MATTHEW B WINGATE

THE STANDARD MODEL

These notes are for use with the Cambridge University Part III lecture course *The Standard Model*, given in the Department of Applied Mathematics and Theoretical Physics in Lent Term 2015. They have grown out of previous versions of the course; in particular the written notes provided to me by Hugh Osborn and Ben Allanach have formed the basis for the present set of notes.

This version was compiled 2 March 2015 (12:05).

Contents

Introduction

Quantum field theory was developed in order to describe natural phenomena that are both relativistic and quantum. The prototype, quantum electrodynamics (QED), described the physics of electrons and photons. Since then, the quantum field theory framework has been found to be rich, forcing on us constraints which allow firm theoretical predictions (e.g. the charm quark) yet allowing the description of a wide range of physical phenomena. Ideas such as renormalization carry over to the theory of critical phenomena. Ongoing research is exploring dualities between field theories and gravity theories (gauge-gravity dualities, of which AdS/CFT correspondence is one).

The Standard Model is the most successful application of quantum field theory when it comes to experimental verification. Over the four decades since its ingredients were combined thousands of measurements have been made, all apparantly consistent with the Standard Model. Even the physics of quark flavour-changing interactions, which could be sensitive through quantum loops to new particles as heavy as 10^5 TeV$/c^2$, is well-described by the Standard Model.

The Standard Model describes the physics of three fundamental forces, each mediated by gauge bosons. The electromagnetic force is described as electrically charged particles exchanging photons, as in QED. The weak force is responsible for particles changing their "flavour" as occurs in neutron β-decay. The W boson is responsible for weak decays; its large mass is the reason for the weakness of this force. (The Z boson also carries the weak force, but does not cause a change in flavour.) The strong force binds quarks into nucleons and nucleons into nuclei; the carrier of the strong force is appropriately called the gluon.

Table 1.1: Standard Model bosons

electromagnetic	γ
weak	$W^{\pm}, Z^0$
strong	g
higgs	h

The matter content (the fermions) of the Standard Model are the neutral leptons (the neutrinos), which feel only the weak force; the charged leptons, like the electron which interact weakly and electromagnetically; and the quarks, which are sensitive to all 3 forces.

Table 1.2: Standard Model fermions

leptons	$\begin{pmatrix}\nu_e\\ e\end{pmatrix}, \begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}, \begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}$
quarks	$\begin{pmatrix}u\\ d\end{pmatrix}, \begin{pmatrix}c\\ s\end{pmatrix}, \begin{pmatrix}t\\ b\end{pmatrix}$

Finally there is the Higgs boson, a consequence of the Higgs mechanism for generating masses for the W and Z, as well as for the fermions. In 2012 we saw convincing evidence that the Higgs boson, or a Higgs-like boson, has been discovered at the LHC.

As in QED, the gauge bosons are manifestations of local symmetries. The standard model gauge group is the direct product of the 3 Lie groups

$$SU(3)_C \times SU(2)_L \times U(1)_Y \,.$$

The electroweak theory is described by the product of the chiral gauge theory $SU(2)$-left times $U(1)$-hypercharge. The separate weak and electromagnetic forces we observe are due to the spontaneous breaking of this symmetry down to just $U(1)_{EM}$. The next lectures will gradually explain the details of what is written above.

$SU(3)$-colour describes the strong force (this part of the theory is called quantum chromodynamics or QCD).[1] We shall not discuss QCD much until later in the course. The strong interactions make for rich but complicated behaviour, difficult to study theoretically. In the first half of this course, we shall focus on the electroweak sector of the Standard Model. The strong sector of the Standard Model is fascinating for its richness, the electroweak sector is intriguing because of its persistent mysteries.

[1] Do not confuse the $SU(3)$-colour *gauge* symmetry with the $SU(3)$-flavour *global* symmetry which gives approximate relations between bound states of u, d, and s quarks.

There are many useful sources of relevant information. Prof Osborn has lecture notes[2] for the version of this course he gave a few years ago. I will mostly follow the conventions used in Prof Tong's notes for Part III Quantum Field Theory.[3] (For example, I will mostly use the chiral representation of the Dirac matrices instead of the nonrelativistic, or Bjorken-Drell, representation.) Romão & Silva[4] have recently performed a service to the community by carefully noting the various sign convensions which appear in texts and reviews. These notes shall (when I'm finished editing) consistently use the convension such that $\eta = \eta_s = \eta' = \eta_Z = \eta_\theta = \eta_Y = \eta_e = +1$. I find the QFT text by Peskin & Schroeder[5] to be a good reference on many topic covered in this course; however, be aware that they differ with the sign convensions here by using $\eta = \eta_s = \eta' = -1$ – in past years I followed their convensions (in most places). The book by Halzen & Martin[6] is pitched at readers unfamiliar with QFT, but it contains physically-motivated discussions and arguments which complement the more field theoretic treatments.

These notes appear online, with updates and additions appearing as I edit them.[7] The notes and the webpage will cite other references if they are particularly helpful with respect to a particular topic.

[2] www.damtp.cam.ac.uk/user/ho/SM.ps

[3] www.damtp.cam.ac.uk/user/tong/qft.html

[4] J C Romão and J P Silva. A resource for signs and Feynman diagrams of the Standard Model. *Int. J. Mod. Phys.*, A27:1230025, 2012. arXiv:1209.6213

[5] M E Peskin and D V Schroeder. *An Introduction to Quantum Field Theory*. Addison Wesley, 1995. ISBN 0-201-50397-2

[6] F Halzen and A D Martin. *Quarks and Leptons*. Wiley & Sons, 1984. ISBN 0-471-88741-2

[7] www.damtp.cam.ac.uk/user/wingate/StdM

Chiral and gauge symmetries

In this chapter we present a few concepts which may have been introduced in last term's courses, but which are crucial to constructing the Standard Model. This also allows us to set our notation and conventions. Throughout, we use natural units, $\hbar = c = 1$.

2.1 *Chiral symmetry*

We begin with spin-$\frac{1}{2}$ fermions. Let $\psi(x)$ be a spinor field satisfying the Dirac equation[8]

$$(i\partial\!\!\!/ - m)\psi \;=\; 0\,.$$

The Dirac-adjoint field $\bar{\psi} = \psi^\dagger\gamma^0$ satisfies[9]

$$\bar{\psi}(-i\overleftarrow{\partial\!\!\!/} - m) \;=\; 0\,.$$

The Dirac matrices satisfy the anticommutation relation

$$\{\gamma^\mu, \gamma^\nu\} \;=\; 2g^{\mu\nu}\mathbb{1} \tag{2.1.1}$$

where $g^{\mu\nu}$ is the Minkowski metric and $\mathbb{1}$ is the 4×4 identity matrix. From now on, we will dispense with the blackboard bold font for identity matrices unless it is crucial for clarity. These notes will use the signature $(+, -, -, -)$ as is usual in particle physics. When we need an explicit representation for the Dirac matrices, we will usually use the chiral representation; written as 2×2 block matrices of 2×2 matrices they are

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{2.1.2}$$

with Pauli matrices σ^i. This choice for γ^0 differs from the Bjorken-Drell (or the nonrelativistic) and Dirac representations, but is useful when concerned with chiral symmetry.

Now we will consider the Dirac equation in the massless limit, $m = 0$. Despite the fact that in Nature, all fermions appear to have mass – in fact a diverse range of masses – the massless limit is the natural foundation (from the theorist's perspective) from which to construct the Standard Model. We will see this when we treat the electroweak theory in depth.

Let us introduce γ^5:

$$\gamma^5 \;=\; -i\,\gamma^0\gamma^1\gamma^2\gamma^3 \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\;; \tag{2.1.3}$$

[8] The slash indicates contraction of a 4-vector with the Dirac γ-matrices: $\partial\!\!\!/ \equiv \gamma^\mu\partial_\mu$.

[9] The arrow above ∂ indicates that the derivative acts to the left.

the last equality holds only in the chiral representation. It is useful later to note here that

$$(\gamma^5)^2 = 1, \qquad \{\gamma^5, \gamma^\mu\} = 0.$$

Since γ^5 anticommutes with γ^μ, if ψ solves the massless Dirac equation, then $\gamma^5\psi$ is also a solution:

$$\partial\!\!\!/\psi = 0 \;\Rightarrow\; \partial\!\!\!/(\gamma^5\psi) = 0.$$

It is useful to work with the linear combinations

$$\begin{aligned} \psi_L(x) &= \frac{1}{2}(1-\gamma^5)\psi(x) = P_L\psi(x) \\ \psi_R(x) &= \frac{1}{2}(1+\gamma^5)\psi(x) = P_R\psi(x). \end{aligned} \tag{2.1.4}$$

where we have implicitly defined the operators $P_{L,R}$. These are projection operators:

$$(P_{L,R})^2 = P_{L,R}, \quad P_L P_R = 0 = P_R P_L, \quad P_L + P_R = 1.$$

With the convention (2.1.3) the lower (upper) 2 components of 4-component spinors contain the left-handed (right-handed) degrees-of-freedom. Because γ^5 anticommutes with γ^0

$$\bar\psi_L(x) = \bar\psi(x)P_R, \quad \text{and} \quad \bar\psi_R(x) = \bar\psi(x)P_L.$$

The field $\psi_{L,R}$ are eigenvectors of γ^5, with eigenvalues ∓ 1, and are said to have definite *chirality* – this is why the representation (2.1.2) is called "chiral." We shall give further justification to the terms "left-" and "right-handed" in the next chapter.

We can see that a massless Dirac fermion possesses a $U(1)_L \times U(1)_R$ chiral symmetry as follows. Writing $\psi(x) = \psi_L(x) + \psi_R(x)$ and similarly for $\bar\psi(x)$ the Dirac Lagrangian $\mathcal{L} = \bar\psi(i\partial\!\!\!/ - m)\psi$ becomes

$$\mathcal{L} = \bar\psi_L\, i\partial\!\!\!/\,\psi_L + \bar\psi_R\, i\partial\!\!\!/\,\psi_R - m\,(\bar\psi_R\psi_L + \bar\psi_L\psi_R). \tag{2.1.5}$$

Performing global rotations independently for the two chiralities

$$\begin{aligned} \psi_L(x) \mapsto e^{i\alpha_L}\psi_L(x), \;\; \bar\psi_L(x) \mapsto \bar\psi_L(x)e^{-i\alpha_L} \\ \psi_R(x) \mapsto e^{i\alpha_R}\psi_R(x), \;\; \bar\psi_R(x) \mapsto \bar\psi_R(x)e^{-i\alpha_R} \end{aligned} \tag{2.1.6}$$

it is clear that the kinetic term is invariant while the mass term is not. The mass term explicitly breaks the chiral symmetry down to a single vector-like one: $U(1)_L \times U(1)_R \to U(1)_V$, corresponding to (2.1.6) with $\alpha_L = \alpha_R$.

2.2 *Gauge symmetry*

If we "gauge" the symmetry (2.1.6)

$$\begin{aligned} \psi_L(x) \mapsto e^{i\alpha_L(x)}\psi_L(x), \;\; \bar\psi_L(x) \mapsto \bar\psi_L(x)e^{-i\alpha_L(x)} \\ \psi_R(x) \mapsto e^{i\alpha_R(x)}\psi_R(x), \;\; \bar\psi_R(x) \mapsto \bar\psi_R(x)e^{-i\alpha_R(x)} \end{aligned} \tag{2.2.1}$$

then the kinetic term in the Lagrangian (2.1.5) is no longer invariant. Dropping the L and R subscripts and thinking generally about $U(1)$ gauge transformations, another term is generated due to the spatial dependence of the transformation parameter(s)

$$\bar{\psi}\, i\partial\!\!\!/\, \psi \;\mapsto\; \bar{\psi}\, i\partial\!\!\!/\, \psi \;-\; (\bar{\psi}\gamma^{\mu}\psi)(\partial_{\mu}\alpha)\,.$$

We need a gauge covariant derivative D_μ which acts on ψ so that

$$D_{\mu}\psi(x) \;\mapsto\; \exp(i\,\alpha(x))\, D_{\mu}\psi(x)\,.$$

We obtain this by introducing gauge fields $A_{(L,R),\mu}(x)$.

$$D_{\mu}\psi \;=\; (\partial_{\mu} + i\, g A_{\mu})\psi$$

with the gauge fields transforming as

$$A_{\mu}(x) \;\mapsto\; A_{\mu}(x) - \frac{1}{g}\partial_{\mu}\alpha\,.$$

Now $\bar{\psi}\, iD\!\!\!\!/\, \psi$ is invariant under gauge transformations.[10]

The gauge fields contribute a kinetic energy term to the Lagrangian. This is given in terms of the field strength

$$F_{\mu\nu} \;=\; \partial_{\mu}A_{\nu} \;-\; \partial_{\nu}A_{\mu}\,,$$

or equivalently by solving

$$[D_{\mu}, D_{\nu}] \;=\; i\, g F_{\mu\nu}\,,$$

as

$$\mathcal{L} \;=\; -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,.$$

QED has a $U(1)$ gauge symmetry which treats left- and right-handed components equivalently, so we set $\alpha_L = \alpha_R$ above. For a theory which only involved interactions between left-handed particles, we would only need to introduce A_L. In fact, the weak gauge bosons do only couple to the left-handed components of particles; however, $U(1)$ is not the appropriate local symmetry. The weak interactions change one particle into another, predominantly a specific partner. This is realized theoretically through an $SU(2)$ gauge symmetry.

[10] Note that one can flip some signs in these formulae by flipping the sign in the gauge transformation ($\alpha \mapsto -\alpha$). Other additional signs can differ due to how the fields are introduced. See the Introduction and J C Romão and J P Silva. A resource for signs and Feynman diagrams of the Standard Model. *Int. J. Mod. Phys.*, A27:1230025, 2012. arXiv:1209.6213

THE INTRODUCTION OF nonabelian gauge fields proceeds along similar lines to those above. The field transformations (2.2.1) are generalized to introduce a transformation for each generator of the gauge group. (In order to avoid cumbersome writing, let us not specify whether the gauge group couples to one or both of the chiral eigenstates.)

$$\begin{aligned}
\psi_i(x) \;&\mapsto\; \exp(i\, t^a\theta^a(x))_{ij}\psi_j(x) \;\equiv\; U_{ij}\psi_j(x)\,, \\
\bar{\psi}_i(x) \;&\mapsto\; \bar{\psi}_j(x)\exp(-i\, t^a\theta^a(x))_{ji} \;\equiv\; \bar{\psi}_j(x)(U^{\dagger})_{ji}
\end{aligned} \qquad (2.2.2)$$

where the t^a are the Hermitian generators in an n-dimensional representation R ($i, j = 1, \ldots, n$) of a given gauge group whose dimension is D ($a = 1, \ldots, D$). (For $SU(N)$, $D = (N^2 - 1)$.) We implicitly specialized to unitary groups and introduced U as an element of the group. The generators form a Lie algebra, i.e.

$$[t^a, t^b] \;=\; i f^{abc} t^c \tag{2.2.3}$$

where f^{abc} are called structure constants, and are normalized such that

$$\mathrm{Tr}(t^a t^b) \;=\; T(R)\, \delta^{ab}\,. \tag{2.2.4}$$

$T(R)$ is the Dynkin index of the representation, e.g. is equal to $\frac{1}{2}$ for the fundamental representation ($n = N$) and $T(R) = N$ for the adjoint representation of $SU(N)$. Recall the adjoint representation is one where $(t^a)^{bc} = -i f^{abc}$, thus $n = D$.[11] In the Standard Model, the fermion fields belong to either the trivial or fundamental representations of the 3 constituent gauge groups.

[11] See Peskin & Schroeder §15.4 for a more thorough introduction.

In the nonabelian case the covariant derivative becomes

$$(D_\mu)_{ij} \;=\; \partial_\mu \delta_{ij} \;+\; ig(t^a A^a_\mu)_{ij} \tag{2.2.5}$$

so that $(D_\mu \psi(x))_i \mapsto (U(x) D_\mu \psi(x))_i$. It is usual to drop the index i which runs from 1 to n. Constructing the field strength from

$$[D_\mu, D_\nu] = igt^a F^a_{\mu\nu} \tag{2.2.6}$$

we find

$$F^a_{\mu\nu} \;=\; \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g f^{abc} A^b_\mu A^c_\nu\,. \tag{2.2.7}$$

The gauge-field-only term in the Lagrangian is then

$$\mathcal{L}_g \;=\; -\frac{1}{4} F^a_{\mu\nu} F^{a,\mu\nu} \;=\; -\frac{1}{4} \mathrm{Tr} F_{\mu\nu} F^{\mu\nu}\,. \tag{2.2.8}$$

2.3 *Symmetry, a synopsis*

Following Donoghue, Golowich, Holstein, §I-5[12] let us enumerate the possible ways symmetries may manifest themselves

[12] J F Donoghue, E Golowich, and B R Holstein. *Dynamics of the Standard Model*. Cambridge University Press, 1992. ISBN 0-521-47652-6

1. The symmetry may be intact. For example, the gauge symmetries $U(1)_{EM}$ and $SU(3)_C$ are manifestly respected in QED and QCD.

2. A symmetry of the Lagrangian may be broken by an *anomaly*. We say the symmetry holds classically, but then is broken by quantum effects. (More on anomalies later, perhaps.) In this case, the symmetry is not a true symmetry of the theory. The global axial symmetry suffers this fate in the Standard Model.

3. The symmetry may hold for some terms in the Lagrangian, but not others. The symmetry is said to be broken *explicitly*. If the symmetry-breaking terms are small, then the *approximate symmetry* still is a leading-order picture which one can refine perturbatively. Isospin symmetry relating u and d quarks is explicitly broken by electromagnetic and quark mass effects, both of which can be treated as small corrections for most purposes.

4. The Lagrangian may possess a certain symmetry, but the vacuum may not respect it. In this case we say the symmetry is spontaneously broken. Actually, there are physical consequences when a symmetry is *spontaneously broken* rather than explicitly broken, so really this is a case of a *hidden symmetry*. The $SU(2)_L \times U(1)_Y$ gauge symmetry is broken down to $U(1)_{EM}$ by spontaneous symmetry breakdown. We shall study this in depth shortly.

The Standard Model contains examples of all of these possibilities. For me, this is one of the most fascinating aspects of particle physics as we understand it. Although much has been written about beauty and symmetry, the ways in which symmetries are broken or hidden are even more interesting.

Discrete symmetries

The preceding chapter reviewed chiral symmetry and gauge theories. This puts us in a good frame of mind to begin discussing the electroweak theory, which turns out to be a chiral gauge theory involving only left-handed particles. This fact has consequences for discrete symmetries. Unlike gauge theories which have vector-like couplings to Dirac fermions, such as QED and QCD, chiral gauge theories are not symmetric under the operations of parity-flip or charge conjugation.

In fact it turns out that the combination of parity and charge conjugation is also violated by the weak interactions. This *CP violation* is one important ingredient in the evolution of the cosmos. As realized by Sakharov, CP violation, nonequilibrium thermodynamics, and particle number violation are all necessary conditions for generating more matter than antimatter in the universe. According to the CPT theorem we will discuss later, CP violation implies a violation of time reversal symmetry.

Before we can appreciate these statements, first we must investigate the consequences for theories which do respect the discrete symmetries P, C, and T; that is, theories which would hold equally well in our laboratories as well as in mirror-world laboratories, laboratories made out of antimatter, and laboratories where time runs backwards.

3.1 Symmetry operators

Before we investigate the behaviour of quantum fields under parity, charge conjugation, and time reversal, we must introduce the quantum operators corresponding to these transformations.

Parity and time-reversal are special cases of Lorentz transformations. Including rotations, boosts, and transformations, a Lorentz transformation is a change of coordinate frame:

$$x^\mu \mapsto x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu \,.$$

If $\det\Lambda = 1$, the Lorentz transformation is said to be proper. Parity and time-reversal are improper transformations, with Λ respectively given by

$$\mathscr{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \text{and} \quad \mathscr{T} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \,. \tag{3.1.1}$$

In order to introduce quantum operators, let Ψ, Ψ' and Φ be generic vectors in some Hilbert space. Wigner showed that if physics is invariant under a transformation $\Psi \mapsto \Psi'$, then there is an operator W such that $\Psi' = W\Psi$ where (for $\alpha, \beta \in \mathbb{C}$) either

$$\begin{cases} (W\Phi, W\Psi) = (\Phi, \Psi) & W \text{ is unitary} \\ W(\alpha\Phi + \beta\Psi) = \alpha W\Phi + \beta W\Psi & \text{and linear} \end{cases} \tag{3.1.2}$$

or

$$\begin{cases} (W\Phi, W\Psi) = (\Phi, \Psi)^* & W \text{ is anti-unitary} \\ W(\alpha\Phi + \beta\Psi) = \alpha^* W\Phi + \beta^* W\Psi & \text{and anti-linear.} \end{cases} \tag{3.1.3}$$

Let W now be the operator corresponding to a Lorentz transformation for a given rotation/boost $\Lambda^\mu{}_\nu$ and translation a^μ,

$$\Psi \mapsto \Psi' = W(\Lambda, a)\Psi\,.$$

Note that the Lorentz transformation operators obey the following composition rule which will need shortly:

$$W(\Lambda_2, a_2)W(\Lambda_1, a_1) = W(\Lambda_2\Lambda_1, \Lambda_2 a_1 + a_2)\,. \tag{3.1.4}$$

We wish to consider the special operators for parity and time-reversal. Let us denote these by

$$\hat{P} = W(\mathscr{P}, 0) \quad \text{and} \quad \hat{T} = W(\mathscr{T}, 0)\,. \tag{3.1.5}$$

Let us look at how these transformations combine with an infinitesimal, proper Lorentz transformation

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu, \quad a^\mu = \varepsilon^\mu$$

with $\omega^\mu{}_\nu$ and ε^μ small parameters. In this case we expand the corresponding quantum operator as[13]

[13] S Weinberg. *The Quantum Theory of Fields, Volume I*. Cambridge University Press, 1995. ISBN 0-521-55001-7

$$W(1 + \omega, \varepsilon) = 1 + \frac{i}{2}\omega_{\mu\nu}J^{\mu\nu} - i\varepsilon_\mu P^\mu + \ldots \tag{3.1.6}$$

where $P^0 = H$ is the energy operator, i.e. the Hamiltonian; $\vec{P} = (P^1, P^2, P^3)$ is the linear momentum operator, generator of translations; $\vec{J} = (J^{23}, J^{31}, J^{12})$ is the angular momentum operator, generator of rotations; and $\vec{K} = (J^{01}, J^{02}, J^{03})$ generates Lorentz boosts. Let us consider how the parity and time-reversal operators act on these operators. Using the composition rule (3.1.4),

$$\begin{cases} \hat{P}\, W(\Lambda, a)\, \hat{P}^{-1} = W(\mathscr{P}\Lambda\mathscr{P}^{-1}, \mathscr{P}a) \\ \hat{T}\, W(\Lambda, a)\, \hat{T}^{-1} = W(\mathscr{T}\Lambda\mathscr{T}^{-1}, \mathscr{T}a) \end{cases} \tag{3.1.7}$$

Inserting (3.1.6) and equating the coefficients of $-\varepsilon_\mu$, we find

$$\begin{cases} \hat{P}\, iP^\mu \hat{P}^{-1} = i\mathscr{P}_\nu{}^\mu P^\nu \\ \hat{T}\, iP^\mu \hat{T}^{-1} = i\mathscr{T}_\nu{}^\mu P^\nu \end{cases} \tag{3.1.8}$$

and focusing on the $\mu = 0$ component

$$\begin{cases} \hat{P}\, iH\hat{P}^{-1} = iH \\ \hat{T}\, iH\hat{T}^{-1} = -iH \end{cases} . \tag{3.1.9}$$

In order to understand the properties of the operators corresponding to parity and time-reversal, let Ψ be an energy eigenstate with energy E: $(\Psi, iH\Psi) = iE$. Assuming P and T are symmetries of our theory, then the energy of the transformed states $\hat{P}\Psi$ and $\hat{T}\Psi$ should also be E. As we show below, this implies that $\hat{P}$ must be linear and unitary, while $\hat{T}$ must be anti-linear and anti-unitary. Treating as $\hat{P}$ as linear

$$(\hat{P}\Psi, iH\hat{P}\Psi) = (\hat{P}\Psi, \hat{P}iH\Psi) = (\hat{P}\Psi, \hat{P}iE\Psi) = iE(\hat{P}\Psi, \hat{P}\Psi) = iE\,.$$

For unitary $\hat{P}$:

$$(\hat{P}\Psi, iH\hat{P}\Psi) = (\hat{P}\Psi, \hat{P}iH\Psi) = (\Psi, iH\Psi) = iE\,.$$

For anti-linear $\hat{T}$:

$$(\hat{T}\Psi, iH\hat{T}\Psi) = -(\hat{T}\Psi, \hat{T}iH\Psi) = -(\hat{T}\Psi, \hat{T}iE\Psi) = -(iE)^*(\hat{T}\Psi, \hat{T}\Psi) = iE\,.$$

For anti-unitary $\hat{T}$:

$$(\hat{T}\Psi, iH\hat{T}\Psi) = -(\hat{T}\Psi, \hat{T}iH\Psi) = -(\Psi, iH\Psi)^* = iE\,.$$

Choosing otherwise would imply that parity or time-reversal mapped positive energy states to negative energy states.

