\rangle = \sum\!\!\!\!\!\!\int dz_n \, \phi(z_n) \, \mathcal{O}(z_n) \, |0\rangle \, . \tag{3.9}$$

The function ϕ is an amplitude in the variables z_n, *i.e.*, a wavefunction; and $\mathcal{O}$ is a collection of Fock space operators that specifies the rotational properties of $|\Phi\rangle$. The integral-sum symbol indicates summation over both continuous and discrete z_n. In this notation, we obtain the result

$$\Delta E = \sum\!\!\!\!\!\!\int dz_f \, dz_i \, \phi^*(z_f)\phi(z_i) f(z_i, z_f) \mathcal{M}(z_i, z_f), \tag{3.10}$$

where $f(z_i, z_f) \equiv \langle 0| \, \mathcal{O}^\dagger(z_f) \mathcal{H} \mathcal{O}(z_i) \, |0\rangle$; here $\mathcal{H}$, the Fock-space operators from the Hamiltonian, serve to constrain the quantum numbers of the composite system. We have written the energy contribution in this very general expression in order to demonstrate the power of the technique.

Now we apply this prescription in detail to the meson of Fig. 3.2, so that we may make use of the usual Feynman rules. Then z_n are quark momenta, ϕ is the mesonic momentum-space wavefunction, and f specifies the spin of the meson, as we see below. The energy contribution is evaluated in the quark center-of-momentum frame (*i.e.*, the meson rest frame); in this frame the momentum of the quark relative to rest, initially and finally, is denoted by $\mathbf{p}$ and $\mathbf{p}'$, respectively, with opposite signs



$$\mathbf{p}_i = \frac{\mathcal{P}}{2} + \mathbf{p} \quad p_i^0 \equiv \varepsilon_i$$
$$\mathbf{P}_i = \frac{\mathcal{P}}{2} - \mathbf{p} \quad P_i^0 \equiv E_i$$
$$\mathbf{p}_f = \frac{\mathcal{P}}{2} + \mathbf{p}' \quad p_f^0 \equiv \varepsilon_f$$
$$\mathbf{P}_f = \frac{\mathcal{P}}{2} - \mathbf{p}' \quad P_f^0 \equiv E_f$$
$$p_i^2 = p_f^2 \equiv m^2$$
$$P_i^2 = P_f^2 \equiv M^2$$

Fig. 3.4

Figure 3.4: Notation and conventions for the mesonic system.

for the antiquark. Fourier transformation of the wavefunctions from momentum space to position space yields

$$\Delta E_{CM} = \int d^3\mathbf{x}_f \int d^3\mathbf{x}_i \, \psi^*(\mathbf{x}_f) K(\mathbf{x}_f, \mathbf{x}_i) \psi(\mathbf{x}_i), \tag{3.11}$$

where

$$K(\mathbf{x}_f, \mathbf{x}_i) = \int d^3\mathbf{p}' \int d^3\mathbf{p} \exp[i(\mathbf{p}' \cdot \mathbf{x}_f - \mathbf{p} \cdot \mathbf{x}_i)] \sum_{\text{spins}} f(\text{spins}) \mathcal{M}(\mathbf{p}', \mathbf{p}, \text{spins}), \tag{3.12}$$

and

$$\int d^3\mathbf{x} \, \psi^*(\mathbf{x}) \psi(\mathbf{x}) = 1. \tag{3.13}$$

As a technical point of fact, it is necessary to keep track of the normalization conventions used for wavefunctions, Fourier transforms, and Feynman rules in order to obtain the true convention-independent ΔE. As it stands, Eq. 3.12 locks us into a particular set of Feynman rule normalizations, which should be made clear in the following expression. The kinematic conventions are established in Fig. 3.4. (Note, however, that $\mathbf{p}$ and $\mathbf{p}'$ are different vectors in different frames, owing to the nonlinear nature of Lorentz boosts.) Then the Feynman amplitude for free external

quark legs and a virtual photon is

$$\mathcal{M} = i\left[\frac{1}{(2\pi)^{3/2}}\right]^4 \sqrt{\frac{M}{E_f}}\sqrt{\frac{M}{E_i}}\sqrt{\frac{m}{\varepsilon_f}}\sqrt{\frac{m}{\varepsilon_i}}$$
$$\left[\overline{v}_{H_i}(\mathbf{P}_i)(-iQe\gamma_\mu)v_{H_f}(\mathbf{P_f})\right]\left(\frac{-ig^{\mu\nu}}{k^2}\right)\left[\overline{u}_{h_f}(\mathbf{p}_f)(-iqe\gamma_\mu)u_{h_i}(\mathbf{p}_i)\right], \quad (3.14)$$

with Qqe^2 replaced by g_s^2 and $SU(3)$ generators for the gluon-mediated diagram. Note the use of helicity rather than spin eigenstate spinors, which is done in order to implement a relativistic description of the mesons. In a nonrelativistic picture in which meson spin originates solely from the spin of the quarks (s waves), spin-0(1) mesons have spin-space wavefunctions described by the usual singlet and triplet quark wavefunction $\overline{Q}q$ combinations:

$$\frac{\overline{Q}_\uparrow q_\downarrow \pm \overline{Q}_\downarrow q_\uparrow}{\sqrt{2}}, \quad \uparrow,\downarrow \text{ spins.} \quad (3.15)$$

The above expression remains true in a relativistic picture if we take the initial and final spin-quantization axes to coincide with the axes of relative momenta $\mathbf{p}$ and $\mathbf{p}'$, respectively, and then take $\uparrow,\downarrow$ as *helicity* eigenstates; this is nothing more than the simplest nontrivial case of the Jacob–Wick formalism [60]. It is then a simple matter to write the constraint function for singlet (triplet) mesons:

$$f(\text{hel.}) = \frac{1}{\sqrt{2}}(\delta_{h_i\uparrow}\delta_{H_i\downarrow} \pm \delta_{h_i\downarrow}\delta_{H_i\uparrow})\,\frac{1}{\sqrt{2}}(\delta_{h_f\uparrow}\delta_{H_f\downarrow} \pm \delta_{h_f\downarrow}\delta_{H_f\uparrow}), \quad (3.16)$$

with an additional factor of $-\frac{4}{3}$ for the gluon-mediated diagram, which arises from the constraint that the initial and final $q\overline{q}$ pairs are combined into a color singlet. The relevant expression is the constrained matrix element $\mathcal{M}_{\text{sing}}$ or $\mathcal{M}_{\text{trip}}$, which is the Feynman amplitude multiplied by the constraint function and summed over spins (or helicities); this is the object being Fourier transformed in Eq. 3.12.

In summary, mass contributions due to a binding interaction in a system of particles may be computed by writing down the Feynman amplitude induced by the interaction Hamiltonian, constraining the component particles to satisfy the symmetry properties of the system, and convolving with the appropriate system wavefunction. The specific implementation of this technique to spin-0 and spin-1 mesons with constituent quarks in a relative $\ell = 0$ state is described by Eqs. 3.12, 3.14, and 3.16.

3.3 The Nonrelativistic Limit

With the method for computing mass contributions in hand, we find ourselves with two possible courses of action. The first is to compute $\mathcal{M}_{\text{sing}}$ or $\mathcal{M}_{\text{trip}}$ from Eq. 3.14 without approximation, and then Fourier transform the result to obtain ΔE_{CM}. The second is to reduce immediately the spinor bilinears via Pauli approximants, thus producing a nonrelativistic expansion. Let us explore both directions for the pseudoscalar case; the vector case is not much different.

Even though the amplitude $\mathcal{M}$ itself is Lorentz invariant, we must evaluate it in the CM frame of the quarks in order to evaluate ΔE. It is convenient to eliminate spinors from the calculation by means of relations like

$$\sum_h u_h(\mathbf{p}_A)\overline{u}_h(\mathbf{p}_B) = \frac{(m_A + \not{p}_A)}{\sqrt{2m_A(E_A + m_A)}} \frac{(1+\gamma_0)}{2} \frac{(m_B + \not{p}_B)}{\sqrt{2m_B(E_B + m_B)}}, \tag{3.17}$$

Once the spinor reductions and the resultant trace are performed, we find the expression

$$\mathcal{M}_{\text{sing}} = -(Qqe^2 - \frac{4}{3}g_s^2)\mathcal{N}\mathcal{T}\frac{1}{k^2}, \tag{3.18}$$

where $\mathcal{N}$ results from the normalization factors, and $\mathcal{T}$ is the gamma-matrix trace. They are given by

$$\mathcal{N} = \frac{1}{(2\pi)^6}\frac{1}{2^5}\left[E_i(E_i+M)E_f(E_f+M)\varepsilon_i(\varepsilon_i+m)\varepsilon_f(\varepsilon_f+m)\right]^{-1/2}$$

and

$$\begin{aligned}
\mathcal{T} = {} & 8\,\{(p_i\cdot P_i)\,[2\varepsilon_f E_f + 3(mE_f + M\varepsilon_f + mM)] \\
& +(p_f\cdot P_f)\,[2\varepsilon_i E_i + 3(mE_i + M\varepsilon_i + mM)] + (p_i\cdot P_i)(p_f\cdot P_f) \\
& -(p_i\cdot p_f)\,[2E_iE_f + M(E_i + E_f + M)] \\
& -(P_i\cdot P_f)\,[2\varepsilon_i\varepsilon_f + m(\varepsilon_i + \varepsilon_f + m)] + (p_i\cdot p_f)(P_i\cdot P_f) \\
& -(p_i\cdot P_f)\,[mE_i + M\varepsilon_f + mM] \\
& -(P_i\cdot p_f)\,[mE_f + M\varepsilon_i + mM] - (p_i\cdot P_f)(P_i\cdot p_f) \\
& +[-2mM(E_i - E_f)(\varepsilon_i - \varepsilon_f) \\
& \quad +mM\,(m(E_i + E_f) + M(\varepsilon_i + \varepsilon_f) + mM) \\
& \quad +2m^2E_iE_f + 2M^2\varepsilon_i\varepsilon_f]\}.
\end{aligned} \tag{3.19}$$

Also,

$$k^2 = (p_i - p_f)^2 = (\varepsilon_i - \varepsilon_f)^2 - (\mathbf{p} - \mathbf{p}')^2. \tag{3.20}$$

It is, in principle, possible to Fourier-transform the product $\mathcal{M}_{\text{sing}}$ of these unwieldy functions to obtain the full relativistic result for ΔE_{CM}; this has not yet been performed. We can also perform the expansion of the energy factors in powers of (p/m), where all such momentum-over-mass quotients that occur are taken to be of the same order.

However, this is unnecessary work if we require only a nonrelativistic expansion; in this case there is a much faster way, namely expansion of the spinor bilinears via the Pauli approximants

$$\begin{aligned}
\overline{u}(\mathbf{p}')\gamma u(\mathbf{p}) &= \langle\chi'|\frac{(\mathbf{p}+\mathbf{p}')}{2m} + i\frac{\boldsymbol{\sigma}\times(\mathbf{p}'-\mathbf{p})}{2m}|\chi\rangle + O\left[\left(\frac{p}{m}\right)^3\right] \\
\overline{u}(\mathbf{p}')\gamma^0 u(\mathbf{p}) &= \langle\chi'|1 + \frac{(\mathbf{p}+\mathbf{p}')^2}{8m^2} + i\frac{\boldsymbol{\sigma}\cdot(\mathbf{p}'\times\mathbf{p})}{4m^2}|\chi\rangle + O\left[\left(\frac{p}{m}\right)^4\right].
\end{aligned} \tag{3.21}$$

Using these expansions in Eq. 3.14 and taking $|\chi\rangle, |\chi'\rangle$ in helicity basis, we quickly find

$$\begin{aligned}
\mathcal{M}_{\text{sing}} \overset{NR}{\Longrightarrow} \frac{Qqe^2 - \frac{4}{3}g_s^2}{(2\pi)^6(\mathbf{p}-\mathbf{p}')^2}\Bigg\{1 + \frac{(\mathbf{p}+\mathbf{p}')^2}{4mM} - \frac{(\mathbf{p}-\mathbf{p}')^2}{8}\left[\frac{1}{m^2} - \frac{4}{mM} + \frac{1}{M^2}\right] \\
+O\left[\left(\frac{p}{m}\right)^4\right]\Bigg\}.
\end{aligned} \tag{3.22}$$

Then Fourier transformation of this result produces

$$\begin{aligned}
\Delta E_{CM,\text{sing}} &= \left(\alpha Qq - \frac{4}{3}\alpha_s\right) \times \\
&\quad \left\{\left\langle\frac{1}{r}\right\rangle + \frac{1}{2mM}\left\langle\frac{1}{r}(\mathbf{p}^2 + \hat{\mathbf{r}}\cdot(\hat{\mathbf{r}}\cdot\mathbf{p})\mathbf{p})\right\rangle\right. \\
&\quad \left. -\frac{\pi}{2}\left(\frac{1}{m^2} - \frac{4}{mM} + \frac{1}{M^2}\right)\left\langle\delta^3(\mathbf{r})\right\rangle\right\} + \cdots
\end{aligned} \tag{3.23}$$

The expectation values cannot be uniquely evaluated until we choose a basis of energy eigenfunctions, which is equivalent to choosing a potential for the Schrödinger equation; we return to this topic momentarily. In comparison to $\Delta E_{CM,\text{sing}}$, the energy contribution from the Breit–Fermi interaction (Eq. 3.1) for a quark-antiquark

pair of masses m, M in the CM reduces to

$$\begin{aligned}\langle H_{BF}\rangle = &\left(\alpha Qq - \frac{4}{3}\alpha_s\right)\times \\ &\left\{\left\langle\frac{1}{r}\right\rangle + \frac{1}{2mM}\left\langle\frac{1}{r}(\mathbf{p}^2 + \hat{\mathbf{r}}\cdot(\hat{\mathbf{r}}\cdot\mathbf{p})\mathbf{p})\right\rangle\right. \\ &-\frac{\pi}{2}\left\langle\delta^3(\mathbf{r})\right\rangle\left[\frac{1}{m^2} + \frac{1}{M^2} + \frac{4}{mM}\left(\mathcal{G} + \delta_{S,1}\delta_{q\bar{Q}}\right)\right] \\ &\left.-\frac{1}{2}\left\langle\frac{1}{r^3}\right\rangle\left\langle\mathbf{L}\cdot\left(\frac{\mathbf{s}_q}{m^2} + \frac{\mathbf{s}_{\bar{Q}}}{M^2} + \frac{2\mathbf{S}}{mM}\right) + \frac{S_{12}}{2mM}\right\rangle\right\}\end{aligned} \tag{3.24}$$

where $\mathcal{G} \equiv \frac{4}{3}\left\langle\mathbf{s}_q\cdot\mathbf{s}_{\bar{Q}}\right\rangle$, which is $-1\left(\frac{1}{3}\right)$ for $S = 0(1)$. Also, $\mathbf{S} \equiv \mathbf{s}_q + \mathbf{s}_{\bar{Q}}$, and S_{12} is the $\Delta L = 2$ tensor operator

$$S_{12} \equiv 3(\boldsymbol{\sigma}_1\cdot\hat{\mathbf{r}})(\boldsymbol{\sigma}_2\cdot\hat{\mathbf{r}}) - \boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2. \tag{3.25}$$

For mesons with differently-flavored quarks in a relative $\ell = 0$ state, many of the terms drop out. Let us define

$$\begin{aligned} B &\equiv \left\langle\frac{1}{r}\right\rangle, \\ C &\equiv \left\langle\frac{1}{r}(\mathbf{p}^2 + \hat{\mathbf{r}}\cdot(\hat{\mathbf{r}}\cdot\mathbf{p})\mathbf{p})\right\rangle, \\ D &\equiv \left\langle\delta^3(\mathbf{r})\right\rangle. \end{aligned} \tag{3.26}$$

Then Eq. 3.24 becomes

$$\langle H_{BF}\rangle = \left(\alpha Qq - \frac{4}{3}\alpha_s\right)\left[B + \frac{1}{2mM}C - \frac{\pi}{2}\left(\frac{1}{m^2} + \frac{1}{M^2} + \frac{4\mathcal{G}}{mM}\right)D\right], \tag{3.27}$$

and this is exactly Eq. 3.23 where $\mathcal{G} = -1$.

We have been up to now considering only the contributions to the mass originating from the binding interaction due to one-gluon and one-photon exchanges; there are, of course, also contributions from the kinetic energy (K) of the quarks. Were we calculating these quantities in a relativistic theory, we would simply compute $K = \left\langle\sqrt{m^2 + \mathbf{p}^2}\right\rangle$. The square root may be formally expanded in nonrelativistic quantum mechanics (NRQM) as well, resulting in an alternating series in $\langle\mathbf{p}^{2n}\rangle$. However, for large enough n in NRQM, these expectation values tend to diverge. For example, in the hydrogen atom, divergence occurs for s waves at $n = 3$. Furthermore, if the system is not highly nonrelativistic, the inclusion of the $\langle\mathbf{p}^4\rangle$ term

may cause us to underestimate grossly the true value of the kinetic energy. The problem is that there is no positive $\langle \mathbf{p}^6 \rangle$ term to balance the large negative $\langle \mathbf{p}^4 \rangle$ term. For these reasons, we incorporate the alternating nature of the series in a computationally simple way by making the *Ansatz*

$$K = \sqrt{m^2 + \langle \mathbf{p}^2 \rangle}. \tag{3.28}$$

In order to evaluate the expectation values in the above equations, we need to choose a potential. In the meantime, let us simply denote it with $U(r)$. Then at last we have the mass formula:

$$M_{\text{meson}} = \sqrt{M^2 + \langle \mathbf{p}^2 \rangle} + \sqrt{m^2 + \langle \mathbf{p}^2 \rangle} + \langle U(r) \rangle + \langle H_{BF} \rangle \,. \tag{3.29}$$

The static potential $U(r)$ takes the place of L, the universal quark binding function, in Eq. 1 of Ref. [53].

3.4 Mass Splitting Formulas

The static potential in which the quarks interact determines the form of the NRQM wavefunction. The strong "Coulombic" (static one-gluon exchange) term gives the largest energy contribution of terms within the Breit–Fermi interaction, and therefore would also be expected to alter substantially the wavefunction in perturbation theory; we thus include this term in the static potential:

$$V(r) \equiv U(r) - \frac{4}{3}\frac{\alpha_s}{r}. \tag{3.30}$$

Then the mass formula Eq. 3.29 becomes, using Eq. 3.27,

$$\begin{aligned} M_{\text{meson}} &= \sqrt{M^2 + \langle \mathbf{p}^2 \rangle} + \sqrt{m^2 + \langle \mathbf{p}^2 \rangle} + \langle V(r) \rangle + \alpha Q q B \\ &\quad + \left(\alpha Q q - \frac{4}{3}\alpha_s\right)\left[\frac{1}{2mM}C - \frac{\pi}{2}\left(\frac{1}{m^2} + \frac{1}{M^2} + \frac{4\mathcal{G}}{mM}\right)D\right], \end{aligned} \tag{3.31}$$

where the expectation values are now evaluated as integrals over solutions to the Schrödinger equation with potential $V(r)$, not $U(r)$ as in the previous Section.

Now at last we are in a position to write explicit formulas for the mass splittings of interest. Denoting the mass of a meson of spin S and valence quarks $\overline{Q}$,q as $M^S(\overline{Q}q)$, we define

$$\begin{aligned}
\Delta_Q^0 &\equiv M^0(\overline{Q}u) - M^0(\overline{Q}d) \\
\Delta_Q^1 &\equiv M^1(\overline{Q}u) - M^1(\overline{Q}d) \\
\Delta_{Qu}^* &\equiv M^1(\overline{Q}u) - M^0(\overline{Q}u) \\
\Delta_{Qd}^* &\equiv M^1(\overline{Q}d) - M^0(\overline{Q}d),
\end{aligned} \tag{3.32}$$

where u and d, the up and down constituent quarks, are nearly degenerate in mass: Defining $\Delta m \equiv m_u - m_d$ and $m \equiv (m_u + m_d)/2$, we have $|\Delta m/m| \ll 1$. Therefore, the differences in Eq. 3.32 are expanded in Taylor series in $(\Delta m/m)$ about m. It is also convenient to define

$$\begin{aligned}
A &\equiv \left\langle \mathbf{p}^2 \right\rangle, \\
\beta &\equiv \frac{1}{1+m/M}, \\
\mu &\equiv \text{usual reduced mass}, \\
\bar{\mu} &\equiv m\beta, \\
D_{\alpha_s} &\equiv \beta \left(\frac{\mu}{\alpha_s} \frac{\partial \alpha_s}{\partial \mu} \right) \bigg|_{\mu=\bar{\mu}}, \\
D_X &\equiv \beta \left(\mu \frac{\partial X}{\partial \mu} \right) \bigg|_{\mu=\bar{\mu}}, \quad X = A, B, C, D, \langle V(r) \rangle .
\end{aligned} \tag{3.33}$$

Then the expressions for isospin mass splitting are

$$\begin{aligned}
\Delta_Q^{0,1} = & \left[\frac{2m^2 + D_A}{\sqrt{m^2 + A}} + \frac{D_A}{\sqrt{M^2 + A}} \right] \frac{\Delta m}{2m} + D_{\langle V \rangle} \frac{\Delta m}{m} \\
& - \frac{4}{3} \alpha_s \Delta m \left\{ \frac{1}{2m^2 M} \left(D_C - C + C D_{\alpha_s} \right) \right. \\
& \left. - \frac{\pi}{2m^3} \left[\left(1 + 4\mathcal{G} \frac{m}{M} + \frac{m^2}{M^2} \right) \left(D_D + D D_{\alpha_s} \right) - 2 \left(1 + 2\mathcal{G} \frac{m}{M} \right) D \right] \right\} \\
& + \alpha Q \left[B + \frac{1}{2mM} C - \frac{\pi}{2m^2} \left(1 + 4\mathcal{G} \frac{m}{M} + \frac{m^2}{M^2} \right) D \right] \\
& + O\left[\left(\frac{\Delta m}{m} \right)^3 \right] + O\left(\alpha \frac{\Delta m}{m} \right).
\end{aligned} \tag{3.34}$$

Note that no derivatives appear in the α_{EM} terms because we take both α_{EM} and $(\Delta m/m)$ (but not α_s) as expansion parameters. Furthermore, the running of $\alpha_s(\mu)$ is explicitly taken into account.

For vector-pseudoscalar splittings, we have

$$\begin{aligned}\Delta^*_{Qq} &= \frac{8\pi}{3mM}\left[\left(\frac{4}{3}\alpha_s - \alpha Qq\right) D \pm \frac{4}{3}\alpha_s \frac{\Delta m}{2m}\left(D_D + DD_{\alpha_s} - D\right)\right] \\ &\quad + O\left[\left(\frac{\Delta m}{m}\right)^2\right] + O\left(\alpha\frac{\Delta m}{m}\right), \quad \text{with } \pm \text{ for } q = u(d). \end{aligned} \tag{3.35}$$

Let us remind ourselves of the physical significance of the terms in the previous two equations. Terms containing A signify kinetic energy contributions, including intrinsic quark masses. The potential term is identified, of course, by V; B, C, and D denote static Coulomb, Darwin, and (generalized) hyperfine terms, respectively.