In §3.4 we will also want to look at the coefficients of $\frac{1}{2}\omega_{\mu\nu}$

$$\begin{aligned} \hat{T}iJ^{\mu\nu}\hat{T}^{-1} &= i\mathscr{T}_\rho{}^\mu \mathscr{T}_\sigma{}^\nu J^{\rho\sigma} \\ -i\hat{T}\vec{J}\hat{T}^{-1} &= i\vec{J} \\ \hat{T}\vec{J}\hat{T}^{-1} &= -\vec{J}\,. \end{aligned} \tag{3.1.10}$$

Thus we see that angular momentum flips sign under T.

3.2 *Parity*

To say that physics is symmetric under parity transformations implies that reflections do not change the laws of physics. We wish to investigate the consequences of moving from our world to a mirror world, mapping left to right, etc, as follows

$$x \;\mapsto\; x_P = (x^0, -\vec{x}) \tag{3.2.1}$$

imposing the requirement that physical results are unchanged. First we consider bosonic fields, then Dirac fermion fields.

Boson field

Let us consider the quantum scalar field $\phi(x)$, which we can write as a sum of plane waves (using relativistic normalization)

$$\phi(x) \;=\; \sum_p \left[a(p)\, e^{-ip\cdot x} \;+\; c^\dagger(p)\, e^{ip\cdot x}\right]. \tag{3.2.2}$$

We have introduced a shorthand notation for the momentum integrals

$$\sum_p \;\equiv\; \int \frac{d^3p}{(2\pi)^3(2E_{\vec{p}})}\,. \tag{3.2.3}$$

The operator $a^\dagger(p)$ creates a particle with momentum p and $c^\dagger(p)$ creates an antiparticle with momentum p. Given the parity transformation (3.2.1), which in momentum space is equivalent to

$$p \mapsto p_P = (p^0, -\vec{p})\,,$$

the unitary operator $\hat{P}$ should map the momentum eigenstate $|p\rangle = a^\dagger(p)|0\rangle$ to $|p_P\rangle = a^\dagger(p_P)|0\rangle$, up to a complex phase η^{a*}, i.e.

$$\hat{P}a^\dagger(p)|0\rangle \;=\; \eta^{a*}a^\dagger(p_P)|0\rangle\,.$$

Using $\hat{P}^{-1}\hat{P} = 1$ and assuming the uniqueness of the vacuum[14] we find

[14] $\hat{P}|0\rangle = |0\rangle$

$$\hat{P}a^\dagger(p)\hat{P}^{-1} \;=\; \eta^{a*}a^\dagger(p_P)\,.$$

Similarly for antiparticle excitations

$$\hat{P}c^\dagger(p)\hat{P}^{-1} \;=\; \eta^{c*}c^\dagger(p_P)\,.$$

In order to conserve normalizations (of wavepackets for example) $\hat{P}$ should be unitary and $\hat{P}a(p)\hat{P}^{-1} = \eta^a_P a(p_P)$. Applying the parity transformation to the scalar field (3.2.2), we find

$$\begin{aligned}
\hat{P}\phi(x)\hat{P}^{-1} \;&=\; \sum_p \left[\hat{P}a(p)\hat{P}^{-1}e^{-ip\cdot x} + \hat{P}c^\dagger(p)\hat{P}^{-1}e^{ip\cdot x}\right] \\
&=\; \sum_p \left[\eta^a\, a(p_P)\, e^{-ip\cdot x} + \eta^{c*}\, c^\dagger(p_P)\, e^{ip\cdot x}\right] \\
&=\; \sum_p \left[\eta^a\, a(p)\, e^{-ip\cdot x_P} + \eta^{c*}\, c^\dagger(p)\, e^{ip\cdot x_P}\right]\,.
\end{aligned}$$

In the last step we note that $p_P \cdot x = p \cdot x_P$ and

$$\sum_p a(p_P)\, e^{-ip\cdot x} = \sum_{p_P} a(p)\, e^{-ip_P\cdot x} = \sum_p a(p)\, e^{-ip\cdot x_P}$$

and that $\sum_p$ is invariant under P ($\int_{-\infty}^{\infty} dp_i \mapsto -\int_{\infty}^{-\infty} dp_i = \int_{-\infty}^{\infty} dp_i$).

Presently, $\phi^P(x) \equiv \hat{P}\phi(x)\hat{P}^{-1}$ looks like a different field from $\phi(x_P)$. This is not what we expect physically; in a P-symmetric theory, we should not need different rules for how to combine particle and antiparticle plane waves into a scalar field. Furthermore, inconsistencies with Lorentz invariance: for general η^a and η^c, $[\phi(x), \phi^P(y)]$ does not vanish for spacelike $x - y$. These problems are solved if $\eta^a = \eta^{c*}_P \equiv \eta_P$, hence

$$\hat{P}\phi(x)\hat{P}^{-1} \;=\; \eta_P\,\phi(x_P) \;=\; \phi^P(x)\,. \tag{3.2.4}$$

The phase η_P is the *intrinsic parity* of ϕ. If ϕ is a real field, then $c(p) = a(p)$. Then

$$\eta^c = \eta^a \implies \eta^*_P = \eta_P \implies \eta_P = \pm 1 \begin{cases} \phi \text{ scalar} \\ \phi \text{ pseudoscalar}\,. \end{cases}$$

In other words, the intrinsic parity of real (and therefore neutral) fields has definite meaning. On the other hand for complex $\phi(x)$ which has a conserved charge Q, using the fact that the corresponding operator $\hat{Q}$, $\hat{P}$, and the Hamiltonian $\hat{H}$ are all mutually commuting, we can redefine $\hat{P}$ by multiplying by a phase $\hat{P}' = \hat{P}e^{-i\alpha Q}$ with α chosen so that $\hat{P}'^2 = 1$. The original η_P is thus expressible in terms of the charge Q. (Further discussion appears in Weinberg §2.2 and §3.3.[15])

[15] S Weinberg. *The Quantum Theory of Fields, Volume I.* Cambridge University Press, 1995. ISBN 0-521-55001-7

For vector fields

$$V^\mu(x) = \sum_{p,\lambda}\left[\varepsilon^{\lambda,\mu}(p)a^\lambda(p)e^{-ip\cdot x} \,+\, \varepsilon^{\lambda,\mu*}(p)c^{\lambda\dagger}(p)e^{ip\cdot x}\right] \tag{3.2.5}$$

where $\varepsilon^{\lambda,\mu}(p)$ are polarization vectors (say, $\lambda = -1, 0, 1$). In order to consider how V^μ is transformed under P we perform the same steps as with the scalar field. Additionally we need the relation between polarization vectors when $p \mapsto p_P$

$$\varepsilon^{\lambda,\mu}(p_P) = -\mathscr{P}^{\mu}{}_{\nu}\varepsilon^{\lambda,\nu}(p)\,. \qquad (3.2.6)$$

This can be shown using Lorentz boosts. For example, one chooses the polarization vectors in the particle's rest frame in some basis, such as

$$\varepsilon^{-1,\mu}(0) = \begin{pmatrix} 0 \\ 1 \\ -i \\ 0 \end{pmatrix}\,, \quad \varepsilon^{0,\mu}(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\,, \quad \varepsilon^{1,\mu}(0) = \begin{pmatrix} 0 \\ 1 \\ i \\ 0 \end{pmatrix}\,,$$

and boosts to a frame moving with respect to the particle with momentum p or p_P, e.g. $\varepsilon^{\lambda,\nu}(p) = L^{\mu}{}_{\nu}(p)\varepsilon^{\lambda,\nu}(0)$, with $L(p)$ a standard Lorentz boost. The conclusion one reaches is that

$$\hat{P}V^{\mu}(x)\hat{P}^{-1} = -\eta_P\mathscr{P}^{\mu}{}_{\nu}V^{\nu}(x_P)\,. \qquad (3.2.7)$$

With the conventions above polar vector fields have $\eta_P = -1$ while axial vector fields have $\eta_P = 1$.

Dirac field

Recall the solution of the free Dirac equation can be expressed as a combination of plane waves

$$\psi(x) \;=\; \sum_{p,s}\left[b^s(p)\,u^s(p)\,e^{-i\,p\cdot x} \;+\; d^{s\dagger}(p)\,v^s(p)\,e^{i\,p\cdot x}\right]\,. \qquad (3.2.8)$$

The operators $b^\dagger(p)$ and $d^\dagger(p)$ respectively create positive and negative frequency particles.[16] $u^s(p)$ and $v^s(p)$ are 4-component spinors satisfying $(\not{p} - m)u = 0$ and $(\not{p} + m)v = 0$ and $s = \pm\frac{1}{2}$.

In the chiral representation (2.1.2) we can write,

$$u^s(p) \;=\; \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\,\xi^s \end{pmatrix} \qquad (3.2.9)$$

where $\sigma \;=\; (1,\vec{\sigma})$ and $\bar{\sigma} \;=\; (1,-\vec{\sigma})$.[17] The ξ^s are 2-component spinors; as a convenient basis we take $\xi^{1/2} \;=\; (1,0)^T$ and $\xi^{-1/2} \;=\; (0,1)^T$. We can also at this stage deduce that

$$v^s(p) \;=\; \begin{pmatrix} \sqrt{p\cdot\sigma}\,\zeta^s \\ -\sqrt{p\cdot\bar{\sigma}}\,\zeta^s \end{pmatrix} \qquad (3.2.10)$$

for some 2-component spinors ζ^s. We will find a relation between ζ^s and ξ^s when we discuss charge conjugation in § 3.3.

To see why we associate this property with a type of handedness (left or right), we will make use of the total angular momentum operator, the sum of orbital and spin angular momentum operators:

$$\vec{J} = -i\,\vec{r}\times\nabla \;+\; \vec{S} \qquad (3.2.11)$$

where

$$S_i = \frac{i}{4}\epsilon_{ijk}\gamma^j\gamma^k = -\frac{1}{2}\begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix} \qquad (3.2.12)$$

[16] We will mostly use relativistically normalized states $|p\rangle \;=\; b^\dagger(p)|0\rangle$. These are related to nonrelativistically normalized states, $|\vec{p}\rangle$, by $|p\rangle \;=\; \sqrt{2E_{\vec{p}}}|\vec{p}\rangle = \sqrt{2E_{\vec{p}}}\,b^\dagger_{\vec{p}}|0\rangle$.

[17] You can check this by showing (3.2.9) solves $\begin{pmatrix} -m & p_\mu\sigma^\mu \\ p_\mu\bar{\sigma}^\mu & -m \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = 0$. It may be useful to note that $(p\cdot\sigma)(p\cdot\bar{\sigma}) = (p^0)^2 - p^ip^j\delta_{ij} = p^2 = m^2$ (recall $\delta_{ij} = -g_{ij}$).

In order to deal with the square root of a matrix like $p\cdot\sigma$, we should work in its eigenvector basis. In this case it amounts to rotating the spatial coordinates so that the 3-momentum is in the $\vec{e}_3$ direction. Then

$$\sqrt{p\cdot\sigma} \;=\; \sqrt{p^0 - p^3}\left(\frac{1+\sigma^3}{2}\right) + \sqrt{p^0+p^3}\left(\frac{1-\sigma^3}{2}\right)\,.$$

is the spin operator and satisfies $\gamma^5 S^i = \frac{1}{2}\gamma^0\gamma^i = S^i\gamma^5$. If we consider the massless spinor $u^s(p)$ and multiply the Dirac equation $\not{p}u = 0$ by γ^0/p^0 we obtain

$$\frac{\gamma^0}{p^0}(\gamma^0 p^0 - \vec{\gamma}\cdot\vec{p})u^s(p) = 0$$
$$(1 - 2\vec{S}\cdot\hat{\vec{p}}\gamma^5)u^s(p) = 0$$

where $\hat{\vec{p}} \equiv \vec{p}/|\vec{p}|$, and $|\vec{p}| = p^0$ in this massless case. By inserting $(P_L + P_R) = 1$ in front of the spinor, using $\gamma^5 P_{L,R} = \mp P_{L,R}$ and the linear independence of u_L and u_R we conclude

$$(1 \pm 2\vec{S}\cdot\hat{\vec{p}})u^s_{L,R}(p) = 0\,. \tag{3.2.13}$$

Note that the scalar product $h \equiv \vec{S}\cdot\hat{\vec{p}} = \vec{J}\cdot\hat{\vec{p}}$ is the projection of angular momentum in the direction of the linear momentum, i.e. the *helicity*.[18] Rewriting (3.2.13) one finds that $u_{L,R}$ are eigenvectors of the helicity operator with eigenvalues $\mp\frac{1}{2}$:

$$hu^s_{L,R} = \mp\frac{1}{2}u^s_{L,R}\,.$$

Therefore we can think of the eigenvectors with positive helicity, u_R, as obeying a right-hand rule for spin; correspondingly we call u_L spinors left-handed. We see that ψ_L is a field which annihilates left-handed particle states. [19]

One can also show the Dirac adjoint field $\bar{\psi}_{L(R)}$ is left-handed (right-handed). Given

$$\bar{\psi}(x) = \sum_{p,s}\left[b^{s\dagger}(p)\bar{u}^s(p)e^{ip\cdot x} + d^s(p)\bar{v}^s(p)e^{-ip\cdot x}\right]$$

we can apply the same steps as before. Looking at the antiparticle spinor, we find

$$\bar{v}^s(p)\not{p}\left(\frac{\gamma^0}{p^0}\right) = 0$$
$$\bar{v}^s(p)(P_R + P_L)(1 + 2\gamma^5\vec{S}\cdot\hat{\vec{p}}) = 0$$
$$\bar{v}^s_{L,R}(p)(1 \pm 2\gamma^5\vec{S}\cdot\hat{\vec{p}}) = 0$$
$$\bar{v}^s_{L,R}(p)(1 \pm 2\vec{S}\cdot\hat{\vec{p}}) = 0$$
$$\bar{v}^s_{L,R}(p)h = \mp\frac{1}{2}\bar{v}^s_{L,R}(p)$$

Thus $\bar{\psi}_{L(R)}$ annihilates a left-handed (right-handed) antiparticle.

Having established the notations and conventions for Dirac fields, we consider their behaviour under parity transformations. The creation operators (and hence annihilation operators) should transform as in the case of bosons: under parity a particle's spatial momentum switches directions; the spin component s is left unchanged. For the moment, we assign a different phase to the 2 operators:

$$\hat{P}b^s(p)\hat{P}^{-1} = \eta^b b^s(p_P)$$
$$\hat{P}d^{s\dagger}(p)\hat{P}^{-1} = \eta^{d*}d^{s\dagger}(p_P)\,. \tag{3.2.14}$$

[18] Above we used $p^0 = |\vec{p}|$, true only for $m = 0$. Nevertheless the definition of helicity is in terms of $\hat{\vec{p}}$ even for $m > 0$.

[19] The handedness indicated by the helicity is not Lorentz invariant for massive particles; one can always boost to a frame where the momentum and hence the helicity flips sign. However for massless particles discussed here, helicity and chirality coincide.

When we transform the Dirac field (3.2.8), we find

$$\begin{aligned}
\hat{P}\hat{\psi}(x)\hat{P}^{-1} &= \sum_{p,s}\left[\hat{P}b^s(p)\hat{P}^{-1}u^s(p)e^{-ip\cdot x} + \hat{P}d^{s\dagger}(p)\hat{P}^{-1}v^s(p)e^{ip\cdot x}\right] \\
&= \sum_{p,s}\left[\eta^b b^s(p_P)u^s(p)e^{-ip\cdot x} + \eta^{d*}d^{s\dagger}(p_P)v^s(p)e^{ip\cdot x}\right] \\
&= \sum_{p,s}\left[\eta^b b^s(p)u^s(p_P)e^{-ip\cdot x_P} + \eta^{d*}d^{s\dagger}(p)v^s(p_P)e^{ip\cdot x_P}\right] .
\end{aligned}$$

Finally we must determine the relationship between the spinors as $p \mapsto p_P$. Using Lorentz boosts, one finds

$$u^s(p_P) = \gamma^0 u^s(p) \quad \text{and} \quad v^s(p_P) = -\gamma^0 v^s(p) . \tag{3.2.15}$$

This can be verified as well in a particular basis such as the chiral basis used in (3.2.9) and (3.2.10). [20] Requiring $\eta^b = -\eta^{d*} \equiv \eta_P$ so that the transformed field takes the form of the original field, we find

[20] Note that $p \cdot \bar{\sigma} = p_P \cdot \sigma$.

$$\psi^P(x) = \hat{P}\psi(x)\hat{P}^{-1} = \eta_P \gamma^0 \psi(x_P) . \tag{3.2.16}$$

Likewise

$$\bar{\psi}^P(x) = \hat{P}\bar{\psi}(x)\hat{P}^{-1} = \eta_P^* \bar{\psi}(x_P)\gamma^0 . \tag{3.2.17}$$

From (3.2.16) and (3.2.17) we can easily show that the parity transformation swaps left-handed and right-handed chiralities:

$$\begin{aligned}
\hat{P}\psi_L(x)\hat{P}^{-1} &= \eta_P \gamma^0 \psi_R(x_P) \\
\hat{P}\bar{\psi}_L(x)\hat{P}^{-1} &= \eta_P^* \bar{\psi}_R(x_P)\gamma^0 .
\end{aligned} \tag{3.2.18}$$

We can check that ψ^P satisfies the Dirac equation, assuming ψ does:

$$\begin{aligned}
(i\partial\!\!\!/ - m)\psi^P(x) &= \eta_P(i\gamma^0\partial_0 - i\vec{\gamma}\cdot\nabla - m)\gamma^0\psi(x_P) \\
&= \eta_P(i\gamma^0\partial_0 + i\vec{\gamma}\cdot\nabla - m)\gamma^0\psi(x) \\
&= \eta_P\gamma^0(i\gamma^0\partial_0 - i\vec{\gamma}\cdot\nabla - m)\psi(x) \\
&= \eta_P\gamma^0(i\partial\!\!\!/ - m)\psi(x) = 0 .
\end{aligned}$$

With (3.2.16) and (3.2.17), one can determine the transformation properties of composite operators such as bilinears:

$$\begin{array}{lcll}
\bar{\psi}(x)\psi(x) & \mapsto & \bar{\psi}(x_P)\psi(x_P) & \text{scalar,} \\
\bar{\psi}(x)\gamma^5\psi(x) & \mapsto & -\bar{\psi}(x_P)\gamma^5\psi(x_P) & \text{pseudoscalar,} \\
\bar{\psi}(x)\gamma^0\psi(x) & \mapsto & \bar{\psi}(x_P)\gamma^0\psi(x_P) & \text{charge density,} \\
\bar{\psi}(x)\gamma^i\psi(x) & \mapsto & -\bar{\psi}(x_P)\gamma^i\psi(x_P) & \text{current density.}
\end{array}$$

Note the last two lines show $\bar{\psi}\gamma^\mu\psi$ transforms as a vector field as discussed earlier in this section. Similarly $\bar{\psi}\gamma^\mu\gamma^5\psi$ transforms as an axial-vector field.

3.3 *Charge conjugation*

We now investigate the behaviour of fields in theories which treat particles and antiparticles on equal footing.

Boson field

Consider how the scalar field ϕ transforms under C. The corresponding unitary operator $\hat{C}$ should transform a momentum eigenstate of a particle to one of its antiparticle. Defining the phase η_C such that[21]

$$\begin{aligned}\hat{C}\,a(p)\,\hat{C}^{-1} &= \eta_C\,c(p)\\ \hat{C}\,c(p)\,\hat{C}^{-1} &= \eta_C^*\,a(p)\end{aligned}$$

we find

$$\hat{C}|p,\text{particle}\rangle = \hat{C}a^\dagger(p)|0\rangle = \eta_C^* c^\dagger(p)|0\rangle = \eta_C^*|p,\text{antiparticle}\rangle\,.$$

In the decomposition of $\phi(x)$ in terms of plane waves (3.2.2), $a(p)$ annihilates particles while $c(p)$ annihilates antiparticles. Then

$$\begin{aligned}\hat{C}\,\phi(x)\,\hat{C}^{-1} &= \eta_C\,\phi^\dagger(x)\\ \hat{C}\,\phi^\dagger(x)\,\hat{C}^{-1} &= \eta_C^*\,\phi(x)\,.\end{aligned} \tag{3.3.1}$$

If $\phi^\dagger = \phi$ then the field has definite intrinsic charge-conjugation parity: $\eta_C = \pm 1$. However, if ϕ is complex, then η_C is arbitrary. Say $\eta_C = e^{2i\beta}$, then we can perform a global $U(1)$ rotation $\phi \mapsto \phi' = e^{-i\beta}\phi$ so that $\hat{C}\phi'\hat{C}^{-1} = (\phi')^\dagger$. That is, we can redefine a complex ϕ so that it has $\eta_C = 1$.

For real fields $\eta_C = \pm 1$ has physical significance. As we shall see, the photon field $A_\mu(x)$ must obey

$$\hat{C}A_\mu(x)\hat{C}^{-1} = -A_\mu(x) \tag{3.3.2}$$

in order for the interactions with Dirac fermions to be C-invariant. As a consequence, a particle with $\eta_C = +1$ such as the π^0 cannot decay to an odd number of photons.[22]

[21] We could introduce different phases for the 2 operators, but Lorentz invariance requires the choice given here, similar to what we saw in the previous section.

[22] We know experimentally that the π^0 can decay to 2 photons, thus we can infer that $\eta_C^{\pi_0} = (-1)^2$.

Dirac field

As with the scalar field, the particle and antiparticle operators are swapped under C

$$\begin{aligned}\hat{C}b^s(p)\hat{C}^{-1} &= \eta_C d^s(p)\\ \hat{C}d^{s\dagger}(p)\hat{C}^{-1} &= \eta_C b^{s\dagger}(p)\,.\end{aligned}$$

Operators on the left-hand sides of the above equations appear in the field ψ, while those on the right-hand sides appear in $\bar{\psi}$.

We will see that the transposes of the Dirac matrices are necessary ingredients in working out the charge conjugated Dirac field. In conventional basis, γ^0 and γ^2 are symmetric while γ^1 and γ^3 are anti-symmetric. Let us define a matrix C which gives symmetric matrices when multiplied by any γ^μ

$$(C\gamma^\mu)^T = C\gamma^\mu\,. \tag{3.3.3}$$

One choice is

$$C = -i\gamma^0\gamma^2 = \begin{pmatrix} i\sigma_2 & 0\\ 0 & -i\sigma_2\end{pmatrix} \tag{3.3.4}$$

with the last equality holding in our γ-matrix basis. One can check that C is anti-symmetric, real, and unitary:

$$C = -C^T = -C^\dagger = -C^{-1} \,.$$

Thus the transposed γ-matrix is given by

$$\gamma^{\mu T} = -C\gamma^\mu C^{-1} \,. \tag{3.3.5}$$

The $\gamma^{\mu T}$ form an equivalent Clifford algebra to the γ^μ:

$$\{\gamma^{\mu T}, \gamma^{\nu T}\} = 2g^{\mu\nu} \,.$$

We also note

$$\gamma^{5T} = C\gamma^5 C^{-1} \,. \tag{3.3.6}$$

Thus the charge conjugated field is

$$\hat{C}\psi(x)\hat{C}^{-1} \;=\; \eta_C \sum_{p,s} \left[d^s(p)\, u^s(p)\, e^{-ip\cdot x} \;+\; b^{s\dagger}(p)\, v^s(p)\, e^{ip\cdot x} \right] . \tag{3.3.7}$$

We can compare this to the transpose of the Dirac-adjoint field

$$\bar{\psi}^T(x) \;=\; \sum_{p,s} \left[b^{s\dagger}(p)\, \bar{u}^{sT}(p)\, e^{ip\cdot x} \;+\; d^s(p)\, \bar{v}^{sT}(p)\, e^{-ip\cdot x} \right] . \tag{3.3.8}$$

In fact for the spinors (3.2.9) and (3.2.10), taking $\zeta^s = i\sigma^2 \xi^{s*}$ allows us to write

$$v^s(p) = C\bar{u}^{sT}(p) \;\text{ and }\; u^s(p) = C\bar{v}^{sT}(p) \,. \tag{3.3.9}$$

Thus

$$\psi^c(x) = \hat{C}\psi(x)\hat{C}^{-1} = \eta_C C\bar{\psi}^T(x) \tag{3.3.10}$$

where we can think if ψ^C as a particle transforming in the complex-conjugated representation representation to ψ, or as an antiparticle. Similar steps as those above lead to

$$\bar{\psi}^c(x) = \hat{C}\bar{\psi}(x)\hat{C}^{-1} = \eta_C^* \psi^T(x) C = -\eta_C^* \psi^T(x) C^{-1} \,. \tag{3.3.11}$$

Finally, we can see that C maps, for example, left-handed particles to left-handed antiparticles

$$\begin{aligned} \hat{C}\psi_L(x)\hat{C}^{-1} &= \eta_C C\bar{\psi}_R^T(x) \equiv \psi_L^c(x) \\ \hat{C}\bar{\psi}_L(x)\hat{C}^{-1} &= -\eta_C^* \psi_R^T(x) C^{-1} \equiv \bar{\psi}_L^c(x) C^{-1} \,. \end{aligned} \tag{3.3.12}$$

Let us explicitly check that $\psi_L^c(x)$ is left-handed as our notation suggests. Given $\gamma^5 \psi_{L,R} = \mp \psi_{L,R}$ (and hence $\bar{\psi}_{L,R}\gamma^5 = \pm \bar{\psi}_{L,R}$), then

$$\gamma^5 C\bar{\psi}_R^T = C\gamma^{5T}\bar{\psi}_R^T = C(\bar{\psi}_R \gamma^5)^T = -C\bar{\psi}_R^T \,.$$

Thus the γ^5-eigenvalue of $\psi_L^c(x)$ is that of a left-handed field.

We can show that $\psi^c(x)$ satisfies the Dirac equation. Consider the Dirac equation for $\bar{\psi}$, times η_C:

$$\eta_C \bar{\psi}(x)(-i\overleftarrow{\not{\partial}} - m) = 0 \,.$$

Take the transpose

$$\eta_c(-i(\gamma^\mu)^T\partial_\mu - m)\bar{\psi}^T(x) = 0\,.$$

Insert $C^{-1}C$ between the Dirac operator and the field and left-multiply the equation by C to find

$$(-iC(\gamma^\mu)^T C^{-1}\partial_\mu - m)\psi^c(x) = 0\,.$$

Applying (3.3.5), we arrive at

$$(i\partial\!\!\!/ - m)\psi^c(x) = 0$$

as required.

Majorana fermions are ones where $b^s(p) = d^s(p)$; that is, the particle is its own antiparticle. In that case $\psi^c = \psi$. It is not yet known whether the only neutral fermions in the Standard Model, the neutrinos, are Majorana fermions.

Let us consider the behaviour of fermion bilinears under C. It is convenient to work with anti-symmetrized operators, as this will make the transformation properties manifest. For example let us take the as the vector current (corresponding to $\bar{\psi}\gamma^\mu\psi$)

$$j^\mu(x) = \frac{1}{2}(\bar{\psi}\gamma^\mu\psi - \psi^T\gamma^{\mu T}\bar{\psi}^T) = \frac{1}{2}(\gamma^\mu)_{ij}[\bar{\psi}_i(x), \psi_j(x)]\,. \qquad (3.3.13)$$

In the last step we label the spin indices explicitly. Then j^μ has well-defined behaviour under C. Using (3.3.10) and (3.3.11) we find

$$\begin{aligned}
\hat{C}j^\mu\hat{C}^{-1} &= \frac{1}{2}(\gamma^\mu)_{ij}[\hat{C}\bar{\psi}_i\hat{C}^{-1}, \hat{C}\psi_j\hat{C}^{-1}] \\
&= -\frac{1}{2}(\gamma^\mu)_{ij}[(\psi^T C^{-1})_i, (C\bar{\psi}^T)_j] \\
&= -\frac{1}{2}(C^{-1}\gamma^\mu C)_{k\ell}[\psi_k, \bar{\psi}_\ell] \\
&= \frac{1}{2}(\gamma^\mu)_{\ell k}[\psi_k, \bar{\psi}_\ell] \;=\; -j^\mu\,.
\end{aligned}$$

(In the above argument, we drop the c-number arising from anti-commuting ψ and $\psi^\dagger$.)

Recalling the transformation of the photon field (3.3.2), we see that the interaction $j^\mu(x)A_\mu(x)$ is C-invariant; it induces transitions only between states of equal charge.

One can undertake a similar calculation to show that the axial-vector current

$$j^{\mu 5}(x) \;=\; \frac{1}{2}(\gamma^\mu\gamma^5)_{ij}[\bar{\psi}_i(x), \psi_j(x)] \qquad (3.3.14)$$

behaves as

$$\hat{C}\,j^{\mu 5}\,\hat{C}^{-1} \;=\; +j^{\mu 5}\,.$$

Therefore (foreshadowing) a linear combination of j^μ and $j^{\mu 5}$, such as the left-handed current $j^{\mu L} = \frac{1}{2}(j^\mu - j^{\mu 5})$ cannot couple to a single field in a C-invariant way.

3.4 Time reversal

The third and final discrete symmetry we investigate here is time reversal

$$x \mapsto x_T = (-x^0, \vec{x})\,. \tag{3.4.1}$$

In T-symmetric theories, the physics would be unchanged if the flow of time were to run backwards.

Boson field

Given (3.2.2) and

$$\begin{aligned}\hat{T}a(p)\hat{T}^{-1} &= \eta_T a(p_T)\\ \hat{T}c^\dagger(p)\hat{T}^{-1} &= \eta_T c^\dagger(p_T)\end{aligned} \tag{3.4.2}$$

we find

$$\begin{aligned}\hat{T}\phi(x)\hat{T}^{-1} &= \sum_p \left[\hat{T}a(p)\hat{T}^{-1}e^{ip\cdot x} + \hat{T}c^\dagger(p)\hat{T}^{-1}e^{-ip\cdot x}\right]\\ &= \eta_T \sum_p \left[a(p)e^{-ip\cdot x_T} + c^\dagger(p)e^{ip\cdot x_T}\right] = \eta_T\phi(x_T)\,.\end{aligned} \tag{3.4.3}$$

In the first step we used anti-linearity of $\hat{T}^{-1}$ to write $e^{\pm ip\cdot x}\hat{T}^{-1} = \hat{T}^{-1}e^{\mp ip\cdot x}$. In the second step we swapped integration variable $p \leftrightarrow p_T$ and noted $ip_T \cdot x = -ip\cdot x_T$.

Dirac field

In addition to taking $p \mapsto p_T = (p^0, -\vec{p})$, we saw from (3.1.10) that T will flip the sign of a particle's angular momentum. Thus the creation/annihilation operators for a particle/antiparticle with spin $s = \pm\frac{1}{2}$ can be taken to transform as

$$\begin{aligned}\hat{T}b^s(p)\hat{T}^{-1} &= \eta_T(-1)^{\frac{1}{2}-s}b^{-s}(p_T)\\ \hat{T}d^{s\dagger}(p)\hat{T}^{-1} &= \eta_T(-1)^{\frac{1}{2}-s}d^{-s\dagger}(p_T)\,.\end{aligned} \tag{3.4.4}$$

One can show that the spinors satisfy

$$\begin{aligned}(-1)^{\frac{1}{2}-s}u^{-s*}(p_T) &= -\gamma^5 C u^s(p)\\ (-1)^{\frac{1}{2}-s}v^{-s*}(p_T) &= -\gamma^5 C v^s(p)\,.\end{aligned} \tag{3.4.5}$$

It is usual to define

$$B = \gamma^5 C = \begin{pmatrix} i\sigma^2 & 0\\ 0 & i\sigma^2\end{pmatrix}\,. \tag{3.4.6}$$

Then the Dirac field transforms as

$$\begin{aligned}\hat{T}\psi(x)\hat{T}^{-1} &= \eta_T\sum_{p,s}(-1)^{\frac{1}{2}-s}\left[b^{-s}(p_T)u^{s*}(p)e^{ip\cdot x} + d^{-s\dagger}(p_T)v^{s*}(p)e^{-ip\cdot x}\right]\\ &= -\eta_T\sum_{p,s}(-1)^{\frac{1}{2}-s}\left[b^{s}(p)u^{-s*}(p_T)e^{-ip\cdot x_T} + d^{s\dagger}(p)v^{-s*}(p_T)e^{ip\cdot x_T}\right]\\ &= \eta_T\gamma^5 C\psi(x_T) = \eta_T B\psi(x_T)\,.\end{aligned} \tag{3.4.7}$$

Similarly

$$\hat{T}\bar{\psi}(x)\hat{T}^{-1} = \eta_T^*\bar{\psi}(x_T)B^{-1}\,. \tag{3.4.8}$$

Exercise: Using $B^{-1}\gamma^{5*}B = \gamma^5$ verify the following:

$$\hat{T}\psi_L(x)\hat{T}^{-1} = \eta_T B\psi_L(x_T)$$
$$\hat{T}\bar{\psi}_L(x)\hat{T}^{-1} = \eta_T^*\bar{\psi}_L(x_T)B^{-1}$$

Bilinears. It is straightforward to check that $\bar{\psi}(x)\psi(x) \mapsto \bar{\psi}(x_T)\psi(x_T)$ under T. For other cases, first let us note

$$\begin{aligned} B^{-1}\gamma^{0*}B &= \gamma^0 \\ B^{-1}\gamma^{i*}B &= -\gamma^i\,. \end{aligned}$$

Thus[23]

$$\begin{aligned} \hat{T}\bar{\psi}(x)\gamma^\mu\psi(x)\hat{T}^{-1} &= \bar{\psi}(x_T)B^{-1}\gamma^{\mu*}B\psi(x_T) \\ &= \begin{cases} \bar{\psi}(x_T)\gamma^0\psi(x_T) & \text{charge density} \\ -\bar{\psi}(x_T)\gamma^i\psi(x_T) & \text{charge current} \end{cases} \end{aligned}$$

[23] Insert $\hat{T}^{-1}\hat{T}$ between $\bar{\psi}$ and γ^μ then complex conjugate γ^μ when commuting it with $\hat{T}$.

To conclude this section, we look at how the S-matrix transforms under T. The S-matrix governing scattering is

$$S = \mathcal{T}\exp\left(-i\int_{-\infty}^{\infty} dt\, V(t)\right) \tag{3.4.9}$$

where

$$V(t) = -\int d^3x\,\mathcal{L}_I(x)$$

is the potential energy term in the Hamiltonian, and $\mathcal{L}_I$ contains the interaction terms of the Lagrangian. For example in QED, the electron-photon interaction term is

$$\mathcal{L}_I(x) = -e\bar{\psi}(x)\gamma^\mu A_\mu(x)\psi(x)\,.$$

We have already investigated the transformations of the Lagrange density under P, C, and T:

$$\hat{P}\mathcal{L}_I(x)\hat{P}^{-1} = \mathcal{L}_I(x_P)\,,\ \ \hat{C}\mathcal{L}_I(x)\hat{C}^{-1} = \mathcal{L}_I(x)\,,\ \ \hat{T}\mathcal{L}_I(x)\hat{T}^{-1} = \mathcal{L}_I(x_T)\,.$$

The consequent effects on V are straightforward to deduce

$$\hat{P}V(t)\hat{P}^{-1} = V(t)\,,\ \ \hat{C}V(t)\hat{C}^{-1} = V(t)\,,\ \ \hat{T}V(t)\hat{T}^{-1} = V(-t)\,.$$

However, while the properties of S under P and C are clear

$$\hat{P}S\hat{P}^{-1} = S\,,\ \ \hat{C}S\hat{C}^{-1} = S\,,$$

more care is needed to examine S under T.

Let us expand the time-ordered exponential as

$$S = \sum_n (-i)^n \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 \cdots \int_{-\infty}^{t_{n-1}} dt_n\, V(t_1)V(t_2)\cdots V(t_n)\,.$$

Then

$$\begin{aligned} S_T &= \hat{T}S\hat{T}^{-1} \\ &= \sum_n i^n \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 \cdots \int_{-\infty}^{t_{n-1}} dt_n\, V(-t_1)V(-t_2)\cdots V(-t_n) \\ &= \sum_n i^n \int_{-\infty}^{\infty} d\tau_1 \int_{-\infty}^{\tau_1} d\tau_2 \cdots \int_{-\infty}^{\tau_{n-1}} d\tau_n\, V(\tau_n)V(\tau_{n-1})\cdots V(\tau_1) \\ &= S^\dagger \end{aligned} \tag{3.4.10}$$

where we changed variables $\tau_i = -t_{(n+1)-i}$ and changed the limits of integration.[24]

Before looking at S-matrix elements, let us introduce some notation which will allow us to use Dirac's bra-ket notation, which is not normally suited to dealing with anti-linear, anti-unitary operators. Recall we need $\hat{T}$ to satisfy (3.1.3):

$$(\hat{T}\Phi, \hat{T}\Psi) = (\Phi, \hat{T}^\dagger \hat{T}\Psi)^* = (\Phi, \Psi)^* = (\Psi, \Phi)\,. \qquad (3.4.11)$$

We also need to introduce some way of distinguishing whether an operator acts on the bra or on the ket (i.e. whether it belongs on the left-hand or right-hand side of the inner product in the $(\cdot,\cdot)$ notation); let's use a semi-colon. Let $|\phi\rangle$ and $|\psi\rangle$ be the kets corresponding to the vectors Φ and Ψ, then the equation above reads

$$\langle\phi|\hat{T}^\dagger;\hat{T}|\psi\rangle = \langle\phi|\hat{T}^\dagger\hat{T}|\psi\rangle^* = \langle\phi|\psi\rangle^* = \langle\psi|\phi\rangle\,. \qquad (3.4.12)$$

(The absence of a semi-colon implies the operators act on the ket.)

Taking the Hermitian conjugate of (3.4.10), i.e. $S = S_T^\dagger$, and introducing the time-reversed partners of two states $|\eta\rangle$ and $|\zeta\rangle$ such that $|\eta_T\rangle = \hat{T}|\eta\rangle$ and $|\zeta_T\rangle = \hat{T}|\zeta\rangle$, we find

$$\begin{aligned}\langle\eta_T|S|\zeta_T\rangle \;=\; \langle\eta_T|S_T^\dagger|\zeta_T\rangle \;&=\; \langle\eta|\hat{T}^\dagger;\hat{T}S^\dagger\hat{T}^\dagger\hat{T}|\zeta\rangle \\ &=\; \langle\eta|S^\dagger|\zeta\rangle^* \;=\; \langle\zeta|S|\eta\rangle\,. \qquad (3.4.13)\end{aligned}$$

We see that, given $\hat{T}\mathcal{L}_I(x)\hat{T}^{-1} = \mathcal{L}_I(x_T)$, S-matrix elements are equal for time-reversed processes, where initial and final states are swapped. Correspondingly, observables such as decay rates and cross-sections are related.

3.5 *CPT theorem*

There is a theorem that says any Lorentz-invariant Lagrangian density should be invariant under the product of P, C, and T.[25] In other words, there is no way to distinguish, for example, a particle propagating forward in time from an antiparticle propagating backward in time. Define the shorthand notation for the product of the 3 transformations

$$\hat{\Theta} = \hat{C}\hat{P}\hat{T}\,. \qquad (3.5.1)$$

Then

$$\hat{\Theta}\mathcal{L}_I(x)\hat{\Theta}^{-1} = \mathcal{L}_I(-x)$$

This will have consequences on the creation and annihilation operators of momentum states, as we saw throughout this section. For example, for Dirac fermions, the operators satisfy

$$\hat{\Theta}b^s(p)\hat{\Theta}^{-1} = (-1)^{\frac{1}{2}-s}d^{-s}(p)\,.$$

In the case of Dirac fermions, CPT maps a left-handed, spin-up, forward-propagating particle to a right-handed, spin-down, backward-propagating antiparticle.

[24] As in the simple example $\int_0^1 dx \int_0^x dy f(x,y) = \int_0^1 dy \int_y^1 dx f(x,y)$.

[25] R.F. Streater and A.S. Wightman. *PCT, Spin and Statistics, and All That*. Addison-Wesley, 1989. ISBN 0-201-09410-X

3.6 Applications

We close by remarking on the role of these discrete symmetries in baryogenesis, the generation of a matter–antimatter asymmetry in the universe. Sakarhov is credited with enumerating 3 necessary conditions in order for such an imbalance to occur.

The first is baryon-number violation. There must be some process between state X and state Y which yields a baryon excess (B): $X \to Y + B$.[26]

The second is nonequilibrium so that the transition is not undone. In equilibrium, one would expect $Y + B \to X$ would be as likely as $X \to Y + B$, i.e.

$$\Gamma(Y + B \to X) \;=\; \Gamma(X \to Y + B) \;\text{ in equilibrium} \tag{3.6.1}$$

whereas

$$\Gamma(Y + B \to X) \;\propto\; \exp(-M_X/T) \;\text{ out of equilibrium}\,. \tag{3.6.2}$$

where T is temperature and M_X is the mass of the state X.

The first part of the third condition is C violation. If the universe starts with equal numbers of X and $\bar{X}$ particles, then baryons are produced at a rate

$$\frac{dB}{dt} \;\propto\; \Gamma(X \to Y + B) - \Gamma(\bar{X} \to \bar{Y} + \bar{B})\,. \tag{3.6.3}$$

C symmetry would imply that the two decay widths should be equal, hence C must be violated in baryogenesis.

The final part of the third condition is CP violation. For simplicity, let's assume that the B is composed of n quarks and that there is no extra Y left over.[27] C-violation implies

$$\Gamma(X \to nq_L) \neq \Gamma(\bar{X} \to n\bar{q}_L) \tag{3.6.4}$$

Under CP, $q_L \;\mapsto\; \bar{q}_R$ So even with C violation, CP symmetry would imply

$$\Gamma(X \to nq_L) + \Gamma(X \to nq_R) \;=\; \Gamma(\bar{X} \to n\bar{q}_L) + \Gamma(\bar{X} \to n\bar{q}_R) \tag{3.6.5}$$

which would preclude baryogenesis. Therefore, baryogenesis requires CP violation.

[26] This is the simplest possibility. Actually it is possible that it is a lepton asymmetry which is initially created, in which case we describe the process as *leptogenesis*. Then excess baryons can be generated through $B + L$ violation, which although conserved at the level of the SM Lagrangian, is violated by nonperturbative effects related to the chiral anomaly; $B - L$ is still conserved.

[27] The argument can be generalized, but we are more interested in the main point which is conveyed most simply here.

Nonrelativistic quantum mechanics

This section is not being lectured in 2015.

Newtonian dynamics is T-invariant: if $\vec{x}(t)$ is a solution to $m\ddot{\vec{x}} = \vec{F}(\vec{x})$ then so is $\vec{x}(-t)$. Care is needed if time or time-derivatives appear in the equations of motion. T-invariance in the presence of a Lorentz force $m\ddot{\vec{x}} = q(\vec{E}(\vec{x}) + \dot{\vec{x}} \times \vec{B}(\vec{x}))$ implies we must flip the sign of the magnetic field: $\vec{x}(t)|_{\vec{B}}$ a solution implies $\vec{x}(-t)|_{-\vec{B}}$ is a solution. In general under T, $\vec{E}(t,\vec{x}) \;\mapsto\; \vec{E}(-t,\vec{x})$ and $\vec{B}(t,\vec{x}) \;\mapsto\; -\vec{B}(-t,\vec{x})$. If time flows backwards, magnetic north becomes south and vice versa.

If we consider the Schrödinger equation

$$i\frac{\partial}{\partial t}\Psi(t,\vec{x}) = \left(-\frac{\nabla^2}{2m}+V(\vec{x})\right)\Psi(t,\vec{x})$$

we can see that the left-hand side gets an extra minus sign relative to the right-hand side when $t \mapsto -t$. Thus in order to have a time-reversed field which satisfies the Schrödinger equation, we make use of the complex conjugate

$$\Psi(t,\vec{x}) \mapsto \eta_T \Psi^{TR}(t,\vec{x}) \quad \text{where} \quad \Psi^{TR}(t,\vec{x}) = \Psi^*(-t,\vec{x})\,. \qquad (3.6.6)$$

(We use the superscript TR to avoid confusion with the transpose operation.) Since two applications of T should return the original wavefunction up to a phase, $|\eta_T| = 1$.

The difference between transformations like (3.6.6) compared to those for P and C is that the time-reversal transformation is *anti-linear*. If we let $\Psi' = \alpha\Psi$ then T transforms Ψ' as

$$\alpha\Psi(t,\vec{x}) \mapsto \alpha^*\eta_T\Psi^{TR}(t,\vec{x})\,.$$

Under T, the direction of momentum should be reversed

$$\hat{T}|\vec{p}\rangle = |-\vec{p}\rangle$$

A general state can be written as a linear combination of momentum eigenstates

$$|\Psi\rangle = \sum_{\vec{p}} \tilde{\Psi}(\vec{p})|\vec{p}\rangle\,.$$

Thus,[28]

$$\hat{T}|\Psi\rangle = \sum_{\vec{p}} \tilde{\Psi}^*(\vec{p})|-\vec{p}\rangle = \sum_{\vec{p}} \tilde{\Psi}^*(-\vec{p})|\vec{p}\rangle \equiv |\Psi^{TR}\rangle\,.$$

[28] In the penultimate step we change the sign of the summed momentum variable.

We must be careful using anti-linear operators with Dirac's bra-ket notation. Let us use a semi-colon to separate operators which act to the left from those which act to the right. Decomposing another state $|\Phi\rangle$ into its momentum components, we find

$$\langle\Phi^{TR}|\Psi^{TR}\rangle = \langle\Phi|\hat{T}^\dagger;\hat{T}|\Psi\rangle = \langle\Phi|;\hat{T}^\dagger\hat{T}|\Psi\rangle^* = \langle\Psi|\Phi\rangle\,.$$

Spontaneous symmetry breaking

In this chapter we present the main ideas behind symmetries which are hidden. That is, symmetries respected by the Lagrangian, but not manifest in physical observables.

Some of this material was presented in Prof Manton's course last term, but the ideas important enough to present again.

4.1 Spontaneous breaking of a discrete symmetry

Consider a real scalar field $\phi(x)$ with Lagrangian

$$\mathcal{L} \;=\; \frac{1}{2}\partial_\mu\phi\,\partial^\mu\phi \;-\; V(\phi)\,. \qquad (4.1.1)$$

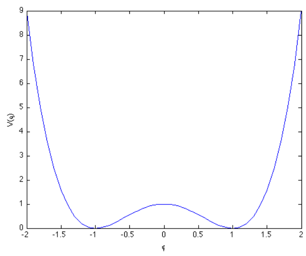

Figure 4.1: Double well potential

Let us assume a potential which is symmetric in ϕ; as a concrete example let us take that corresponding to ϕ^4-theory

$$V(\phi) \;=\; \frac{1}{2}m^2\phi^2 \;+\; \frac{\lambda}{4}\phi^4\,;\;\; \lambda > 0\,.$$

The theory has a discrete symmetry: $\mathcal{L}$ is invariant under $\phi \mapsto -\phi$.

In the usual case describing the physics of a massive scalar field, $m^2 > 0$ and $V(\phi)$ has a minimum at $\phi = 0$. For small λ, we can develop a perturbative expansion about the minimum of the potential.

If, on the other hand, if we allow ourselves to consider what happens if $m^2 < 0$, we see quite different behaviour. Let us complete the square in V, defining $v = \sqrt{-m^2/\lambda}$ and dropping the unimportant constant

$$V(\phi) \;=\; \frac{\lambda}{4}\left(\phi^2 - v^2\right)^2\,.$$

With this "double-well" potential, now $\phi = 0$ corresponds to an unstable vacuum. Instead there are two degenerate minima of the potential, $\phi = \pm v$, and hence two degenerate vacua. We say that ϕ has acquired a nonzero vacuum expectation value (VEV). Without loss of generality, let us study a theory of small excitations from the vacuum where $\phi = v$, writing $\phi(x) = v + f(x)$.

$$\mathcal{L} \;=\; \frac{1}{2}\partial_\mu f\partial^\mu f - \lambda\left(v^2 f^2 + v f^3 + \frac{1}{4}f^4\right)\,.$$

From the quadratic term, we see $f(x)$ is a scalar field representing massive excitations with mass-squared $m_f^2 = 2\lambda v^2$. This Lagrangian, describing the leading-order behaviour in a perturbative expansion, is not invariant under sign flip, $f \mapsto -f$. The symmetry of the original Lagrangian is broken by the VEV of ϕ.

The fact that the vacuum is not unique can lead to interesting nonperturbative consequences, but we will not explore these here.

4.2 *Spontaneous breaking of a continuous symmetry*

Let us begin our generalization to the case of spontaneous breaking of continuous symmetries with a simple example, that of an N-component real scalar field $\phi = (\phi_1, \ldots, \phi_N)^T$. The Lagrangian is

$$\mathcal{L} \;=\; \frac{1}{2}\partial_\mu\phi \cdot \partial^\mu\phi \;-\; V(\phi) \tag{4.2.1}$$

with[29]

$$V(\phi) \;=\; \frac{1}{2}m^2\phi^2 \;+\; \frac{\lambda}{4}\phi^4\,; \;\; \lambda > 0\,.$$

and is invariant under global $O(N)$ transformations of the field.

[29] $\phi^2 = \phi \cdot \phi$; $\phi^4 = (\phi^2)^2$.

We are primarily interested in the case $m^2 < 0$. We can replace the potential (up to an irrelevant constant term) by

$$V(\phi) \;=\; \frac{\lambda}{4}\left(\phi^2 - v^2\right)^2\;;\;\; v^2 = -\frac{m^2}{\lambda} > 0\,.$$

This potential is often called the *sombrero*, or Mexican hat, potential even though it clearly bears a much closer resemblance to the bottom of a wine bottle (Fig. 4.2, right). Now there are a continuum of vacua satisfying $\phi^2 = v^2$.