3.5 Quantum-mechanical Theorems

In order to apply the preceding results, we need to evaluate the expectation values A,B,C,D, and $\langle V(r)\rangle$ for a chosen potential $V(r)$. Following Quigg and Rosner [61], we present two quantum-mechanical theorems that make the evaluation of these expectation values and their mass derivatives simpler.

Theorem 3.1 (Feynman–Hellmann theorem) *For normalized eigenstates of a Hamiltonian depending on a parameter λ,*

$$\frac{\partial E}{\partial \lambda} = \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle. \tag{3.36}$$

In the particular case that $\lambda = \mu$,

$$\frac{\partial E}{\partial \mu} = -\frac{1}{\mu}\left[E - \langle V(r)\rangle\right] + \left\langle \frac{\partial V}{\partial \mu} \right\rangle. \tag{3.37}$$

The other result may be less familiar. For reasons that will become clear, let us call it the *generalized virial theorem* (GVT).

Theorem 3.2 (Generalized virial theorem) *Consider bound eigenstates $u_\ell(r)$ in a spherically symmetric potential $V(r)$ such that*

$$\lim_{r\to 0} r^2 V(r) = 0.$$

Then, writing the Schrödinger equation as

$$u_\ell''(r) + \frac{2\mu}{\hbar^2}\left[E - V(r) - \frac{\hbar^2\ell(\ell+1)}{2\mu r^2}\right]u_\ell(r) = 0,$$

and defining a_ℓ by

$$\lim_{r\to 0}\frac{u_\ell(r)}{r^{\ell+1}} \equiv a_\ell,$$

then

i) a_ℓ is a nonzero constant;

ii) for $q \geq -2\ell$,

$$\begin{aligned}(2\ell+1)^2 a_\ell^2 \delta_{q,-2\ell} &= -\frac{2\mu}{\hbar^2}\left\langle r^{q-1}\left(2q[E-V(r)] - r\frac{dV}{dr}\right)\right\rangle \\ &\quad +(q-1)\left[2\ell(\ell+1) - \frac{1}{2}q(q-2)\right]\left\langle r^{q-3}\right\rangle. \qquad (3.38)\end{aligned}$$

Clearly this theorem proves most useful for potentials easily expressed as polynomials in r. But in fact there are some interesting general results included. For example, the $q = \ell = 0$ case generates the well-known result for s waves,

$$|\Psi(0)|^2 = \frac{\mu}{2\pi\hbar^2}\left\langle\frac{dV}{dr}\right\rangle, \qquad (3.39)$$

whereas the $q = 1$ case produces

$$E - \langle V(r)\rangle = \frac{1}{2}\left\langle r\frac{dV}{dr}\right\rangle, \qquad (3.40)$$

the quantum-mechanical virial theorem.

Using partial integration, the Schrödinger equation, and the GVT, it is possible to show the following ($\hbar = 1$):

$$\begin{aligned} A &= 2\mu[E - \langle V(r)\rangle], \\ C &= 4\mu\left[E\left\langle\frac{1}{r}\right\rangle - \left\langle\frac{V(r)}{r}\right\rangle - \frac{1}{4}\left\langle\frac{dV}{dr}\right\rangle(1+\delta_{\ell,0})\right], \\ D &= \frac{\mu}{2\pi}\left\langle\frac{dV}{dr}\right\rangle\delta_{\ell,0}, \\ &\int_0^\infty\left(\frac{du_\ell(r)}{dr}\right)^2 dr = A - \ell(\ell+1)\left\langle\frac{1}{r^2}\right\rangle. \qquad (3.41)\end{aligned}$$

In addition, we must also uncover what we can about the μ-dependence of expectation values. For a general potential this is actually an unsolved problem; unless the potential has very special μ-dependence, one can show that the only case in which one may scale away all dimensionful parameters in the Schrödinger equation is when $V(r) = V_0 r^\nu$. In that case, the μ-dependence is entirely contained in the scaling factors, and computing D_X is trivial. Unfortunately, for the potential we consider in the next Section, this is not true, and we must resort to subterfuge to obtain the required information.

3.6 Example: $V(r) = r/a^2 - \kappa/r$

The potential $V(r) = r/a^2 - \kappa/r$, where $\kappa = \frac{4}{3}\alpha_s$, is interesting because it phenomenologically includes quark confinement via the linear term. This potential was considered in greatest detail by Eichten *et al.* [62] to describe the mass splitting structure of the charmonium system (and was later applied to bottomonium). The Schrödinger equation was solved numerically, but it is possible to extract a great deal of information from their tabulated results.

This is possible because of the GVT. If we rescale the Schrödinger equation with the linear-plus-Coulomb potential to

$$\left(\frac{d^2}{d\rho^2} - \frac{\ell(\ell+1)}{\rho^2} + \frac{\lambda}{\rho} + \zeta - \rho\right) w_\ell(\rho) = 0, \tag{3.42}$$

where

$$\begin{aligned} \rho &\equiv \left(\tfrac{2\mu}{a^2}\right)^{1/3} r, & \lambda &\equiv \kappa(2\mu a)^{2/3}, \\ \zeta &\equiv (2\mu a^4)^{1/3} E, & w_\ell(\rho) &\equiv u_\ell(r)\left(\tfrac{a^2}{2\mu}\right)^{1/6}, \end{aligned} \tag{3.43}$$

then the GVT gives

$$\begin{aligned} (q=0) \quad a_0^2\,\delta_{0,\ell} &= \left(\frac{2\mu}{a^2}\right)\left[1 + \lambda\left\langle\frac{1}{\rho^2}\right\rangle - 2\ell(\ell+1)\left\langle\frac{1}{\rho^3}\right\rangle\right], \\ (q=1) \qquad 0 &= 3\langle\rho\rangle - 2\zeta - \lambda\left\langle\frac{1}{\rho}\right\rangle. \end{aligned} \tag{3.44}$$

Also, defining

$$\left\langle v^2\right\rangle \equiv \int_0^\infty \left(\frac{dw_\ell(\rho)}{d\rho}\right)^2 d\rho, \tag{3.45}$$

we find

$$\left\langle v^2 \right\rangle = -\langle \rho \rangle + \zeta + \lambda \left\langle \frac{1}{\rho} \right\rangle - \ell(\ell+1) \left\langle \frac{1}{\rho^2} \right\rangle . \tag{3.46}$$

It is a happy accident of this potential that all of the quantities in the expectation values we need, for any ℓ, may be expressed in terms of the three quantities ζ, $\langle 1/\rho^2 \rangle$, and $\langle v^2 \rangle$. These are exactly the values tabulated for the $1s$ state, as functions of λ, in Eichten *et al.* Table I. Defining $\sigma \equiv (2\mu/a^2)^{1/3}$ and taking $\ell = 0$ (for our mesonic model), we find

$$\begin{aligned} A &= \sigma^2 \left\langle v^2 \right\rangle , \\ B &= \frac{\sigma}{2\lambda} \left[3 \left\langle v^2 \right\rangle - \zeta \right] , \\ C &= \sigma^2 \left[2B\zeta + \sigma \left(-3 + \lambda \left\langle \frac{1}{\rho^2} \right\rangle \right) \right] , \\ D &= \frac{\sigma^3}{4\pi} \left[\lambda \left\langle \frac{1}{\rho^2} \right\rangle + 1 \right] . \end{aligned} \tag{3.47}$$

So now we can compute all of the necessary expectation values numerically. The superficial singularity in $B(\lambda = 0)$ is not real: Note that $B(0)$ is just the ground state expectation value $\langle 1/r \rangle$ for a pure linear potential. Then item *i*) in Theorem 3.2 guarantees that the integral around $r = 0$ converges, and the normalization condition of the wavefunction assures convergence of the rest of the integral.

The mass derivatives must be handled in a different fashion. We begin by defining

$$\tilde{D}_\zeta \equiv \mu \frac{\partial \zeta}{\partial \mu}, \quad \tilde{D}_v \equiv \mu \frac{\partial \langle v^2 \rangle}{\partial \mu}, \quad \tilde{D}_\rho \equiv \mu \frac{\partial \langle 1/\rho^2 \rangle}{\partial \mu},$$

and

$$\tilde{D}_{\alpha_s} \equiv \frac{\mu}{\alpha_s} \frac{\partial \alpha_s}{\partial \mu}. \tag{3.48}$$

From the Feynman–Hellmann theorem (Eq. 3.37) we may show

$$\tilde{D}_\zeta = \left(\frac{\zeta}{3} - \left\langle v^2 \right\rangle \right) \left(1 + \frac{3}{2} \tilde{D}_{\alpha_s} \right) . \tag{3.49}$$

As mentioned in the previous Section, scaling of the Schrödinger equation can be accomplished for μ-independent potentials that are monomials. In the case $\lambda = 0$ (a

purely linear potential), the scaling would be perfect, and ζ, $\langle 1/\rho^2 \rangle$, and $\langle v^2 \rangle$ would be μ-independent. In the $\lambda \neq 0$ case, the derivatives must be found numerically. Again, we fortunately have a table of numerical values of the desired expectation values, as functions of $\lambda(\mu)$. We fit the expectation values Y ($= \langle 1/\rho^2 \rangle$, $\langle v^2 \rangle$) to the functional form

$$Y(\lambda) = Y_0 + K\lambda^{n_Y}. \tag{3.50}$$

Then, using Eq. 3.43, we find

$$\tilde{D}_Y = \left(\frac{2}{3} + \tilde{D}_{\alpha_s}\right) n_Y \left(Y - Y_0\right). \tag{3.51}$$

Finally, define

$$\tilde{D}_X \equiv \mu \frac{\partial X}{\partial \mu} \text{ for } X = A, B, C, D, \tag{3.52}$$

so that

$$D_X = \beta\, \tilde{D}_X\Big|_{\mu=\bar{\mu}}. \tag{3.53}$$

Then we find

$$\begin{aligned}
\tilde{D}_A &= \frac{2}{3}A + \sigma^2 \tilde{D}_v, \\
\tilde{D}_B &= \frac{3\sigma}{2\lambda}\tilde{D}_v - \frac{1}{2}B\tilde{D}_{\alpha_s}, \quad (\lambda \neq 0), \\
\tilde{D}_C &= \frac{5}{3}C + 2\sigma^2 \left\{ \left(-\zeta + \tilde{D}_\zeta\right) B + \zeta \tilde{D}_B + \sigma \left[\frac{\lambda}{2} \left(\tilde{D}_\rho + \tilde{D}_{\alpha_s} \left\langle \frac{1}{\rho^2} \right\rangle \right) + 1 \right] \right\}, \\
\tilde{D}_D &= \frac{\sigma^3}{4\pi} \left\{ \lambda \left[\left(\frac{5}{3} + \tilde{D}_{\alpha_s} \right) \left\langle \frac{1}{\rho^2} \right\rangle + \tilde{D}_\rho \right] + 1 \right\}.
\end{aligned} \tag{3.54}$$

In the exceptional case of $\tilde{D}_B$, we simply note that the wave equation may be exactly scaled when $\lambda = 0$, and then we can quickly show that $\tilde{D}_B\Big|_{\lambda=0} = \frac{1}{3}B\Big|_{\lambda=0}$. This provides us with everything we need to produce numerical results.

Before leaving the topic, let us mention that many complications of μ-derivatives of expectation values vanish if the potential itself has the appropriate μ-dependence, for then scaling of the wave equation is possible. For example, one can scale the Schrödinger equation for the potential

$$V(r) = c\mu^2 r - \frac{\kappa}{r}, \tag{3.55}$$

where c is a pure number.

3.7 Numerical Results

The method of obtaining results from the theory requires us to choose several numerical inputs, most of which are believed known to within a few percent. Let us choose the following inputs to the model:

$$m = 340 \text{ MeV}, \quad M_s = 540 \text{ MeV},$$
$$M_c = 1850 \text{ MeV}, \quad M_b = 5200 \text{ MeV}, \quad a = 1.95 \text{ GeV}^{-1}. \tag{3.56}$$

The light quark constituent mass is obtained by assuming that nucleons consist of quarks with Dirac magnetic moments only, which may be added nonrelativistically to provide the full nucleonic magnetic moment. Likewise, the strange quark mass issues from the same considerations applied to strange baryons [53]. The c and b quark masses are simply found by dividing the threshold energy value for charm and bottom mesons by two (however, smaller masses have been predicted using semileptonic decay results in addition to meson masses [63]). The confinement constant is inferred from charmonium levels [62].

One important input not yet mentioned is Δm, the up-down quark mass difference. Traditionally, values of $\Delta m \approx -3$ to -8 MeV have been inferred to account for the isospin mass splittings of the lighter hadrons. Using our model with the inputs listed in Eq. 3.56, we find that the experimental splittings for the D and B mesons (both vector and pseudoscalar) as of 1992 can be satisfied within one standard deviation of experimental uncertainty for values of Δm in the narrow range of -4.05 to -4.10 MeV. In contrast, it is found that for no choice of Δm can one simultaneously fit D- and K-meson data, as was done in the earlier models; this conclusion remains true even with the restrictions of the new CLEO data.

Before exhibiting the quantitative results, let us describe the method by which they are obtained. Once particular inputs for the above variables are chosen, one can compute the various mass splittings for the values of $\lambda \propto \alpha_s$ that occur in Table I of Ref. [62], and in-between values may be interpolated. We then fit vector-pseudoscalar splittings (computed via Eq. 3.35) to the corresponding experimental data (since these numbers have the smallest relative uncertainties of the splittings

Table I: Contributions to mass splittings of heavy mesons: Isospin pairs

	D mesons	B mesons
α_s	0.363	0.312
Source	(MeV)	(MeV)
$\Delta_Q^{0,1}$		
Kinetic energy	-4.109	-3.523
Potential energy	1.057	-1.645
Strong Darwin	-0.834	-0.635
EM Darwin	-0.769	0.147
Static Coulomb	-2.442	1.252
Δ_Q^0		
Strong hyperfine	2.148	4.075
EM hyperfine	0.424	-0.561
Total Δ_Q^0	-4.525	-0.889
Δ_Q^1		
Strong hyperfine	3.683	5.244
EM hyperfine	1.817	-0.825
Total Δ_Q^1	-1.596	0.017

we consider) and thus obtain a value of α_s. For the three systems K, D, and B, we use the three values of α_s to estimate graphically (and, admittedly, rather crudely) its mass derivative. Applying the values of the strong coupling constant and its derivative to the splittings in Eq. 3.34, we generate all of the other values. If the resultant numbers do not fall within the experimental uncertainties for such splittings, we vary the input parameters (most importantly, Δm) until a simultaneous fit is achieved.

Tables I and II display the various contributions to mass splittings derived in this fashion for B and D mesons. Although the kinetic term (which includes the explicit difference Δm) and the static Coulomb term are expectedly large, a significant contribution to the mass splitting arises in the strong hyperfine term. That strong contributions to the so-called electromagnetic mass splittings could be important was observed by Chan [50], and was exploited in the subsequent literature. It is exactly this term that is most significant in driving the B splittings

Table II: Contributions to mass splittings of heavy mesons: $(1^-, 0^-)$ pairs

	D mesons	B mesons
α_s	0.363	0.312
Source	(MeV)	(MeV)
$\Delta^*_{Qu,Qd}$		
Strong hyperfine		
(leading)	141.30	46.04
(subleading)	$\pm$ 0.77	$\pm$ 0.58
Δ^*_{Qu}		
EM hyperfine	0.93	-0.18
Total Δ^*_{Qu}	143.00	46.45
Δ^*_{Qd}		
EM hyperfine	-0.46	0.09
Total Δ^*_{Qd}	140.07	45.54

toward zero. Note also the decrease in the derived value of α_s as the reduced mass of the system increases when we move from the D system to the B system, consistent with asymptotic freedom in QCD. It was this running that motivated the inclusion of mass derivatives of the strong coupling constant in this model. If they are not included, one actually obtains a value of $\Delta m > 0$, in contrast with all estimates from both nonrelativistic and chiral models.

The net result is that, using 1992 data, one can satisfactorily fit the data for the D and B systems simultaneously in the most natural nonrelativistic model with a physically reasonable potential. The comparison of the results of this calculation for $\Delta m = -4.10$ MeV to experimental data is presented in Table III. (Note, however, comments on effects of new data in the following Section.)

However, the table also exhibits very poor agreement for the K system (despite the fact that the fit to vector-pseudoscalar splittings yields the value $\alpha_s = 0.424$, which runs in the correct direction). One may view this as a failure of the nonrelativistic assumptions of the model in a variety of ways. Most obvious are the *Ansatz* Eq. 3.28, which is certainly not an airtight assumption in even

Table III: Meson mass splittings compared to experiment

Mass splitting	Notation	Pred. (MeV)	Expt. (MeV)
$K^+ - K^0$	Δ_s^0	-0.98	-4.024 ± 0.032
$K^{*+} - K^{*0}$	Δ_s^1	-0.15	-4.51 ± 0.37[a]
$K^{*+} - K^+$	Δ_{su}^*	398.6	397.94 ± 0.24[a]
$K^{*0} - K^0$	Δ_{sd}^*	397.8	398.43 ± 0.28[a]
$D^0 - D^+$	Δ_c^0	-4.53	-4.77 ± 0.27
$D^{*0} - D^{*+}$	Δ_c^1	-1.60	-2.9 ± 1.3
$D^{*0} - D^0$	Δ_{cu}^*	143.0	142.5 ± 1.3
$D^{*+} - D^+$	Δ_{cd}^*	140.1	140.6 ± 1.9[a]
$B^+ - B^0$	Δ_b^0	-0.89	-0.1 ± 0.8
$B^{*+} - B^{*0}$	Δ_b^1	0.02	NA
$B^{*+} - B^+$	Δ_{bu}^*	46.5	46.0 ± 0.6[b]
$B^{*0} - B^0$	Δ_{bd}^*	45.5	46.0 ± 0.6[b]

[a]Obtained as a difference of world averages.

[b]Average of charged and neutral states.

the best of circumstances, and the crudeness of the estimate of $\partial\alpha_s/\partial\mu$. Other possible problems include the assumption that the quarks occur only in a relative $\ell = 0$ state (relevant for K^*-mesons), and the assumption that the strong effects are dominated by a confining potential and one-gluon exchange, since at the lower energies associated with the K system, $O(\alpha_s^2)$ terms and more complex models of confinement may be required. The failure of these assumptions can drastically alter the strong hyperfine interaction, which determines the size of α_s, and hence the other mass splittings.

Some may find the small size of α_s somewhat puzzling. This is primarily the result of the confining term of the model potential: It causes the wavefunction to be large at the origin, and thus a small α_s is required to give the same experimentally measured vector-pseudoscalar splitting (see Eq. 3.35). Such small values for the strong coupling constant might lead to excessively small values of Λ_{QCD} and large values for mesonic decay constants $f_{\bar{Q}q}$. We use the naive expressions for these

quantities,

$$\alpha_s(\mu) = \frac{12\pi}{(33 - 2n_f)\ln(\mu^2/\Lambda^2)}, \tag{3.57}$$

and, assuming the relative momenta of the quarks is small,

$$f^2_{\bar{Q}q} = \frac{12}{M_{\bar{Q}} + m_q} |\Psi(0)|^2 . \tag{3.58}$$

In the D system, for example, $\alpha_s = 0.363$ and $\mu = 287$ MeV, and with three flavors of quark, we calculate $\Lambda_{QCD} = 42$ MeV (roughly consistent with $\Lambda_{QCD} = 34$ MeV from the B system) and $f_D = 342$ MeV. However, one may state the following objections: First, Λ_{QCD} is computed from the full theory of QCD, but the nonrelativistic potential approach includes the confinement in an *ad hoc* fashion, by including a confinement constant a, which is independent of α_s. Furthermore, choosing Λ_{QCD} as the renormalization point forces an artificial Landau singularity at $\mu = \Lambda_{QCD}$. The problem is that little is known about the low-energy behavior of strong interactions. At low energies the computation and interpretation of Λ_{QCD} requires a more careful consideration of confinement. With respect to the decay constant, the assumption that the quarks are relatively at rest leads to the evaluation of the wavefunction at zero separation. Inclusion of nonzero relative momentum presumably results in the necessity of considering separations of up to a Compton wavelength $r \approx \bar{\mu}^{-1}$, for which the wavefunction is smaller in the $1s$-state. Thus decay constants may be smaller than computed in the naive model.

There is one further qualitative success of this model, a partial explanation of the experimental facts that $D_s^* - D_s = 141.5 \pm 1.9$ MeV $\approx D^* - D$, and $B_s^* - B_s = 47.0 \pm 2.6$ MeV $\approx B^* - B$, namely, the approximate independence of vector-pseudoscalar splitting on the light quark mass. In our model, the leading term of the splitting is, using Eqs. 3.35 and 3.47,

$$\Delta^*_{Qq} \approx \frac{16}{9Ma^2}\alpha_s\beta\left[\lambda\left\langle\frac{1}{\rho^2}\right\rangle + 1\right]. \tag{3.59}$$

Inasmuch as β, $\lambda\langle 1/\rho^2\rangle$, and α_s are slowly varying in the light quark mass m, the full expression reflects this insensitivity, in accord with experiment. In fact, we may

fit the experimental values above to obtain more running values of α_s:

$$\begin{aligned} \Delta^*_{cs} &= 141.5\ \text{MeV} \quad \text{for} \quad \alpha_s = 0.351, \\ \Delta^*_{bs} &= 47.0\ \text{MeV} \quad \text{for} \quad \alpha_s = 0.295, \end{aligned} \tag{3.60}$$

and again these decrease as the mass scale increases. Note, however, one kink in this interpretation: The heavy-strange mesons all have larger reduced masses than their unflavored counterparts, yet the corresponding values of α_s are nearly the same.

3.8 Conclusions

In this Chapter, we have seen how mass contributions to a bound system of particles are derived from an interaction Hamiltonian in field theory, and how this calculation is then reduced to a problem in nonrelativistic quantum mechanics. For the system of a quark and antiquark bound in a meson, the exchange of one mediating vector boson reduces to the Breit–Fermi interaction in the nonrelativistic limit. It is also important to consider contributions to the total energy from the kinetic energy and the long-range potential of the system; in fact, the higher-order momentum expectation values can be so large that it is necessary to impose an *Ansatz* (Eq. 3.28) in order to estimate their combined effect. Future work may suggest better estimates.

It is found in the case of a linear-plus-Coulomb potential that the largest contributions to electromagnetic mass splittings originate in the kinetic energy, static Coulomb, and strong hyperfine terms. However, it is likely that similar results hold for other *Ansätze* and potentials. As in other models, vector-pseudoscalar mass differences are determined by strong hyperfine terms.