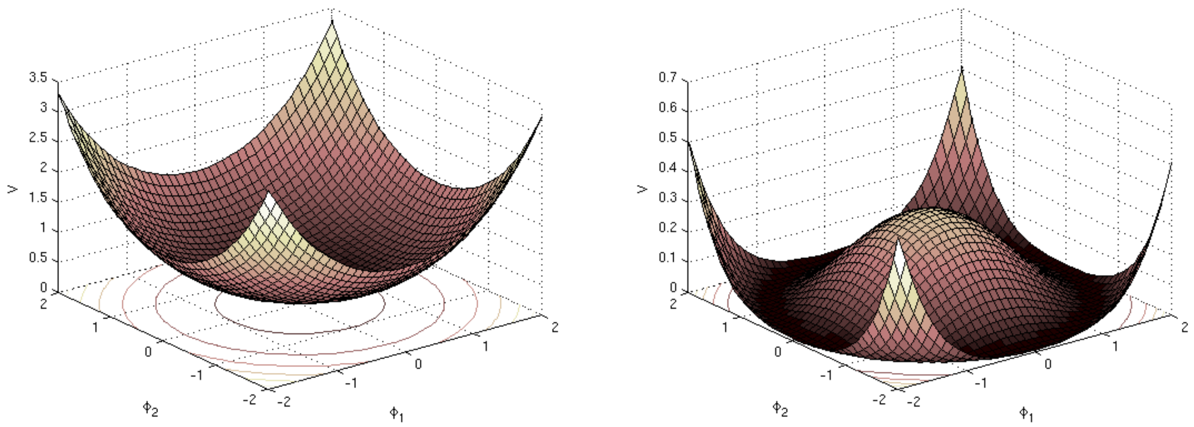

Figure 4.2: Symmetric ($m^2 > 0$) and spontaneously broken ($m^2 < 0$) potentials.

Without loss of generality let us choose the vacua such that $\phi_0 = (0, 0, \ldots, 0, v)^T$. Studying small fluctuations from this field configuration, let us define shifted fields, the $N-1$ component $\pi(x)$ and 1 component $\sigma(x)$ so that

$$\phi(x) \;=\; \begin{pmatrix} \pi_1(x) \\ \pi_2(x) \\ \vdots \\ \pi_{N-1}(x) \\ v + \sigma(x) \end{pmatrix}.$$

In terms of these, the Lagrangian is

$$\mathcal{L} \;=\; \frac{1}{2}\partial_\mu \pi \cdot \partial^\mu \pi \;+\; \frac{1}{2}\partial_\mu \sigma \partial^\mu \sigma \;-\; V(\pi,\sigma) \tag{4.2.2}$$

with

$$V(\pi,\sigma) \;=\; \frac{1}{2}m_\sigma^2\sigma^2 \;+\; \lambda v(\sigma^2+\pi^2)\sigma \;+\; \frac{\lambda}{4}(\sigma^2+\pi^2)^2\,.$$

Note that the σ field has a mass $m_\sigma = \sqrt{2\lambda v^2}$, but the $N-1$ π fields are massless. This makes sense intuitively from the wine-bottle shape of the potential (Fig. 4.2): radial excitations come with a large energy penalty, whereas excitations in the field which locally seek to transform the field to another of the equivalent vacua can be made to have arbitrarily small energy difference from the vacuum.

Let us generalize to a symmetry group G of the Lagrangian, which is broken down to a subgroup $H \subset G$ by the vacuum.[30] That is, for a transformation $\phi(x) \mapsto g\phi(x)$ with $g \in G$ (in some representation)

$$V(g\phi) = V(\phi)\;\forall(g \in G)\,.$$

[30] We will always be interested in the case that H is a *normal* subgroup, $H \triangleleft G$.

Let us assume though that G is spontaneously broken and hence the vacuum is not unique but is described by a manifold $\Phi_0 = \{\phi_0 : V(\phi_0) = V_{\min}\}$, the collection of all field configurations which minimize V. The invariant subgroup (or stability group) $H \subset G$ is such that

$$H \;=\; \{h : h\phi_0 = \phi_0\}\,.$$

The different vacua in Φ_0 are related by group transformations[31]

$$\phi_0' = g\phi_0 \;\text{ for some } g \in G\,.$$

[31] In other words we assume that G acts transitively on Φ_0. This is true in the cases of physical interest, but one can concoct counterexamples.

Then the stability groups for the different vacua are isomorphic: the invariant subgroup for ϕ_0' is $H' \simeq gHg^{-1}$. In fact the group elements which map one vacuum to another belong to the coset space G/H and fall into equivalence classes: we say $g_1 \sim g_2$ if $\exists h \in H$ such that $g_1 = g_2 h$. Correspondingly $\phi_0' = g_1\phi_0 = g_2\phi_0$ implies $g_2^{-1}g_1 \in H$, so with each $\phi_0' \in \Phi_0$ we can associate an equivalence class. Thus we say

$$\Phi_0 \simeq G/H\,. \tag{4.2.3}$$

Let us consider infinitesimal transformations

$$g\phi \;=\; \phi \;+\; \delta\phi\,, \quad \text{with } \delta\phi = i\alpha^a t^a \phi$$

where $a = 1,\ldots,\mathrm{dim}G$, t^a are the generators of the Lie algebra of G (in the representation of ϕ) and the α^a are small parameters. G-invariance of the theory implies $V(\phi+\delta\phi) = V(\phi)$, or expanding $V(\phi+\delta\phi)$ about $\delta\phi = 0$

$$V(\phi+\delta\phi) - V(\phi) \;=\; i\alpha^a (t^a\phi)_r \frac{\partial V}{\partial \phi_r} = 0\,. \tag{4.2.4}$$

neglecting contributions higher order in the α^a. Here $r = 1 \dots N$ is the index for components of ϕ in its representation of G.

Let ϕ_0 be a minimum of V. Considering small departures from this particular minimum, we have

$$V(\phi) - V(\phi_0) \;=\; \frac{1}{2}(\phi - \phi_0)_r \left.\frac{\partial^2 V}{\partial\phi_r \partial\phi_s}\right|_{\phi_0} (\phi - \phi_0)_s \,.$$

The matrix of second derivatives can be equated with a mass-squared matrix M^2_{rs}. If we differentiate (4.2.4) and evaluate at ϕ_0, then we find[32]

$$\begin{aligned} \frac{\partial}{\partial\phi_s}\left[(t^a\phi)_r \frac{\partial V}{\partial\phi_r}\right]\bigg|_{\phi_0} \;&=\; 0 \\ (t^a\phi_0)_r \left.\frac{\partial^2 V}{\partial\phi_s \partial\phi_r}\right|_{\phi_0} \;&=\; 0\,. \end{aligned} \tag{4.2.5}$$

[32] The second term in the product rule vanishes because $V(\phi_0)$ is a minimum and so the first derivative vanishes there.

If the symmetry is unbroken and the vacuum is unique, i.e. $g\phi_0 = \phi_0$ for all $g \in G$, then $\delta\phi = 0$ and $(t^a\phi_0) = 0$ for all a. Otherwise, if there are some $g \in G$ such that there exists some a with $(t^a\phi_0) \neq 0$, then $t^a\phi_0$ is an eigenvector of the mass-squared matrix with zero eigenvalue: $(t^a\phi_0)_r M^2_{rs} = 0$.

We wish to find out how many eigenvectors of M^2 have vanishing eigenvalue. Let us denote using a tilde the generators of G which satisfy $\tilde{t}^i\phi_0 \;=\; 0$. These generate a subgroup $H \;\subset\; G$, and hence there are dim$H$ of them. If G is compact and semi-simple (as it usually is in cases interesting to us) then we can define a group-invariant scalar product and therefore define the notion of orthogonality. Then we can choose a basis for the Lie algebra of G to be

$$t^a = (\tilde{t}^i, \theta^{\tilde{a}})$$

with the generators $\theta^{\tilde{a}}$ orthogonal to the $\tilde{t}^i$: $\mathrm{Tr}\tilde{t}^i\theta^{\tilde{a}} \;=\; 0$. Each vector $(\theta^{\tilde{a}}\phi_0)$ is then a unique zero eigenvector and implies there are dim$G$ $-$ dim$H$ massless modes, one for each broken generator $\theta^{\tilde{a}}$. These massless modes are called Goldstone bosons.[33] Since M^2 is rank N, then we should expect $N - (\mathrm{dim}G - \mathrm{dim}H)$ massive modes.

[33] or Nambu-Goldstone bosons.

In the $O(N)$ model we saw at the beginning of this subsection, the nonzero VEV broke the symmetry $O(N) \;\to\; O(N-1)$. Given that there are $N(N-1)/2$ generators of $O(N)$ (corresponding to the same number of planes of rotation) and $(N-1)(N-2)/2$ generators of $O(N-1)$, we should expect $N-1$ massless modes – one for each broken generator. This is what we found for the $N-1$ π fields.

We have just seen a classical proof of *Goldstone's theorem*. We will outline a fully quantum proof next.

4.3 *Goldstone's theorem*

In this section, we discuss spontaneous symmetry breaking at the fully quantum level, not resorting to arguments based on small departures from the minimum of the classical potential.

Let us assume that the symmetry group G of the Lagrangian is broken spontaneously down to a subgroup H. That is, a scalar field gets a nonzero vacuum expectation value [34]

$$\langle 0|\phi(x)|0\rangle \;=\; \phi_0 \;\neq\; 0\,. \tag{4.3.1}$$

[34] Here we consider scalar field theory, but in general ϕ could be a composite local operator constructed out of a different type of fundamental field, e.g. $\phi(x) = \bar{\psi}(x)\psi(x)$ in a gauge theory with Dirac fermion field $\psi(x)$.

The VEV is invariant under transformations in H: $\langle 0|h\phi(x)|0\rangle = \phi_0$ for $h \in H$. However the VEV is not invariant under transformations $\tilde{g} \in G$ where $\tilde{g} \notin H$. Let us distinguish between generators of the Lie algebra of G: t^a where $a \in [1, \dim G]$; and those of the Lie algebra of H: $\tilde{t}^i$ where $i \in [1, \dim H]$. Now, if G is a symmetry group of the Lagrangian, there are conserved currents $j^{a\mu}(x)$ and charges $Q^a = \int d^3x\, j^{a0}(x)$ associated with each generator. These charges induce a representation of the Lie algebra of G on ϕ

$$[Q^a, \phi(0)] \;=\; -it^a\phi(0)\,. \tag{4.3.2}$$

In order to investigate excitations which result from the spontaneous breaking of the global symmetry, we consider the VEV of the commutator of the conserved current with the scalar field $\langle 0|[j^{a\mu}(x), \phi(0)]|0\rangle$. It will be most instructive to use the *Källén-Lehmann spectral representation* of the two-point function. First define the spectral density functions as

$$\begin{aligned} i\rho^{a\mu}(k) \;&=\; (2\pi)^3 \sum_n \delta^{(4)}(k - p_n)\langle 0|j^{a\mu}(0)|n\rangle\langle n|\phi(0)|0\rangle \\ i\tilde{\rho}^{a\mu}(k) \;&=\; (2\pi)^3 \sum_n \delta^{(4)}(k - p_n)\langle 0|\phi(0)|n\rangle\langle n|j^{a\mu}(0)|0\rangle\,. \end{aligned} \tag{4.3.3}$$

Between the operators $j(x)$ and $\phi(0)$, we insert a complete set of states, $\sum_n |n\rangle\langle n| = 1$, and use the translation operator to write $j^{a\mu}(x) = e^{i\hat{P}\cdot x} j^{a\mu}(0) e^{-i\hat{P}\cdot x}$, obtaining

$$\langle 0|[j^{a\mu}(x), \phi(0)]|0\rangle \;=\; i\int \frac{d^4k}{(2\pi)^3}\left[\rho^{a\mu}(k)e^{-ik\cdot x} - \tilde{\rho}^{a\mu}(k)e^{ik\cdot x}\right]\,. \tag{4.3.4}$$

One can check this by inserting the expressions for $i\rho^{a\mu}(k)$ and $i\tilde{\rho}^{a\mu}(k)$ into the righthand side above and doing the integral over k.

The spectral density function can be further simplified. Lorentz invariance implies proportionality with k^μ, physical states imply $k^0 > 0$:

$$\begin{aligned} \rho^{a\mu}(k) \;&=\; k^\mu \Theta(k^0)\rho^a(k^2) \\ \rho^{a\mu}(k) \;&=\; k^\mu \Theta(k^0)\tilde{\rho}^a(k^2)\,. \end{aligned}$$

After substituting these into the integral (4.3.4) we can equate it to the derivative of another integral

$$\langle 0|[j^{a\mu}(x), \phi(0)]|0\rangle \;=\; -\partial^\mu \int \frac{d^4k}{(2\pi)^3}\,\Theta(k^0)\left[\rho^a(k^2)e^{-ik\cdot x} + \tilde{\rho}^a(k^2)e^{ik\cdot x}\right]\,.$$

Recalling the propagator given by

$$\langle 0|\phi(z)\phi(y)|0\rangle = \int \frac{d^3p}{(2\pi)^3 2p^0}\, e^{-ip\cdot(z-y)}\bigg|_{p^0=\sqrt{|\vec{p}|^2+\sigma}}$$
$$= \int \frac{d^4p}{(2\pi)^3}\Theta(p^0)\delta(p^2-\sigma)e^{-ip\cdot(z-y)} \equiv D(z-y;\sigma) \tag{4.3.5}$$

We can replace $\rho(k^2)$ by $\int d\sigma\,\rho(\sigma)\delta(k^2-\sigma)$. We can write

$$\langle 0|[j^{a\mu}(x),\phi(0)]|0\rangle \;=\; -\partial^\mu\int d\sigma\;[\rho^a(\sigma)D(x;\sigma)+\tilde{\rho}^a(\sigma)D(-x;\sigma)]\;.$$

For spacelike x^2 (i.e. $x^2<0$) then $D(x;\sigma)=D(-x;\sigma)$.[35] Requiring that $\langle 0|[j^{a\mu}(x),\phi(0)]|0\rangle|_{x^2<0}=0$ implies that we must have $\rho^a(\sigma) = -\tilde{\rho}^a(\sigma)$. Thus

$$\langle 0|[j^{a\mu}(x),\phi(0)]|0\rangle \;=\; -\partial^\mu\int d\sigma\rho^a(\sigma)i\Delta(x;\sigma) \tag{4.3.6}$$

where

$$i\Delta(x;\sigma) = D(x;\sigma)\;-\;D(-x;\sigma)$$
$$= \int \frac{d^4k}{(2\pi)^3}\delta(k^2-\sigma)\,\epsilon(k^0)\,e^{-ik\cdot x} \tag{4.3.7}$$

with $\epsilon(k^0) = \mp 1$ for $k^0 \lessgtr 0$. In obtaining the last line above, we changed integration variable $k^\mu \mapsto -k^\mu$ in $D(-x;\sigma)$.

Current conservation $\partial_\mu j^{a\mu}=0$ and the Klein-Gordan equation $(\partial^2+\sigma)\Delta=0$ imply that when we differentiate (4.3.6) we obtain

$$0 \;=\; \int d\sigma\,\sigma\rho^a(\sigma)i\Delta(x;\sigma)$$

The fact that this is true for all x (in particular timelike x where $\Delta(x;\sigma)\neq 0$) implies

$$\sigma\rho^a(\sigma) \;=\; 0\,. \tag{4.3.8}$$

There are 2 cases for the $\dim G$ spectral densities.

1. $\rho^a(\sigma)=0$. This implies $\langle 0|[j^{a\mu}(x),\phi(0)]|0\rangle=0$, i.e. that t^a is not a broken generator.
2. $\rho^a(\sigma)=N^a\delta(\sigma)$, with N^a a dimensionful nonzero constant.

It is the second case which is interesting here. Write

$$\langle 0|[j^{a\mu}(x),\phi(0)]|0\rangle = -\partial^\mu\int d\sigma\,N^a\delta(\sigma)\,i\Delta(x;\sigma) = -iN^a\partial^\mu\Delta(x;0)\,.$$

To get an expression for the commutator of the charge Q^a with the field, we will want to integrate the equation above for $\mu=0$. First let us prove an useful identity, $\int d^3x\,\Delta(x,0) \;=\; -x_0$. Since $\int d^3x\exp(i\vec{k}\cdot\vec{x}) = (2\pi)^3\delta^{(3)}(\vec{k})$,

$$\int d^3x\,i\Delta(x,0) \;=\; \lim_{\sigma\to 0}\int dk^0\,\delta((k^0)^2-\sigma)\,\epsilon(k^0)\,e^{-ik^0x_0}$$
$$=\lim_{\sigma\to 0}\int dk^0\left[\frac{\delta(k^0-\sqrt{\sigma})}{|2\sqrt{\sigma}|}+\frac{\delta(k^0+\sqrt{\sigma})}{|-2\sqrt{\sigma}|}\right]\epsilon(k^0)\,e^{-ik^0x_0}$$
$$=\lim_{\sigma\to 0}\frac{1}{2\sqrt{\sigma}}\left(e^{-i\sqrt{\sigma}\,x_0}-e^{i\sqrt{\sigma}\,x_0}\right) \;=\; -ix_0\,.$$

[35] E.g. see §5.2 of Weinberg.
S Weinberg. *The Quantum Theory of Fields, Volume I*. Cambridge University Press, 1995. ISBN 0-521-55001-7

Thus we find (4.3.2) becomes

$$t^a\phi_0 \;=\; i\langle 0|[Q^a,\phi(0)]|0\rangle \;=\; N^a\int d^3x\,\partial^0\Delta(x;0) \;=\; -N^a\,. \qquad (4.3.9)$$

The fact that $N^a \neq 0$ and $\phi_0 \neq 0$ implies some of the states in the sums (4.3.3) have nonvanishing matrix elements. Let us label those states by B and their momentum by p. Dimensional analysis and Lorentz covariance implies we can parametrize the matrix elements of the current and field as

$$\langle 0|j^{a\mu}(0)|B(p)\rangle \;=\; iF^a_B p^\mu \qquad (4.3.10)$$

$$\langle B(p)|\phi(0)|0\rangle \;=\; Z^B \qquad (4.3.11)$$

where the F^a_B are dimension-1 constants and Z^B are dimensionless constants. Note the states $|B(p)\rangle$ are spinless since $\phi(0)|0\rangle$ is rotationally invariant and massless since $\rho^a(\sigma)\delta(\sigma)$ only contributes for $\sigma = p^2 = 0$.

Inserting $\rho^a(\sigma) = N^a\delta(\sigma)$ into (4.3.3), where now the sum over complete states is an integral over the momenta of the Goldstone boson states $|B(p)\rangle$

$$ik^\mu\Theta(k^0)N^a\delta(k^2) \;=\; \sum_B\int\frac{d^3p}{2|\vec{p}|}\,\delta^{(4)}(k-p)\langle 0|j^{a\mu}(0)|B(p)\rangle\langle B(p)|\phi(0)|0\rangle\,.$$

Simplifying the right-hand side and expressing the left-hand side as an integral we find

$$\int\frac{d^3p}{2|\vec{p}|}\,\delta^{(4)}(k-p)\,ik^\mu N^a \;=\; \int\frac{d^3p}{2|\vec{p}|}\,\delta^{(4)}(k-p)\,ip^\mu\sum_B F^a_B Z^B$$

which implies

$$N^a \;=\; \sum_B F^a_B Z^B\,.$$

Since there are $\dim H$ generators of H which are unbroken, there are exactly $d = \dim G - \dim H$ broken generators, and the same number of densities $\rho^a(\sigma)$ which have nonzero contributions at $\sigma = 0$. Therefore (F^a_B) is a matrix of rank n and there are n Goldstone bosons.

Nota bene, we assumed that we were working with a Lorentz invariant theory in spacetime dimensions greater than 2. The counting of the number of Goldstone modes is more subtle in nonrelativistic theories. The Coleman-Merman-Wagner theorem trumps the Goldstone theorem in 1 and 2 dimensional theories. Finally, the proof of Goldstone's theorem requires the space of states to have positive definite norm, therefore gauge theories are exempt as we should expect from the success of the Higgs mechanism and electroweak theory.

4.4 *Higgs mechanism*

Or the Anderson-Brout-Englert-Guralnik-Hagen-Higgs-Kibble-'t Hooft mechanism.[36]

[36] F Close. *The Infinity Puzzle*. Oxford University Press, 2011. ISBN 978-0199593507

Gauge theories do not satisfy all of the axioms supposed in Goldstone's theorem; depending on the choice of gauge, one of the axioms must be violated. Taking QED as an example, if we quantize imposing a Lorentz invariant gauge condition, such as Lorenz gauge,[37] then the theory contains negative-norm states. On the other hand, one can quantize in a gauge, e.g. radiation gauge, which yields a theory without negative-norm states, but at the expense of breaking Lorentz invariance.

[37] The Dutch Hendrik Antoon Lorentz (1853-1928) and the Danish Ludvig Lorenz (1829-1891) are different people. Neither should be confused with chaos pioneer Edward Norton Lorenz (1917-2008, Lorenz attractor), nor with economist Max Otto Lorenz, (1880-1962, Lorenz curve).

Let us consider a theory of *scalar electrodynamics*, i.e. a complex scalar field $\phi(x)$ which interacts with a photon $A_\mu(x)$ and with itself. The Lagrangian is

$$\mathcal{L} \;=\; -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \;+\; (D_\mu\phi)^*(D^\mu\phi) \;-\; V(\phi^*\phi) \qquad (4.4.1)$$

where $F_{\mu\nu} \;=\; \partial_\mu A_\nu - \partial_\nu A_\mu$ and $D_\mu \;\equiv\; \partial_\mu + iqA_\mu$.[38] U(1) gauge invariance implies the fields transform as

[38] If we write $\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$, with real ϕ_1, ϕ_2 we see the kinetic term here has the canonical normalization.

$$\begin{aligned} \phi(x) \;&\mapsto\; e^{i\alpha(x)}\phi(x) \\ A_\mu(x) \;&\mapsto\; A_\mu(x) - \frac{1}{q}\partial_\mu\alpha(x)\,. \end{aligned} \qquad (4.4.2)$$

Let us take for the scalar potential

$$V(\phi^*\phi) \;=\; \mu^2|\phi|^2 \;+\; \lambda|\phi|^4\,, \quad \text{with } \lambda > 0\,. \qquad (4.4.3)$$

If $\mu^2 > 0$ then the quadratic term is a usual mass term, the potential has a unique minimum, the $U(1)$ symmetry is preserved by the vacuum, and the physics is that of a massless photon and a massive complex scalar.

Now consider the case that $\mu^2 < 0$. The minima of the potential (4.4.3) satisfy

$$|\phi_0|^2 \;=\; -\frac{\mu^2}{2\lambda} \;\equiv\; \frac{v^2}{2}$$

or

$$\phi_0 \;=\; \frac{v}{\sqrt{2}}\,e^{i\zeta_0/v}\,, \quad \text{with } \; v > 0 \text{ and } \zeta_0 \in \mathbb{R}\,.$$

Without loss of generality, let us choose to expand about the vacuum with $\zeta_0 = 0$. Then the field can fluctuation in modulus and in phase

$$\phi(x) \;=\; \frac{e^{i\zeta(x)/v}}{\sqrt{2}}\,(v + \eta(x))$$

Assuming small fluctuations about the VEV, we can expand the exponential

$$\phi \;\simeq\; \frac{1}{\sqrt{2}}\,(v + \eta + i\zeta)$$

to obtain the Lagrangian

$$\mathcal{L} \;=\; \frac{1}{2}(\partial_\mu\eta\partial^\mu\eta \;+\; 2\mu^2\eta^2) \;+\; \frac{1}{2}\partial_\mu\zeta\partial^\mu\zeta \;-\; \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \;+\; qvA_\mu\partial^\mu\zeta \;+\; \frac{q^2v^2}{2}A_\mu A^\mu \;+\; \mathcal{L}_{\text{int}} \qquad (4.4.4)$$

where the last term represents the contributions from $\lambda|\phi|^4$. We appear to have mass terms for the η and A_μ fields, but none for the

ζ, the would-be Goldstone bosons. Note the unusual term $A_\mu \partial^\mu \zeta$, however. We can rewrite the terms containing A_μ and ζ as

$$\frac{q^2 v^2}{2}\left(A_\mu + \frac{1}{qv}\partial_\mu \zeta\right)\left(A^\mu + \frac{1}{qv}\partial^\mu \zeta\right).$$

Then, we can transform gauge, making a very specific choice of gauge called *unitary gauge*. We can work with new fields which differ from the original ones by a gauge transformation, i.e. (4.4.2) with $\alpha(x) = -\frac{1}{v}\zeta(x)$:

$$\begin{aligned} A'_\mu &= A_\mu + \frac{1}{qv}\partial_\mu \zeta \\ \phi' &= e^{-i\zeta/v}\phi = \frac{1}{\sqrt{2}}(v+\eta)\,. \end{aligned}$$

We could have begun working in unitary gauge simply by performing a $U(1)$ gauge transformation at the start which ensured the fluctuations in $\phi(x)$ remained real. The unitary gauge Lagrangian is $\mathcal{L} = \mathcal{L}^{\text{quad}} + \mathcal{L}^{\text{int}}$ plus an irrelevant constant:[39]

[39] Now dropping the primes from the A_μ field.

$$\mathcal{L}^{\text{quad}} = \frac{1}{2}(\partial_\mu \eta \partial^\mu \eta + 2\mu^2\eta^2) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{q^2v^2}{2}A_\mu A^\mu$$

The photon now has a mass, $m_A^2 = q^2v^2$. There is a massive scalar[40] $m_\eta^2 = -2\mu^2 = 2\lambda v^2$. The Goldstone mode ζ has been "eaten" to become the longitudinal polarization of the A_μ.