With typical values for quark masses, the confinement constant, and the up-down quark mass difference, we have obtained agreement using 1992 numbers for the mass splittings of the D and B mesons. The failure of the model for K mass splittings is attributed to the collapse of the nonrelativistic assumptions in that case. The model also qualitatively explains the similarity of heavy-strange to heavy-unflavored vector-pseudoscalar splittings, although additional work is needed

to explain why these numbers are nearly equal, despite the expected inequality of α_s at the two different energy scales.

However, data published after the publication of Ref. [56] may serve to render the simple nonrelativistic linear-plus-Coulomb potential model less fit to describe isospin splittings. The numerical fit performed in the previous Section is notable in that it was just barely possible with the 1992 data to fit simultaneously the isospin splittings $(D^0 - D^+)$, $(D^{*+} - D^{*0})$, and $(B^+ - B^0)$ within one standard deviation of their central values (see Table III). Recently, these experimental uncertainties were reduced by the CLEO collaboration [57]:

$$\begin{aligned} D^+ - D^0 &= +4.80 \pm 0.10 \pm 0.06 \text{ MeV}, \\ D^{*+} - D^{*0} &= +3.32 \pm 0.08 \pm 0.05 \text{ MeV}, \\ B^+ - B^0 &= -0.41 \pm 0.25 \pm 0.19 \text{ MeV}, \end{aligned} \tag{3.61}$$

where the two uncertainties are statistical and systematic, respectively. These numbers, particularly the new D^* splitting, upset the fit of the previous Section. The problem, as before, is easy to state: The B splitting is surprisingly small compared to the D splitting. If we wish to keep the nonrelativistic potential model, we might simply have to abandon the original linear-plus-Coulomb potential once and for all.

For completeness, we list other relevant recent CLEO measurements [57]:

$$\begin{aligned} D^{*+} - D^+ &= 140.64 \pm 0.08 \pm 0.06 \text{ MeV}, \\ D^{*0} - D^0 &= 142.12 \pm 0.05 \pm 0.05 \text{ MeV}, \\ D_s^{*+} - D_s^+ &= 144.22 \pm 0.47 \pm 0.37 \text{ MeV}. \end{aligned} \tag{3.62}$$

Another interesting problem is the running of α_s itself at low energies. As mentioned in the previous Section, this running cannot be neglected if we are to obtain sensible results, and yet our approximation of this running is based on crude assumptions. The size of α_s also enters into another possible development, namely, whether terms of $O(\alpha_s^2)$ are important, particularly for the K system. More reliable estimates are required.

In addition to the explicit formulas derived in this Chapter, the techniques employed here may be applied to later efforts: In particular, the explicit consid-

eration of the mass-dependence of expectation values and the use of quantum-mechanical theorems to reduce the number of such expectation values may prove useful and even necessary in subsequent models.

Chapter 4

Baryon Masses 1: Group Theory

4.1 Introduction

The final two Chapters are dedicated to a study of the mass spectrum of the lightest baryon multiplets (*i.e.*, the octet and decuplet) in an $SU(3)_L \times SU(3)_R$ chiral Lagrangian formalism. Specifically, we study baryon mass relations and their corrections, and consider the determination of light quark mass parameters.

Historically, the chiral Lagrangian [64] approach was developed in the late 1960's and early 1970's to incorporate in one effective field theory the specific techniques of the nonlinear sigma model, current algebra, and soft-pion theorems, but the essential physics behind all of these particular approaches lies in just two elements: symmetry and dynamics. One begins with a Lagrangian possessing a large symmetry ($SU(3)_L \times SU(3)_R$), which is spontaneously broken to a smaller symmetry ($SU(3)_V$) approximately obeyed by nature; perturbative terms that explicitly break the smaller symmetry are then added by hand. The Nambu–Goldstone bosons of the spontaneous symmetry breaking form the $SU(3)_V$ octet of light pseudoscalar mesons, and the explicit symmetry breaking provides them with masses. Finally, since the relevant dynamical degrees of freedom in this Lagrangian are the Nambu–Goldstone bosons, one has now a dynamical theory of the pseudoscalar octet with an approximate $SU(3)_V$ symmetry built in.

These two elements also provide a natural means of studying the struc-

ture of the baryon mass spectrum in a chiral theory. In this Chapter we explore the relations of baryon masses purely as consequences of $SU(3)$ group theory (suppressing here and henceforth the subscript V) because any theory possessing chiral symmetry must respect the appropriate group-theoretical constraints; here the algebraic structure of the theory is emphasized. In the next Chapter we introduce the specifics of our chiral theory and use its dynamical properties to make numerical predictions of corrections to baryon mass relations.

This Chapter is organized as follows: In Sec. 4.2 we exhibit the full $SU(3)$ structure of octet and decuplet baryon masses as organized by the representations of $SU(3)$ symmetry-breaking operators. In Sec. 4.3 we study how such representations may come about in a chiral Lagrangian and consequently find, to second order in flavor breaking, four relations holding among the decuplet masses [65] and one relation among the octet masses [66]. We discuss the independent quark mass parameters of an arbitrary chiral theory in Sec. 4.4 and examine their importance in determining the viability of $m_u = 0$. In Sec. 4.5 we digress to discuss the problem of uncertainties in decuplet masses, and then, as an application of the group theory without dynamics, compute numerical values for the decuplet chiral coefficients.

4.2 The Structure of $SU(3)$ Breaking

We begin with a systematic classification of mass terms of the octet and decuplet baryons within $SU(3)$ group theory. Consider, within the effective Lagrangian, any term contributing to the mass of a field multiplet transforming under an R-dimensional representation:

$$\delta\mathcal{L} = \overline{R}\mathcal{O}R, \tag{4.1}$$

where $\mathcal{O}$ is some operator. The pattern of $SU(3)$ breaking in the masses is exhibited by the decomposition of $(\overline{R} \times R)$ into combinations transforming under all possible irreducible representations. For the octet and decuplet, these representations are

$$\mathbf{8} \otimes \mathbf{8} = \mathbf{1} \oplus \mathbf{8}_1 \oplus \mathbf{8}_2 \oplus \mathbf{10} \oplus \overline{\mathbf{10}} \oplus \mathbf{27} \tag{4.2}$$

and

$$\mathbf{10} \otimes \overline{\mathbf{10}} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{27} \oplus \mathbf{64}, \tag{4.3}$$

respectively. The projections of $\mathcal{O}$ forming the coefficients of these combinations, called mass operators, are labeled with the $SU(3)$ indices of the corresponding combinations. Furthermore, this analysis assumes negligible mixing from heavier states with the same quantum numbers.

A further restriction on mass terms in the Lagrangian is that they not only form bilinears of the desired *multiplet* field with its conjugate, but in fact connect the same *states* within the multiplet. Because all additive quantum numbers of fields are the opposites of those of their conjugates, we see that the additive quantum numbers of mass operators must be zero. In $SU(3)$ this means mass operators possess the additional properties $\Delta I_3 = 0$ and $\Delta Y = 0$. Note, however, that we still have "mixing" terms for any states with the same values of I_3 and Y; in our cases, no decuplet states mix, whereas in the octet, Σ^0-Λ mixing can occur.

It remains only to distinguish degenerate $\Delta I_3 = \Delta Y = 0$ operators within a representation. As usual, we use the standard notation of labeling with the isospin Casimir $I(I+1)$, as in $\mathcal{O} = c_I^R \mathcal{O}_I^R$. It then becomes a straightforward exercise with $SU(3)$ Clebsch–Gordan coefficients to decompose multiplet masses into the forms

$$\begin{aligned} \mathbf{M}_8 &= \mathcal{C}_8 \mathbf{a}, \\ \mathbf{M}_{10} &= \mathcal{C}_{10} \mathbf{b}, \end{aligned}$$

where

$$
\mathcal{C}_8 =
\begin{pmatrix}
\frac{1}{2\sqrt{2}} & +\frac{1}{2\sqrt{5}} & -\frac{1}{2}\sqrt{\frac{3}{5}} & +\frac{1}{2} & +\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{3}} & +\frac{1}{2\sqrt{3}} & +\frac{1}{2}\sqrt{\frac{3}{10}} & +\frac{1}{\sqrt{10}} & 0 \\
\frac{1}{2\sqrt{2}} & +\frac{1}{2\sqrt{5}} & +\frac{1}{2}\sqrt{\frac{3}{5}} & +\frac{1}{2} & -\frac{1}{2\sqrt{3}} & +\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{3}} & +\frac{1}{2}\sqrt{\frac{3}{10}} & -\frac{1}{\sqrt{10}} & 0 \\
\frac{1}{2\sqrt{2}} & -\frac{1}{\sqrt{5}} & 0 & 0 & +\frac{1}{\sqrt{3}} & +\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{30}} & 0 & +\frac{1}{\sqrt{6}} \\
\frac{1}{2\sqrt{2}} & -\frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2\sqrt{30}} & 0 & -\sqrt{\frac{2}{3}} \\
\frac{1}{2\sqrt{2}} & +\frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 & -\frac{3}{2}\sqrt{\frac{3}{10}} & 0 & 0 \\
\frac{1}{2\sqrt{2}} & -\frac{1}{\sqrt{5}} & 0 & 0 & -\frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} & +\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{30}} & 0 & +\frac{1}{\sqrt{6}} \\
\frac{1}{2\sqrt{2}} & +\frac{1}{2\sqrt{5}} & +\frac{1}{2}\sqrt{\frac{3}{5}} & -\frac{1}{2} & +\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{3}} & +\frac{1}{2\sqrt{3}} & +\frac{1}{2}\sqrt{\frac{3}{10}} & -\frac{1}{\sqrt{10}} & 0 \\
\frac{1}{2\sqrt{2}} & +\frac{1}{2\sqrt{5}} & -\frac{1}{2}\sqrt{\frac{3}{5}} & -\frac{1}{2} & -\frac{1}{2\sqrt{3}} & +\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{3}} & +\frac{1}{2}\sqrt{\frac{3}{10}} & +\frac{1}{\sqrt{10}} & 0 \\
0 & 0 & -\frac{1}{\sqrt{5}} & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & 0 & -\sqrt{\frac{3}{10}} & 0 \\
0 & 0 & -\frac{1}{\sqrt{5}} & 0 & 0 & +\frac{1}{2} & +\frac{1}{2} & 0 & -\sqrt{\frac{3}{10}} & 0
\end{pmatrix},
$$

$$
\mathcal{C}_{10} =
\begin{pmatrix}
\frac{1}{\sqrt{10}} & +\frac{1}{\sqrt{10}} & +\sqrt{\frac{3}{10}} & +\sqrt{\frac{3}{70}} & +\frac{3}{\sqrt{70}} & +\sqrt{\frac{3}{14}} & +\frac{1}{2\sqrt{35}} & +\frac{1}{2}\sqrt{\frac{3}{35}} & +\frac{1}{2\sqrt{7}} & +\frac{1}{2\sqrt{5}} \\
\frac{1}{\sqrt{10}} & +\frac{1}{\sqrt{10}} & +\frac{1}{\sqrt{30}} & +\sqrt{\frac{3}{70}} & +\frac{1}{\sqrt{70}} & -\sqrt{\frac{3}{14}} & +\frac{1}{2\sqrt{35}} & +\frac{1}{2\sqrt{105}} & -\frac{1}{2\sqrt{7}} & -\frac{3}{2\sqrt{5}} \\
\frac{1}{\sqrt{10}} & +\frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{30}} & +\sqrt{\frac{3}{70}} & -\frac{1}{\sqrt{70}} & -\sqrt{\frac{3}{14}} & +\frac{1}{2\sqrt{35}} & -\frac{1}{2\sqrt{105}} & -\frac{1}{2\sqrt{7}} & +\frac{3}{2\sqrt{5}} \\
\frac{1}{\sqrt{10}} & +\frac{1}{\sqrt{10}} & -\sqrt{\frac{3}{10}} & +\sqrt{\frac{3}{70}} & -\frac{3}{\sqrt{70}} & +\sqrt{\frac{3}{14}} & +\frac{1}{2\sqrt{35}} & -\frac{1}{2}\sqrt{\frac{3}{35}} & +\frac{1}{2\sqrt{7}} & -\frac{1}{2\sqrt{5}} \\
\frac{1}{\sqrt{10}} & 0 & +\sqrt{\frac{2}{15}} & -\sqrt{\frac{5}{42}} & -\frac{3}{\sqrt{70}} & +\frac{1}{\sqrt{42}} & -\frac{2}{\sqrt{35}} & -\sqrt{\frac{5}{21}} & -\frac{1}{\sqrt{7}} & 0 \\
\frac{1}{\sqrt{10}} & 0 & 0 & -\sqrt{\frac{5}{42}} & 0 & -\sqrt{\frac{2}{21}} & -\frac{2}{\sqrt{35}} & 0 & +\frac{2}{\sqrt{7}} & 0 \\
\frac{1}{\sqrt{10}} & 0 & -\sqrt{\frac{2}{15}} & -\sqrt{\frac{5}{42}} & +\frac{3}{\sqrt{70}} & +\frac{1}{\sqrt{42}} & -\frac{2}{\sqrt{35}} & +\sqrt{\frac{5}{21}} & -\frac{1}{\sqrt{7}} & 0 \\
\frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{10}} & +\frac{1}{\sqrt{30}} & -\sqrt{\frac{3}{70}} & -\frac{2\sqrt{2}}{\sqrt{35}} & 0 & +\frac{3}{\sqrt{35}} & +\sqrt{\frac{5}{21}} & 0 & 0 \\
\frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{30}} & -\sqrt{\frac{3}{70}} & +\frac{2\sqrt{2}}{\sqrt{35}} & 0 & +\frac{3}{\sqrt{35}} & -\sqrt{\frac{5}{21}} & 0 & 0 \\
\frac{1}{\sqrt{10}} & -\sqrt{\frac{2}{5}} & 0 & +\frac{3\sqrt{3}}{\sqrt{70}} & 0 & 0 & -\frac{2}{\sqrt{35}} & 0 & 0 & 0
\end{pmatrix},
$$

and

$$\mathbf{a} \equiv \begin{pmatrix} a_0^1 \\ a_0^{8_1} \\ a_1^{8_1} \\ a_0^{8_2} \\ a_1^{8_2} \\ a_1^{10} \\ a_1^{\overline{10}} \\ a_0^{27} \\ a_1^{27} \\ a_2^{27} \end{pmatrix}, \quad \mathbf{M}_8 \equiv \begin{pmatrix} \overline{p}p \\ \overline{n}n \\ \overline{\Sigma^+}\Sigma^+ \\ \overline{\Sigma^0}\Sigma^0 \\ \overline{\Lambda}\Lambda \\ \overline{\Sigma^-}\Sigma^- \\ \overline{\Xi^0}\Xi^0 \\ \overline{\Xi^-}\Xi^- \\ \overline{\Sigma^0}\Lambda \\ \overline{\Lambda}\Sigma^0 \end{pmatrix}, \quad \mathbf{b} \equiv \begin{pmatrix} b_0^1 \\ b_0^8 \\ b_1^8 \\ b_0^{27} \\ b_1^{27} \\ b_2^{27} \\ b_0^{64} \\ b_1^{64} \\ b_2^{64} \\ b_3^{64} \end{pmatrix}, \quad \mathbf{M}_{10} \equiv \begin{pmatrix} \overline{\Delta^{++}}\Delta^{++} \\ \overline{\Delta^+}\Delta^+ \\ \overline{\Delta^0}\Delta^0 \\ \overline{\Delta^-}\Delta^- \\ \overline{\Sigma^{*+}}\Sigma^{*+} \\ \overline{\Sigma^{*0}}\Sigma^{*0} \\ \overline{\Sigma^{*+}}\Sigma^{*-} \\ \overline{\Xi^{*0}}\Xi^{*0} \\ \overline{\Xi^{*-}}\Xi^{*-} \\ \overline{\Omega^-}\Omega^- \end{pmatrix}. \tag{4.4}$$

Here the $\mathbf{8} \otimes \mathbf{8}$ representations $\mathbf{8}_{1,2}$ are distinguished by the symmetry properties of their components under reflection through the origin in weight space (*i.e.*, exchanging the component transforming with quantum numbers (I,I_3,Y) with that transforming under $(I,-I_3,-Y)$). $\mathbf{8}_{1,2}$ is symmetric (antisymmetric) under this exchange, giving, for instance, the same (opposite) contributions to the masses of the p and Ξ^-.

With the above normalization of the chiral coefficients a_I^R and b_I^R, the matrices $\mathcal{C}_{8,10}$ are orthogonal. The phase conventions of the Clebsch–Gordan coefficients (see, *e.g.*, Ref. [67]) are chosen so that the coefficient of the term $\overline{\psi}\psi$ is truly m_ψ, so that, for example, each octet term has the same singlet coefficient $a_0^1/2\sqrt{2}$.

It is easy to understand the number of chiral coefficients appearing in the octet and decuplet. With arbitrary $SU(3)$ breaking, one may clearly supply each baryon with a distinct arbitrary mass; hence the decuplet must have at least ten chiral coefficients and the octet eight. But because the octet supports Σ^0-Λ mixing, there must be at least one further coefficient to parametrize a mixing angle θ. In the above matrices there are two, corresponding to the bilinears $\overline{\Sigma^0}\Lambda$ and $\overline{\Lambda}\Sigma^0$. However, hermiticity of the Lagrangian reduces these to one, imposing the constraint $a_1^{10} = -a_1^{\overline{10}}$.

Now it is a simple matter to extract numerical predictions for the chiral

coefficients in terms of the baryon masses (and the Σ^0-Λ mixing, in the octet case). In fact, the lack of decuplet mixing terms allows one to represent compactly the entire C_{10} matrix equation in terms of a single formula in I_3 and Y. First renormalize the coefficients by removing roots: Define the vector $\mathbf{c}$ by multiplying the corresponding entries of the vector $\mathbf{b}$ in Eq. 4.4 by

$$\left(\tfrac{1}{\sqrt{10}},\ \tfrac{1}{\sqrt{10}},\ \sqrt{\tfrac{2}{15}},\ \tfrac{1}{\sqrt{210}},\ \tfrac{1}{\sqrt{70}},\ \tfrac{1}{\sqrt{42}},\ \tfrac{1}{2\sqrt{35}},\ \tfrac{1}{2\sqrt{105}},\ \tfrac{1}{2\sqrt{7}},\ \tfrac{1}{2\sqrt{5}} \right),$$

so that, for example, $c_1^{27} \equiv b_1^{27}/\sqrt{70}$. Then the mass formula reads

$$\begin{aligned} M_{10} \;=\; & c_0^1 + c_0^8 Y + c_1^8 I_3 + c_0^{27}\left(5Y^2 + 3Y - 5\right) \\ & + c_1^{27} I_3 \left(5Y - 3\right) + \frac{1}{4} c_2^{27}\left(12 I_3^2 - Y^2 - 6Y - 8\right) \\ & + \frac{1}{6} c_0^{64}\left(35Y^3 + 45Y^2 - 50Y - 24\right) + c_1^{64} I_3 \left(21Y^2 - 9Y - 10\right) \\ & + \frac{1}{12} c_2^{64}\left[12 I_3^2 \left(7Y - 6\right) - 7Y^3 - 36Y^2 - 20Y + 48\right] \\ & + \frac{1}{6} c_3^{64} I_3 \left(20 I_3^2 - 3Y^2 - 18Y - 20\right), \end{aligned} \tag{4.5}$$

where hypercharge Y is normalized by $Q = I_3 + \frac{1}{2}Y$.

4.3 Baryon Mass Relations

In the strictest sense, there could not exist *exact* baryon mass relations, even if the lightest observed baryons truly formed an exact octet and decuplet of $SU(3)$; as pointed out in the previous Section, there are at least as many independent chiral coefficients as baryons in each multiplet. However, if we can find *physical* reasons that operators transforming under certain representations do not appear, then the corresponding chiral coefficients vanish, indicating (by Eq. 4.4) a relation among baryon masses.

When Gell-Mann and Okubo [68, 69], and Coleman and Glashow [70] derived the famous hadron mass relations named for them, the physical nature of the symmetry breaking was not well understood. One could only assume that operators in a few representations were responsible for hadron masses, tally the

resulting relations, and check them against experiment for self-consistency. We have come to believe that the chiral Lagrangian holds the answer: $SU(3)$ breaking is accomplished by inequality of masses and electric charges of the three light quarks u, d, and s.

In terms of $SU(3)$ flavor indices, the quark mass and charge operators M_q and Q_q are 3×3 matrices; in terms of $SU(3)$ representations, such a matrix X may be decomposed into singlet ($(\mathrm{Tr}X)\mathbf{1}$) and octet ($X - \frac{1}{3}(\mathrm{Tr}X)\mathbf{1}$) portions. Thus, to *first* order in flavor breaking, any combination of baryon masses with no singlet or octet piece forms a mass relation.

Let us consider some examples, first supposing that splittings within isospin multiplets are negligible. Then all chiral coefficients of the form c_I^R with $I > 0$ must also be negligible. In this case, the only independent octet masses are N, Σ, Λ, and Ξ, whereas the only nontrivial chiral coefficients are a_0^1, $a_0^{8_1}$, $a_0^{8_2}$, and a_0^{27}. If we only work to first order in flavor breaking, the last of these is identically zero, and we find

$$\Delta_{\mathrm{GMO}} \equiv \tfrac{1}{2}\sqrt{\tfrac{10}{3}}a_0^{27} = \tfrac{3}{4}\Lambda + \tfrac{1}{4}\Sigma - \tfrac{1}{2}(N + \Xi) = 0, \tag{4.6}$$

the Gell-Mann–Okubo relation [68]. For the decuplet, the independent masses are Δ, Σ^*, Ξ^*, and Ω, whereas the nontrivial chiral coefficients are c_0^1, c_0^8, c_0^{27}, and c_0^{64}. To first order in flavor breaking, the vanishing of the last two coefficients gives rise to two nontrivial relations, which may be written

$$\begin{aligned} 0 &= 5(2c_0^{27} + c_0^{64}) = (\Delta - \Sigma^*) - (\Sigma^* - \Xi^*), \\ 0 &= 10(c_0^{27} - 2c_0^{64}) = (\Sigma^* - \Xi^*) - (\Omega - \Xi^*), \end{aligned} \tag{4.7}$$

Gell-Mann's famous equal-spacing rule [69]. The equal spacing is also clear from Eq. 4.5, because the only surviving coefficients in this case are c_0^1 and c_0^8.

Now consider second-order terms in flavor breaking. *A priori* we might expect to find that all of the representations within the product $\mathbf{8} \times \mathbf{8}$ occur, but we show that this is not the case. Because of charge conjugation symmetry of the strong interaction, the mass Lagrangian contains no terms with an odd number of Q_q factors. Thus the only second-order terms in flavor breaking are of the forms

$(M_q \times M_q)$ and $(Q_q \times Q_q)$. Consider the product of two identical arbitrary matrices: $(X \times X)_{ij}{}^{kl}$, which contains such terms as $X_i{}^k X_j{}^l$, $X_i{}^l X_j{}^k$, and various traces of X, where i, j, k, l are flavor indices in the usual notation. It is readily seen that this product has no piece transforming under a **10**, for such a tensor with the given indices has the form $A_{ijm}\epsilon^{mkl}$, and is symmetric under permutation of $\{i, j, m\}$. If we attempt to construct a product with these symmetry properties from two identical matrices, we quickly see that such a term vanishes. Similarly, the product of two identical matrices may contain no piece of a $\overline{\mathbf{10}}$.

We conclude that, to second order in flavor breaking, the octet chiral coefficients $a_1^{10} = -a_1^{\overline{10}}$ are zero. The baryon mass relation corresponding to the vanishing of these coefficients is

$$\Delta_{\mathrm{CG}} \equiv 2\sqrt{3}a_1^{10} = (n-p) + (\Sigma^+ - \Sigma^-) - (\Xi^0 - \Xi^-) = 0, \tag{4.8}$$

the Coleman–Glashow relation [70]. For the decuplet, the analysis is even easier: $\mathbf{8} \times \mathbf{8}$ contains no **64** for *arbitrary* pairs of 3×3 matrices, and so we have four mass relations good to second-order in flavor breaking, corresponding to the vanishing of ${c^{64}}_{0,1,2,3}$:

$$\Delta_1 \equiv \quad 20c_3^{64} = \Delta^{++} - 3\Delta^+ + 3\Delta^0 - \Delta^-, \tag{4.9}$$

$$\Delta_2 \equiv \quad 28c_2^{64} = \left(\Delta^{++} - \Delta^+ - \Delta^0 + \Delta^-\right) - 2\left(\Sigma^{*+} - 2\Sigma^{*0} + \Sigma^{*-}\right), \tag{4.10}$$

$$\Delta_3 \equiv 6(7c_1^{64} - c_3^{64}) = \left(\Delta^+ - \Delta^0\right) - \left(\Sigma^{*+} - \Sigma^{*-}\right) + \left(\Xi^{*0} - \Xi^{*-}\right), \tag{4.11}$$

$$\begin{aligned}\Delta_4 \equiv \quad 35c_0^{64} &= \frac{1}{4}\left(\Delta^{++} + \Delta^+ + \Delta^0 + \Delta^-\right) - \left(\Sigma^{*+} + \Sigma^{*0} + \Sigma^{*-}\right) \\ &\quad + \frac{3}{2}\left(\Xi^{*0} + \Xi^{*-}\right) - \Omega^-\end{aligned} \tag{4.12}$$

are four vanishing combinations. Notice that the first three of these are isospin-breaking, and only the fourth remains in the limit that isospin is a good symmetry.

These relations are not unknown in the literature. In fact, the first three can be trivially derived from quark model calculations over a quarter of a century old [71, 72]. The relations found in these papers connect the masses of the octet to decuplet baryons, for when these were first derived, the octet masses were already much better known than the decuplet masses. Eqs. 4.9–4.12 can be obtained

from the earlier relations by eliminating octet baryon masses in the latter. Such intermultiplet relations do not occur in the form of chiral perturbation theory that we use, as the octet and decuplet are treated as two independent multiplets, unrelated by the physical fact that they are both three-quark states. Here the average decuplet-octet mass splitting $\delta \approx 300$ MeV is an independent parameter.

Eq. 4.12 is the remnant of Gell-Mann's equal-spacing rule, as may be seen from Eq. 4.7. It was pointed out by Okubo [73] as early as 1963 that the only surviving such relation at second order in flavor breaking is

$$\Delta - 3\Sigma^* + 3\Xi^* - \Omega = 0, \tag{4.13}$$

which is, neglecting isospin splitting, the same as Eq. 4.12. It was also derived in this form by Jenkins [86], using the heavy baryon effective field theory described in the next Chapter.

Equivalent forms of all four relations have been derived in a very general quark model, the general parametrization method of Morpurgo [75]. Again, the relations were written in terms of equations connecting octet to decuplet masses.

Eq. 4.9 is somewhat special because it is the only relation good to second order among baryons from only one isomultiplet. Indeed, it is seen to be simply a consequence of the $SU(2)$ Wigner–Eckart theorem applied to an isospin-3/2 multiplet, when all Lagrangian mass terms transform as $I = 0, 1, 2$. $I = 3$ terms require at least three octet operators, as the largest isomultiplet contained in an **8** is $I = 1$. Similarly, if we consider relations good to first order in flavor breaking, the $SU(2)$ Wigner–Eckart theorem supplies us a number of additional relations, including the Σ equal-spacing rule [70]:

$$\Delta_\Sigma \equiv \sqrt{6}a_2^{27} = (\Sigma^+ - \Sigma^0) - (\Sigma^0 - \Sigma^-), \tag{4.14}$$

which is clearly broken only by $I = 2$ operators. We caution that Σ^0 in this equation refers to the isospin $I = 1$ eigenstate rather than the mass eigenstate; this relation is explored in Sec. 5.7.

An amusing implication of the insensitivity of Eqs. 4.9–4.12 to the form of the chiral breaking, combined with the symmetry of the decuplet field under

permutation of flavor indices, is that the set of four decuplet relations is invariant under permutations of the three isospin axes T_3, U_3, V_3 in the weight diagram of the decuplet in flavor space; for example, Eq. 4.9 becomes

$$0 = \Delta^- - 3\Sigma^{*-} + 3\Xi^{*-} - \Omega^-, \tag{4.15}$$

under $T_3 \mapsto U_3$, $U_3 \mapsto -V_3$, $V_3 \mapsto -T_3$, or

$$0 = \Omega^- - 3\Xi^{*0} + 3\Sigma^{*0} - \Delta^{++}, \tag{4.16}$$

under $T_3 \mapsto -V_3$, $U_3 \mapsto T_3$, $V_3 \mapsto -U_3$, both of which lie within the linear span of Eqs. 4.9–4.12. That is, the set of relations is unaffected if we permute, for example, the up and strange quarks in all decuplet wavefunctions.

4.4 Quark Mass Parameters

One natural problem we may attempt to solve with the chiral Lagrangian is the determination of the light quark masses m_u,m_d,m_s (which we relabel for convenience in this Section as u,d,s). This turns out to be impossible for a number of reasons varying from trivial to subtle, as we now discuss.

First, consider any effective theory that takes some related set of undetermined parameters as inputs; in our case, this set is the matrix of current quark masses $M_q \equiv diag(u, d, s)$. Each term in the Lagrangian then contains a certain number of factors of M_q. The coefficients c_i of the various terms are *a priori* unrelated parameters (unlike in a renormalized field theory), and have significance only in a product with a power of M_q. Given a particular term with n_i powers of M_q, one readily sees that the term is invariant under the transformation

$$M_q \mapsto kM_q, \qquad c_i \mapsto k^{-n_i} c_i, \tag{4.17}$$

where k is arbitrary. Thus, in any chiral theory one cannot hope to obtain quark masses, but only ratios of quark masses. For three light flavors we have two ratios; we find it convenient to use two particular combinations,

$$q, r \equiv \frac{d \pm u}{s - \frac{1}{2}(u + d)}. \tag{4.18}$$

Note that both parameters are small, inasmuch as $u,d \ll s$, and that q can appear with isospin-conserving operators, whereas r can only appear with isospin-breaking operators.

The heavy baryon theory introduced in the next Chapter eliminates the common mass of the baryon multiplet as a Lagrangian parameter, and in such a model we may reduce the set of quark mass parameters even further. We simply observe that the tree-level Lagrangian is insensitive to transformations of the type $M_q \mapsto M_q + c\mathbf{1}$, where $\mathbf{1}$ is the identity matrix and c is arbitrary. This follows because each insertion of $\mathbf{1}$ in a term of $O(M_q^n)$ is equivalent to a redefinition of the coefficient of the terms of lower orders in M_q by simple binomial expansion. Eventually, we generate singlet terms in this way, which may be ignored using the argument above. In particular, this tells us that the Lagrangian is sensitive only to differences of quark masses; or, combining this with the previous result, it is sensitive only to *ratios of differences* of quark masses. For three light flavors, only one parameter remains, which we choose to be the parameter r.

Beyond tree level, as we see in Sec. 5.3, the operator M_q is replaced by $M = M_q + O(\Pi^2)$, where Π is the pseudoscalar octet field. In this case, the above argument fails to hold because the coefficients no longer shift simply by c-numbers. We must then use both quark mass parameters, q and r. However, the field Π appears in the mass computation only in loop effects. For all examples computed in this model, we find that the loop diagrams either vanish because they transform under $SU(3)$ representations satisfying mass relations, or the quark mass dependence may be removed by expressing coefficients in loop expressions in terms of hadron masses alone. Thus we never need to consider q in these calculations.

Even the determination of the quark mass parameter r is not unique. This is a consequence of the Cayley–Hamilton theorem for any 3×3 matrix X, which reads .

$$X^3 - (\mathrm{Tr}X)X^2 + \frac{1}{2}[(\mathrm{Tr}X)^2 - \mathrm{Tr}X^2]X - (\det X)\mathbf{1} = 0. \tag{4.19}$$

For the quark mass matrix, this implies

$$(\det M_q)M_q^{-1} \;=\; diag\,(ds, us, ud)$$

$$= \quad M_q^2 - (\mathrm{Tr}\, M_q) M_q + \frac{1}{2}[(\mathrm{Tr}\, M_q)^2 - \mathrm{Tr}\, M_q^2]1, \qquad (4.20)$$

so that in the chiral Lagrangian one may generate an *effective* quark mass matrix $M_q{}'$ physically indistinguishable from M_q merely by shifting the Lagrangian coefficients of terms one order higher in M_q:

$$\begin{aligned} M_q{}' &\equiv M_q + \frac{\lambda}{\Lambda_\chi}\left\{M_q^2 - (\mathrm{Tr}\, M_q) M_q + \frac{1}{2}[(\mathrm{Tr}\, M_q)^2 - \mathrm{Tr}\, M_q^2]1\right\} \\ &= diag\left(u + \frac{\lambda ds}{\Lambda_\chi},\ d + \frac{\lambda us}{\Lambda_\chi},\ s + \frac{\lambda ud}{\Lambda_\chi}\right), \end{aligned} \qquad (4.21)$$

where Λ_χ is the scale of the chiral symmetry breaking, and λ is an arbitrary dimensionless coefficient. If we now suppose that the current quark mass $u = 0$, then the effective quark mass matrix possesses unaltered d- and s-quark masses, but an *effective* u-quark mass

$$u_{\mathrm{eff}} = \frac{\lambda}{\Lambda_\chi} ds. \qquad (4.22)$$

The possibility of a massless up-quark generated in this manner was explored by Kaplan and Manohar [76] and Leutwyler [77], who used light hadron masses to determine the relevant chiral coefficients. It is of considerable theoretical interest, because setting $u = 0$ provides an economical solution to the strong CP problem [78].

If λ were truly arbitrary, we would be able to conclude absolutely nothing about any quark masses from a chiral Lagrangian; however, the assumption of perturbativity in the effective Lagrangian imposes a *naturalness criterion* [64, 79]: Dimensionless parameters are all $O(1)$ unless there is a special reason for them to be larger or smaller. It is thus only natural to choose a massless up-quark if the desired value for u_{eff} is $O(ds/\Lambda_\chi)$.

The parameter r is not immune to this ambiguity; let us define

$$\varepsilon \equiv \frac{\lambda}{\Lambda_\chi}\left(s - \frac{1}{2}(u + d)\right). \qquad (4.23)$$

With $s \approx 150$ MeV and $\Lambda_\chi \approx 1$ GeV, we see that $\lambda = O(1)$ makes ε a reasonably small parameter. Under the transformation of Eq. 4.21, we find

$$r \mapsto r\frac{\left[1 - \varepsilon + \frac{1}{2}\varepsilon q\right]}{\left[1 - \frac{1}{2}\varepsilon q - \frac{1}{4}\varepsilon r^2\right]}. \qquad (4.24)$$

Because $q \ll 1$, even with the mass ambiguity it is still possible in principle to determine r to within a few tens of percent. We have more to say about the determination of r in Sec. 5.4.

4.5 Decuplet Mass Measurements

In this Chapter, we have focused somewhat more upon the decuplet mass structure than on that of the octet because the problem is theoretically cleaner: There are four relations to check, and one need not disentangle $SU(3)$-mixed mass eigenstates. If the Σ^0-Λ mixing angle θ were known, we could easily compute all of the octet chiral coefficients a_I^R in Eq. 4.4. By similar reasoning, we should be able to compute immediately all ten decuplet chiral coefficients c_I^R; however, the status of current experimental results is not yet up to this task. Little experimental refinement of the decuplet masses has occurred in the past fifteen years, and decuplet mass differences, particularly isospin splittings, have large relative uncertainties. The mass of one decuplet baryon, the Δ^-, has not even been measured directly, but only deduced from a comparison of pion-nucleon and pion-deuteron cross sections.

Because the mass of the Δ^- is not known independently, we can either treat the relations Eqs. 4.9–4.12 as predictions of its mass, or we can eliminate it from three of the relations using the fourth; since all four of the relations result from group theory alone, any linear combination of them is also a valid relation. We choose to eliminate the Δ^- using Eq. 4.9 (and its corrections), because it is isospin-breaking and is the only one involving Δ masses alone. After this elimination, the other three relations depend only on measured quantities and thus may be checked against our calculations in the next Chapter.

The central problem with the data is that the Δ masses, as presented in the Particle Data Group's (PDG's) *Review of Particle Properties* [48], rely on data fifteen years old, which generally have substantial uncertainties relative to the isospin breaking of the multiplet. The statistical averages of the accepted independent

measurements in the PDG are

$$\begin{aligned} \Delta^{++} &= 1230.86 \quad \pm 0.13 \quad \text{MeV}, \\ \Delta^{+} &= 1234.9 \quad \pm 1.4 \quad \text{MeV}, \\ \Delta^{0} &= 1233.42 \quad \pm 0.16 \quad \text{MeV}. \end{aligned} \tag{4.25}$$

We may ask about the reliability of these data. A recent discussion of the status of baryon isospin splitting measurements is found in a paper by Cutkosky [80]; in particular, the author points out that it is very difficult to accommodate the PDG values for the Δ-mass splittings in a quark-model fit. One explanation, of course, is that the quark model is inadequate; the other, that it is the old measurements that are inadequate. The Virginia Polytechnic Institute (VPI) group [81] currently recommends the value

$$\Delta^0 - \Delta^{++} = 1.3 \pm 0.5 \text{ MeV}. \tag{4.26}$$

Here the uncertainty depends on the measurement of scattering lengths and the pion-nucleon coupling constant, and is expected to fall as their fit is refined. Even so, this value with its current uncertainty is in disagreement with Eq. 4.25.

It would also be very helpful to bring down the large uncertainty in the PDG Δ^+ mass measurement. There are a few pieces of information, experimental and theoretical, in disagreement with the value in Eq. 4.25. This particular number is based on one measurement [82], and is in discrepancy with three other independent values quoted by PDG, which have the simple average

$$\Delta^+ = 1231.5 \pm 0.3 \text{ MeV}. \tag{4.27}$$

These were not used in the PDG fit because uncertainties of the individual measurements were not estimated; the uncertainty given here is the statistical variance, not experimental uncertainty. However, we may use this information as an alternative to the PDG value to demonstrate dependence of the results on the measurement of Δ masses. To support this choice, there are additional predictions of a smaller Δ^+ mass: Using the general parametrization method of Morpurgo [75], one predicts

$$\Delta^+ = \Delta^{++} - (p - n) - (\Sigma^+ - 2\Sigma^0 + \Sigma^-) \tag{4.28}$$

$$= 1230.45 \pm 0.27 \text{ MeV}, \tag{4.29}$$

where we have used the PDG number for the Δ^{++} mass.

Furthermore, pion-deuterium scattering data taken by Pedroni *et al.* [83] produce the mass combination measurement

$$\left(\Delta^- - \Delta^{++}\right) + \frac{1}{3}\left(\Delta^0 - \Delta^+\right) = 4.6 \pm 0.2 \text{ MeV}. \tag{4.30}$$

However, the uncertainty is statistical only and does not reflect a number of theoretical corrections made in the processing of the data. If we combine this number with the relation Eq. 4.9, the PDG value of Δ^{++}, and the VPI result, we find

$$\Delta^+ = 1230.78 \pm 0.52 \text{ MeV}. \tag{4.31}$$

Because we are using the relation Eq. 4.9, this is only a prediction, rather than a true piece of data. Nevertheless, these numbers all appear to be roughly consistent and quite different from the PDG value.

In summary, to demonstrate the dependence of results on Δ masses, we define two sets of experimental measurements of Δ masses. Data set A consists of only the PDG numbers from Eq. 4.25, whereas data set B consists of the PDG number for Δ^{++} in Eq. 4.25, the VPI result in Eq. 4.26, and the alternate value for the Δ^+ mass given in Eq. 4.27. Set B is expected to represent more accurately the true mass values (if we believe the preceding arguments for a smaller Δ^+ mass), and we see in Sec. 5.7 that this set indeed provides a better fit to the corrections to the relations Eq. 4.9–4.12.

We can now fit to the ten chiral coefficients used in Eq. 4.5; with nine known decuplet masses, one parameter remains, which we chose to be c_3^{64}. We show in Sec. 5.7 that the relation Eq. 4.9 remains unbroken by the lowest-order loop contributions we have considered, and that its third-order tree-level contributions are estimated to be tiny, so that we expect c_3^{64} to be quite small ($\ll 0.1$ MeV). With

data set B, we obtain the following fit (all numbers in MeV):

$$
\begin{aligned}
c_0^1 &= 1382.03 \pm 0.24 \quad -2c_3^{64}, & c_2^{27} &= +0.06 \pm 0.15 \quad -\tfrac{10}{7}c_3^{64}, \\
c_0^8 &= -148.43 \pm 0.21 \quad -2c_3^{64}, & c_0^{64} &= +0.17 \pm 0.05 \quad -\tfrac{1}{7}c_3^{64}, \\
c_1^8 &= -1.24 \pm 0.38 \quad +4c_3^{64}, & c_1^{64} &= +0.01 \pm 0.03 \quad +\tfrac{1}{7}c_3^{64}, \\
c_0^{27} &= -0.64 \pm 0.04 \quad -\tfrac{2}{7}c_3^{64}, & c_2^{64} &= -0.18 \pm 0.16 \quad -\tfrac{5}{7}c_3^{64}. \\
c_1^{27} &= +0.28 \pm 0.09 \quad +\tfrac{6}{7}c_3^{64},
\end{aligned}
\tag{4.32}
$$

From Eq. 4.5, the coefficient c_0^1 is clearly the common decuplet mass and c_0^8 is the equal-spacing parameter. The interesting feature of this fit is that the uncertainties in the decuplet mass differences are so large that it becomes increasingly difficult to determine reliably the coefficients of the larger $SU(3)$ representations. If we used data set A for the fit, we would find nearly identical values for c_0^R, the coefficients insensitive to isospin splittings, whereas the other coefficients would vary substantially. Precise measurements of baryon decuplet mass differences would allow us to pin down these coefficients and thus constrain any particular operators with the same transformation properties.

Chapter 5

Baryon Masses 2: Chiral Dynamics

5.1 Introduction

With the theoretical preliminaries of the previous Chapter in hand, we are ready to construct our chiral Lagrangian of mesons and baryons. Using the experimentally measured baryon masses as inputs, we compute corrections to second-order tree-level baryon mass relations and consider corrections to various first-order mass relations.

There are two broad categories of findings we may hope to make in the following pages. The first regards the chiral coefficients that do not appear in the tree-level chiral Lagrangian to second order in $SU(3)$ breaking, namely, those associated with the baryon mass relations derived in the last Chapter. In the usual perturbative expansion of the chiral Lagrangian, we achieve a consistent expansion in the symmetry-breaking parameters by including all loop effects with orders no higher than the highest tree-level order we are considering. We find that, to include consistently tree-level terms to second order in quark masses and charges, we need to compute only one-meson loop diagrams. Even so, many loop diagrams are shown to have the group-theoretical structure of those $SU(3)$ representations that respect the relations previously derived; this serves to explain why the relations experimentally

have such small corrections. In order to perform the necessary calculations, we use the Heavy Baryon Effective Field Theory (HBEFT) developed by Jenkins and Manohar [84, 85, 86], which is convenient in describing the physics of nearly on-shell baryons.

The second category of findings regards what we may learn from those chiral coefficients that do appear in the second-order tree-level Lagrangian, specifically the extraction of the quark mass parameters discussed in Sec. 4.4. It is clear that knowledge about the parameter r would help us to constrain the value of the up-quark mass and determine whether $m_u = 0$ is admissible with current low-energy information. As we show, however, limitations of the chiral Lagrangian formalism serves to keep this parameter out of reach for the time being. Another nonvanishing parameter is that associated with the Σ equal-spacing rule (Eq. 4.14), which we show to be dominated by electromagnetic effects. Furthermore, we show that there are lowest-order corrections to the Gell-Mann–Okubo formula (Eq. 4.6) that cannot be computed from measured matrix elements.

This Chapter is organized as follows: In Sec. 5.2, we review the motivation and construction of HBEFT. Within this formalism, the Lagrangian is constructed in Sec. 5.3 to second order in both the flavor-breaking operators M_q and Q_q, and derivative operators. We examine in Sec. 5.4 the free parameters of the theory and point out redundancies, which provides us with a means of counting the number of independent mass relations, alternate to the approach in Chap. 4. We also discuss the fate of the determination of r in this method. In Sec. 5.5, we demonstrate that one-loop corrections to the mass relations Eqs. 4.8 (the Coleman–Glashow relation) and 4.9–4.12 are calculable, and show why many of the loop corrections vanish. In addition, we consider corrections to the Gell-Mann–Okubo relation, as discussed above. Sec. 5.6 presents the details of how loop calculations are performed in HBEFT (the full loop corrections are presented in Appendices A and B). Numerical evaluations of these calculations are presented in Sec. 5.7, and are found to be in excellent agreement with experiment. We also estimate the size of third-order terms and show that they do not alter our results, and explain the size of violation of the Σ equal-spacing rule. We summarize our conclusions in Sec. 5.8.

5.2 Heavy Baryon Theory and the Effective Lagrangian

The success of the chiral Lagrangian formalism in computing physical quantities rests upon two principles. First, the effective Lagrangian using the meson octet fields as the fundamental degrees of freedom, and the QCD Lagrangian written in terms of the fundamental quark fields, are *formally equivalent* if one includes in the former the infinite set of all terms possessing the same symmetries as the latter [64]. Second, this infinite set of terms in the effective Lagrangian may be organized perturbatively in successively higher orders of the chiral symmetry-breaking parameters and the derivative operator (divided by the dimensionally appropriate powers of the chiral symmetry breaking scale $\Lambda_\chi \approx 1$ GeV [79]), such that their numerical importance in physical processes decreases as the order of the terms increases. For this second requirement to be satisfied, it is necessary not only that the parameters that break chiral symmetry explicitly, namely current quark masses and their electric charges, give contributions sufficiently small (which is guaranteed by $m_{u,d,s} \ll \Lambda_\chi$ and $\alpha_{EM} \ll 1$), but that the meson momenta are also small compared to Λ_χ.