[40] the Higgs boson for this theory

The interacting part of the Lagrangian, in unitary gauge is

$$\mathcal{L}^{\text{int}} = \frac{q^2}{2}A_\mu A^\mu \eta^2 + q m_A A_\mu A^\mu \eta - \frac{\lambda}{4}\eta^4 - m_\eta\sqrt{\frac{\lambda}{2}}\eta^3\,.$$

Feynman rules. . . .

4.5 *Nonabelian theories*

In the next chapter we come to the full electroweak theory, which employs the Higgs mechanism to break $SU(2)_L \times U(1)_Y$ down to $U(1)_{EM}$ in order to give mass to the weak gauge bosons. In the examples sheets you will consider other gauge theories, e.g. where $SU(2)$ is broken to $U(1)$.

Usually the symmetry breaking follows a similar pattern to that seen in § 4.2: the potential V is minimized when some components of the scalar field ϕ are nonzero. This breaks a symmetry G of the Lagrangian. The difference is that G is a gauge symmetry, so that ϕ is coupled to a gauge field through the covariant derivative

$$D_\mu \phi = (\partial_\mu + igt^a A^a_\mu)\phi\,.$$

The steps of § 4.4 are repeated, taking care with the noncommuting generators t^a.

In the examples sheet you consider not only the case where the scalar field transforms in the fundamental (2-dimensional) representation of $SU(2)$ (where $t^a = \sigma^a/2$), but also where ϕ is in the adjoint (3-dimensional) representation. A convenient explicit matrix representation of the generators in this latter case is given by

$$(t^a)_{jk} = -i\epsilon_{ajk} \,.$$

Electroweak theory

The electroweak theory is attributed primarily to Steven Weinberg and Abdus Salam with important work earlier by Sheldon Glashow. The goal here is to construct a gauge theory for the weak interactions of Nature utilizing the Higgs mechanism to give the weak gauge bosons mass. Throughout this chapter we will make a number of *choices* in order to *construct* a theory which is capable of describing experimental data. One could write down consistent field theories making different choices, but then these would likely give predictions which contradict existing measurements.

5.1 Gauge theory

We work with the direct product of $SU(2) \times U(1)$.[41] This gauge symmetry will be broken by the Higgs mechanism in order to explain how the weak gauge bosons become massive.

[41] We probably should not say that we have a *unified* theory of electroweak interactions, since we have to introduce 2 gauge couplings. It is at most a unified framework.

Introduce a scalar field (the Higgs field) in the doublet representation of $SU(2)$ with a $U(1)$ (hyper)charge $\frac{1}{2}$. Under $SU(2) \times U(1)$ gauge transformations

$$\phi(x) \mapsto e^{i\alpha^a(x)\tau^a} e^{i\beta(x)/2} \phi(x) \tag{5.1.1}$$

with the $SU(2)$ generators $\tau^a = \frac{\sigma^a}{2}$. In the unbroken theory, we would have $3+1$ massless gauge bosons. Let the scalar acquire a VEV, without loss of generality choose

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} .$$

This VEV breaks the $SU(2)_L \times U(1)_Y$ symmetry down to only a $U(1)_{EM}$: the theory is still invariant under (5.1.1) with $\alpha^1 = \alpha^2 = 0$ and $\alpha^3(x) = \beta(x)$.

The covariant derivative for the electroweak theory is

$$D_\mu \phi = \left(\partial_\mu + i g W_\mu^a \tau^a + \tfrac{i}{2} g' B_\mu \right) \phi \tag{5.1.2}$$

The W_μ^a are the three $SU(2)$ gauge bosons and B_μ is the $U(1)_Y$ gauge boson. The part of the electroweak Lagrangian concerned solely with the gauge and scalar fields is then

$$\mathcal{L} = -\frac{1}{4} \mathrm{Tr} F^W_{\mu\nu} F^{W,\mu\nu} - \frac{1}{4} F^B_{\mu\nu} F^{B,\mu\nu} + (D_\mu \phi)^\dagger (D^\mu \phi) - \mu^2 |\phi|^2 - \lambda |\phi|^4 \tag{5.1.3}$$

where the $SU(2)$ and $U(1)$ field strength tensors are[42]

$$F^{W,a}_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g\epsilon^{abc} W^b_\mu W^c_\nu$$
$$F^B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \,.$$

[42] Using $f^{abc} = \epsilon^{abc}$ in the fundamental representation of $SU(2)$.

Spontaneous symmetry breaking is assumed to occur, letting $\mu^2 = -\lambda v^2 < 0$.

As in § 4.4, the Higgs VEV generates mass as follows. From the term in the Lagrangian $(D_\mu\phi)^\dagger(D^\mu\phi)$ we have

$$\begin{aligned}\frac{1}{2}(0,v)\left(-igW^a_\mu\tau^a - \frac{i}{2}g'B_\mu\right)\left(igW^{a\mu}\tau^a + \frac{i}{2}g'B^\mu\right)\begin{pmatrix}0\\ v\end{pmatrix}\\ = \frac{1}{2}\frac{v^2}{4}\left[g^2(W^1)^2 + g^2(W^2)^2 + (-gW^3 + g'B)^2\right] \qquad (5.1.4)\end{aligned}$$

This term in the Lagrangian evidently generates 3 mass terms. Let us define 4 new gauge boson fields in terms of the linear combinations which appear above (as well as one which doesn't). Let

$$\begin{aligned} W^\pm_\mu &= \frac{1}{\sqrt{2}}\left(W^1_\mu \mp iW^2_\mu\right)\\ Z^0_\mu &= \frac{1}{\sqrt{g^2+g'^2}}\left(gW^3_\mu - g'B_\mu\right)\\ A_\mu &= \frac{1}{\sqrt{g^2+g'^2}}\left(g'W^3_\mu + gB_\mu\right) \qquad (5.1.5)\end{aligned}$$

Then (5.1.4) endows these fields with with masses $m_W = vg/2$ and $m_Z = v\sqrt{g^2+g'^2}/2$. Corresponding to the A_μ field, we have the massless photon $m_\gamma = 0$ of electromagnetism.

The mixing of the $SU(2)$ and $U(1)$ gauge bosons is governed by the weak mixing angle, or the Weinberg angle, θ_W defined by

$$\begin{pmatrix}Z^0\\ A\end{pmatrix} = \begin{pmatrix}\cos\theta_W & -\sin\theta_W\\ \sin\theta_W & \cos\theta_W\end{pmatrix}\begin{pmatrix}W^3\\ B\end{pmatrix} \qquad (5.1.6)$$

and

$$\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}\,, \quad \sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}}$$

In terms of the Weinberg angle, $m_W = m_Z\cos\theta_W$.[43]

[43] Experimentally, $m_W = 80.385(15)$ GeV and $m_Z = 91.1876(21)$ GeV, while $m_\gamma < 10^{-18}$ eV.

J Beringer et al. Review of Particle Physics (RPP). *Phys. Rev.*, D86:010001, 2012

5.2 *Coupling to matter*

Now we discuss how fermions participate in electroweak interactions. We explicitly deal with lepton interactions in this section although the general steps are the same for quarks. However, additional ingredients are necessary in the latter case, so we defer some of those details to § 5.3.

In terms of the physical gauge bosons in the spontaneously broken theory the covariant derivative may be written

$$\begin{aligned} D_\mu &= \partial_\mu + igW^a_\mu T^a + ig'YB_\mu\\ &= \partial_\mu + \frac{ig}{\sqrt{2}}\left(W^+_\mu T^+ + W^-_\mu T^-\right) + \frac{igZ_\mu}{\cos\theta_W}\left(\cos^2\theta_W T^3 - \sin^2\theta_W Y\right) + ig\sin\theta_W A_\mu(T^3+Y)\end{aligned}$$

where $T^{\pm} = T^1 \pm iT^2$. We identify A_μ as the photon, so to identify the physical couplings to fermions, e.g. the electron, we write

$$\begin{aligned} Q &= T^3 + Y \quad \text{is the } U(1)_{EM} \text{ charge matrix} \\ e &= g \sin\theta_W \quad \text{is the electron charge (magnitude).} \end{aligned} \tag{5.2.1}$$

Therefore

$$D_\mu \;=\; \partial_\mu \;+\; \frac{ig}{\sqrt{2}} \left(W_\mu^+ T^+ + W_\mu^- T^- \right) \;+\; \frac{igZ_\mu}{\cos\theta_W} \left(T^3 - \sin^2\theta_W Q \right) \;+\; ieA_\mu Q \tag{5.2.2}$$

In the Standard Model, the left-handed components of leptons and quarks transform in the fundamental representation of $SU(2)_L$. The corresponding generators are related to the Pauli matrices $T^a = \tau^a = \sigma^a/2$. Let us introduce a doublet to describe the left-handed electron and electron neutrino

$$L(x) \;=\; \begin{pmatrix} \nu_e(x) \\ e_L(x) \end{pmatrix} \tag{5.2.3}$$

where $e_L(x) = \frac{1}{2}(1-\gamma^5)e(x)$. Guided by experiment, we do not couple the right-handed components of the electrons to the weak bosons. This means that $R(x) = e_R(x)$ transforms trivially under $SU(2)_L$, or that it transforms in the trivial representation where the generator is $T = 0$.

At this stage in the discussion we assume that the neutrinos are massless and strictly left-handed. This assumption was consistent with experimental results until 1998-2001 when neutrino oscillations were conclusively observed. We will return to a discussion of how to amend the Standard Model to account for neutrino oscillations later, but for most purposes we can still treat the neutrinos as massless in this course.

We know that the electron has a negative electric charge and that the neutrinos are neutral:

$$QL(x) \;=\; \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} L(x) \;\text{ and }\; QR(x) = -R(x)$$

From this we can infer the hypercharges from $Q = T^3 + Y$ as

$$\begin{aligned} T = \tfrac{1}{2} \;&\Rightarrow\; Q = \tau^3 + Y \quad \text{so} \quad Y = -\tfrac{1}{2} \;\text{ for }\; L(x) \\ T = 0 \;&\Rightarrow\; Q = 0 + Y \quad \text{so} \quad Y = -1 \;\text{ for }\; R(x) \end{aligned}$$

With these assignments, we can use (5.2.2) to write the lepton-gauge boson part of the electroweak Lagrangian as

$$\mathcal{L}_{\text{lept}} \;=\; \bar{L}\, i\not{D}\, L \;+\; \bar{R}\, i\not{D}\, R\,. \tag{5.2.4}$$

The electron's heavy cousins (μ, τ) and their neutrinos can be included simply by introducing 2 more $SU(2)_L$ doublets and corresponding right-handed fields.

Lepton mass: $m_e(\bar{e}_L e_R + \bar{e}_R e_L)$ cannot appear in $\mathcal{L}$ in an $SU(2)_L \times U(1)_Y$-invariant way. Instead the Higgs mechanism gives the charged leptons mass.

Higgs doublet ϕ with $Y = \frac{1}{2}$.

$$\mathcal{L}_{\text{lept},\phi} \;=\; -\sqrt{2}\lambda_e(\bar{L}\phi R + \bar{R}\phi^\dagger L)$$

Given that the symmetry is broken, we can expand about

$$\phi(x) \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}$$

where we implicitly imposed the unitary gauge condition.[44]

$$\mathcal{L}_{\text{lept},\phi} \;=\; -\lambda_e(v+h)(\bar{e}_L e_R + \bar{e}_R e_L) \;=\; -m_e\bar{e}e - \lambda_e h\bar{e}e \tag{5.2.5}$$

with the electron mass given by the product of the Higgs VEV and the Yukawa coupling λ_e.[45]

Writing out the covariant derivative (5.2.2), we get the gauge interactions of the leptons

$$\begin{aligned}\mathcal{L}^{\text{int}}_{\text{lept}} &= g\bar{L}\gamma^\mu\tau^a W^a_\mu L - g'\left(\frac{1}{2}\bar{L}\gamma^\mu L + \bar{R}\gamma^\mu R\right) B_\mu \\ &= \frac{g}{2\sqrt{2}}\left(J^\mu W^+_\mu + J^{\mu\dagger} W^-_\mu\right) + e j^\mu_{EM} A_\mu + \frac{g}{2\cos\theta_W} J^\mu_n Z_\mu\,. \end{aligned}\tag{5.2.6}$$

We have introduced the charged weak current, the neutral weak current, and the electromagnetic current:

$$\begin{aligned} J^\mu &= 2\bar{L}\gamma^\mu\sigma^+ L \;=\; \bar{\nu}_e\gamma^\mu(1-\gamma^5)e \\ J^\mu_n &= \bar{L}\gamma^\mu\left(\cos^2\theta_W\sigma^3 \,+\, \sin^2\theta_W\right)L \;+\; 2\sin^2\theta_W\bar{R}\gamma^\mu R \\ &= \frac{1}{2}\left[\bar{\nu}_e\gamma^\mu(1-\gamma^5)\nu_e \,-\, \bar{e}\gamma^\mu(1-\gamma^5-4\sin^2\theta_W)e\right] \\ j^\mu_{EM} &= \frac{1}{2}\bar{L}\gamma^\mu(\sigma^3-1)L \,-\, \bar{R}\gamma^\mu R \;=\; -\bar{e}\gamma^\mu e \end{aligned}\tag{5.2.7}$$

We can include more generations of leptons. In the Standard Model, the electron has 2 heavy cousins, the μ and the τ.

$$\begin{array}{ccc} L^1 = \begin{pmatrix}\nu_e \\ e\end{pmatrix}_L & L^2 = \begin{pmatrix}\nu_\mu \\ \mu\end{pmatrix}_L & L^3 = \begin{pmatrix}\nu_\tau \\ \tau\end{pmatrix}_L \\ R^1 = e_R & R^2 = \mu_R & R^3 = \tau_R \end{array}\,.$$

Then the coupling of the leptons to the Higgs field is via

$$\mathcal{L}_{\text{lept},\phi} \;=\; -\sqrt{2}\left(\lambda^{ij}\bar{L}^i\phi R^j \,+\, (\lambda^\dagger)^{ij}\bar{R}^i\phi^\dagger L^j\right)\,. \tag{5.2.8}$$

The matrix λ may be diagonalized as follows. $\lambda\lambda^\dagger$ and $\lambda^\dagger\lambda$ Hermitian implies there exist unitary matrices U, S such that

$$\lambda\lambda^\dagger = U\Lambda^2 U^\dagger \;\text{ and }\; \lambda^\dagger\lambda = S\Lambda^2 S^\dagger$$

with Λ^2 a diagonal matrix possessing nonnegative eigenvalues.[46] Then $\lambda = U\Lambda S^\dagger$.

We can transform the lepton fields

$$L^i \mapsto U^{ij}L^j\,,\;\; \bar{L}^i \mapsto \bar{L}^j(U^\dagger)^{ji}\,,\;\; R^i \mapsto S^{ij}R^j\,,\;\; \bar{R}^i \mapsto \bar{R}^j(S^\dagger)^{ji}$$

which diagonalizes $\mathcal{L}_{\text{lept},\phi}$ while leaving $\mathcal{L}_{\text{lept}}$ (the generalization of (5.2.4)) invariant. The fact that we can perform this simultaneous diagonalization means that the freely propagating leptons (sometimes called mass eigenstates) are also eigenstates of the weak Hamiltonian which we will define next chapter. At least within present experimental precision, the weak interactions do not induce mixing between lepton generations.

[44] We could have chosen $\phi(x) = U(x)(0, v+h(x))^T/\sqrt{2}$ for $U(x) \in SU(2)$. Then transformation to unitary gauge is the one which exactly cancels out the $U(x)$.

[45] In tribute to Hideki Yukawa's theory of nucleons interacting with a (pseudo)scalar pion, we refer to any local coupling of fermions to a scalar field as a Yukawa interaction.

[46] Let v be a normalized eigenvector of λ with eigenvalue α. $\lambda v = \alpha v \Rightarrow v^\dagger\lambda^\dagger = \alpha^* v^\dagger$. Thus $v^\dagger\lambda^\dagger\lambda v = |\alpha|^2 \geq 0$.

5.3 Quark flavour

As far as we know, there are 6 quarks in Nature, each its own "flavour."[47] The weak interactions couple them in pairs through $SU(2)_L$ doublets

$$Q_L^i \;=\; \begin{pmatrix} u^i \\ d^i \end{pmatrix}_L \;=\; \left(\begin{pmatrix} u \\ d \end{pmatrix}_L , \begin{pmatrix} c \\ s \end{pmatrix}_L , \begin{pmatrix} t \\ b \end{pmatrix}_L \right) .$$

[47] Model builders often investigate the consequences of adding another pair of quarks to the Standard Model. So far, while this addition can sometimes explain some things, specific models have so far run into conflict with experimental data. Nevertheless, both theorists and experimentalists actively keep in mind the possibility of a "fourth generation" of quarks and leptons. (The number of quark and lepton doublets should be equal in order to be free of a gauge anomaly.)

Each $SU(2)_L$ doublet is assigned hypercharge $Y = \frac{1}{6}$ in order that we get the correct electric charges after the symmetry is broken. We also have right-handed $SU(2)_L$ singlets $u_R^i = (u_R, c_R, t_R)$ and $d_R^i = (d_R, s_R, b_R)$ with hypercharges $\frac{2}{3}$ and $-\frac{1}{3}$ respectively. The quarks couple to the $SU(2) \times U(1)$ gauge bosons via the Lagrangian

$$\mathcal{L}_{\text{quark}}^{\text{weak}} \;=\; \bar{Q}_L\, i\not{D}\, Q_L \;+\; \bar{u}_R\, i\not{D}\, u_R \;+\; \bar{d}_R\, i\not{D}\, d_R\,. \tag{5.3.1}$$

The quark-Higgs couplings take the general form

$$\mathcal{L}_{\text{quark},\phi} \;=\; -\sqrt{2}\left[\lambda_d^{ij}\bar{Q}_L^i \phi d_R^j \;+\; \lambda_u^{ij}\epsilon^{\alpha\beta}\bar{Q}_L^{\alpha,i}\phi^{\dagger\beta}u_R^j \;+\; \text{h.c.}\right] \tag{5.3.2}$$

where $i, j = 1, 2, 3$ are generation indices and $\alpha, \beta = 1, 2$ are $SU(2)$ representation indices. In order to have a hypercharge-neutral term in the Lagrangian coupling $\bar{Q}_L$ and u_R, we need $\phi^\dagger$ instead of ϕ. Given Q_L transforms in the fundamental representation of $SU(2)$, $\bar{Q}_L$ transforms in the antifundamental representation obtained by complex-conjugating the generators of the group in the fundamental representation. In order to have an invariant Lagrangian, the $\bar{Q}_L$ must be multiplied by a field which transforms in the fundamental representation. We can achieve this with $\epsilon^{\alpha\beta}\phi^{\dagger\beta} \;=\; (\phi^c)^\alpha$, a field which transforms in the fundamental representation.[48]

[48] The representations of $SU(2)$ are pseudoreal. "Pseudoreal" means $\exists V$ such that $-(T^a)^* \;=\; V^{-1}T^aV$; in fact $V^{\alpha\beta} \;=\; (i\sigma^2)^{\alpha\beta} \;=\; \epsilon^{\alpha\beta}$ does the job for the τ^a.

Note that while $\mathcal{L}_{\text{quark}}^{\text{weak}}$ does not respect C and P invariance, the product CP as well as T are symmetries of $\mathcal{L}_{\text{quark}}^{\text{weak}}$. $\mathcal{L}_{\text{gauge}}$ and $\mathcal{L}_\phi$ are invariant under each of C, P, and T. $\mathcal{L}_{\text{quark},\phi}$ is CP invariant if and only if

$$\lambda_q^{ij} = (\lambda_q^{ij})^* \quad \text{with } q = u, d\,,$$

that is, if and only if the matrix elements λ_q^{ij} are real.

We expect $\mathcal{L}_{\text{quark},\phi}$ to contain mass terms for the quarks when we expand about the Higgs VEV. We find these by diagonalizing the Yukawa matrices λ_u and λ_d:

$$\lambda_u \;=\; U_u\Lambda_u S_u^\dagger \;\;,\;\; \lambda_d \;=\; U_d\Lambda_d S_d^\dagger$$

with $U, S \;\in\; U(3)$ and Λ_u, Λ_d diagonal. We transform the quark fields as

$$u_L \mapsto U_u u_L \;\;,\quad d_L \mapsto U_d d_L \;\;,\quad u_R \mapsto S_u u_R \;\;,\quad d_R \mapsto S_d d_R \tag{5.3.3}$$

and the Dirac adjoint fields with the corresponding Hermitian adjoint matrices. Then we have, for example,

$$\lambda_d^{ij}\bar{Q}_L^i\phi d_R^j \;\mapsto\; \bar{Q}_L\phi U_d\Lambda_d S_d^\dagger S_d d_R = \bar{Q}_L\phi U_d\Lambda_d d_R\,.$$

The terms in $\mathcal{L}_{\text{quark},\phi}$ are diagonal in quark generation. Setting the Higgs field to its VEV $\phi = \frac{1}{\sqrt{2}}(0, v)^T$ we obtain mass terms for the quarks[49]

$$\begin{aligned}\mathcal{L}_{\text{quark},\phi}|_{\phi=\phi_0} &= -v\Lambda_d^{ij}\bar{d}_L^i d_R^j - v\Lambda_u^{ij}\bar{u}_L^i u_R^j + \text{h.c.}\\ &= -\sum_i \left(m_d^i \bar{d}_L^i d_R^i + m_u^i \bar{u}_L^i u_R^i + \text{h.c.} \right) . \end{aligned} \tag{5.3.4}$$

[49] Note the quark masses, or equivalently the Yukawa couplings, are not determined by the Standard Model; they are free parameters which must be inferred empirically. The disparate range of quark masses is something of a mystery: the up, down, strange, charm, bottom, and top quark masses are respectively 0.002, 0.005, 0.1, 1, 4, 170 GeV (precise values first require specifying a regularization scheme and renormalization scale).

In this basis the mass and Higgs coupling terms in the Standard Model are C, P, and T invariant.

The transformation (5.3.3) has not left the rest of the Lagrangian (5.3.1) alone. The latter 2 terms $\bar{u}_R i\not{D} u_R$ and $\bar{d}_R i\not{D} d_R$ are invariant, but $\bar{Q}_L i\not{D} Q_L$ is not. Specifically the interaction terms involving quarks coupling to $W^\pm$, the charged currents $J^\pm$ which appear in the Lagrangian as $\frac{g}{2\sqrt{2}} J^{\pm,\mu} W^\pm_\mu$ are transformed

$$J^{\mu,+} = \bar{u}_L^i \gamma^\mu d_L^i \mapsto \bar{u}_L^i \gamma^\mu (U_u^\dagger U_d)^{ij} d_L^j .$$

Writing the Lagrangian in terms of fields which have diagonal mass terms exposes inter-generational couplings between weak doublets. Thinking in terms of corresponding terms in the quantum mechanical Hamiltonian, we say the weak eigenstates are linear combinations of the mass eigenstates.

The mixing matrix is called the Cabibbo-Kobyashi-Maskawa matrix[50] Conventionally it is written

[50] Cabibbo first described this mixing for 2 generations. Kobyashi and Maskawa generalized to 3 generations *before* either the b or t quark were discovered.

$$U_u^\dagger U_d = V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} . \tag{5.3.5}$$

The matrix is not determined by the Standard Model. Its matrix elements must be determined experimentally.

Not every CKM matrix element is independent. V_{CKM} is unitary, which imposes relations between matrix elements. Furthermore, we can use global phase-invariance of the quark fields to eliminate seemingly free parameters. We examine two cases here, those for 2 and 3 generations of quarks.

In the 2 generation case, unitarity alone implies V has 4 free parameters which can be express as an angle and 3 phases

$$V = \begin{pmatrix} \cos\theta_c \, e^{i\alpha} & \sin\theta_c \, e^{i\beta} \\ -\sin\theta_c \, e^{i(\alpha+\gamma)} & \cos\theta_c \, e^{i(\beta+\gamma)} \end{pmatrix} .$$

Terms in the Lagrangian are invariant under global $U(1)$ transformations of any quark field, say q_L^i where $q^i \in \{u, d, s, c\}$:

$$q_L^i \mapsto e^{i\alpha^i} q_L^i \quad (\text{no sum on } i) . \tag{5.3.6}$$

By transforming one field relative to the others we can eliminate a phase in V. Since we have 4 fields, we can perform 3 such transformations to eliminate α, β, and γ.

$$V = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \tag{5.3.7}$$

where we call θ_c the Cabibbo angle.[51] The charged weak current is then

$$J^{\mu,+} = \cos\theta_c \bar{u}_L \gamma^\mu d_L + \sin\theta_c \bar{u}_L \gamma^\mu s_L - \sin\theta_c \bar{c}_L \gamma^\mu d_L + \cos\theta_c \bar{c}_L \gamma^\mu s_L\,.$$

[51] From experiment we infer that $\sin\theta_c \approx 0.22$.

The same arguments apply in the 3-generation case we see in Nature, with one important difference. A 3×3 unitary matrix has 9 independent parameters; these can be written as 3 angles and 6 phases. However, we only have 6 quark fields which we can transform as in (5.3.6), only 5 phases can be eliminated (corresponding to the 5 phase differences). Therefore, the CKM matrix in the Standard Model has 4 free parameters, which can be written as 3 angles and 1 phase.

It is usual to follow Wolfenstein and make the following parametrization instead of using angles. Making the empirical observation that $\lambda \equiv V_{us} \approx \sin\theta_c \approx 0.22 \ll 1$ one can expand

$$V_{CKM} = \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \ldots \quad (5.3.8)$$

Experimental (and theoretical) precision is good enough that in practice the next order terms are included and new parameters $\bar{\rho}$ and $\bar{\eta}$, which include some of these are defined. Present values for the 4 free parameters of the SM are $\lambda = 0.2254(9)$, $A = 0.80(2)$, $\bar{\rho} = 0.14(2)$, and $\bar{\eta} = 0.343(15)$.

Although not exact, this parametrization exposes the interesting hierarchy observed for the mixing of mass and weak eigenstates.

The fact that V_{CKM} has a phase means that the Yukawa matrices λ^{ij} cannot be real. Therefore the Standard Model Lagrangian violates CP.

The flavour mixing described in this section for quarks actually is relevant for the lepton sector, as neutrino oscillations were discovered at the start of the millennium. These experiments detecting neutrinos emitted from the Sun or in collisions of cosmic rays in the Earth's atmosphere demonstrated that electron neutrinos can turn into muon neutrinos, etc. Therefore the mass eigenstates and weak eigenstates are not equivalent for the neutrinos. The mixing matrix is attributed to Pontecorvo, Maki, Nakagawa and Sakata. Defining it as the transformation matrix acting on mass eigenstates ν_1, ν_2, ν_3 to give the flavour eigenstates ν_e, ν_μ, ν_τ, U can be parametrized in terms of 3 angles and up to 3 phases.

$$U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \times \mathrm{diag}\left(1, e^{i\alpha_{21}/2}, e^{i\alpha_{31}/2}\right) \quad (5.3.9)$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$, for the 3 angles θ_{12}, θ_{23}, and θ_{13}. If the phase associated with δ is complex, then there is CP violation in the neutrino sector, as in the quark sector. The phases due to α_{21} and α_{31} arise only in the case that neutrinos are Majorana fermions (see below), because in this case we cannot perform as many of the phase transformations (5.3.6) as we did for quarks (which are Dirac fermions). The 3 angles have been measured to be approximately

$\theta_{12} \approx 35°$ (solar neutrinos), $\theta_{23} \approx 45°$ (atmospheric neutrinos), and $\theta_{13} \approx 9°$ (reactor neutrinos). The solar neutrino experiments observed fewer-than-expected electron neutrinos emitted from the sun. Muon neutrinos are byproducts of cosmic muons decaying in the atmosphere and are produced isotropically, but detectors measured a deficit of neutrinos coming through the earth upward vs. coming downward from the sky; even accounting for interactions with matter, it was shown that muon neutrinos were oscillating to another flavour. The reactor experiment at Daya Bay China followed by RENO just measured θ_{13} in 2012. The goal now is to discover CP violation by inferring a phase for this matrix. Searches are also underway for lepton flavour violating decays at the LHC.

As alluded to above, there is more than one way to introduce neutrino mass into the Standard Model. The first possibility is that neutrinos are Dirac fermions like the charged leptons and quarks. In this case there must be a right-handed $SU(2)_L$ singlet for each lepton generation

$$N^i = \nu_R^i = (\nu_{eR}, \nu_{\mu R}, \nu_{\tau R})\,.$$

and the lepton-Higgs terms in the Standard Model Lagrangian (5.2.8) gets modified to become

$$\mathcal{L}_{\text{lept},\phi} \;=\; -\sqrt{2}\left(\lambda^{ij}\bar{L}^i\phi R^j \;+\; \lambda_\nu^{ij}\bar{L}^i\phi^c N^j \;+\; \text{h.c.}\right)\,. \qquad (5.3.10)$$

where $(\phi^c)^\alpha \;=\; \epsilon^{\alpha\beta}\phi^{\dagger\beta}$. Then we proceed by diagonalizing these terms to obtain the mass eigenstates, yielding a mixing matrix with 3 angles and 1 phase – the θ_{ij} and δ in (5.3.9). The neutrinos get mass terms

$$\mathcal{L}_{m_\nu,D} \;=\; -\sum_i m_\nu^i(\bar{\nu}_R^i\nu_L^i + \bar{\nu}_L^i\nu_R^i) \qquad (5.3.11)$$

just like the charged leptons and quarks.

Since the neutrinos are electrically neutral particles, another possibility exists: they could be Majorana fermions. These are spin $\frac{1}{2}$ fields which are their own antiparticle. The operator annihilating an antiparticle is identically equal to the operator annihilating a particle $d^s(p) = b^s(p)$, so the quantum field is written

$$\nu(x) = \sum_{p,s}\left[b^s(p)u^s(p)e^{-ip\cdot x} + b^{s\dagger}(p)v^s(p)e^{ip\cdot x}\right] \qquad (5.3.12)$$

(compare to (3.2.8)). We can show that in this case

$$\nu^c(x) = C\bar{\nu}^T(x) = C(C^{-1}\nu(x)) = \nu(x) \qquad (5.3.13)$$

using (3.3.10) and (5.3.12) and letting the intrinsic parity be 1. Then the right-handed neutrino field $\nu_R(x) \;=\; P_R\nu(x)$ is not an independent field; it is the charge-conjugate of the left-handed field

$$\nu_R(x) = \nu_L^c(x) = C\bar{\nu}_L^T(x)\,. \qquad (5.3.14)$$

Therefore Majorana mass terms look like

$$\mathcal{L}_{m_\nu,M} \;=\; -\frac{1}{2}\sum_i m_\nu^i(\bar{\nu}_L^{i,c}\nu_L^i + \bar{\nu}_L^i\nu_L^{i,c})\,. \qquad (5.3.15)$$

The factor of $\frac{1}{2}$ is introduced implicitly to avoid double-counting since ν_L and ν_L^c are not independent.

Just like (5.3.15), the term (5.3.15) explicitly breaks $SU(2)_L$. We must find $SU(2)_L \times U(1)_Y$ invariant terms which will generate (5.3.15) when the Higgs field ϕ aqcuires a nontrivial VEV. It turns out that the simplest term (the term with lowest mass-dimension) is

$$\mathcal{L}_{\nu\phi M} = -\frac{Y^{ij}}{M}(L^{iT}\phi^c)C(\phi^{c\,T}L^j) + \text{h.c.} \tag{5.3.16}$$

This dimension-5 operator is nonrenormalizable. However, as we will discuss in § 8.1 this is fine as long as we think of the Standard Model as an effective field theory describing physics at scales well below the scale of some "new physics."

Weak decays

In this chapter we investigate several processes which occur due to the weak interactions. Since the energies and momenta involved here are much smaller in magnitude than the masses of the Z and W bosons,[52] we do not need to use the full electroweak theory. To a high level of precision we can use the Fermi weak Lagrangian. This is the first time we will see the utility of effective field theory, which we will address more systematically later in the course.

[52] $m_Z = 91.1876(21)$ GeV and $m_W = 80.385(15)$ GeV. High precision studies of weak decays typically involve processes with energy scales of a few GeV or lower.

6.1 Effective Lagrangian

Recall the weak part of the Lagrangian (5.2.6) is

$$\mathcal{L}_W = \frac{g}{2\sqrt{2}}\left(J^\mu W^+_\mu + J^{\mu\dagger} W^-_\mu\right) + \frac{g}{2\cos\theta_W} J^\mu_n Z_\mu \, .$$

The interaction Hamiltonian $H_I(t)$ or $V(t)$ is equal to $-\int d^3x \mathcal{L}_W$. Thus the S-matrix describing scattering of in-states $|i\rangle$ to out-states $|f\rangle$ is given by the time-ordered exponential

$$S = \mathcal{T} \exp\left[i\int d^4x \, \mathcal{L}_W(x)\right] .$$

For small g we can expand the exponential in a Taylor series. Assuming we do not have a W or Z in either the initial or final state then we find

$$\begin{aligned} \langle f|S|i\rangle &= \langle f|\mathcal{T}\Bigg\{1 - \frac{g^2}{8}\int d^4x d^4x' \Big[J^{\mu\dagger}(x) D^W_{\mu\nu}(x-x') J^\nu(x') \\ &+ \frac{1}{\cos^2\theta_W} J^{\mu\dagger}_n(x) D^Z_{\mu\nu}(x-x') J^\nu_n(x') \Big] + O(g^4)\Bigg\} |i\rangle \quad (6.1.1) \end{aligned}$$

having used Wick's theorem and the following contractions to give the W and Z propagators $\mathcal{T}[W^-_\mu(x) W^+_\nu(x')] = D^W_{\mu\nu}(x-x')$ and $\mathcal{T}[Z_\mu(x) Z_\nu(x')] = D^Z_{\mu\nu}(x-x')$.

Let us derive an expression for the Z propagator. Focus on the free part of the electroweak Lagrangian involving the Z boson

$$\mathcal{L}^{\text{free}}_Z = -\frac{1}{4}(\partial_\mu Z_\nu - \partial_\nu Z_\mu)(\partial^\mu Z^\nu - \partial^\nu Z^\mu) + \frac{1}{2} m_Z^2 Z_\mu Z^\mu$$

The Euler-Lagrange equation

$$\partial^\sigma\left(\frac{\partial\mathcal{L}}{\partial(\partial^\sigma Z^\rho)}\right) - \frac{\partial\mathcal{L}}{\partial Z^\rho} = 0$$

yields

$$\begin{aligned}\partial^\sigma(\partial_\sigma Z_\rho - \partial_\rho Z_\sigma) + m_Z^2 Z_\rho &= 0\\ \partial^2 Z_\rho - \partial_\rho \partial\cdot Z + m_Z^2 Z_\rho &= 0\,.\end{aligned} \tag{6.1.2}$$

Take the divergence of the above to find $m_Z^2\,\partial\cdot Z = 0$, or $\partial\cdot Z = 0$ since $m_Z^2 \neq 0$ after spontaneous symmetry breaking.[53] Now (6.1.2) reads

$$(\partial^2 + m_Z^2) Z_\rho = 0\,,$$

the Klein-Gordon equation.

[53] Here we see that $m_Z \neq 0$ implies $\partial\cdot Z = 0$ *in any gauge*. This is not so for massless gauge boson fields A_μ (like the photon), in which case imposing $\partial\cdot A = 0$ is a gauge-fixing condition (Lorentz gauge).

To find the propagator, introduce an external current j^μ which couples to Z_μ. The Lagrangian is appended by $j^\mu(x)Z_\mu(x)$ and the equations of motion (6.1.2) become

$$\partial^2 Z_\rho - \partial_\rho\partial\cdot Z + m_Z^2 Z_\rho = -j_\rho\,. \tag{6.1.3}$$

As before, we can apply ∂^ρ to the above implying $m_Z^2\,\partial\cdot Z = -\partial\cdot j$. Substituting into (6.1.3) we find

$$(\partial^2 + m_Z^2) Z_\mu = -\left(g_{\mu\nu} + \frac{\partial_\mu\partial_\nu}{m_Z^2}\right) j^\nu$$

We can obtain the solution by the method of Green's functions

$$Z_\mu(x) = \int d^4y\, D^Z_{\mu\nu}(x-y)\, j^\nu(y)$$

where

$$\begin{aligned} D^Z_{\mu\nu}(x-y) &= \int \frac{d^4p}{(2\pi)^4}\, e^{-ip\cdot(x-y)} \tilde{D}^Z_{\mu\nu}(p) \quad \text{with}\\ \tilde{D}^Z_{\mu\nu}(p) &= \frac{i}{p^2 - m_Z^2 + i\epsilon}\left(-g_{\mu\nu} + \frac{p_\mu p_\nu}{m_Z^2}\right)\,. \end{aligned} \tag{6.1.4}$$

The discussion above can be repeated for the quantum W_μ field, using The propagator is the same as (6.1.4) except with m_Z replaced by m_W.

For reference: The quantum Z field is

$$Z_\mu(x) = \sum_{p,\lambda}\Big[a_Z(p,\lambda)\varepsilon_\mu(p,\lambda)e^{-ip\cdot x} + a_Z^\dagger(p,\lambda)\varepsilon^*_\mu(p,\lambda)e^{ip\cdot x}\Big]$$

with the operators satisfying $[a_Z(p,\lambda), a_Z^\dagger(p',\lambda')] = \delta_{pp'}\delta_{\lambda\lambda'}$ and $\lambda\in\{-1,0,1\}$. The quantum W field

$$W_\mu(x) = \sum_{p,\lambda}\Big[a_W(p,\lambda)\varepsilon_\mu(p,\lambda)e^{-ip\cdot x} + c_W^\dagger(p,\lambda)\varepsilon^*_\mu(p,\lambda)e^{ip\cdot x}\Big]$$

where $a_W^\dagger$ creates a W^+ and $c_W^\dagger$ creates a W^-. The polarization vectors satisfy $p\cdot\varepsilon(p,\lambda) = 0$ and $\varepsilon^*(p,\lambda)\cdot\varepsilon(p,\lambda') = -\delta_{\lambda\lambda'}$ (orthonormal by convention). Another identity is obtained by contracting both sides with the linearly independent set $\{p, \varepsilon(p,\lambda)\}$,

$$\sum_\lambda \varepsilon_\mu(p,\lambda)\varepsilon^*_\nu(p,\lambda) = -g_{\mu\nu} + \frac{p_\mu p_\nu}{m_Z^2}\,.$$

At the low energies involved in weak decays of leptons and quarks (except the top quark), the weak boson masses are much larger than any combination of initial and final momentum components. Therefore the propagators can be approximated

$$\begin{aligned} \tilde{D}^W_{\mu\nu}(p) &\simeq -\frac{i}{m_W^2}(-g_{\mu\nu})\\ D^W_{\mu\nu}(x-y) &\simeq \frac{i}{m_W^2}\, g_{\mu\nu}\delta^{(4)}(x-y) \end{aligned} \tag{6.1.5}$$

and similarly for the Z propagator. We see that the propagator reduces to a δ-function in spacetime; the weak decays can essentially be described by a 4-fermion interaction. In this limit, we cease describing the weak interactions as mediated by a gauge boson, but instead via a 4-fermion (or a "Fermi") coupling:

$$\frac{g^2}{8} J^{\mu\dagger}(x) D^W_{\mu\nu}(x-x') J^\nu(x') \to \frac{ig^2}{8m_W^2} J^{\mu\dagger}(x) J^\nu(x') g_{\mu\nu}\delta(x-x')\,.$$

Carrying out the x'-integration in (6.1.1) we can define an effective weak Lagrangian

$$\mathcal{L}_W^{\text{eff}}(x) = -\frac{G_F}{\sqrt{2}}\left[J^{\mu\dagger}(x)J_\mu(x) \;+\; \rho J_n^{\mu\dagger}(x)J_{n\mu}(x)\right] \qquad (6.1.6)$$

where we define the conventional Fermi coupling G_F and ρ-parameter

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2}\,, \quad \rho = \frac{m_W^2}{m_Z^2\cos^2\theta_W}\,. \qquad (6.1.7)$$

In the Standard Model, $\rho = 1 + \Delta\rho$, with $\Delta\rho \approx 0.008$ due to quantum loops, but of course we must keep an eye out for experimental hints of physics beyond the Standard Model which might cause ρ to deviate from its Standard Model value.

We can re-exponentiate the expression for the S matrix element to see that (6.1.6) can really be interpreted as an interaction Lagrangian

$$S \;=\; \mathcal{T}\left[1 + i\int d^4x\mathcal{L}_W^{\text{eff}}(x)\right] \;=\; \mathcal{T}\exp\left(i\int d^4x\mathcal{L}_W^{\text{eff}}(x)\right)\,.$$

We make the following observation here, leaving the consequences to be investigated in the EFT Chapter. Note that the Fermi coupling G_F has dimensions of inverse mass squared to compensate for the 2 extra mass dimensions of the dimension-6 operator $J^{\mu\dagger}J_\mu$.[54] This implies that the effective weak theory is nonrenormalizable. We will see later that this is not an impediment to accurate calculations at scales well-below m_W. The appearance of m_W in the denominator of G_F indicates that the theory breaks down when energies reach that scale. Of course we know that we need to use the full electroweak theory with its propagating W and Z bosons for physics at the electroweak scale and above.

[54] The fermion fields in the bilinear J are dimension $\frac{3}{2}$, and $\int d^4x\mathcal{L}$ must be dimensionless.

6.2 *Decay rates, cross sections*

In the next few sections we will show how to calculate a few quantities which experimentalists can measure. Particle physics experiments are some of the most ambitious and technically complex activities we undertake. Yet the questions we ask them must be ultimately be formulated as a counting question: e.g. "Given N collisions between beams of A and B particles, how many times do we produce particle X? And then how frequently does X decay to products $\alpha + \beta + \gamma$?" From these results, we have to precisely determine the free parameters of the Standard Model and/or observe something unexpected.

Let us consider the second type of question first. The X particle's decay rate Γ_X is simply a measurement of the number of X decays per unit time, divided by the number of X particles present. By convention we quote the result in the rest frame of the X particle, as the result will change, due to time dilation, in moving reference frames. Typically X will decay in a variety of ways; it is simplest to consider partial decay rates $\Gamma_{X\to f}$ to specific final states, labelled by f say, and sum them up to get the total decay rate at the end. A particle's lifetime, is the reciprocal $\tau = 1/\Gamma$.

The relevant quantity is the S matrix for scattering between an initial state i and final state f. In the case of decays just described, $i = X$ and we are interested in inelastic scattering, where f has different particle content than i. In general, the S matrix elements are given by Dyson's formula

$$\langle f|S|i\rangle = \lim_{t_\pm\to\pm\infty}\langle f|U(t_+,t_-)|i\rangle \tag{6.2.1}$$

with

$$U(t_+,t_-) = \mathcal{T}\exp\left(-i\int_{t_-}^{t_+} dt'\, H_I(t')\right).$$

The S matrix can be separated into a boring part (where nothing happens) and an interesting part: $S = 1 + iT$. Employing the principle that momentum is conserved, we factor out an explicit momentum-conserving Dirac δ-function and define the invariant amplitude $\mathcal{M}$ through

$$\langle f|S-1|i\rangle = (2\pi)^4\delta^{(4)}(p_f - p_i)i\mathcal{M}_{fi}\,. \tag{6.2.2}$$

The probability that we measure $i \to f$ is given by the relevant S matrix element squared over the norm-squared for the initial and final states

$$\mathscr{P} = \frac{|\langle f|S-1|i\rangle|^2}{\langle f|f\rangle\langle i|i\rangle} \tag{6.2.3}$$

where

$$\begin{aligned}\langle i|i\rangle &= (2\pi)^3\, 2p_i^0\delta^{(3)}(0) = 2p_i^0 V\\ \langle f|f\rangle &= \prod_r (2p_r^0 V)\end{aligned}$$

where we have resorted to working in finite spatial volume V to avoid dealing with subtleties regarding with non-normalizable states.[55] The probability the decay will occur is

$$\mathscr{P} = \frac{|\mathcal{M}_{fi}|^2}{2mV}(2\pi)^4\delta^{(4)}\Big(p_i - \sum_r p_r\Big)\, VT\prod_r \frac{1}{2p_r^0 V} \tag{6.2.4}$$

where the factor VT in the numerator comes from one factor of the δ-function squared in $|\langle f|S|i\rangle|^2$.

Experimentalists never measure final state momentum with infinite precision; momentum is always integrated over some region corresponding to the precision of the detector or over all possible values. The integration measure for final state f is

$$d\rho_f = (2\pi)^4\delta^{(4)}\Big(p_i - \sum_{r\in f} p_r\Big)\prod_{r\in f}\left(\frac{d^3p_r}{(2\pi)^3 2p_r^0}\right). \tag{6.2.5}$$

Therefore we obtain the partial decay rate for $i \to f$ by dividing the probability (6.2.4) by T and integrating over momenta. Since the number of 1-particle states in the box with momentum in a

[55] Normally we have an inner product between momentum eigenstates which is $\langle q|p\rangle = (2\pi)^3 2\sqrt{q^0p^0}\delta^{(3)}(\vec{p}-\vec{q})$. In a finite volume, e.g. a cubic box $V = L^3$ with periodic boundary conditions, the momenta are discretized: $\vec{p} = 2\pi\vec{n}/L$, $n \in \mathbb{Z}^3$ and the δ-function becomes

$$\begin{aligned}\delta_V^{(3)}(\vec{p}-\vec{q}) &= \frac{1}{(2\pi)^3}\int_V d^3x\, e^{-i(\vec{p}-\vec{q})\cdot\vec{x}}\\ &= \frac{V}{(2\pi)^3}\delta_{\vec{p}\vec{q}}\,.\end{aligned}$$

momentum-space volume d^3p is $Vd^3p/(2\pi)^3$, the volume factors in numerator and denominator cancel to give us the partial decay rate

$$\Gamma(i \to f) \;=\; \frac{1}{2m}\int |\mathcal{M}_{fi}|^2\, d\rho_f \tag{6.2.6}$$

and the total decay rate

$$\Gamma_i \;=\; \frac{1}{2m}\sum_f \int |\mathcal{M}_{fi}|^2\, d\rho_f\,. \tag{6.2.7}$$

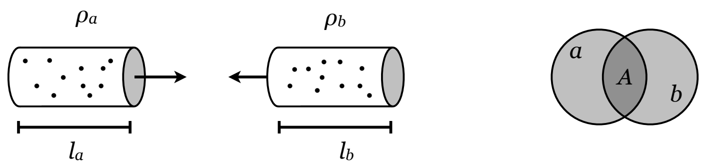

Figure 6.1: Two bunches of particles, with densities ρ_a, ρ_b and lengths ℓ_a, ℓ_b collide with cross-sectional area A.

The companion question to "How often does a particle decay?" is "How often does a particle collide?" We quantify the answer by defining the cross section. Imagine 2 bunches of particles (in particle beams) colliding together (Fig. 6.1). We count the number of scattering events and divide by the densities and lengths of the bunches, as well as the cross-sectional area of the collision region. The cross section is

$$\sigma \;=\; \frac{\#\text{ scattering events}}{\rho_a \ell_a \rho_b \ell_b A}$$

The cross-section has dimensions of area, and the traditional unit is the *barn*. By definition 1 barn $= 10^{-28}\text{m}^2$.[56] Usually we are more interested in more specific questions that require the differential cross sections. Again we need the differential probability per unit time of an event or transition $i \to f$, this time divided by the flux of particles through the interaction region. Since the particles are moving in the lab frame, the prefactor $1/2m$ in Γ becomes $1/2E$ for each bunch of particles.

[56] Despite not being able to fit much farm equipment through a door in such an area, nuclear physicists at Purdue University in Indiana must have been impressed by the size of the uranium nucleus when they declared, "Well golly, it's as big as a barn!"

$$d\sigma \;=\; \frac{1}{F}\,\frac{1}{4E_aE_bV}\,|\mathcal{M}_{fi}|^2\, d\rho_f\,.$$

The flux F is given by the relative velocities of particles in the 2 bunches, divided by volume V. Thus we find

$$d\sigma \;=\; \frac{1}{|\vec{v}_a - \vec{v}_b|}\,\frac{1}{4E_aE_b}\,|\mathcal{M}_{fi}|^2\, d\rho_f\,. \tag{6.2.8}$$

If the discussion in this section seems less than rigorous, you can take comfort in the fact that even Weinberg resorts to hand-wavery:

> In what follows we will instead give a quick and easy derivation of the main results, actually more of a mnemonic than a derivation, with the excuse that (as far as I know) no interesting open problems in physics hinge on getting the fine points right regarding these matters.[57]

[57] S Weinberg. *The Quantum Theory of Fields, Volume I.* Cambridge University Press, 1995. ISBN 0-521-55001-7

6.3 μ decay

As a first example let us consider the purely leptonic decay $\mu \to e\bar{\nu}_e\nu_\mu$. This proceeds through the weak current derived in § 5.2

$$J^\alpha \;=\; \bar{\nu}_e\gamma^\alpha(1-\gamma^5)e \,+\, \bar{\nu}_\mu\gamma^\alpha(1-\gamma^5)\mu \,+\, \bar{\nu}_\tau\gamma^\alpha(1-\gamma^5)\tau\,.$$

In fact this is the only decay channel for the muon. Since the muon mass $m_\mu = 105.658367(4)$ MeV is much less than $m_W = 80.385(15)$ GeV, we may use the effective Lagrangian of § 6.1

$$\mathcal{L}_W^{\text{eff}} \;=\; -\frac{G_F}{\sqrt{2}}\left(J^{\alpha\dagger}J_\alpha + J_n^{\alpha\dagger}J_{n,\alpha}\right)\,.$$

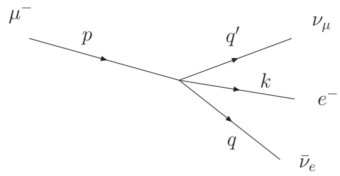

Figure 6.2: Momentum assignments for $\mu^- \to e^-\bar{\nu}_e\nu_\mu$.