Because the masses of the lowest-lying baryons are already about 1 GeV, the assumptions of chiral perturbativity are violated by a chiral Lagrangian naively including baryons fields as degrees of freedom. However, an indication of how one may avoid this problem is suggested by the heavy quark effective field theory [3]–[8] used in Chap. 1. That such effective theories are formally equivalent to QCD is demonstrated in Ref. [87]. We again adopt the formalism developed by Georgi [8]. One considers a chiral baryon multiplet to be a collection of heavy, nearly on-shell particles degenerate in mass m_B and unit-norm four-velocity v^μ that is approximately conserved, and having momentum

$$p^\mu = m_B v^\mu + k^\mu, \tag{5.1}$$

where k^μ is the residual, off-shell momentum of the baryon. The statement that the baryons are nearly on-shell is expressed by the constraint $k \cdot v \ll m_B$; thus we

have a momentum small enough to use as an expansion parameter in a perturbative chiral theory. This is implemented by the transformation on the baryon fields in any multiplet B via

$$B_v(x) \equiv e^{im_B \not{v} v_\mu x^\mu} B(x). \tag{5.2}$$

The positive- and negative-energy solutions are separated by means of the projection operators

$$P_\pm \equiv \frac{1}{2}(1 \pm \not{v}), \tag{5.3}$$

and we work only with the positive solutions. Whereas the original free fields B satisfy the usual Dirac equation $(i\not{\partial} - m_B)B = 0$, the new free fields B_v satisfy a massless Dirac equation $i\not{\partial}B_v = 0$. This means that a derivative acting upon B_v pulls down a factor of k rather than p, producing the perturbative expansion we require. Henceforth we work only in perturbations about these effectively massless fields, and drop the subscript v. This method, called heavy baryon effective field theory (HBEFT), has been developed as a useful calculational tool by Jenkins and Manohar [84, 85, 86], although the general method could be applied to effective field theories with other heavy degrees of freedom. However, it must be stressed that we have lost some information in HBEFT, namely the baryon multiplet mass. The parameter m_B is nowhere present, so the Lagrangian is sensitive only to baryon mass *differences*.

That baryon multiplets are assumed to be degenerate in lowest-order mass and four-velocity makes it convenient to write a Lagrangian expression $\mathcal{L}_v$ in the baryon fields for each velocity v. As with the fields, we henceforth suppress the index v on $\mathcal{L}_v$. The use of this description has the effect of greatly simplifying the Dirac algebra; in particular, particles and antiparticles no longer mix, reducing four-spinors to two-spinors. It is easily seen [84, 85] that any gamma matrix structure in baryon bilinears may be replaced with the c-number velocity v^μ and a generalized spin operator S^μ defined by the properties

$$[S^\mu, S^\nu]_+ = \frac{1}{2}\left(v^\mu v^\nu - g^{\mu\nu}\right), \qquad [S^\mu, S^\nu]_- = i\epsilon^{\mu\nu\rho\sigma} v_\rho S_\sigma, \tag{5.4}$$

where $\epsilon_{0123} = +1$. A specific representation of the operator satisfying these relations

is

$$S^\mu = \frac{1}{2}\left(\not{v}\gamma^\mu - v^\mu\right)\gamma_5. \tag{5.5}$$

In the rest frame ($v^\mu = (1, \mathbf{0})$), the operator S^μ reduces to the usual Pauli matrices $(0, \boldsymbol{\sigma}/2)$.

The prescriptions listed above are useful for any baryon multiplet included in our theory; in the current model we must consider both baryon decuplet and octet degrees of freedom. Whereas the octet baryon fields are taken to be ordinary Dirac fields, the decuplet is taken to be represented by Rarita–Schwinger fields T^μ_{ijk}, symmetric under permutations of the $SU(3)$ flavor indices i, j, k, with the spin-1/2 portion projected out via the usual constraint $\gamma_\mu T^\mu = 0$. In HBEFT, this translates to the two conditions

$$v_\mu T^\mu = 0, \qquad S_\mu T^\mu = 0. \tag{5.6}$$

The consequences of these conditions for the Feynman rules of this theory are summarized in Reference [86].

One point that should be made at this time is that the octet and decuplet in this theory are taken to have different multiplet common masses, $m_B = m_8, m_{10}$. Both fields, $B^i{}_j$ and T^μ_{ijk}, transform under the rule given by Eq. 5.2; since the values of m_B in the phases are different, the intermultiplet spacing $\delta \equiv m_{10} - m_8$ is a parameter which appears in the Lagrangian and must be placed into the theory by hand.

5.3 Constructing the Lagrangian

5.3.1 Field Transformation Properties

The hadron field multiplets may be compactly represented by matrices in flavor space. The baryon and meson octets have the familiar forms

$$B = \begin{pmatrix} \frac{1}{\sqrt{2}}\Sigma^0 + \frac{1}{\sqrt{6}}\Lambda & \Sigma^+ & p \\ \Sigma^- & -\frac{1}{\sqrt{2}}\Sigma^0 + \frac{1}{\sqrt{6}}\Lambda & n \\ \Xi^- & \Xi^0 & -\frac{2}{\sqrt{6}}\Lambda \end{pmatrix}, \tag{5.7}$$

and

$$\Pi = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{1}{\sqrt{2}}\pi^0 + \frac{1}{\sqrt{6}}\eta & \pi^+ & K^+ \\ \pi^- & -\frac{1}{\sqrt{2}}\pi^0 + \frac{1}{\sqrt{6}}\eta & K^0 \\ K^- & \overline{K}^0 & -\frac{2}{\sqrt{6}}\eta \end{pmatrix}. \tag{5.8}$$

The baryon decuplet in this notation, a $3 \times 3 \times 3$ array, may be represented as (suppressing Lorentz indices) a collection of three matrices:

$$T_{ijk} = \begin{pmatrix} \Delta^{++} & \frac{1}{\sqrt{3}}\Delta^+ & \frac{1}{\sqrt{3}}\Sigma^{*+} \\ \frac{1}{\sqrt{3}}\Delta^+ & \frac{1}{\sqrt{3}}\Delta^0 & \frac{1}{\sqrt{6}}\Sigma^{*0} \\ \frac{1}{\sqrt{3}}\Sigma^{*+} & \frac{1}{\sqrt{6}}\Sigma^{*0} & \frac{1}{\sqrt{3}}\Xi^{*0} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}}\Delta^+ & \frac{1}{\sqrt{3}}\Delta^0 & \frac{1}{\sqrt{6}}\Sigma^{*0} \\ \frac{1}{\sqrt{3}}\Delta^0 & \Delta^- & \frac{1}{\sqrt{3}}\Sigma^{*-} \\ \frac{1}{\sqrt{6}}\Sigma^{*0} & \frac{1}{\sqrt{3}}\Sigma^{*-} & \frac{1}{\sqrt{3}}\Xi^{*-} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}}\Sigma^{*+} & \frac{1}{\sqrt{6}}\Sigma^{*0} & \frac{1}{\sqrt{3}}\Xi^{*0} \\ \frac{1}{\sqrt{6}}\Sigma^{*0} & \frac{1}{\sqrt{3}}\Sigma^{*-} & \frac{1}{\sqrt{3}}\Xi^{*-} \\ \frac{1}{\sqrt{3}}\Xi^{*0} & \frac{1}{\sqrt{3}}\Xi^{*-} & \Omega^- \end{pmatrix}. \tag{5.9}$$

One may assign any particular permutation of indices i,j,k to denote row, column, and sub-matrix in this representation, because the decuplet is completely symmetric under rearrangement of flavor indices.

We require in particular that the chiral Lagrangian with baryons contains the usual nonlinear sigma model. To this end, we define the field

$$\xi \equiv e^{i\Pi/f}, \tag{5.10}$$

where the choice of pion decay constant normalization is $f \approx 93$ MeV. Then, with L and R specifying the left- and right-handed chiral transformations respectively, the field ξ is chosen to transform under $SU(3)_L \times SU(3)_R$ according to the rule

$$\xi \mapsto L\xi U^\dagger = U\xi R^\dagger, \tag{5.11}$$

a mapping that implicitly defines the transformation U, and that implies the usual transformation (ξ^2 is often called Σ in the literature)

$$\xi^2 \mapsto L\xi^2 R^\dagger. \tag{5.12}$$

Under our transformation choice of ξ, we may define Hermitian vector and axial vector currents:

$$V^\mu \equiv \frac{i}{2}\left(\xi\partial^\mu\xi^\dagger + \xi^\dagger\partial^\mu\xi\right), \qquad A^\mu \equiv \frac{i}{2}\left(\xi\partial^\mu\xi^\dagger - \xi^\dagger\partial^\mu\xi\right). \tag{5.13}$$

So that the couplings of these currents to the baryon fields may be chirally invariant, we require the baryonic transformation properties to be

$$B \mapsto UBU^\dagger, \qquad T^\mu_{ijk} \mapsto U^l{}_i U^m{}_j U^n{}_k T^\mu_{lmn}. \tag{5.14}$$

Then we may define the chirally-covariant derivatives

$$\begin{aligned} \mathcal{D}^\nu B &\equiv \partial^\nu B - i\,[V^\nu, B]_-\,, \\ (\mathcal{D}^\nu T^\mu)_{ijk} &\equiv \partial^\nu T^\mu_{ijk} - i(V^\nu)_i{}^l T^\mu_{ljk} - i(V^\nu)_j{}^l T^\mu_{ilk} - i(V^\nu)_k{}^l T^\mu_{ijl}, \end{aligned} \tag{5.15}$$

which have the same transformation properties (Eq. 5.14) as the baryon fields upon which they act.

The explicitly chiral symmetry-breaking operators, namely

$$M_q \equiv \begin{pmatrix} m_u & & \\ & m_d & \\ & & m_s \end{pmatrix}, \qquad Q_q \equiv \begin{pmatrix} +\frac{2}{3} & & \\ & -\frac{1}{3} & \\ & & -\frac{1}{3} \end{pmatrix}, \tag{5.16}$$

are included by the usual spurion procedure applied to the QCD Lagrangian. That is, every symmetry-breaking operator is assigned a spurious $SU(3)_L \times SU(3)_R$ transformation property conjugate to the real transformation property of the Lagrangian term in which it appears. The relevant pieces of the QCD Lagrangian in a chiral basis are

$$\delta\mathcal{L} = -\left(\overline{\psi}_L M_q \psi_R + \text{h.c.}\right) - e\mathcal{A}^\mu \left(\overline{\psi}_L Q_L \gamma_\mu \psi_L + \overline{\psi}_R Q_R \gamma_\mu \psi_R\right), \tag{5.17}$$

where

$$\psi \equiv \begin{pmatrix} u \\ d \\ s \end{pmatrix}, \quad Q_L \equiv Q_R \equiv Q_q. \tag{5.18}$$

From this we obtain the usual spurion rules

$$M_q \mapsto L M_q R^\dagger, \tag{5.19}$$

$$Q_L \mapsto L Q_L L^\dagger, \tag{5.20}$$

$$Q_R \mapsto R Q_R R^\dagger. \tag{5.21}$$

One can quickly see in the theory with the field ξ that the matrices M_q and Q_q always occur in the combinations

$$M_+ \equiv \frac{1}{2}\left(\xi^\dagger M_q \xi^\dagger + \xi M_q \xi\right), \qquad M_- \equiv \frac{i}{2}\left(\xi^\dagger M_q \xi^\dagger - \xi M_q \xi\right), \tag{5.22}$$

$$Q_+ \equiv \frac{1}{2}\left(\xi^\dagger Q_L \xi + \xi Q_R \xi^\dagger\right), \qquad Q_- \equiv \frac{i}{2}\left(\xi^\dagger Q_L \xi - \xi Q_R \xi^\dagger\right), \tag{5.23}$$

which are designed to be Hermitian, have definite parity properties as indicated by the subscript, and transform appropriately under $SU(3)_L \times SU(3)_R$:

$$X \mapsto UXU^\dagger, \text{ for } X = A^\mu, M_\pm, Q_\pm. \tag{5.24}$$

5.3.2 Lagrangian Terms

Now it is a simple matter to construct the most general Lagrangian. For example, the lowest-order terms in the meson Lagrangian respecting C, P, and T are

$$\mathcal{L}_\Pi = f^2 \operatorname{Tr}(A^\mu A_\mu) + a f^2 \Lambda_\chi \operatorname{Tr} M_+ + b \frac{\alpha}{4\pi} f^2 \Lambda_\chi^2 \operatorname{Tr} Q_+^2. \tag{5.25}$$

The remaining coefficients here and below are normalized to be dimensionless (once factors of the pion decay constant f and the chiral symmetry-breaking scale Λ_χ are extracted) and expected to be of order unity (as demanded by naturalness). The factor $\alpha/4\pi$ multiplying the electromagnetic term follows from the fact that such terms, if computed from the quark Lagrangian, would arise from photon loop diagrams.

This model is constructed to include all terms to two orders in the perturbation operators we have discussed, namely all terms with a total of two of the following operators: ∂^μ, M_q, and Q_q. Because the derivative operators ultimately generate meson masses and hence quark masses, the series should be thought of as one in quark masses and charges alone. The physics of the expansion becomes more lucid when we distinguish our results by the number of powers of m_s, $m_{u,d}$, and α_{EM}, namely, organizing them according to their $SU(3)$ and isospin-$SU(2)$ properties. Since we are representing strong and electromagnetic interactions only, we include only those terms that respect the same symmetries. Charge conjugation

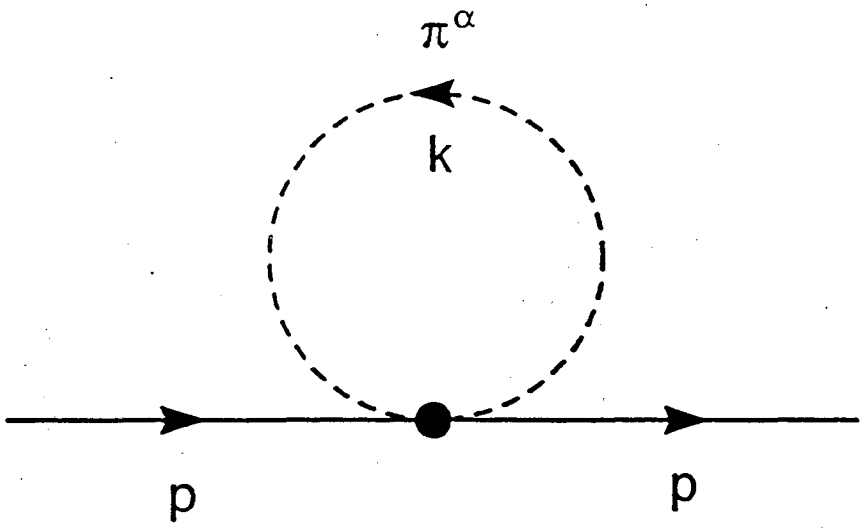


Figure 5.1: "Keyhole" (quartic vertex) diagram contributing to baryon masses.

symmetry eliminates all terms with an odd total number of $Q_\pm$, whereas parity conservation requires that M_- and Q_- only occur in an even total number (before we include derivative terms), that is, not before second order. Furthermore, since M_- and Q_-, when expanded in powers of the meson fields, give no contribution before $O(\Pi^2)$, these terms would not appear in mass relations until we include loops of *second*-order terms. Thus the parity-odd combinations do not occur at this order, and we may therefore suppress the subscript (+).

Next, all terms with derivatives produce factors of meson momenta and thus contribute to masses only through diagrams with meson loops. Because higher-order loops require more meson fields and hence produce more powers of the quark masses, we compute only those loop diagrams with one meson loop. Derivatives appear in the Lagrangian through the covariant derivative $\mathcal{D}^\mu(= \partial^\mu + O(\Pi^2))$ and through the axial vector current $A^\mu(= \partial^\mu\Pi/f + O(\Pi^3))$. Meson fields also occur in $M = M_q + O(\Pi^2)$ and $Q = Q_q + O(\Pi^2)$. We quickly learn that all one-meson loop diagrams are either of the "keyhole" variety (Fig. 5.1) or have two separated insertions of A^μ (Fig. 5.2).

There is a bewildering proliferation of terms at second order once we in-

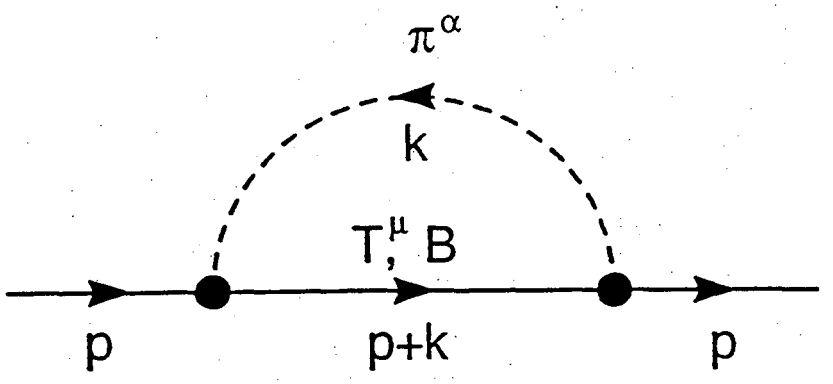


Figure 5.2: Trilinear vertex diagram contributing to baryon masses. The internal baryon line may be either octet or decuplet.

clude derivative terms, for example,

$$\overline{T}(v \cdot A)(S \cdot \mathcal{D})T, \qquad \overline{B}(v \cdot A)M_B, \tag{5.26}$$

and many others. Two constraints, however, simplify the situation: The first is that we are only computing diagrams with one meson loop, and the second is that terms with one covariant derivative $\mathcal{D}$ and some other operator X may be transformed away by a suitable redefinition of the baryon field and the addition of new $O(X^2)$ terms; as it stands, such terms contribute to baryon wavefunction renormalizations. The physical reasoning behind this transformation is that the only one-derivative terms we allow in the fermion Lagrangian are the kinetic terms; if we find other one-derivative terms, it means that we then have non-canonical equations of motion for the baryons and unfamiliar forms for the Feynman propagators. Under these transformations, the only remaining terms at second order and including meson fields are of the forms

$$\overline{T}AAT, \qquad \overline{T}\mathcal{D}\mathcal{D}T, \tag{5.27}$$

with appropriate factors of v and S thrown in to give the terms the correct Lorentz structure. Clearly, the only one-loop graphs possible from these terms are keyhole

diagrams; this proves to have special significance when we consider loop corrections to certain mass combinations.

It may seem confusing that we include *loop* effects of two-derivative operators, which are already at *second* order in the chiral expansion. However, we justify this inclusion by pointing out, using simple power-counting, that operators in Eq. 5.27 give rise to baryon mass corrections of the form

$$\delta M_B \approx \frac{m_\Pi^4}{16\pi^2 f^2 \Lambda_\chi} \ln m_\Pi^2 = O(m_q^2 \ln m_q), \tag{5.28}$$

which we show below, are of the same order in quark masses as one-loop diagrams with one insertion of M_q, and therefore must be considered in a consistent expansion in m_q.

The light-flavor symmetry $SU(3)_V$, which includes strong isospin, is broken in this model only by the inequality of quark masses and charges; all other operators and coefficients are assumed to obey chiral symmetry. Under these restrictions, the most general octet Lagrangian we need to consider in this model is

$$\begin{aligned}
\mathcal{L}_8 \;=\;& \mathrm{Tr}\overline{B}\,(iv\cdot\mathcal{D})\,B + 2D\,\mathrm{Tr}\overline{B}S^\mu\,[A_\mu, B]_+ + 2F\,\mathrm{Tr}\overline{B}S^\mu\,[A_\mu, B]_- \\
&+ 2\sigma\,(\mathrm{Tr}\,M)\,\mathrm{Tr}\overline{B}B + 2b_D\,\mathrm{Tr}\overline{B}\,[M,B]_+ + 2b_F\,\mathrm{Tr}\overline{B}\,[M,B]_- \\
&+ \frac{\alpha}{4\pi}\Lambda_\chi \Big\{ d_0(\mathrm{Tr}\,Q^2)\,\mathrm{Tr}\overline{B}B \\
&\qquad + d_D\mathrm{Tr}\overline{B}[Q^2,B]_+ + d_F\mathrm{Tr}\overline{B}[Q^2,B]_- \\
&\qquad + d_1\mathrm{Tr}\overline{B}QBQ + d_2(\mathrm{Tr}\overline{B}Q)(\mathrm{Tr}\,QB)\Big\} \\
&+ \frac{1}{\Lambda_\chi}\Big\{ \sigma_1\,(\mathrm{Tr}\,M^2)\,\mathrm{Tr}\overline{B}B + \sigma_2\,(\mathrm{Tr}\,M)^2\,\mathrm{Tr}\overline{B}B \\
&\qquad + \ell_D\,(\mathrm{Tr}\,M)\,\mathrm{Tr}\overline{B}\,[M,B]_+ + \ell_F\,(\mathrm{Tr}\,M)\,\mathrm{Tr}\overline{B}\,[M,B]_- \\
&\qquad + c_D\mathrm{Tr}\overline{B}[M^2,B]_+ + c_F\mathrm{Tr}\overline{B}[M^2,B]_- \\
&\qquad + c_1\mathrm{Tr}\overline{B}MBM + c_2(\mathrm{Tr}\overline{B}M)(\mathrm{Tr}\,MB)\Big\} \\
&+ \text{terms of the form of Eq. 5.27,}
\end{aligned} \tag{5.29}$$

whereas the corresponding decuplet Lagrangian is

$$\mathcal{L}_{10} \;=\; -i\overline{T}^\mu_{ijk} v_\nu\,(\mathcal{D}^\nu T_\mu)^{ijk} + \delta\overline{T}^\mu_{ijk} T^{ijk}_\mu + 2\mathcal{H}\overline{T}^\mu_{ijk} S_\nu\,(A^\nu)^k{}_l T^{ijl}_\mu$$

$$
\begin{aligned}
&- 2\tilde{\sigma}\,(\mathrm{Tr}\,M)\,\overline{T}^{\mu}_{ijk}T^{ijk}_{\mu} + 2c\,\overline{T}^{\mu}_{ijk}M^{k}{}_{l}T^{ijl}_{\mu} \\
&+ \frac{\alpha}{4\pi}\Lambda\chi\left\{f_0\,(\mathrm{Tr}\,Q^2)\,\overline{T}^{\mu}_{ijk}T^{ijk}_{\mu}\right. \\
&\qquad\qquad \left. + f_1\overline{T}^{\mu}_{ijk}Q^{k}{}_{l}Q^{l}_{m}T^{ijm}_{\mu} + f_2\overline{T}^{\mu}_{ijk}Q^{j}{}_{l}Q^{k}{}_{m}T^{ilm}_{\mu}\right\} \\
&+ \frac{1}{\Lambda\chi}\left\{\tilde{\sigma}_1(\mathrm{Tr}\,M^2)\,\overline{T}^{\mu}_{ijk}T^{ijk}_{\mu} + \tilde{\sigma}_2\,(\mathrm{Tr}\,M)^2\,\overline{T}^{\mu}_{ijk}T^{ijk}_{\mu}\right. \\
&\qquad + e_0(\mathrm{Tr}\,M)\,\overline{T}^{\mu}_{ijk}M^{k}{}_{l}T^{ijl}_{\mu} \\
&\qquad\qquad \left. + e_1\overline{T}^{\mu}_{ijk}M^{k}{}_{l}M^{l}{}_{m}T^{ijm}_{\mu} + e_2\overline{T}^{\mu}_{ijk}M^{j}{}_{l}M^{k}{}_{m}T^{ilm}_{\mu}\right\} \\
&+ \text{terms of the form of Eq. 5.27.}
\end{aligned}
\tag{5.30}
$$

The sign of the kinetic term follows from the fact that the Rarita–Schwinger spinor solutions are spacelike. Note also that terms involving $\mathrm{Tr}\,Q = 0$ do not appear, a result which follows from the definition of $Q = Q_+$ in Eq. 5.23 and $\mathrm{Tr}\,Q_q = 0$.