Sandwiching this between the initial and final states of interest, assigning labels for the 4 particles' momenta as indicated in Fig. 6.2, we find for the invariant amplitude

$$\begin{aligned}\mathcal{M} &= \langle e^-(k)\bar{\nu}_e(q)\nu_\mu(q')|\mathcal{L}_W^{\text{eff}}|\mu^-(p)\rangle\\ &= -\frac{G_F}{\sqrt{2}}\langle e^-(k)\bar{\nu}_e(q)|\bar{e}\gamma^\alpha(1-\gamma^5)\nu_e|0\rangle\langle\nu_\mu(q')|\bar{\nu}_\mu\gamma_\alpha(1-\gamma^5)\mu|\mu^-(p)\rangle\\ &= -\frac{G_F}{\sqrt{2}}\bar{u}_e(k)\gamma^\alpha(1-\gamma^5)v_{\nu_e}(q)\bar{u}_{\nu_\mu}(q')\gamma_\alpha(1-\gamma^5)u_\mu(p)\,. \qquad (6.3.1)\end{aligned}$$

In the interest of tidiness, the spin indices on the spinors have not been written above; however, it may be instructive to include these when writing down an expression for $|\mathcal{M}|^2$. For many observables, we are not interested in specific spin states for any of the particles. If we wish to measure the partial decay rate for this channel (in this example, this is the total decay rate) we should *sum* over the final state spins, since any combination will count as a decay in this channel. We should also *average* over initial state spins; unmeasured, we will not know which spin state decayed.

Therefore, when we take the modulus-squared of the invariant amplitude, and perform the spin sum/average, we find

$$\begin{aligned}\frac{1}{2}\sum_{\text{spins}}|\mathcal{M}|^2 &= \frac{1}{2}\sum_{\text{spins}}\frac{G_F^2}{2}\left[\bar{u}_e(k)\gamma^\alpha(1-\gamma^5)v_{\nu_e}(q)\bar{v}_{\nu_e}(q)\gamma^\beta(1-\gamma^5)u_e(k)\right]\\ &\quad\times\left[\bar{u}_{\nu_\mu}(q')\gamma_\alpha(1-\gamma^5)u_\mu(p)\bar{u}_\mu(p)\gamma_\beta(1-\gamma^5)u_{\nu_\mu}(q')\right]\\ &= \frac{G_F^2}{4}S_1^{\alpha\beta}S_{2,\alpha\beta} \qquad (6.3.2)\end{aligned}$$

where[58]

$$\begin{aligned}S_1^{\alpha\beta} &= \mathrm{Tr}\left[(\not{k}+m_e)\gamma^\alpha(1-\gamma^5)\not{q}\gamma^\beta(1-\gamma^5)\right]\\ S_{2,\alpha\beta} &= \mathrm{Tr}\left[(\not{p}+m_\mu)\gamma_\beta(1-\gamma^5)\not{q}'\gamma_\alpha(1-\gamma^5)\right]\,. \qquad (6.3.3)\end{aligned}$$

[58] Reminders: $\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0$, $\gamma^{5\dagger} = \gamma^5$, $\{\gamma^5, \gamma^\mu\} = 0$, and $\bar{u} = u^\dagger\gamma^0$.

having used the expressions (with $m^2 = p^2$)

$$\sum_s u^s(p)\bar{u}^s(p) \;=\; \not{p}+m \quad\text{and}\quad \sum_s v^s(p)\bar{v}^s(p) \;=\; \not{p}-m\,. \qquad (6.3.4)$$

The following useful trace formulae can be shown to be correct

$$\begin{aligned}\mathrm{Tr}(\gamma^{\mu_1}\cdots\gamma^{\mu_n}) &= 0 \quad \text{for } n \text{ odd}\\ \mathrm{Tr}(\gamma^{\mu}\gamma^{\nu}\gamma^{\rho}\gamma^{\sigma}) &= 4\left(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma} + g^{\mu\sigma}g^{\nu\rho}\right)\\ \mathrm{Tr}(\gamma^5\gamma^{\mu}\gamma^{\nu}\gamma^{\rho}\gamma^{\sigma}) &= 4i\epsilon^{\mu\nu\rho\sigma}\,. \end{aligned} \tag{6.3.5}$$

Therefore

$$\begin{aligned} S_1^{\alpha\beta} &= 8\left(k^\alpha q^\beta + k^\beta q^\alpha - k\cdot q g^{\alpha\beta} + i\epsilon^{\alpha\beta\sigma\rho}k_\sigma q_\rho\right)\\ S_{2,\alpha\beta} &= 8\left(p_\alpha q'_\beta + p_\beta q'_\alpha - p\cdot q' g_{\alpha\beta} - i\epsilon_{\alpha\beta\lambda\tau}p^\lambda q'^\tau\right)\end{aligned}$$

Contracting these together we find[59]

[59] Note $\epsilon^{\alpha\beta\sigma\rho}\epsilon_{\alpha\beta\lambda\tau} = 2!\,\epsilon^{\gamma\sigma\rho}\epsilon_{\gamma\lambda\tau}$.

$$\frac{1}{2}\sum_{\text{spins}}|\mathcal{M}|^2 = 64G_F^2\,(p\cdot q)(k\cdot q')\,. \tag{6.3.6}$$

We can check that this result is consistent with our physical intuition in the following limiting case. Consider the final state where the electron and muon neutrino fly away in the z-direction, and the electron antineutrino in the $-z$-direction. In this case

$$k\cdot q' = \sqrt{m_e^2 + k_z^2}\,q'_z - k_z\,q'_z\,.$$

Notice that this dot product vanishes, and thus so does $|\mathcal{M}|^2$, in the limit $m_e \to 0$. We can understand this from conservation of angular momentum. In this scenario we have two massless, left-handed particles, each with helicity $-\frac{1}{2}$, moving in the $+z$-direction along with one massless, right-handed antiparticle, with helicity $+\frac{1}{2}$, moving in the $-z$-direction. The z-components of spin then sum to $-\frac{3}{2}$, which is not equal in magnitude to the spin of the muon. Therefore, this scenario is forbidden in the $m_e \to 0$ limit. For nonzero electron mass, the left-handed and right-handed components of the electron are coupled, and this scenario may occur, parametrized by m_e. We say such scenarios are *helicity suppressed*.

Now to obtain the total decay rate we integrate over the final state momenta

$$\begin{aligned}\Gamma &= \frac{1}{2m_\mu}\int\frac{d^3k}{(2\pi)^3 2k^0}\int\frac{d^3q}{(2\pi)^3 2q^0}\int\frac{d^3q'}{(2\pi)^3 2q'^0}(2\pi)^4\delta^{(4)}(p-k-q-q')\,\frac{1}{2}\sum_{\text{spins}}|\mathcal{M}|^2\\ &= \frac{G_F^2}{8\pi^5 m_\mu}\int\frac{d^3k}{k^0}\frac{d^3q}{q^0}\frac{d^3q'}{q'^0}\,\delta^{(4)}(p-k-q-q')\,(p\cdot q)(k\cdot q')\,.\end{aligned} \tag{6.3.7}$$

Let us introduce $Q = p - k$ and consider the integrals

$$I_{\mu\nu}(Q) = \int\frac{d^3q}{|\vec{q}|}\frac{d^3q'}{|\vec{q}'|}\,\delta^{(4)}(Q-q-q')\,q_\mu q'_\nu\,. \tag{6.3.8}$$

Given that $I_{\mu\nu}$ is symmetric under exchange of indices and carries dimensions of momentum squared, it can have only 2 possible terms

$$I_{\mu\nu}(Q) = aQ_\mu Q_\nu + bg_{\mu\nu}Q^2\,.$$

Contract both sides of the equation above with $g^{\mu\nu}$ and $Q^\mu Q^\nu$ in order to find $a + 4b = \frac{I}{2}$ and $a + b = \frac{I}{4}$, where[60]

[60] Tricks used: $q^2 = q'^2 = 0$ for the massless neutrinos; with the δ-function factor present, we can replace $(q+q')^2 = 2q\cdot q'$ by Q^2 in the integrand without changing the integral.

$$I \;=\; \int \frac{d^3q}{|\vec{q}|}\frac{d^3q'}{|\vec{q}'|}\,\delta^{(4)}(Q-q-q')\,. \tag{6.3.9}$$

Since I is Lorentz-invariant, we can choose the convenient center-of-mass frame $Q=(\sigma,\vec{0})$ in which to evaluate it:

$$I = \int \frac{d^3q}{|\vec{q}|^2}\,\delta(\sigma-2|\vec{q}|) \;=\; 4\pi\int_0^\infty dq\,\delta(\sigma-2q) \;=\; 2\pi$$

and thus $a=\frac{\pi}{3}$ and $b=\frac{\pi}{6}$. Inserting these results into (6.3.7) we find

$$\Gamma \;=\; \frac{G_F^2}{3m_\mu(2\pi)^4}\int\frac{d^3k}{k^0}\left[2p\cdot(p-k)k\cdot(p-k)\;+\;p\cdot k(p-k)^2\right]. \tag{6.3.10}$$

By convention, a particle's decay rate is quoted in its rest frame. Here $p\cdot k = m_\mu E$, using $E=k^0$ to represent the electron energy. Since $\frac{m_e}{m_\mu}=0.0048\ll 1$, it is sufficient for us to treat the electron in its massless limit. Using $|\vec{k}|=E$ and the fact that the integrand is independent of the direction of $\vec{k}$, we can express (6.3.10) as

$$\Gamma \;=\; \frac{2G_F^2 m_\mu}{3(2\pi)^3}\int_0^{m_\mu/2} dE\,E^2(3m_\mu-4E) \;=\; \frac{G_F^2 m_\mu^5}{192\pi^3}\,. \tag{6.3.11}$$

Note the electron energy must be in $[0,\frac{m_\mu}{2}]$, the extremes of which correspond to the the neutrinos recoiling back-to-back and the electron at rest $E=0$, and to the electron moving in the opposite direction to both the neutrinos (momentum conservation implying $E=\frac{m_\mu}{2}$).

This width Γ is actually the total width for the muon, since $\mu\to e\bar{\nu}_e\nu_\mu$ is the only decay channel. Given a measurement of the muon lifetime $\tau=1/\Gamma=2.1970\times10^{-6}$ seconds, one infers for the Fermi coupling $G_F=1.1638\times10^{-5}\,\text{GeV}^2$. One-loop corrections affect G_F only at the per-mille level. Experimentally one finds consistent values for G_F inferred from the decays $\tau\to e\bar{\nu}_e\nu_\tau$ and $\tau\to\mu\bar{\nu}_\mu\nu_\tau$, a property frequently referred to as *lepton universality*.

6.4 π decay

Next we consider a decay similar to the muon's, the decay of the π^- meson to electron and antineutrino. This is similar in that it proceeds through the charged weak current, with a down quark and an up antiquark annihilating to a W^- boson, which then promptly decays into electron and antineutrino. The main difference, is that the d and $\bar{u}$ do not ever propagate freely; they are strongly bound together as a hadronic state, the π^- meson, or pion.

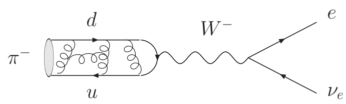

Figure 6.3: Weak decay of a pion to electron and antineutrino.

In addition to the leptonic weak current of the last section

$$J^\alpha_{\text{lept}} \;=\; \bar{\nu}_e\gamma^\alpha(1-\gamma^5)e\;+\;\ldots$$

we need the hadronic weak current. We will see shortly that it is convenient to separate the current into terms of definite parity

$$J^\alpha_{\text{had}} \;=\; V^\alpha_{\text{had}}\;-\;A^\alpha_{\text{had}} \tag{6.4.1}$$

with the vector and axial-vector currents

$$\begin{aligned} V^\alpha_{\text{had}} &= \bar{u}\gamma^\alpha (V_{ud} d \,+\, V_{us} s \,+\, V_{ub} b) \,+\, \ldots \\ A^\alpha_{\text{had}} &= \bar{u}\gamma^\alpha \gamma^5 (V_{ud} d \,+\, V_{us} s \,+\, V_{ub} b) \,+\, \ldots\,. \end{aligned}$$

The ellipses represent the higher-generation terms which do not play a role in π decay.

The invariant amplitude

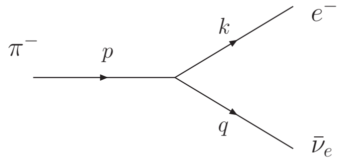

Figure 6.4: Momentum assignments for $\pi \to e\bar{\nu}_e$.

$$\begin{aligned} \mathcal{M} &= \langle e^-(k)\bar{\nu}_e(q)|\mathcal{L}^{\text{eff}}_W|\pi^-(p)\rangle \\ &= -\frac{G_F}{\sqrt{2}}\langle e^-(k)\bar{\nu}_e(q)|\bar{e}\gamma_\alpha(1-\gamma^5)\nu_e|0\rangle\langle 0|J^\alpha_{\text{had}}|\pi^-(p)\rangle \\ &= \frac{G_F}{\sqrt{2}}\bar{u}_e(k)\gamma_\alpha(1-\gamma^5)v_{\nu_e}(q)\langle 0|A^\alpha_{\text{had}}|\pi^-(p)\rangle\,. \end{aligned} \tag{6.4.2}$$

In the last step, we use the fact that the matrix element of V^α_{had} between the QCD vacuum and the parity-odd pseudoscalar meson vanishes, $\langle 0|V^\alpha_{\text{had}}|\pi\rangle = 0$ because QCD is a P-invariant theory.[61]

Since QCD is strongly interacting – in fact free quarks are forbidden to be free by a mysterious mechanism called confinement – we cannot perturbative approximate hadronic matrix elements. Instead, we package our uncertainty into a single dimensionful parameter called the pion decay constant F_π, so that[62]

$$\langle 0|V_{ud}\bar{u}\gamma^\alpha\gamma^5 d|\pi^-(p)\rangle \;=\; iV_{ud}\sqrt{2}F_\pi p^\alpha\,. \tag{6.4.3}$$

Pre-emptively using the fact that we know momentum will be conserved, we write $p = k + q$, and then

$$\bar{u}_e(k)\not{k} \;=\; \bar{u}_e(k)m_e \quad\text{and}\quad \not{q}v_{\nu_e}(q) \;=\; 0$$

(since the neutrino is massless), we have

$$\mathcal{M} \;=\; iG_F F_\pi m_e V_{ud}\bar{u}_e(k)(1-\gamma^5)v_{\nu_e}(q)\,. \tag{6.4.4}$$

As in § 6.3 we see helicity suppression, this time in all decays since we have just a 2-body final state, which in the pion rest frame consists of the electron and antineutrino flying away back-to-back. The decay is suppressed as $m_e \to 0$ since the final state with net spin-component equal to 1 in the direction of flight is inconsistent with the spinless initial state.

The decay rate is an integral over the invariant amplitude squared, with a sum over the electron and neutrino spins, so we need

$$\sum_{\text{spins}}|\mathcal{M}|^2 \;=\; 2(G_F F_\pi m_e V_{ud})^2\,\mathrm{Tr}\left[(\not{k}+m)(1-\gamma^5)\not{q}\right] \tag{6.4.5}$$

having used $(1-\gamma^5)\gamma^\mu(1+\gamma^5) = 2(1-\gamma^5)\gamma^\mu$. We use the following trace identities

$$\begin{aligned} \mathrm{Tr}\,\not{k}\not{q} &= 4\,k\cdot q \\ \mathrm{Tr}\,\gamma^5\not{k}\not{q} &= 0 \\ \mathrm{Tr}\,\gamma^\mu &= \mathrm{Tr}\,\gamma^\mu\gamma^5 \;=\; 0 \end{aligned}$$

[61] To elaborate, we take as an experimental fact that the pion is a spin-0, parity-odd state (a pseudoscalar meson), and of course the vacuum is parity-even. We showed in § 3.2 how vector and axial-vector currents under P (3.2.7). Since the initial state is parity-odd the matrix elements $\langle 0|V^\alpha_{\text{had}}|\pi(p)\rangle$ and $\langle 0|A^\alpha_{\text{had}}|\pi(p)\rangle$ transform respectively as an axial-vector and a vector under P. The only physical variable carrying a Lorentz index in the QCD part of the problem is the momentum vector p; we cannot construct an axial vector. Therefore vector current matrix element must vanish.

[62] There is an equally prevalent convention for the decay constant in the literature: $f^{\text{alt}}_\pi = \sqrt{2}F_\pi$.

to write

$$\sum_{\text{spins}} |\mathcal{M}|^2 \;=\; 8(G_F F_\pi m_e V_{ud})^2\, k\cdot q\,. \tag{6.4.6}$$

Therefore the decay rate in the π rest frame is

$$\begin{aligned}\Gamma_{\pi\to e\bar\nu} &= \frac{1}{2m_\pi}\int \frac{d^3k}{(2\pi)^3 2k^0}\frac{d^3q}{(2\pi)^3 2q^0}\,(2\pi)^4\,\delta^{(4)}(p-k-q)\sum_{\text{spins}}|\mathcal{M}|^2\\ &= (G_F F_\pi m_e V_{ud})^2\,\frac{1}{4\pi^2 m_\pi}\int\frac{d^3k}{E|\vec k|}\,\delta(m_\pi - E - |\vec k|)(E+|\vec k|)|\vec k|\,.\end{aligned}$$

We used the fact we are working in the π rest frame. Performing the integration over $\vec q$ yields $q = (|\vec k|, -\vec k)$ so that $k\cdot q = E|\vec k| + |\vec k|^2$ ($E = k^0$). Next we use the composition rule for the δ-function: $\delta(f(k)) = \sum_i \delta(k-k_0^i)/|f'(k_0^i)|$, where k_0^i are the roots of $f(k) = 0$. In this case

$$k_0 = \frac{m_\pi^2 - m_e^2}{2m_\pi}\quad\text{and}\quad f'(k_0) = 1 + \frac{k_0}{E}$$

so

$$\begin{aligned}\Gamma_{\pi\to e\bar\nu} &= (G_F F_\pi m_e V_{ud})^2\,\frac{1}{4\pi^2 m_\pi}\int_0^\infty \frac{4\pi k^2 dk}{E}(E+k)\frac{\delta(k-k_0)}{1+\frac{k_0}{E}}\\ &= \frac{G_F^2 F_\pi^2 V_{ud}^2}{4\pi}\,m_e^2 m_\pi\left(1-\frac{m_e^2}{m_\pi^2}\right)^2.\end{aligned} \tag{6.4.7}$$

A similar calculation for $\pi\to\mu\bar\nu_\mu$ yields

$$\Gamma_{\pi\to\mu\bar\nu} \;=\; \frac{G_F^2 F_\pi^2 V_{ud}^2}{4\pi}\,m_\mu^2 m_\pi\left(1-\frac{m_\mu^2}{m_\pi^2}\right)^2. \tag{6.4.8}$$

One can take the ratio where the least well-known quantities cancel

$$\frac{\Gamma(\pi\to e\bar\nu_e)}{\Gamma(\pi\to\mu\bar\nu_\mu)} \;=\; \frac{m_e^2}{m_\mu^2}\left(\frac{m_\pi^2-m_e^2}{m_\pi^2-m_\mu^2}\right)^2 \;=\; 1.28\times10^{-4}\,. \tag{6.4.9}$$

Experimentally, the ratio is measured to be 1.23×10^{-4}, exposing quantum effects arising from loop diagrams. In this case virtual photon effects need to be included in order to agree with experiment.

6.5 $K^0 - \bar K^0$ *mixing*

Kaons are pseudoscalar mesons containing either a strange quark or a strange antiquark. The neutral kaons K^0 and $\bar K^0$ denote mesons with valence quark content $\bar s d$ and $\bar d s$, respectively. (In addition to the *valence* quarks familiar from simple quark models, there are quark-antiquark pairs popping in and out of existence in the *sea* of hadrons. More about this when we discuss QCD.) There are also charged kaons K^+ and K^- with an up quark in place of the neutral kaons' down quarks.

We knew when we constructed the electroweak Lagrangian that the W should only couple to left-handed particles and right-handed antiparticles. Thus, the weak interactions violate P and

C maximally. However, while CP violation is a possibility in a 3-flavour theory, as we saw when we introduced the CKM matrix, it is not necessary that Nature oblige. In fact, it does turn out to be an empirical fact, evident in the behaviour of neutral kaons, that CP *is* violated.

K^0 and $\bar{K}^0$ are C conjugates of each other. Here we are more interest in the combined transformation CP. Under CP we can take the phases so that, for kaons at rest

$$\hat{C}\hat{P}|K^0\rangle = -|\bar{K}^0\rangle\,, \quad \hat{C}\hat{P}|\bar{K}^0\rangle = -|K^0\rangle\,. \tag{6.5.1}$$

Construct CP eigenstates

$$|K_1^0\rangle = \frac{1}{\sqrt{2}}\left(|K^0\rangle - |\bar{K}^0\rangle\right)\,, \quad |K_2^0\rangle = \frac{1}{\sqrt{2}}\left(|K^0\rangle + |\bar{K}^0\rangle\right)\,. \tag{6.5.2}$$

With these labels, K_1^0 is CP-even and K_2^0 is CP-odd.

Now let us consider decays of neutral kaons to 2 pions, either $\pi^+\pi^-$ or $\pi^0\pi^0$ (Fig. 6.5).[63] Recall the pions are pseudoscalar mesons just like the kaons (pseudo = parity-odd, scalar = spin-zero, i.e. $J^P = 0^-$). In the centre-of-mass frame, after the decay the pions fly away back-to-back. The action of parity is the swap the particles' positions and momenta ($\vec{x} \mapsto -\vec{x}$, $\vec{p} \mapsto -\vec{p}$), but the action of charge conjugation is to swap the particle charges. Thus, if the pions were classical objects, we would already see that the combined transformation CP leaves the 2-pion system invariant. Quantum mechanically, however, we still have to consider whether the 2-pion wavefunction has orbital angular momentum ℓ. In general, it could (e.g. as in the strong decay of the vector meson $\rho^0 \to \pi^+\pi^-$). However, since the kaon at rest had no angular momentum, orbital or spin, then the pions must be in an $\ell = 0$ state. Thus

[63] Note this must proceed by the weak interactions, the only mechanism for changing quark flavour, strangeness in this case.

$$\hat{C}\hat{P}|\pi^+\pi^-\rangle = (-1)^\ell|\pi^+\pi^-\rangle\,, \quad \hat{C}\hat{P}|\pi^0\pi^0\rangle = (-1)^\ell|\pi^0\pi^0\rangle \tag{6.5.3}$$

with $\ell = 0$ tells us the final state of $K^0 \to \pi\pi$ is CP-even. If CP is respected by the weak interactions, then there should be one neutral kaon, the K_1^0, which can decay into 2 pions (and thus is short-lived due to the large available phase space) and another neutral kaon, the K_2^0, which cannot decay into 2 pions, but may decay into 3 pions or other final states (and thus is longer-lived).

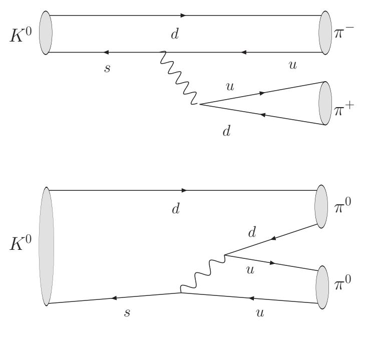

Figure 6.5: Decay of K^0 to $\pi^+\pi^-$ (top) and to $\pi^0\pi^0$ (bottom).

Experimentally, it is true that there are 2 neutral kaons, K_S^0 and K_L^0 which respectively have short (0.89×10^{-10} s) and long (5.18×10^{-8} s) lifetimes. However, occasionally one sees $K_L^0 \to \pi\pi$. Quantifying the relative likelihoods of the relevant decays, define the ratios

$$\eta_{+-} = \frac{\langle\pi^+\pi^-|H_W|K_L^0\rangle}{\langle\pi^+\pi^-|H_W|K_S^0\rangle}\,, \quad \eta_{00} = \frac{\langle\pi^0\pi^0|H_W|K_L^0\rangle}{\langle\pi^0\pi^0|H_W|K_S^0\rangle}\,. \tag{6.5.4}$$

Experimentally it is found that $\eta_{+-} = \eta_{00} = 2.28 \times 10^{-3}$. We can conclude the weak interactions violate CP. However, there are 2 possible ways: direct CP violation of the underlying $s \to u$

transition (due to a complex phase in the CKM matrix), or indirect CP violation due to a K^0 turning into a $\bar{K}^0$ before decaying (or vice versa). It is the latter effect which turns out to be responsible here. This oscillation between the 2 weak eigenstates is due to loop effects, the dominant ones given by the so-called box-diagrams shown in Figure 6.6.