5.4 Parameter Counting

Naively, the octet Lagrangian $\mathcal{L}_8$ contains sixteen undetermined coefficients and the three quark mass parameters; the coefficients of the derivative terms, D and F, can be measured in baryonic decay processes. Thus the prospect of extracting useful information from the Lagrangian seems hopeless. The situation is not much better for the decuplet Lagrangian $\mathcal{L}_{10}$, where one counts ten undetermined coefficients in addition to the quark mass parameters, the decuplet-octet splitting δ, and the measurable derivative coefficient $\mathcal{H}$. Each Lagrangian also contains a host of two-derivative terms of the form of Eq. 5.27, which again could be measured in principle if we possessed a sufficiently precise set of data from baryon-meson scattering; as it stands, however, only D and F, and to a lesser accuracy $\mathcal{H}$, are currently measured. In addition, the decuplet and octet are connected at lowest order by the Lagrangian term

$$
\delta\mathcal{L} = \mathcal{C}\left(\epsilon^{ilm}\overline{T}^{\mu}_{ijk}(A^{\mu})^{j}{}_{l}B^{k}{}_{m} + \text{h.c.}\right). \tag{5.31}
$$

Numerically, it is believed that the nonanalytic loop corrections to D and F are not excessively large [88], and so one may use the lowest-order fit values of these

parameters (the uncertainties are estimates of higher-order corrections):

$$D = 0.85 \pm 0.06, \qquad F = 0.52 \pm 0.04, \tag{5.32}$$

whereas the most current values from decuplet strong decays and loop effects, respectively, are [89]

$$|\mathcal{C}| = 1.2 \pm 0.1, \qquad \mathcal{H} = -2.2 \pm 0.6. \tag{5.33}$$

There still remains the problem of a multitude of undetermined coefficients. As one might expect, many of these turn out to be redundant, as we now show; the following methods may be thought of as using "poor folks' group theory" to determine the independent parameters of the Lagrangian, as opposed to the full treatment presented in Chap. 4. Let us first consider terms that appear as singlets in the Lagrangian. In the usual theories with massive baryons, the octet and decuplet mass terms are just

$$\delta\mathcal{L}_8 = -m_8 \overline{B}^i{}_j B^j{}_i, \tag{5.34}$$

$$\delta\mathcal{L}_{10} = +m_{10} \overline{T}^\mu_{ijk} T^{ijk}_\mu. \tag{5.35}$$

These terms do not occur in HBEFT, for we have eliminated the common mass terms through field redefinition; only the term proportional to δ remains. If a term that is of the same form appears at higher order in perturbation theory, that is, an $SU(3)$ singlet multiplying the fully contracted $\overline{B}B$ or $\overline{T}T$ fields, it too contributes only to the overall multiplet mass, and hence may have been defined away at the outset. Thus singlets may be neglected in this theory; the coefficients σ, d_0, $\sigma_{1,2}$ in $\mathcal{L}_8$ and $\tilde{\sigma}$, f_0, and $\tilde{\sigma}_{1,2}$ are ignorable in this model.

The parameter δ itself is an exception. If we worked within a theory in which only one multiplet were present, we would simply transform away δ along with all other singlets; however, this model includes both octet and decuplet states [85, 86]. From the theoretical point of view, both multiplets arise as ground states of the quark-model $SU(3)$ product $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$ and are equally fundamental, whereas the physical point of view also suggests incorporating both because the splitting of the multiplets is not large ($\delta/\Lambda_\chi \approx 0.3$), so that intermediate decuplet states are

not numerically suppressed. Either way, both multiplets should be included in the theory.

The isolation of singlets is just the first step in a formal group-theoretical reduction of couplings through eliminating traces of operators. For example, the next step is to realize that the operators with coefficients $\ell_{D,F}$ and e_0 may be absorbed into those with coefficients $b_{D,F}$ and c, respectively. As a fine point, the traces also include field-dependent pieces, but to consistent order in this model, only the tree-level portions of operators second order in M and Q appear.

The baryon Lagrangian at second order is also simplified by an application of the Cayley–Hamilton theorem. The theorem assumes a particularly elegant form in $SU(3)$ for any traceless 3×3 matrix X, given only that B and $\overline{B}$ are also 3×3 and traceless:

$$\mathrm{Tr}\,\overline{B}[X^2, B]_+ + \mathrm{Tr}\,\overline{B}XBX - (\mathrm{Tr}\,\overline{B}X)(\mathrm{Tr}\,XB) - \frac{1}{2}(\mathrm{Tr}\,X^2)\,\mathrm{Tr}\,\overline{B}B = 0. \qquad (5.36)$$

Inasmuch as Q is traceless, we immediately see that one of the terms with coefficient $d_{D,1,2}$, say d_2, may be absorbed by the others and an ignorable d_0 piece. Likewise, all factors of $\mathrm{Tr}\,M$ may first be extracted from $c_{D,1,2}$ terms and absorbed into $b_{D,F}$ and singlet terms; then Eq. 5.36 eliminates one of the terms (*e.g.*, c_2) second order in the *traceless* part of M.

Now note that the operators associated with b_D, c_D, d_D and b_F, c_F, d_F all share the same structure (in the notation of Chap. 4, those subscripted with $D(F)$ transform according to the $\mathbf{8}_{1(2)}$ representation). Thus we may replace them with the operators $X_{D(F)}$, defined through

$$\begin{aligned} \delta\mathcal{L}_{8\,D(F)} &= \mathrm{Tr}\overline{B}[X_{D(F)}, B]_\pm, \\ &\equiv \mathrm{Tr}\overline{B}\left[\left(2b_{D(F)}M + \frac{\alpha}{4\pi}\Lambda_\chi d_{D(F)}Q^2 + \frac{1}{\Lambda_\chi}c_{D(F)}M^2\right), B\right]_\pm . \end{aligned} \qquad (5.37)$$

As before, the singlet (trace) portion of this operator simply contributes to the common octet mass; off-diagonal entries (which appear in loop effects) are $SU(3)$ ladder operators, which do not create mass terms at tree level. Thus we are left only with operators proportional to the Cartan subalgebra of isospin and hypercharge. No matter how many operators we have of the form of $X_{D(F)}$, they can

always be redefined into only two for each of $X_{D(F)}$, one proportional to T^3 and one proportional to T^8.

The decuplet Lagrangian also admits a similar simplification. Consider any 3×3 matrix X as an operator in the mass term

$$\overline{T}^{\mu}_{ijk} X^k{}_l T^{ijl}_{\mu}. \tag{5.38}$$

Because T is symmetric in its flavor indices, the placement of the contraction is irrelevant. Precisely the same argument as before follows, so the set of all operators X may be redefined into only two. In the decuplet case, these operators obey equal-spacing rules in isospin and hypercharge. We find three such operators in $\mathcal{L}_{10}$, those with coefficients c, f_1, and e_1.

Finally, recall our observations on quark mass parameters from Sec. 4.4. There we learned that $SU(3)$ chiral Lagrangians possess only two quark mass parameters, q and r, and we not only showed that q in HBEFT occurs only in loop diagrams, but anticipated the result (demonstrated in Sec. 5.5) that we are able to remove all explicit q dependence from our expressions. Therefore, only r holds importance for us.

We are again ready to count parameters. For the octet, the only remaining parameters are two associated with $X_{D,F}$ (two each), c_1, d_1, and the parameter r, making seven. As for known quantities, we begin with eight octet masses; but again, HBEFT tells us that we can discover nothing of the common multiplet mass, leaving seven mass differences. In addition, the Σ^0-Λ mixing angle θ is in principle a measurable quantity; its determination depends upon the discovery of an interaction that produces neutral baryon pure isospin eigenstates [90]. Thus we have eight "known" quantities, and predict one relation, which turns out to be the Coleman–Glashow relation (Eq. 4.8, which is fortuitously independent of θ).

For the decuplet, there are the two operators proportional to isospin and hypercharge, one quark mass parameter r, and one remaining $O(M^2)$ and one $O(Q^2)$ term, making five unknowns. On the other side of the equation, we have nine mass differences, implying that there are four nontrivial mass relations between the decuplet baryons (Eqs. 4.9–4.12).

This construction teaches us one more sobering fact, that one cannot determine the quark mass parameter r solely from the measured baryon masses. The electromagnetic terms labeled by $d_{D,F}$ for the octet and f_1 for the decuplet can be made to mimic exactly shifts in the $O(M)$ terms proportional to $c_{D,F}$ and c, respectively, using the redundancy of the operator forms $X_{D,F}$ and X discussed above. Thus the determination of r becomes tangled with a determination of the electromagnetic contributions to baryon mass differences. The computation of such terms has been attempted in the past [91], and will be reconsidered in a future publication [92].

5.5 Loop Corrections

We have seen in Sec. 4.3 that tree-level mass relations holding to second order in flavor breaking occur because operators in certain representations of $SU(3)$ do not arise at this order. For the octet, only operators transforming under the **10** (or equivalently, the $\overline{\mathbf{10}}$) break the Coleman–Glashow relation (Eq. 4.8), whereas for the decuplet, only the four components of the **64** representation break Eqs. 4.9–4.12. Relations that hold to second order in quark masses are particularly interesting, because any loop corrections that have the form $O(m_q^0, m_q^1,$ or $m_q^2)$ are absorbed through the usual renormalization procedure by the appropriate counterterms, and therefore must also satisfy the relations. Furthermore, it has long been known that the leading corrections to lowest-order predictions in a spontaneously broken chiral theory are nonanalytic in the chiral limit, owing to the presence of massless Nambu–Goldstone bosons in the intermediate states; in particular, the leading corrections to baryon masses are of the forms $O(m_q^{3/2})$ and $O(m_q^2 \ln m_q)$ [93]. As a consequence, the loop corrections below $O(m_q^3)$ are calculable and nonanalytic in quark masses, and independent of the arbitrary renormalization point μ (which in this theory appears only in logarithms multiplying analytic terms). In one-loop diagrams involving less than two powers of M, all corrections are of this form. Thus, in this model, such diagrams give calculable corrections to $\Delta_{\rm CG}$ or $\Delta_{1,2,3,4}$ if and only if they contain a piece transforming under a **10** or **64**, respectively. Here we wish to consider the

various one-meson loop corrections to baryon masses and classify them by their transformation properties.

The one-loop diagrams fall into two basic categories: those with one quartic vertex (Fig. 5.1), which we have called keyhole diagrams, and those with two trilinear vertices (Fig. 5.2). Keyhole diagrams in this model occur in two forms: those from the meson-field expansion of the operator M (which transforms as singlet plus octet, and therefore its contributions satisfy the relations), and those with the structure

$$f(\alpha)\overline{T}\,\Pi^\alpha\Pi^\alpha T, \qquad f(\alpha)\overline{B}\,\Pi^\alpha\Pi^\alpha B, \tag{5.39}$$

where flavor indices can appear in all possible contractions, α is the meson octet index, and $f(\alpha)$ is the function that appears upon integration over meson momentum. That the two meson field indices are the same simply implies that the meson loop must close; π^0-η mixing does not change this result, because we may rediagonalize the meson $SU(3)$ generators so that they refer now to mass eigenstates. All that is important is that both the unrotated and rotated generators are octet. Because the largest representation obtained from the two octet fields is a **27**, such loop expressions for the decuplet do not contribute to $\Delta_{1,2,3,4}$; because no **10** appears in the $\mathbf{8}\otimes\mathbf{8}$ product of two *identical* octet operators (Sec. 4.3), such loop expressions for the octet contribute neither to $\Delta_{\rm CG}$. Note also that this result holds for *each* value of α, not just the sum. Hence all keyhole diagrams respect the second-order relations.

This result has valuable consequences for the Gell-Mann–Okubo relation, Eq. 4.6. It is a first-order relation, nonvanishing in the isospin limit, and thus has $O(m_s^2)$ tree-level corrections. Jenkins [74] computed nonanalytic corrections to this famous relation in HBEFT, and found $O(m_s^{3/2})$ corrections of about 15 MeV and $O(m_s^2 \ln m_s)$ corrections of order 5 MeV. We now demonstrate the appearance of additional corrections of $O(m_s^2 \ln m_s)$, which were not computed in that work and cannot at present be evaluated. In Sec. 4.3 we saw that the Gell-Mann–Okubo relation is broken by operators transforming under a **27**. As we see from Eq. 5.39, some two-derivative operators in Eq. 5.27 are of this form, and as discussed in Sec.

5.4, their coefficients are not currently measured. Furthermore, as pointed out in Eq. 5.28, power-counting shows the baryon mass contributions from such terms to have the same dependence, $O(m_q^2 \ln m_q)$, as those computed in Ref. [74], and these terms clearly do not vanish when we set $m_u = m_d = 0$. In any case, they are not expected to be much larger than the tree-level $O(m_s^2)$ corrections, which may be estimated by naive power-counting at $O(20\text{ MeV})$. Since experimentally $\Delta_{\text{GMO}} \approx +6.5$ MeV, we see that these numbers are in satisfactory agreement with nature, but not precise enough to make any predictions.

Next consider loop diagrams with two trilinear vertices. In this model, the intermediate state may be either octet or decuplet, the mixing made possible through the inclusion of the interaction Eq. 5.31; the resultant mixed diagrams are labeled by *trans*. First, however, consider the case of what we call *cis* diagrams, in which the intermediate and external baryons are in the same multiplet. The general structure of *cis* diagrams is

$$\begin{aligned} (D^2, DF, F^2) \quad &\times \quad f(\alpha) \sum_j \left(\overline{B}\,\Pi^\alpha B_j\right) \mathcal{O}_j \left(\overline{B}_j \Pi^\alpha B\right), \\ \mathcal{H}^2 \quad &\times \quad f(\alpha) \sum_j \left(\overline{T}\,\Pi^\alpha T_j\right) \mathcal{O}_j \left(\overline{T}_j \Pi^\alpha T\right), \end{aligned} \tag{5.40}$$

where j refers to the allowed intermediate baryon states, and $\mathcal{O}$ is an operator that may appear on the intermediate line (for example, the $O(M)$ tree-level term). Now if $\mathcal{O}$ is arbitrary, there is little that can be said using only group theory. However, if we consider only those diagrams for which $\mathcal{O} = \mathbf{1}$, then we see that the j-dependence is trivial, for then we have completeness relations over baryon generators:

$$\sum_j B_j \overline{B}_j \propto \mathbf{1}, \qquad \sum_j T_j \overline{T}_j \propto \mathbf{1}. \tag{5.41}$$

We find that such diagrams have exactly the same group-theoretic structure as in Eq. 5.39 and therefore also do not contribute to Δ_{CG} or $\Delta_{1,2,3,4}$. Since each loop contributes a factor of $(16\pi^2 f^2)^{-1}$, dimensional analysis shows these loops to have quark mass dependence $O(m_q^{3/2})$.

The *trans* diagrams have an analogous property: If the baryons of the internal multiplet are taken degenerate in mass (but the masses of the external

multiplet need not be constrained), we again have a completeness sum over internal multiplet generators, and conclude that diagrams with $\mathcal{O} = \mathbf{1}$ satisfy the relations for the external multiplet. For example, decuplet mass diagrams with internal octet lines do not contribute to $\Delta_{1,2,3,4}$ if the octet baryons are all taken to have the same mass; equivalently, the *trans*-decuplet diagrams break the decuplet relations only through octet baryon mass *differences*.

However, there is a complication in the presence of the intermultiplet spacing δ. Neglecting the mass splittings within the internal multiplet seems unreasonable when many of these splittings are comparable in size to δ. The most thorough means of including this information would be to solve simultaneously for both the octet and decuplet chiral coefficients using all experimental baryon mass values, and to include the corresponding operators $\mathcal{O}$ in the loop diagrams. However, it is clearly more direct and convenient to use instead the physical masses for the internal multiplet.

The remaining *cis* diagrams (those with $\mathcal{O} \neq \mathbf{1}$) may also be rewritten in terms of the corresponding internal baryon masses, rather than the chiral coefficients; because the loop coefficients are proportional to powers of quark masses and charges, the difference between using chiral coefficients and baryon masses within loops is higher order still (at least $O(m_q^3 \ln m_q)$) in quark masses and charges. It is thus not unreasonable to present not only *trans* but also *cis* loop expressions in this manner.

5.6 Method of Calculation

Here we present important details in the calculation of loop diagrams to enable the reader to understand the nature of the computations. As discussed in Sec. 5.5, we need to compute only those diagrams with two trilinear vertices, since keyhole diagrams respect the relations Eqs. 4.8–4.12. The mass contribution, as computed using the Feynman rules of HBEFT as outlined in Ref. [86], is obtained

from the generic loop expression

$$\Sigma(p_i \cdot v) \;=\; \frac{i}{f^2}(Clebsch)^2_{i,j,\alpha} \left\{ \begin{array}{c} \overline{T}^\mu \mathcal{O}_\mu{}^\nu T_\nu \\ \overline{B}\mathcal{O}_\mu{}^\mu B \end{array} \right\}, \tag{5.42}$$

with

$$\mathcal{O}_\mu{}^\nu \equiv \int \frac{d^4k}{(2\pi)^4} \frac{1}{(k+p_i)\cdot v - \delta_j} \frac{1}{k^2 - m_\alpha^2} k_\mu k^\nu, \tag{5.43}$$

where the index i refers to a particular external baryon state, j refers to the intermediate baryon, and α refers to the loop meson, with an implicit sum over the last two indices. The external momentum p_i is residual, the common octet mass m_8 having been removed through HBEFT. The group-theoretical couplings *Clebsch* are readily computed by constructing matrix representations of the generators for the octet meson and baryon, and decuplet baryon states (as in Eqs. 5.7–5.9) and taking appropriate traces. Furthermore, factors from contraction of spin and projection operators and the couplings D^2, DF, F^2, $\mathcal{H}^2$, and $\mathcal{C}^2$ are suppressed.

The term δ_j is designed to implement our program of including full physical baryon masses instead of chiral coefficients for internal lines; it represents the amount by which the mass of baryon j exceeds the common octet mass m_8. If we worked only with chiral coefficients, this term would read $0, \delta$ for octet(decuplet) intermediates. Also, in the rest frame, $p_i^\mu = (\delta_i, \mathbf{0})$ and $v^\mu = (1, \mathbf{0})$, so that $p_i \cdot v = \delta_i$ in all frames. As a result, we find that the integral depends on baryon masses through their differences only, for $p_i \cdot v - \delta_j = m_i - m_j$. One further point is that the mass contribution for octet(decuplet) baryons issue from $\pm\Sigma(p_i \cdot v)$ with the external momentum as chosen above, where the sign difference between the two cases comes from the fact that the spin-3/2 kinetic term has the opposite sign to that for the usual spin-1/2 case.

There is a small amount of sleight of hand here, for had we used the masses of intermediate lines in *cis* diagrams from the outset, we would not have obtained the cancellation of the diagrams as described in Sec. 5.5 with $\mathcal{O} = \mathbf{1}$ (the $O(m_q^{3/2})$ corrections). Instead, inclusion of the full decuplet masses would mix the $\mathcal{O} = \mathbf{1}$ terms with all the others. But this is exactly what we do with the *trans* diagrams. The difference is the presence of the octet-decuplet splitting parameter δ

in the latter. Although experimentally *intra*multiplet splittings may be of the same magnitude as *inter*multiplet splittings, they are formally two different phenomena in the chiral Lagrangian, where the octet and decuplet are taken as independent. As a result, the functional forms of *trans* diagrams are more complicated.

We now present the expressions for contributions to baryon mass m_i that do not vanish in the relations Eqs. 4.8–4.12. For *cis* diagrams,

$$\begin{aligned} \delta m_i &= (Clebsch)^2_{i,j,\alpha} \frac{1}{16\pi^2 f^2} m_\alpha^2 \ln\left(\frac{m_\alpha^2}{\mu^2}\right)(m_j - m_i)\,, \\ &\times \begin{cases} \frac{5}{6}\mathcal{H}^2 & \text{for } cis\text{-decuplet diagrams} \\ \frac{3}{4}(D^2, DF, F^2) & \text{for } cis\text{-octet diagrams} \end{cases}, \end{aligned} \tag{5.44}$$

and for both *trans*-octet and *trans*-decuplet diagrams,

$$\delta m_i = \frac{1}{3}(Clebsch)^2_{i,j,\alpha} \frac{C^2}{16\pi f^2} m_\alpha^3 \; H(\xi_{ij}^\alpha), \tag{5.45}$$

with the notation

$$\xi_{ij}^\alpha \equiv \frac{m_i - m_j}{m_\alpha}, \tag{5.46}$$

$$\begin{aligned} F(\xi) &\equiv \int_0^\infty dx \left(x^2 + 2\xi x + 1\right)^{-1} \\ &= \begin{cases} \frac{1}{\sqrt{\xi^2-1}}(\operatorname{sgn}\xi)\cosh^{-1}|\xi|\,, & |\xi| > 1 \\ \frac{1}{\sqrt{1-\xi^2}}\cos^{-1}\xi, & |\xi| < 1 \end{cases}, \end{aligned} \tag{5.47}$$

and

$$H(\xi) \equiv \frac{3}{2\pi}\left[\xi\left(\frac{2}{3}\xi^2 - 1\right)\ln\frac{m_\alpha^2}{\mu^2} - \frac{4}{3}\left(\xi^2 - 1\right)^2 F(-\xi)\right]. \tag{5.48}$$

The explicit forms of these corrections are detailed in Appendix A (for $\Delta_{1,2,3,4}$) and Appendix B (for $\Delta_{\rm CG}$). Note that, as $\xi \to 0$ (corresponding to $\delta \to 0$), these corrections assume the standard forms $m_q^{3/2}$ (ξ^0 term) and $m_q^2 \ln m_q$ (ξ^1 term). We demonstrate these limiting forms explicitly for $\Delta_{\rm CG}$ in Appendix B.

The *cis*-decuplet ($\mathcal{H}^2$) and the F^2 *cis*-octet corrections to the relations Eqs. 4.8–4.12 possess a remarkable property: The portion linear in decuplet and octet masses in Eq. 5.44, for *each* meson α, may be written in terms of the $O(m_q^3)$

mass combinations $\Delta_{1,2,3,4}$ and $\Delta_{\rm CG}$, respectively. Consequently, although these corrections are *a priori* $O(m_q^2 \ln m_q)$, they are in fact $O(m_q^4 \ln m_q)$, and are thus formally smaller than the third-order tree-level terms we are neglecting. This cancellation occurs because the meson operators in the $\mathcal{H}$ and F terms, for each meson α, have the form of infinitesimal generators of group rotations; they rotate the **64** and **10** irreducible representations only into themselves, preserving the combinations of baryon masses ($\Delta_{1,2,3,4}$ and $\Delta_{\rm CG}$) indicative of these representations. On the other hand, the D term acts through an anticommutator, which does not generate $SU(3)$ rotations, and hence does not preserve the same baryon mass combinations. As a result, $\Delta_{1,2,3,4}$ have no $\mathcal{H}^2$ terms in this model; nor has $\Delta_{\rm CG}$ any F^2 terms, but DF and D^2 contributions still appear.

5.7 Results and Predictions

5.7.1 Estimating Parameters

The expressions derived in the previous Section and exhibited in the Appendices generally depend upon the decuplet and octet baryon masses, the meson octet masses, the axial-current couplings in Eqs. 5.32 and 5.36, and the quark mass parameter r. As pointed out in the Appendices, this last factor arises from a consistent treatment of π^0-η and Σ^0-Λ mixing; however, as uncovered in Sec. 5.4, r cannot be determined at present. To accommodate all currently suggested up-quark masses, we will adopt the range $0.025 \leq r \leq 0.043$, which allows $m_u = 0$ (upper value) [76, 77] as well as the value from lowest-order chiral perturbation theory (lower value) [94].

In order to judge the reliability of the predictions to follow, we must also be able to estimate the size of Lagrangian terms neglected in this model, namely two-loop effects and third-order tree-level terms. Beyond what we have calculated, the next contributions are two-loop effects of formal orders $m_q^{5/2}$ and $m_q^3 \ln m_q$, and then the tree-level terms at $O(m_q^3)$. We expect the two-loop corrections to be numerically small compared to the one-loop contributions, so we are led to consider the size of

third-order contributions to $\Delta_{\rm CG}$ and $\Delta_{1,2,3,4}$. For the decuplet, there are only two nontrivial such terms,

$$\frac{1}{\Lambda_\chi}\overline{T}^\mu_{ijk}M^i{}_l M^j{}_m M^k{}_n T^{lmn}_\mu, \qquad \frac{\alpha}{4\pi}\overline{T}^\mu_{ijk}M^i{}_l Q^j{}_m Q^k{}_n T^{lmn}_\mu, \tag{5.49}$$

whereas $\Delta_{\rm CG}$ receives contributions from terms like

$$\mathrm{Tr}\overline{B}M^2BM, \qquad \mathrm{Tr}\overline{B}QMBQ. \tag{5.50}$$

We may now estimate the generic third-order contributions to the relations using the isospin transformation properties of the relevant operators, because the $\Delta I = 1$ portion of M_q is proportional to the small parameter r. With a unit-size chiral coefficient, $r \approx 0.03$, $\Lambda_\chi \approx 1$ GeV, $m_s \approx 200$ MeV, we have the following estimates for $O(M^3)$ and $O(MQ^2)$ contributions, respectively (in MeV):

$$\begin{array}{lrr} \Delta I = 3: & 2\cdot 10^{-4}, & 3\cdot 10^{-3}, \\ \Delta I = 2: & 7\cdot 10^{-3}, & 0.1, \\ \Delta I = 1: & 0.2, & 0.1, \\ \Delta I = 0: & 8, & 0.1. \end{array} \tag{5.51}$$

5.7.2 Decuplet Predictions: $\Delta_{1,2,3,4}$

We consider now numerical corrections to the decuplet relations Eqs. 4.9–4.12. In Sec. 5.6 we found that no *cis*-decuplet ($\mathcal{H}^2$) diagrams contribute until $O(m_q^4 \ln m_q)$; this is fortunate, considering the large current uncertainties on $\mathcal{H}$ (Eq. 5.33). Furthermore, we found in App. A that *trans*-decuplet ($\mathcal{C}^2$) diagrams do not contribute to Δ_1 until $O(m_q^3 \ln m_q)$, and thus are negligible in this model. We therefore predict

$$\Delta_1 \equiv \Delta^{++} - 3\Delta^+ + 3\Delta^0 - \Delta^- = 0, \tag{5.52}$$

which remains unbroken at least until two-loop effects of $O(m_q^{5/2})$. As discussed in Sec. 4.5, our choice makes the Δ_1 constraint a prediction of the Δ^- mass, which we eliminate from the other three relations. Thus we also predict $(\Delta_2 + \Delta_1)$, Δ_3, and $(\Delta_4 + \frac{1}{4}\Delta_1)$ and compare them against their experimental values.

The results are presented as follows: For each set of Δ masses defined in Sec. 4.5, we first obtain a prediction for the mass of the Δ^-. In the subsequent pairs of expressions, the first line indicates which combination of decuplet masses is considered and its experimentally measured value. The second line exhibits the computed value of corrections to this combination, after the substitution of hadron masses alone (left side) and after the substitution of the measured values of $\mathcal{C}$ (Eq. 5.33) and r as estimated above (right side). Experimental uncertainties of the decuplet masses are included. All numbers are in MeV. For data set A,

$$\Delta^- = 1226.42 \pm 4.23; \tag{5.53}$$

$$\left.\begin{aligned} \Delta_2 + \Delta_1 &= -16.24 \pm 7.01, \\ \mathcal{C}^2\left[(+1.26 - 1.4r) \pm 1.74\right] &= +1.88 \pm 2.53; \end{aligned}\right\} \tag{5.54}$$

$$\left.\begin{aligned} \Delta_3 &= +2.68 \pm 1.69, \\ \mathcal{C}^2\left[(-0.01 + 20.5r) \pm 0.43\right] &= +0.87 \pm 0.87; \end{aligned}\right\} \tag{5.55}$$

$$\left.\begin{aligned} \Delta_4 + \tfrac{1}{4}\Delta_1 &= +5.47 \pm 1.75, \\ \mathcal{C}^2\left[(-0.61 - 0.4r) \pm 0.25\right] &= -0.89 \pm 0.39; \end{aligned}\right\} \tag{5.56}$$

and for data set B,

$$\Delta^- = 1232.84 \pm 1.81; \tag{5.57}$$

$$\left.\begin{aligned} \Delta_2 + \Delta_1 &= -5.16 \pm 4.50, \\ \mathcal{C}^2\left[(-0.70 + 0.5r) \pm 0.90\right] &= -0.98 \pm 1.30; \end{aligned}\right\} \tag{5.58}$$

$$\left.\begin{aligned} \Delta_3 &= +0.54 \pm 1.11, \\ \mathcal{C}^2\left[(+0.19 - 18.9r) \pm 0.34\right] &= -0.54 \pm 0.74; \end{aligned}\right\} \tag{5.59}$$

$$\left.\begin{aligned} \Delta_4 + \tfrac{1}{4}\Delta_1 &= +5.91 \pm 1.69, \\ \mathcal{C}^2\left[(+1.85 - 0.4r) \pm 0.59\right] &= +2.64 \pm 0.95. \end{aligned}\right\} \tag{5.60}$$

We see that, for both cases, the experimental values of $(\Delta_2 + \Delta_1)$, Δ_3, and $(\Delta_4 + \frac{1}{4}\Delta_1)$ are numerically roughly comparable to their calculated loop corrections. Compared with other chiral loop calculations, the size of these corrections is surprisingly small. As to the particular results, set B appears to give a better

fit, but to be certain we must refer to the estimates of higher-order corrections in the theory (Eq. 5.51). Noting the isospin properties of the combinations $\Delta_{1,2,3,4}$ in Eqs. 4.9–4.12, we see that these contributions are indeed of the right numerical orders to explain the differences between the one-loop contributions and the experimental breakings, with the exception of the quantity $(\Delta_2 + \Delta_1)$ for set A; this leads us to favor set B. Another distinction between the two sets is the prediction for the Δ^- mass; once it is measured, it will probe the validity of the prediction of this model that $\Delta_1 = 0$ to $O(m_q^{5/2})$.

5.7.3 Octet Prediction: $\Delta_{\rm CG}$

The Coleman–Glashow relation is known to hold extremely well. The experimental value for its breaking is $\Delta_{\rm CG} = -0.3 \pm 0.6$ MeV. Substituting hadron masses into the expressions in App. B, we find

$$\Delta_{\rm CG}^{\rm theor} = -2.2D^2 + 1.3DF + C^2(0.5 \pm 0.5 + 8.1r) \text{ MeV}. \tag{5.61}$$

The numerical uncertainty in the coefficient of C^2 is due to the uncertainty in the decuplet isospin splittings. Using the values for D and F in Eq. 5.32 and the above bounds on r, we find

$$\Delta_{\rm CG}^{\rm theor} = 0.2 \pm 0.7 \text{ MeV}, \tag{5.62}$$

where the quoted error is dominated by the uncertainty of the decuplet isospin splittings. This prediction is in agreement with experiment, and the higher-order $(\Delta I = 1)$ corrections discussed above are about the same size. Note the substantial cancellation between the octet and decuplet contributions: The octet contribution alone would give $\Delta_{\rm CG}^{\rm theor} = -1.0 \pm 0.2$.

5.7.4 Octet Prediction: Δ_Σ

Finally, let us consider corrections to the Σ equal-spacing rule, Eq. 4.14. Corrections to this relation, as discussed in Sec. 4.3, arise necessarily as $\Delta I = 2$ operators, and thus in this model originate in the $O(Q^2)$ terms or have the coefficient

$r^2 \propto O((m_u - m_d)^2)$. The fact that the physical Σ^0 is not an isospin eigenstate actually does not alter the numerical result, for the mass and isospin eigenstates differ by an isosinglet admixture of $O(r)$, and under diagonalization of the mass matrix, the difference becomes $O(r^2)$.

Comparing the relative sizes of terms,

$$\Delta_\Sigma^{(r^2)} \approx r^2 \frac{m_s^2}{\Lambda_\chi} \approx 0.04 \text{ MeV}, \tag{5.63}$$

whereas

$$\Delta_\Sigma^{(Q^2)} \approx \frac{\alpha}{4\pi} \Lambda_\chi \approx 0.5 \text{ MeV}. \tag{5.64}$$

We therefore conclude that Δ_Σ is utterly dominated by the electromagnetic contribution. Because experimentally $\Delta_\Sigma = 1.7 \pm 0.2$ MeV, the electromagnetic term is the right order of magnitude to naturally explain the measured value.

5.8 Conclusions

Let us now summarize our findings of the past two Chapters. In Chap. 4 we learned that the origin of baryon mass relations is determined by the presence or absence of operators transforming under various representations of $SU(3)$. We found which relations are expected to hold most accurately (*i.e.*, to highest order in chiral perturbation theory) and how to determine the chiral coefficients in general. In particular, we found four relations holding at second order in flavor breaking for the decuplet, and one (the Coleman–Glashow relation) for the octet. We also examined the issue of determining quark mass parameters, and found that only one (r) is in principle easy to determine.

In this Chapter, we reviewed a chiral Lagrangian formalism (HBEFT) that permitted us to perform reliable perturbative calculations with baryons. We constructed the second-order Lagrangian and proceeded to count independent parameters; after removing all redundancies, we found an alternate explanation for the number of relations found in Chap. 4. We then pointed out that r cannot be determined without information on electromagnetic mass contributions. Next we

considered one-loop corrections to the mass relations and found that many of them vanish for group-theoretical reasons, and that the Gell-Mann–Okubo relation has a number of $O(m_s^2 \ln m_s)$ loop corrections with unknown coefficients. The nonvanishing loop corrections to the second-order relations were then computed, numerically evaluated, and compared to experiment. We found, in addition to predictions for the (unmeasured) mass of the Δ^-, agreement within errors in all cases, especially when estimates of higher-order corrections were included. Finally, we found that corrections to the Σ equal-spacing rule are dominated by electromagnetic contributions.

Bibliography

[1] F. Gürsey and L. A. Radicati, Phys. Rev. Lett. **13**, 173 (1964); B. Sakita, Phys. Rev. **136B**, 1756 (1964).

[2] S. Coleman and J. Mandula, Phys. Rev. **159**, 1251 (1967).

[3] G. P. Lepage and B. A. Thacker, in *Field Theory on the Lattice*, Proceedings of the International Symposium, Seillar, France, 1987, edited by A. Billoire *et al.* [Nucl. Phys. B (Proc. Suppl.) **4**, 199 (1988)].

[4] W. E. Caswell and G. P. Lepage, Phys. Lett. B **167**, 437 (1986).

[5] H. D. Politzer and M. B. Wise, Phys. Lett. B **206**, 681 (1988); **208**, 504 (1988).

[6] N. Isgur and M. B. Wise, Phys. Lett. B **232**, 113 (1989); **237**, 527 (1990).

[7] E. Eichten and B. Hill. Phys. Lett. B **234**, 511 (1990).

[8] H. Georgi, Phys. Lett. B **240**, 447 (1990).

[9] S. Abachi *et al.* (D0 Collaboration), Fermilab preprint Fermilab Pub-94/004-E.

[10] M. Suzuki, Nucl. Phys. **B258**, 553 (1985).

[11] M. Wirbel, B. Stech, and M. Bauer, Z. Phys. C **29**, 637 (1985).

[12] B. Grinstein, M. B. Wise, and N. Isgur, Phys. Rev. Lett. **56**, 298 (1986); T. Altomari and L. Wolfenstein, Phys. Rev. Lett. **58**, 1583 (1987).

[13] M. B. Voloshin and M. A. Shifman, Yad. Fiz. **45**, 463 (1987) [Sov. J. Nucl. Phys. **45**, 292 (1987)]; **47**, 801 (1988) [**47**, 511 (1988)].

[14] N. Isgur, D. Scora, B. Grinstein, and M. B. Wise, Phys. Rev. D **39**, 799 (1989).

[15] N. Isgur and M. B. Wise, Nucl. Phys. **B348**, 276 (1991); H. Georgi, Nucl. Phys. **B348**, 293 (1991).

[16] A. Falk, H. Georgi, B. Grinstein, and M. B. Wise, Nucl. Phys. **B343**, 1 (1990).

[17] A. Falk and B. Grinstein, Phys. Lett. B **247**, 406 (1990).

[18] M. E. Luke, Phys. Lett. B **252**, 447 (1990).

[19] A. Falk, B. Grinstein, and M. E. Luke, Nucl. Phys. **B357**, 185 (1991).

[20] M. Golden and B. Hill, Phys. Lett. B **254**, 225 (1991).

[21] J. M. Flynn, O. F. Hernandez, and B. R. Hill, Phys. Rev. D **43**, 3709 (1991).

[22] C. G. Boyd and D. E. Brahm, Phys. Lett. B **257**, 393 (1991).

[23] M. Ademollo and R. Gatto, Phys. Rev. Lett. **13**, 264 (1964).

[24] R. F. Lebed and M. Suzuki, Phys. Rev. D **44**, 829 (1991).

[25] A. Pais and G. E. Uhlenbeck, Phys. Rev. **79**, 145 (1950).

[26] P. A. M. Dirac, *Lectures on Quantum Mechanics* (Belfer Graduate School, Yeshiva University, New York, 1964).

[27] H. D. Politzer, Phys. Lett. B **250**, 128 (1990).

[28] Y. Nambu, in *New Theories in Physics*, Proceedings of the XI Warsaw Symposium on Elementary Particle Physics, Kazimierz, Poland, 1988, edited by Z. A. Ajduk *et al.* (World Scientific, Singapore, 1989), p. 1; in *New Trends in Strong Coupling Gauge Theories*, Proceedings of the 1988 International Workshop , Nagoya, Japan, edited by M. Bando *et al.* (World Scientific, Singapore 1989), p. 2; in *Proceedings of the 1989 Workshop on Dynamical Symmetry Breaking*, edited by T. Muti and K. Yamawaki (Nagoya University, Nagoya, 1990), p. 2; University of Chicago Report No. EFI 89-08 (unpublished).

[29] V. A. Miransky, M. Tanabashi, and K. Yamawaki, Mod. Phys. Lett. A **4**, 1043 (1989); Phys. Lett. B **221**, 177 (1989).

[30] Y. Nambu and G. Jona-Lasinio, Phys. Rev. **122**, 345 (1960).

[31] W. A. Bardeen, C. T. Hill, and M. Lindner, Phys. Rev. D **41**, 1647 (1990).

[32] M. Suzuki, Mod. Phys. Lett. A **5**, 1205 (1990).

[33] A. Hasenfratz, P. Hasenfratz, K. Jansen, J. Kuti, and Y. Shen, Nucl. Phys. **B365**, 79 (1991).

[34] W. J. Marciano, Phys. Rev. Lett. **62**, 2793 (1989).

[35] W. A. Bardeen, in *Present and Future of High-Energy Physics*, Proceedings of the 5th Nishinomiya Yukawa Memorial Symposium, Nishinomiya City, Japan, 1990, edited by K. I. Aoki and M. Kobayashi (Springer-Verlag, New York 1992), p. 94.

[36] M. Suzuki, Phys. Rev. D **44**, 3628 (1991).

[37] R. F. Lebed and M. Suzuki, Phys. Rev. D **45**, 1744 (1992).

[38] M. Suzuki, Phys. Rev. D **41**, 3457 (1990).

[39] J. Frölich and L. Lavoura, Phys. Lett. B **253**, 218 (1991).

[40] M. Luty, Phys. Rev. D **41**, 2893 (1990).

[41] Y. Nagoshi, K. Nakanishi, and S. Tanaka, Prog. Theor. Phys. **85**, 131 (1991); M. Bando, T. Kugo, N. Sasakura, Y. Watabiki, and K. Suehiro, Phys. Lett. B **246**, 466 (1990); M. Bando, T. Kugo, and K. Suehiro, Prog. Theor. Phys. **85**, 1299 (1991).

[42] S. King and M. Suzuki, Phys. Lett. B **277**, 153 (1992).

[43] C. T. Hill, C. N. Leung, and S. Rao, Nucl. Phys. **B262**, 517 (1985).

[44] M. S. Chanowitz, J. Ellis, and M. K. Gaillard, Nucl. Phys. **B128**, 506 (1977); A. J. Buras, J. Ellis, M. K. Gaillard, and D. V. Nanopoulos, *ibid.* **B135**, 66 (1978).

[45] S. L. Glashow and S. Weinberg, Phys. Rev. D **15**, 1958 (1977).

[46] W. J. Marciano, Phys. Rev. D **41**, 219 (1990).

[47] A. Zee, Phys. Rep. **3C**, 127 (1972).

[48] Particle Data Group, *Review of Particle Properties*, Phys. Rev. D **45**, VII.1 (mesons) and VIII.1 (baryons) (1992).

[49] C. P. Singh, A. Sharma, and M. P. Khanna, Phys. Rev. D **24**, 788 (1981).

[50] L.-H. Chan, Phys. Rev. Lett. **51**, 253 (1983).

[51] D. Y. Kim and S. N. Sinha, Ann. Phys. (N.Y.) **42**, 47 (1985).

[52] D. Flamm, F. Schöberl, and H. Uematsu, Nuov. Cim. A **98**, 559 (1987).

[53] A. De Rújula, H. Georgi, and S. L. Glashow, Phys. Rev. D **12**, 147 (1975).

[54] K. P. Tiwari, C. P. Singh, and M. P. Khanna, Phys. Rev. D **31**, 642 (1985).

[55] J. L. Goity and W.-S. Hou, Phys. Lett. B **282**, 243 (1992).

[56] R. F. Lebed, Phys. Rev. D **47**, 1134 (1993).

[57] D. Bortoletto *et al.* (CLEO collaboration), Phys. Rev. Lett. **69**, 2046 (1992); M. S. Alam *et al.* (CLEO collaboration), Cornell Univ. preprint CLNS-94-1270; D. N. Brown *et al.* (CLEO collaboration), Cornell Univ. preprint CLNS-94-1271.

[58] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Plenum, New York, 1977).

[59] S. N. Gupta, *Quantum Electrodynamics* (Gordon and Breach, New York, 1977), pp. 191-206.

[60] M. Jacob and G. C. Wick, Ann. Phys. (N.Y.) **7**, 404 (1959).

[61] C. Quigg and J. L. Rosner, Phys. Rep. **56C**, 167 (1979).

[62] E. Eichten, K. Gottfried, T. Kinoshita, K. D. Lane, and T.-M. Yan, Phys. Rev. D **17**, 3090 (1978).

[63] M. G. Olsson, University of Wisconsin - Madison Report No. MAD/PH/656, June, 1991 (unpublished).

[64] Traditional references include:
J. Schwinger, Phys. Lett. B **24**, 473 (1967);
S. Weinberg, Phys. Rev. Lett. **18**, 188 (1967);
S. Coleman, J. Wess and B. Zumino, Phys. Rev. **177**, 2239 (1969);
C. G. Callan, S. Coleman, J. Wess, and B. Zumino, Phys. Rev. **177**, 2247 (1969);
R. Dashen and M. Weinstein, Phys. Rev. **183**, 1261 (1969);
L.-F. Li and H. Pagels, Phys. Rev. Lett. **26**, 1204 (1971) and **27**, 1089 (1971);
P. Langacker and H. Pagels, Phys. Rev. D **8**, 4595 (1973);
S. Weinberg, Physica **96A**, 327 (1979);
J. Gasser and H. Leutwyler, Ann. Phys. (N.Y.) **158**, 142 (1984).

[65] R. F. Lebed, Lawrence Berkeley Laboratory preprint LBL-34704 and U.C. Berkeley preprint UCB-PTH-93/27.

[66] R. F. Lebed and M. A. Luty, Lawrence Berkeley Laboratory preprint LBL-34779 and U.C. Berkeley preprint UCB-PTH-93/28.

[67] J. J. deSwart, Rev. Mod. Phys. **35**, 916 (1963).

[68] M. Gell-Mann, Phys. Rev. **125**, 1067 (1962); S. Okubo, Prog. Theor. Phys. **27**, 949 (1962).

[69] M. Gell-Mann, in *Proceedings of the 1962 High-Energy Conference at CERN*, edited by J. Prentki (CERN, Geneva, 1962), p. 805.

[70] S. Coleman and S. L. Glashow, Phys. Rev. Lett. **6**, 423 (1961).

[71] S. Ishida, K. Konno and H. Shimodaira, Nuov. Cim. A **46**, 194 (1966).

[72] A. Gal and F. Scheck, Nucl. Phys. **B2**, 110 (1967).

[73] S. Okubo, Phys. Lett. **4**, 14 (1963).

[74] E. Jenkins, Nucl. Phys. **B368**, 190 (1992).

[75] G. Morpurgo, Phys. Rev. D **40**, 2997 (1989); Phys. Rev. Lett. **68**, 139 (1992); Phys. Rev. D **45**, 1686 (1992).

[76] D. B. Kaplan and A. V. Manohar, Phys. Rev. Lett. **56**, 2004 (1986).

[77] H. Leutwyler, Nucl. Phys. **B337**, 108 (1990).

[78] N. V. Krasnikov, V. A. Matveev, and A. N. Tavkhelidze, Fiz. Elem. Chastits At. Yadra, **12**, 100 (1980) [Sov. J. Part. Nucl. **12**, 38 (1981)] is the first review article on strong CP violation; the first specific attempt to solve the $U(1)_A$ problem by setting $m_u = 0$ appears to be in A. Zapeda, Phys. Rev. Lett. **41**, 139 (1978), although earlier authors were aware of the possibility.

[79] H. Georgi and A. V. Manohar, Nucl. Phys. **B234**, 189 (1984).

[80] R. E. Cutkosky, Phys. Rev. C **47**, 367 (1992).

[81] R. A. Arndt, private communication.

[82] I. I. Miroshnichenko *et al.*, Yad. Fiz. **29**, 188 (1979) [Sov. J. Nucl. Phys. **29**, 94 (1979)].

[83] E. Pedroni *et al.*, Nucl. Phys. **A300**, 321 (1978).

[84] E. Jenkins and A. V. Manohar, Phys. Lett. B **255**, 558 (1991).

[85] E. Jenkins and A. V. Manohar, Phys. Lett. B **259**, 353 (1991).

[86] E. Jenkins and A. V. Manohar, in *Proceedings of the Workshop on Effective Field Theories of the Standard Model*, Dobogókő, Hungary, 1991, edited by U.-G. Meissner (World Scientific, River Edge, New Jersey, 1992), pp. 113–137.

[87] T. Mannel, W. Roberts, and Z. Ryzak, Nucl. Phys. **B368**, 204 (1992).

[88] M. A. Luty and M. White, Lawrence Berkeley Laboratory preprint LBL-34040 and Center for Particle Astrophysics preprint CfPA-TH-93-10, to be published in Phys. Lett. **B**.

[89] M. N. Butler, M. J. Savage and R. P. Springer, Nucl. Phys. **B399**, 69 (1993).

[90] J. Schechter and A. Subbaraman, Int. J. Mod. Phys. A **7**, 7135 (1992).

[91] J. Gasser, Ann. Phys. (N.Y.) **136**, 62 (1981); J. Gasser and H. Leutwyler, Phys. Rep. **87C**, 77 (1982).

[92] M. A. Luty and R. Sundrum, in preparation.

[93] L.-F. Li and H. Pagels, Phys. Rev. Lett. **26**, 1204 (1971); P. Langacker and H. Pagels, Phys. Rev. D **10**, 2904 (1974).

[94] J. Gasser and H. Leutwyler, Nucl. Phys. **B94**, 269 (1975); S. Weinberg, in *A Festschrift for I. I. Rabi*, edited by L. Motz (New York Acad. Sci. 1977), p. 185.

Appendix A

Loop Corrections: Decuplet

We here tabulate the breakings of the relations Eqs. 4.9–4.12, which have been labeled $\Delta_{1,2,3,4}$. In Sec. 5.6 we demonstrated that no *cis*-decuplet ($\mathcal{H}^2$) diagrams would contribute until $O(m_q^4 \ln m_q)$. Thus we need consider only *trans*-decuplet ($\mathcal{C}^2$) contributions. As it stands, however, the expressions are quite cumbersome unless we compactify our notation. We denote

$$\frac{m_\alpha^3}{16\pi f^2} H(\xi_{ij}^\alpha) \to \tilde{\alpha}(i,j), \tag{A.1}$$

where the indices are replaced with the particles they represent. We then have

$$\begin{aligned}
\Delta_1 = & \\
& \frac{1}{3}\mathcal{C}^2 \Big\{ +\tilde{\pi}^+ \Big((\Delta^{++},p) - (\Delta^+,n) + (\Delta^0,p) - (\Delta^-,n) \Big) \\
& \quad -2\tilde{\pi}^0 \Big((\Delta^+,p) - (\Delta^0,n) \Big) \\
& \quad +\tilde{K}^+ \Big((\Delta^{++},\Sigma^+) - 2(\Delta^+,\Sigma^0) + (\Delta^0,\Sigma^-) \Big) \\
& \quad -\tilde{K}^0 \Big((\Delta^+,\Sigma^+) - 2(\Delta^0,\Sigma^0) + (\Delta^-,\Sigma^-) \Big) \Big\},
\end{aligned} \tag{A.2}$$

$$\begin{aligned}
\Delta_2 = & \\
& \frac{1}{3}\mathcal{C}^2 \Big\{ +\frac{1}{3}\tilde{\pi}^+ \Big[3(\Delta^{++},p) - (\Delta^+,n) - (\Delta^0,p) + 3(\Delta^-,n) \\
& \quad - \Big(1 - \frac{3}{2}r\Big)(\Sigma^{*+},\Sigma^0) - 3\Big(1 - \frac{1}{2}r\Big)(\Sigma^{*+},\Lambda) \\
& \quad +2\Big((\Sigma^{*0},\Sigma^+) + (\Sigma^{*0},\Sigma^-) \Big)
\end{aligned}$$

$$
\begin{aligned}
&\quad - \left(1 - \frac{3}{2}r\right)(\Sigma^{*-}, \Sigma^0) - 3\left(1 + \frac{1}{2}r\right)(\Sigma^{*-}, \Lambda)\Big] \\
&-\frac{1}{3}\tilde{\pi}^0 \Big[2\big((\Delta^+, p) + (\Delta^0, n)\big) - 6(\Sigma^{*0}, \Lambda) \\
&\quad + \left(1 + \frac{3}{2}r\right)(\Sigma^{*+}, \Sigma^+) + \left(1 - \frac{3}{2}r\right)(\Sigma^{*-}, \Sigma^-)\Big] \\
&- \tilde{\eta} \quad \Big[-2(\Sigma^{*0}, \Sigma^0) + \left(1 - \frac{1}{2}r\right)(\Sigma^{*+}, \Sigma^+) + \left(1 + \frac{1}{2}r\right)(\Sigma^{*-}, \Sigma^-)\Big] \\
&+\frac{1}{3}\tilde{K}^+ \Big[3(\Delta^{++}, \Sigma^+) - 2(\Delta^+, \Sigma^0) - (\Delta^0, \Sigma^-) \\
&\quad -2\big((\Sigma^{*+}, \Xi^0) - (\Sigma^{*0}, p) - (\Sigma^{*0}, \Xi^-) + (\Sigma^{*-}, n)\big)\Big] \\
&-\frac{1}{3}\tilde{K}^0 \Big[(\Delta^+, \Sigma^+) + 2(\Delta^0, \Sigma^0) - 3(\Delta^-, \Sigma^-) \\
&\quad +2\big((\Sigma^{*+}, p) - (\Sigma^{*0}, n) - (\Sigma^{*0}, \Xi^0) + (\Sigma^{*-}, \Xi^-)\big)\Big]\Big\}, \qquad \text{(A.3)}
\end{aligned}
$$

$$
\begin{aligned}
&\Delta_3 = \\
&\frac{1}{3}\mathcal{C}^2 \Big\{+\frac{1}{6}\tilde{\pi}^+ \quad \Big[2\big((\Delta^+, n) - (\Delta^0, p) + (\Xi^{*0}, \Xi^-) - (\Xi^{*-}, \Xi^0)\big) \\
&\quad - \left(1 + \frac{3}{2}r\right)(\Sigma^{*+}, \Sigma^0) - 3\left(1 - \frac{1}{2}r\right)(\Sigma^{*+}, \Lambda) \\
&\quad + \left(1 - \frac{3}{2}r\right)(\Sigma^{*-}, \Sigma^0) + 3\left(1 + \frac{1}{2}r\right)(\Sigma^{*-}, \Lambda)\Big] \\
&+\frac{1}{6}\tilde{\pi}^0 \quad \Big[4\big((\Delta^+, p) - (\Delta^0, n)\big) \\
&\quad - \left(1 + \frac{3}{2}r\right)\big((\Sigma^{*+}, \Sigma^+) - (\Xi^{*0}, \Xi^0)\big) \\
&\quad + \left(1 - \frac{3}{2}r\right)\big((\Sigma^{*-}, \Sigma^-) - (\Xi^{*-}, \Xi^-)\big)\Big] \\
&+\frac{1}{2}\tilde{\eta} \quad \Big[-\left(1 - \frac{1}{2}r\right)\big((\Sigma^{*+}, \Sigma^+) - (\Xi^{*0}, \Xi^0)\big) \\
&\quad + \left(1 + \frac{1}{2}r\right)\big((\Sigma^{*-}, \Sigma^-) - (\Xi^{*-}, \Xi^-)\big)\Big] \\
&+\frac{1}{6}\tilde{K}^+ \Big[4(\Delta^+, \Sigma^0) + \\
&\quad -2\big((\Delta^0, \Sigma^-) + (\Sigma^{*+}, \Xi^0) - (\Sigma^{*-}, n) - (\Xi^{*0}, \Sigma^+)\big) \\
&\quad - \left(1 - \frac{3}{2}r\right)(\Xi^{*-}, \Sigma^0) - 3\left(1 + \frac{1}{2}r\right)(\Xi^{*-}, \Lambda)\Big] \\
&-\frac{1}{6}\tilde{K}^0 \Big[4(\Delta^0, \Sigma^0) - 2\big((\Delta^+, \Sigma^+) - (\Sigma^{*+}, p) + (\Sigma^{*-}, \Xi^-) - (\Xi^{*-}, \Sigma^-)\big) \\
&\quad - \left(1 + \frac{3}{2}r\right)(\Xi^{*0}, \Sigma^0) - 3\left(1 - \frac{1}{2}r\right)(\Xi^{*0}, \Lambda)\Big]\Big\}, \qquad \text{(A.4)}
\end{aligned}
$$

$$
\begin{aligned}
\Delta_4 = & \\
& \frac{1}{3}\mathcal{C}^2 \Big\{ +\frac{1}{12}\tilde{\pi}^+ \Big[3(\Delta^{++},p) + (\Delta^+,n) + (\Delta^0,p) + 3(\Delta^-,n) \\
& \qquad -2\left(1+\frac{3}{2}r\right)(\Sigma^{*+},\Sigma^0) - 6\left(1-\frac{1}{2}r\right)(\Sigma^{*+},\Lambda) \\
& \qquad -2\left((\Sigma^{*0},\Sigma^+) + (\Sigma^{*0},\Sigma^-)\right) \\
& \qquad -2\left(1-\frac{3}{2}r\right)(\Sigma^{*-},\Sigma^0) - 6\left(1+\frac{1}{2}r\right)(\Sigma^{*-},\Lambda) \\
& \qquad +6\left((\Xi^{*0},\Xi^-) + (\Xi^{*-},\Xi^0)\right)\Big] \\
& +\frac{1}{12}\tilde{\pi}^0 \Big[2\left((\Delta^+,p) + (\Delta^0,n)\right) - 6(\Sigma^{*0},\Lambda) \\
& \qquad -\left(1+\frac{3}{2}r\right)\left(+2(\Sigma^{*+},\Sigma^+) - 3(\Xi^{*0},\Xi^0)\right) \\
& \qquad -\left(1-\frac{3}{2}r\right)\left(+2(\Sigma^{*-},\Sigma^-) - 3(\Xi^{*-},\Xi^-)\right)\Big] \\
& +\frac{1}{4}\ \tilde{\eta}\ \Big[-2(\Sigma^{*0},\Sigma^0) \\
& \qquad -\left(1-\frac{1}{2}r\right)\left(+2(\Sigma^{*+},\Sigma^+) - 3(\Xi^{*0},\Xi^0)\right) \\
& \qquad -\left(1+\frac{1}{2}r\right)\left(+2(\Sigma^{*-},\Sigma^-) - 3(\Xi^{*-},\Xi^-)\right)\Big] \\
& +\frac{1}{12}\tilde{K}^+ \Big[3(\Delta^{++},\Sigma^+) + 2(\Delta^+,\Sigma^0) + (\Delta^0,\Sigma^-) \\
& \qquad -2\left(2(\Sigma^{*+},\Xi^0) + (\Sigma^{*0},p) + (\Sigma^{*0},\Xi^-) + 2(\Sigma^{*-},n)\right) \\
& \qquad +3\left(\left(1-\frac{3}{2}r\right)(\Xi^{*-},\Sigma^0) + 9\left(1+\frac{1}{2}r\right)(\Xi^{*-},\Lambda)\right) \\
& \qquad +6\left((\Xi^{*0},\Sigma^+) - 2(\Omega^-,\Xi^0)\right)\Big] \\
& +\frac{1}{12}\tilde{K}^0 \Big[(\Delta^+,\Sigma^+) + 2(\Delta^0,\Sigma^0) + 3(\Delta^-,\Sigma^-) \\
& \qquad -2\left(2(\Sigma^{*+},p) + (\Sigma^{*0},n) + (\Sigma^{*0},\Xi^0) + 2(\Sigma^{*-},\Xi^-)\right) \\
& \qquad +3\left(\left(1+\frac{3}{2}r\right)(\Xi^{*0},\Sigma^0) + 9\left(1-\frac{1}{2}r\right)(\Xi^{*0},\Lambda)\right) \\
& \qquad +6\left((\Xi^{*-},\Sigma^-) - 2(\Omega^-,\Xi^-)\right)\Big]\Big\}. \qquad \text{(A.5)}
\end{aligned}
$$

The factors of r arise from π^0-η and Σ^0-Λ mixing, which we incorporate into the calculation by rotating the $SU(3)$ generators to $O(r)$; this eliminates the mass-mixing terms to order $(O(r^2))$, consistent with the model. However, within baryon loop corrections, the $O(r^2)$ terms are numerically insignificant, so they have been

suppressed in the above expressions.

In Sec. 5.5 we indicated that the *trans*-decuplet contributions would vanish if we took the octet baryons degenerate in mass. As a special case to demonstrate this point, let us take the value of *all* decuplet-octet splittings to be δ; then $\xi_{ij}^{\alpha} = \delta/m_{\alpha}$ depends only on α, the meson index. We may then trivially verify that the coefficient of each $\tilde{\pi}_{\alpha}$ in the loop corrections vanishes if we replace each $\tilde{\alpha}(i,j)$ with the same factor $f(\alpha)$.

Upon expanding all meson and baryon masses in terms of quark masses and charges, we find that the $\mathcal{C}^2$ term is also formally too small ($O(m_q^3 \ln m_q)$) to keep in the current calculation, because there are two-loop effects at $O(m_q^{5/2})$ that have already been neglected. This cancellation is a result of the fact that Δ_1 originates in $\Delta I = 3$ terms, which require three powers of quark masses in this model; at least two of these factors in the loop contributions must come from the decuplet masses. Thus we find in this model that $\Delta_1 = 0$, so that the tree-level relation Eq. 4.9 remains uncorrected.

Appendix B

Loop Corrections: Octet

We here tabulate the breaking of the Coleman–Glashow relation (Eq. 4.8), labeled $\Delta_{\rm CG}$ in the text. We here employ a different format for expressing our results from that in Appendix A in order to explicitly isolate isospin-breaking factors, as described below.

The contribution from octet intermediate states (*cis*-octet diagrams) is computed to be

$$\Delta^8_{\rm CG} = \frac{(K^+)^2 - (K^0)^2}{8\pi^2 f^2}\left[D^2(\Xi - N) + 3DF(\Lambda - \Sigma)\right] + O(m^2_{u,d}). \tag{B.1}$$

This expression is $O(m_s(m_d - m_u))$; in particular, it is *analytic* in the quark masses, which arises as follows: The loop corrections to $\Delta_{\rm CG}$ are $O(m_s(m_d - m_u)\ln m_s)$, where the logarithm involves the renormalization scale μ. We now expand the logarithms in meson mass differences, giving rise to an analytic result; the key fact is that we are taking the difference of a nonanalytic function about two closely-spaced points, which turns out in our case to give an analytic result. Furthermore, because there are no $O(m_q^2)$ counterterms for $\Delta_{\rm CG}$, changing μ changes the result by $O(m_q^3)$. We may therefore choose $\mu = K^0$, which corresponds to neglecting $O(m_q^3 \ln m_s)$ contributions, to eliminate the logarithm altogether.

We now consider contributions from decuplet intermediate states (*trans*-octet diagrams). One finds

$$
\begin{aligned}
\Delta^{10}_{\rm CG} = & \\
\frac{\mathcal{C}^2}{32\pi^2 f^2} & \Big\{(n-p)\left[G_1(\Sigma^*-N,K)+4G_1(\Delta-N,\pi)\right] \\
& +(\Xi^- - \Xi^0)\left[G_1(\Sigma^*-\Xi,K)+2G_1(\Omega-\Xi,K)\right. \\
& \qquad\qquad \left.+G_1(\Xi^*-\Xi,\eta)+G_1(\Xi^*-\Xi,\pi)\right] \\
& +\frac{1}{3}(\Sigma^+-\Sigma^-)\left[8G_1(\Delta-\Sigma,K)+2G_1(\Xi^*-\Sigma,K)\right. \\
& \qquad\qquad \left.+3G_1(\Sigma^*-\Sigma,\eta)+2G_1(\Sigma^*-\Sigma,\pi)\right] \\
& +\frac{1}{3}(\Sigma^{*+}-\Sigma^{*-})\left[2G_1(\Sigma^*-N,K)+2G_1(\Sigma^*-\Xi,K)\right. \\
& \qquad\qquad \left.-3G_1(\Sigma^*-\Sigma,\eta)-G_1(\Sigma^*-\Sigma,\pi)\right] \\
& +\frac{20}{3}(\Xi^{*0}-\Xi^{*-}-\Sigma^{*+}+\Sigma^{*-})\left[G_1(\Delta-\Sigma,K)-G_1(\Delta-N,\pi)\right] \\
& +\frac{1}{3}(\Xi^{*-}-\Xi^{*0})\left[2G_1(\Xi^*-\Sigma,K)-3G_1(\Xi^*-\Xi,\eta)\right. \\
& \qquad\qquad \left.+G_1(\Xi^*-\Xi,\pi)\right] \\
& +\frac{1}{3}\left((K^0)^2-(K^+)^2\right)\left[4G_2(\Delta-\Sigma,K)+G_2(\Sigma^*-N,K)\right. \\
& \qquad\qquad \left.-G_2(\Sigma^*-\Xi,K)+2G_2(\Xi^*-\Sigma,K)\right. \\
& \qquad\qquad \left.-6G_2(\Omega-\Xi,K)\right] \\
& +r\left[G_3(\Sigma^*-\Sigma,\eta)-G_3(\Sigma^*-\Sigma,\pi)\right. \\
& \quad \left.-G_3(\Xi^*-\Xi,\eta)+G_3(\Xi^*-\Xi,\pi)\right]\Big\}. \qquad \text{(B.2)}
\end{aligned}
$$

In writing this result, we have used the decuplet relations Eqs. 4.9–4.12 to eliminate to consistent order ($O((m_d-m_u)m_s)$ is all we require) the dependence on the poorly-measured Δ isospin splittings. As before, terms proportional to r arise from π^0-η mixing. Here

$$
G_1(M,m) \equiv +\frac{4M}{m}(M^2-m^2)F(M/m)+(2M^2-m^2)\ln\frac{m^2}{\mu^2}, \qquad \text{(B.3)}
$$

$$
G_2(M,m) \equiv -\frac{1}{2m}(M^2-m^2)F(M/m)-M\ln\frac{m^2}{\mu^2}, \qquad \text{(B.4)}
$$

$$
G_3(M,m) \equiv +\frac{4}{3m}(M^2-m^2)^2F(M/m)+M(\frac{2}{3}M^2-m^2)\ln\frac{m^2}{\mu^2}, \qquad \text{(B.5)}
$$

where F is the same function as defined in Eq. 5.47. The functions G_1 and G_2 are

simply proportional to the nonanalytic portions of the baryonic and mesonic derivatives, respectively, of $(m_\alpha^3 H(\xi_{ij}^\alpha))$, where the function $H(\xi)$ is defined in Eq. 5.48; G_3 is proportional to $(m_\alpha^3 H(\xi_{ij}^\alpha))$ itself. In these expressions, μ is the renormalization scale.

We now explore the limiting forms of these corrections. In the limit $M \gg m$, we have

$$G_1(M,m) \rightarrow +(2M^2 - m^2)\ln\frac{4M^2}{\mu^2} - \frac{1}{2}m^2, \tag{B.6}$$

$$G_2(M,m) \rightarrow -M\ln\frac{4M^2}{\mu^2}, \tag{B.7}$$

$$G_3(M,m) \rightarrow +M(\frac{2}{3}M^2 - m^2)\ln\frac{4M^2}{\mu^2} - \frac{1}{6}Mm^2, \tag{B.8}$$

up to terms that vanish as $M \rightarrow \infty$. We thus see that the decuplet contributions decouple in the limit where the octet-decuplet splitting δ becomes large, since the only terms that do not vanish as $\delta \rightarrow \infty$ are analytic in the quark masses. In this limit, the decuplet contributions can be absorbed into counterterms in an effective Lagrangian that does not contain decuplet fields, and the μ dependence in Eqs. B.6–B.8 simply renormalizes the couplings in this effective Lagrangian. Physically, such a decoupling *must* occur, because the effects of a massive field on low-energy parameters must vanish as the field becomes infinitely massive.

In the opposite limit $M \ll m$, we have

$$G_1(M,m) \rightarrow -m^2\ln\frac{m^2}{\mu^2}, \tag{B.9}$$

$$G_2(M,m) \rightarrow +\frac{\pi m}{2}, \tag{B.10}$$

$$G_3(M,m) \rightarrow +\frac{2\pi m^3}{3}, \tag{B.11}$$

up to terms that vanish as $M \rightarrow 0$. In this limit, $\Delta_{\rm CG}^{10}$ has the same nonanalytic dependence on quark masses as the contributions from octet intermediate states, as discussed in Sec. 5.6.

Changing the renormalization scale μ in Eqs. B.3–B.5 changes $\Delta_{\rm CG}^{10}$ only by $O(m_q^3)$, so we may again take $\mu = K^0$ for purposes of numerical evaluation.

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
TECHNICAL INFORMATION DEPARTMENT
BERKELEY, CALIFORNIA 94720

AAT836
LBL Libraries