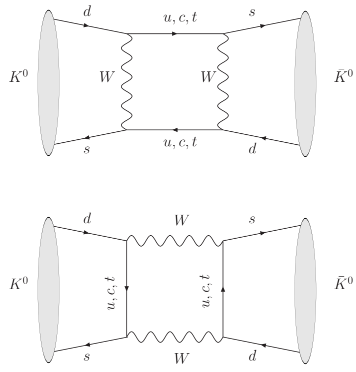

Figure 6.6: $K^0 - \bar{K}^0$ mixing in the Standard Model occurs at loop level, via these "box" diagrams.

Given this apparent violation of CP, we must assume the states which propagate (the mass eigenstates) are combinations of the CP eigenstates:

$$
\begin{aligned}
|K_S^0\rangle &= \frac{1}{\sqrt{1+|\epsilon_1|^2}}\left(|K_1^0\rangle + \epsilon_1 |K_2^0\rangle\right) \\
|K_L^0\rangle &= \frac{1}{\sqrt{1+|\epsilon_2|^2}}\left(|K_2^0\rangle + \epsilon_2 |K_1^0\rangle\right) .
\end{aligned}
\tag{6.5.5}
$$

with the ϵ_k complex in general.

Under the mild assumptions: (1) that K_S and K_L are linear combinations of K^0 and $\bar{K}^0$ alone and not additional, excited states; (2) that we can ignore details of the strong interactions in considering the mixing; then the Wigner-Weisskopf approximation is that as they propagate, the K_S and K_L states will be an oscillating mixture of weak eigenstates:

$$
\begin{aligned}
|K_S(t)\rangle &= a_S(t)|K^0\rangle + b_S(t)|\bar{K}^0\rangle \\
|K_L(t)\rangle &= a_L(t)|K^0\rangle + b_L(t)|\bar{K}^0\rangle .
\end{aligned}
\tag{6.5.6}
$$

Using Heisenberg's prescription for time-evolution: $|\psi(t)\rangle = \exp(-iHt)|\psi(0)\rangle$ so that $i(d/dt)|\psi(t)\rangle = H|\psi(t)\rangle$, the time-dependent coefficients in (6.5.6) obey

$$
\begin{aligned}
i\frac{d}{dt}\begin{pmatrix} a \\ b \end{pmatrix} &= \begin{pmatrix} \langle K^0|H'|K^0\rangle & \langle K^0|H'|\bar{K}^0\rangle \\ \langle \bar{K}^0|H'|K^0\rangle & \langle \bar{K}^0|H'|\bar{K}^0\rangle \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} \\
&= \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} .
\end{aligned}
\tag{6.5.7}
$$

The Hamiltonian H' is the weak Hamiltonian at next-to-leading order in perturbation theory:

$$
H' = H_W - \sum_n \frac{H_W|n\rangle\langle n|H_W}{E_n - m_0 - i\varepsilon}
\tag{6.5.8}
$$

with m_0 being the unperturbed mass of the neutral kaons (any splitting due to this mixing at next-to-leading order). Writing H' this way essentially represents the processes depicted in the box diagrams (Fig. 6.6) as a local $\Delta(\text{strangeness}) = 2$ interaction.[64]

[64] As before, the large mass of the W means that low energy processes appear to be governed by local interactions. It also can be shown that the dominant contribution comes from the top quark in the loop, and $m_t = 173$ GeV.

Since kaons do not oscillate indefinitely, but decay in finite time, the matrix R in (6.5.7) is not Hermitian. We can write it as the combination of Hermitian and antihermitian matrices, $R = M - \frac{i}{2}\Gamma$, where M is referred to as the mass matrix and Γ as the decay matrix, and both M and Γ are Hermitian. Commonly one refers the the dispersive and absorptive parts of the matrix R, given respectively by

$$
M_{12} = \frac{1}{2}(R_{12} + R_{21}^*) \quad \text{and} \quad \Gamma_{12} = i(R_{12} - R_{21}^*) .
$$

Under CPT, $\hat{\Theta}H'\hat{\Theta}^{-1} = H'^{\dagger}$. For kaons at rest we can take $\hat{T}|K^0\rangle = |K^0\rangle$ and $\hat{T}|\bar{K}^0\rangle = |\bar{K}^0\rangle$, i.e. time reversal does nothing to a free particle in its rest frame. So under CPT $\hat{\Theta}|K^0\rangle = -|\bar{K}^0\rangle$ and $\hat{\Theta}|\bar{K}^0\rangle = -|K^0\rangle$. This means the diagonal matrix elements of R must be equal:

$$R_{11} = \langle K^0|(\hat{\Theta}^{-1}\hat{\Theta}); H'(\hat{\Theta}^{-1}\hat{\Theta})|K^0\rangle = \langle \bar{K}^0|H'^{\dagger}|\bar{K}^0\rangle^* = R_{22} \quad (6.5.9)$$

where the complex conjugation appeared due to the anti-linearity of $\hat{\Theta}$.

CPT-invariance leaves R_{12} and R_{21} unconstrained. If T-invariance were respected (which, by the CPT theorem, implies CP would be conserved) then $\hat{T}H'\hat{T}^{-1} = H'^{\dagger}$ and

$$R_{12} = \langle K^0|\hat{T}^{-1}\hat{T}; H'\hat{T}^{-1}\hat{T}|\bar{K}^0\rangle = \langle \bar{K}^0|H'|K^0\rangle = R_{21}\,. \quad (6.5.10)$$

We will see next that $R_{12} \neq R_{21}$ and thus CP must be violated.

By design (from (6.5.2) and (6.5.5)) the (unnormalized) eigenvectors of R corresponding to $|K_S\rangle$ and $|K_L\rangle$ are respectively $(1+\epsilon_1, -1+\epsilon_1)^T$ and $(1+\epsilon_2, 1-\epsilon_2)^T$ in the $|K^0\rangle$, $|\bar{K}^0\rangle$ basis, with corresponding eigenvalues denoted $M_S - \frac{i}{2}\Gamma_S$ and $M_L - \frac{i}{2}\Gamma_L$.[65] Knowing these are the eigenvectors (be definition of the short and long states) allows us to determine ϵ_1 and ϵ_2 in terms of the matrix elements of R. To save some work, let us assume $\epsilon_1 = \epsilon_2 = \epsilon$ (which turns out to work) and abbreviate $1+\epsilon = p$ and $1-\epsilon = q$. Then by construction the following similarity transformation diagonalizes R

$$\frac{1}{2pq}\begin{pmatrix} q & -p \\ q & p \end{pmatrix}\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{11} \end{pmatrix}\begin{pmatrix} p & p \\ -q & q \end{pmatrix}$$

Requiring this to be equal to $\mathrm{diag}(M_S - \frac{i}{2}\Gamma_S, M_L - \frac{i}{2}\Gamma_L)$ implies that $R_{12}q^2 = R_{21}p^2$ or

$$\epsilon = \frac{\sqrt{R_{12}} - \sqrt{R_{21}}}{\sqrt{R_{12}} + \sqrt{R_{21}}}\,. \quad (6.5.11)$$

Thus we see that if CP were a good symmetry, and consequently $R_{12} = R_{21}$, then ϵ would vanish and the CP eigenstates K_1 and K_2 would not mix.

[65] The states evolve in time (in the rest frame) as $|K_{S,L}(t)\rangle = \exp(-i(M_{S,L} - \frac{i}{2}\Gamma_{S,L}))|K_{S,L}(0)\rangle$

QCD

Finally we come to the strong interactions. The existence of quarks as constituents of protons, neutrons and other subatomic particles grew out of idea that the particles could be grouped together based on observed similarities. By analogy with $SU(2)$ spin doublets representing the two spin components of nonrelativistic spin-$\frac{1}{2}$ fermions, proton and neutron were supposed to be the *isospin* components ($I_3^{(p)} = +\frac{1}{2}$ and $I_3^{(n)} = -\frac{1}{2}$, respectively) of an isospin symmetry group $SU(2)_I$. The fact that the proton and neutron have slightly different masses (differing by about 1 MeV) suggests that isospin is slightly broken. Interactions between protons and neutrons (and each other) could be described by exchange of pions. The three pions, π^+, π^0, π^-, transform as an $SU(2)_I$ triplet.

The discovery of "strange" particles necessitated extending the isospin group to "flavour" $SU(3)$ or $SU(3)_F$. This symmetry is not as good a symmetry – the mass-splittings between strange and nonstrange particles is usually of order a few hundred MeV. Nevertheless the observed particles fell into groups which could be understood as multiplets of $SU(3)_F$, for example the octet of pseudoscalar mesons (Fig. 7.1), the octet of spin-$\frac{1}{2}$ baryons (Fig. 7.2), and the decuplet of spin-$\frac{3}{2}$ baryons (Fig. 7.3). These principles led to the successful prediction of the triply-strange Ω^- baryon.

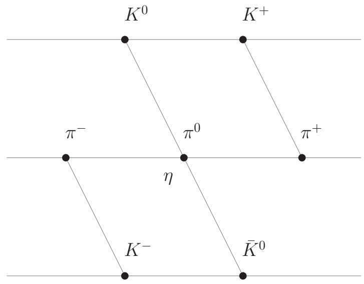

Figure 7.1: Meson octet. Particles on the same horizontal line have the same strangeness; those with the same charge along the diagonal; those with the same isospin component I_3 are aligned vertically.

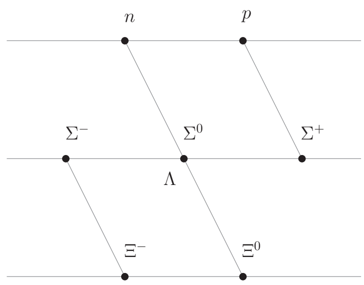

Figure 7.2: Baryon octet, axes as in Fig. 7.1.

The success of $SU(3)_F$ led to the hypothesis that the many hadrons were composites of 3 fundamental particles, the u, d, and s quarks. The up and down quarks were assigned isospin $\pm\frac{1}{2}$ respectively, with no strangeness, while the strange quark carried no isospin and has strangeness -1, unfortunately, in order to coincide with the definition of strangeness earlier assigned to the hadrons.[66] In this quark model, the baryons are bound states of 3 quarks, and the mesons bound states of quark plus antiquark. In order to give the hadrons their observed masses, the up quark must have electron charge $+\frac{2}{3}$ while the down and strange quarks have electron charge $-\frac{1}{3}$.

[66] Thus Murray Gell-Mann and Kazuhiko Nishijima join Benjamin Franklin in the ranks of accidentally saddling us with a minus sign associated with a charge carrier.

The quark model as described has two obvious problems. The first is in describing the Δ^{++} baryon. The charge implies it consists of 3 up quarks. The fact that the spin of the Δ^{++} is $\frac{3}{2}$ implies the 3 quarks have their spins aligned. However if both the spin and flavour degrees-of-freedom for the 3 quarks are identical, then its wavefunction appears totally symmetric. The solution of this problem is to suppose the existence of another quantum number; in this case it was called colour. Then the up quarks in the Δ^{++} could be in an antisymmetric combination of red-green-blue. Following the colour analogy, antibaryons are said to be a combination of cyan-magenta-yellow.

The second problem is that not all combinations of quarks are seen in nature. For example, we do not see any experimental evidence for free quarks or diquarks (qq). Quarks seem to exhibit a phenomenon called *confinement*. That is only "colourless" states, such as red+green+blue baryons or red+cyan mesons can exist as observable initial or final states. The search for a precise explanation for or understanding of confinement is an ongoing one.

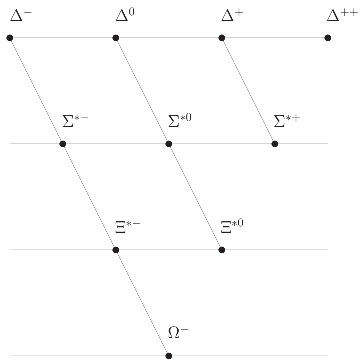

Figure 7.3: Baryon decuplet, axes as in Fig. 7.1.

7.1 *QCD Lagrangian*

The modern description of the strong interactions builds upon the early quark model. As with the weak and electromagnetic forces, the strong force between quarks is mediated by gauge bosons, the gluons. The need for 3 colours implies the gauge group should be $SU(3)_C$ (the subscript just distinguishes the gauge symmetry from the approximate global symmetry $SU(3)_F$). By analogy with quantum electrodynamics, the theory of the strong interactions is called quantum chromodynamics.

The QCD Lagrangian is packaged to look just like the QED Lagrangian

$$\mathcal{L}_{\text{QCD}} \;=\; -\frac{1}{4}F^{a,\mu\nu}F^a_{\mu\nu} \;+\; \sum_f \bar{q}_f(i\not{D} - m_f)q_f \tag{7.1.1}$$

where $D_\mu = \partial_\mu + igA^a_\mu T^a$. The T^a are the generators of $SU(3)$ in the fundamental representation, and satisfy $[T^a, T^b] = if^{abc}T^c$. They are related to the Gell-Mann matrices: $T^a = \frac{1}{2}\lambda^a$.[67] The eight gauge field A^a_μ transform in the adjoint representation of $SU(3)$ and have field strength tensor

$$F^a_{\mu\nu} \;=\; \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - gf^{abc}A^b_\mu A^c_\nu\,.$$

[67] Here are the Gell-Mann matrices:

$$\lambda^1 = \begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix} \quad \lambda^2 = \begin{pmatrix}0&-i&0\\i&0&0\\0&0&0\end{pmatrix}$$

$$\lambda^3 = \begin{pmatrix}1&0&0\\0&-1&0\\0&0&0\end{pmatrix} \quad \lambda^4 = \begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix}$$

$$\lambda^5 = \begin{pmatrix}0&0&-i\\0&0&0\\i&0&0\end{pmatrix} \quad \lambda^6 = \begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix}$$

$$\lambda^7 = \begin{pmatrix}0&0&0\\0&0&-i\\0&i&0\end{pmatrix} \quad \lambda^8 = \frac{1}{\sqrt{3}}\begin{pmatrix}1&0&0\\0&1&0\\0&0&-2\end{pmatrix}$$

The completely anti-symmetric structure constants are: $f^{123} = 1$, $f^{458} = f^{678} = \frac{\sqrt{3}}{2}$, $f^{147} = f^{165} = f^{246} = f^{257} = f^{345} = f^{376} = \frac{1}{2}$; those not related to these by permutation of indices are zero.

7.2 *Renormalization*

The idea of renormalization is presented more thoroughly in the Advanced Quantum Field Theory course, and the renormalization group is most clearly introduced in the context of statistical field theory. Nevertheless, we give a very brief review here because the behaviour of the strong coupling constant under renormalization has a greater impact than for the couplings in the electroweak theory. This is simply due to the fact that the strong coupling constant is numerically much larger, and therefore its "running" under renormalization is more consequential. The discussion here follows §9.2 of Georgi's book.[68]

[68] H Georgi. *Weak interactions and modern particle theory*. Benjamin/Cummings, 1984. ISBN 0-8053-3163-8

Say we have a Lagrangian which contains a set of coupling constants g_i. For massless QCD we have only one coupling; each quark mass introduced may be considered here as another coupling. For each of these, we need a physical or derived quantity g_i^0 in order to define the renormalized theory; for example the physical observable could be a scattering amplitude, or it could be a derived quantity like a bare coupling in a regularized Lagrangian. We then need to calculate an expression for each of the g_i^0. While it is possible to do this nonperturbatively and numerically using a spacetime lattice regulator, here we consider performing a perturbative calculation. Thus we find a function $G_i^0(g(\mu),\mu)$ for each g_i^0 which is a power series in renormalized couplings $g(\mu) = \{g_j(\mu)\}$.[69] Clearly this depends on the particular value of the renormalization point μ we choose. Our renormalization condition consists of demanding for a given μ that the set of $g_j(\mu)$ are such that our expression $G_i^0(g(\mu),\mu)$ is equal to the observed (or deduced) quantities

$$g_i^0 \;=\; G_i^0(g(\mu),\mu)\,. \tag{7.2.1}$$

[69] For brevity we use the subscriptless $g(\mu)$ to denote the set of running couplings.

The renormalization group is concerned with how the renormalized couplings $\{g_j(\mu)\}$ change as we vary the renormalization point μ. It is common to talk about a β-function corresponding to each coupling; the name just comes from the definition

$$\beta_j(g(\mu),\mu) \;=\; \mu\,\frac{d}{d\mu}\,g_j(\mu)\,. \tag{7.2.2}$$

Since physical quantities and bare parameters g_i^0 do not depend on μ, we find upon differentiating (7.2.1)

$$\mu\,\frac{d}{d\mu}\,G_i^0(g(\mu),\mu) \;=\; \left(\mu\,\frac{\partial}{\partial\mu} + \beta_j\,\frac{\partial}{\partial g_j}\right)G_i^0(g(\mu),\mu) \;=\; 0\,. \tag{7.2.3}$$

In so-called mass-independent renormalization schemes, the β-function does not depend explicitly on μ: $\beta(g(\mu))$. These are often convenient schemes in which to work. However in problems involving several mass scales, one is sometimes forced to work in a mass-dependent scheme.

In AQFT, you saw that one can derive similar equations for renormalized Green's functions, in which case one has to also include effects due to the anomalous dimensions of the fields in the theory. Equations of this type are called Callan-Symanzik equations.

In this Chapter on QCD, we are concerned with the running of the QCD couping g. Again in the AQFT course, you may in the end repeat the calculation which earned Gross, Wilczek, and Politzer the Nobel Prize: the one-loop determination of the QCD β-function

$$\mu\,\frac{d}{d\mu}\,g(\mu) \;=\; \beta(g(\mu)) \;=\; bg^3 + O(g^5)\,. \tag{7.2.4}$$

It was remarkable at the time that the coefficient

$$b \;=\; -\frac{1}{16\pi^2}\left(11 - \tfrac{2}{3}N_F\right)\,,$$

for QCD with N_F flavours of quarks, is *negative*.[70] In fact, the β-function for simple (single-coupling) nonabelian gauge groups generically has a negative 1-loop contribution

$$\beta(g) \;=\; -\beta_0\,\frac{g^3}{16\pi^2} + O(g^5)\,. \tag{7.2.5}$$

[70] Here is a minus sign we can be grateful for!

Assuming the gauge field couples to the fermions ψ_f (f for flavour) through covariant derivatives $D_\mu \psi_f = (\partial_\mu + igA^a_\mu t^a_f)\psi_f$, with generators which satisfy $[t^a_f, t^b_f] = if^{abc}t^c_f$, and that the coupling treats left-handed and right-handed components equally, then the coefficient is determined solely by group theory:

$$\beta_0 = \frac{11}{3}C - \frac{4}{3}\sum_f T_f$$

where C and T_f are determined from

$$f^{acd}f^{bcd} = C\delta^{ab} \quad \text{and} \quad \mathrm{Tr}(t^a_f t^b_f) = T_f\,\delta^{ab}\,.$$

For $SU(N)$, $C = N$. If the fermions transform in the fundamental representation of the gauge group, as the quarks do in QCD, then $T_f = \frac{1}{2}$.

Although there are 6 quarks (as far as we know), the number of *active* quarks depends on the energy scale at which we wish to calculate. Since the top quark is so massive, it is not treated as an active quark flavour for energies well below 173 GeV, so we would use the $N_f = 5$ β-function. Similarly, if we were interested in the physics of a few 100 MeV, we should use the $N_f = 3$ β-function. Matching between QCD with different numbers of active quark flavours is something which requires care when working at higher-than-leading order.

In analogy with the fine structure constant of QED, it is convenient to introduce

$$\alpha_s = \frac{g^2}{4\pi}\,; \tag{7.2.6}$$

however in QCD we call this (as well as g) the strong coupling. Multiplying (7.2.5) by $2g$ and neglecting higher orders, we have

$$\mu\frac{d\alpha_s}{d\mu} = \frac{d\alpha_s}{d\log\mu} = -\frac{\beta_0}{2\pi}\alpha_s^2\,.$$

This is easily integrated

$$\int_{\alpha_s(\mu_0)}^{\alpha_s(\mu)} \frac{d\alpha_s}{\alpha_s^2} = -\frac{\beta_0}{2\pi}\int_{\mu_0}^{\mu} d\log\mu$$

to give

$$\alpha_s(\mu) = \frac{2\pi}{\beta_0}\frac{1}{\log\frac{\mu}{\mu_0} + \frac{2\pi}{\beta_0\,\alpha_s(\mu_0)}}\,. \tag{7.2.7}$$

Let us define an energy scale[71] Λ_{QCD} by

[71] Not a cut-off!

$$\log\mu_0 - \frac{2\pi}{\beta_0\,\alpha_s(\mu_0)} = \log\Lambda_{\mathrm{QCD}}$$

Then Λ_{QCD} is the scale $\mu = \Lambda_{\mathrm{QCD}}$ at which α_s diverges:

$$\alpha_s(\mu) = \frac{2\pi}{\beta_0\log(\mu/\Lambda_{\mathrm{QCD}})}\,. \tag{7.2.8}$$

Note that $\alpha_s(\mu)$ *decreases* for increasing μ. If we look at a process where μ corresponds a physical energy, then we see that the QCD coupling gets weaker as the energy gets higher. This phenomenon is called *asymptotic freedom* and is just what was sought after to describe high energy data at the time.

In the absence of quark masses, $\mathcal{L}_{\text{QCD}}$ is scale invariant; the coupling g, and hence α_s are dimensionless. The fact that a characteristic scale emerges from the quantized theory, Λ_{QCD}, is referred to as *dimensional transmutation*. This characteristic scale gives a good estimate for the border between perturbative and nonperturbative physics. Unfortunately the scale is regularization and renormalization-scheme dependent. For QCD, one might estimate $\Lambda_{\text{QCD}} \approx 200 - 500$ MeV.

7.3 $e^+e^- \to$ *hadrons*

Since QCD is asymptotically free, we are able to treat the strong coupling constant α_s as small for high energy processes. Nevertheless the phenomenon of confinement complicates things: since free quarks are never seen, we must always confront or circumvent the nonperturbative dynamics of *hadronization*, the process by which would-be free quarks create *jets* of quarks, anti-quarks, and gluons in order to end up with colour-singlet final states.

Here we consider the annihilation of electron and positron to a virtual photon which then decays to hadronic states. We are interested in the fully-inclusive cross section for $e^+e^- \to$ hadrons; that is we simply count up all events which result in hadronic final states.

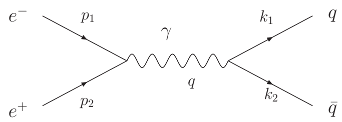

Figure 7.4: Electron-positron annihilation to quark-antiquark.

At the level of the Standard Model fields, we know that the leading process is $e^+e^- \to \gamma^* \to \bar{q}q$ (Fig. 7.4). The invariant amplitude for quarks with electric charge Q is

$$\mathcal{M} \;=\; (-ie)^2 Q\, \bar{u}_q(k_1)\gamma^\mu v_q(k_2)\, \frac{-ig_{\mu\nu}}{q^2}\, \bar{v}_e(p_2)\gamma^\nu u_e(p_1)\,.$$

Neglecting quark and electron masses, summing over quark spins, and averaging over $e^\pm$ spins we obtain

$$\begin{aligned}
\frac{1}{4}\sum_{\text{spins}}|\mathcal{M}|^2 \;&= \frac{e^4Q^2}{4q^4}\,\text{Tr}\,(\not{k}_1\gamma^\mu\not{k}_2\gamma^\nu)\;\text{Tr}\,(\not{p}_1\gamma_\mu\not{p}_2\gamma_\nu) \\
&= \frac{8e^4Q^2}{q^4}\,[p_1\cdot k_1\, p_2\cdot k_2 \;+\; p_2\cdot k_1\, p_1\cdot k_2] \\
&= e^4Q^2(1+\cos^2\theta)\,.
\end{aligned}$$

In the last step we worked in the center-of-momentum frame, writing $p_1 = (|\vec{p}|,\vec{p})$, $p_2 = (|\vec{p}|,-\vec{p})$, $k_1 = (|\vec{k}|,\vec{k})$, $k_2 = (|\vec{k}|,-\vec{k})$, then using momentum conservation we have $q = (2|\vec{p}|,0)$, $|\vec{k}| = |\vec{p}|$. Finally the angle θ is defined so that $p_1\cdot k_1 = p_2\cdot k_2 = \frac{q^2}{4}(1-\cos\theta)$ and $p_1\cdot k_2 = p_1\cdot k_2 = \frac{q^2}{4}(1+\cos\theta)$.

The differential cross section, from (6.2.8), is

$$d\sigma \;=\; \frac{1}{|\vec{v}_1-\vec{v}_2|}\frac{1}{4p_1^0p_2^0}\frac{d^3k_1}{(2\pi)^3 2k_1^0}\frac{d^3k_2}{(2\pi)^3 2k_2^0}\,(2\pi)^4\delta^{(4)}(q-k_1-k_2)\frac{1}{4}\sum_{\text{spins}}|\mathcal{M}|^2\,.$$

In the center-of-momentum frame $|\vec{v}_1 - \vec{v}_2| = 2$ and we may write

$$d\sigma \;=\; \frac{e^4 Q^2}{2(2\pi)^2 q^4}\, d^3k\, \delta(\sqrt{q^2} - 2|\vec{k}|)(1+\cos^2\theta)\,.$$

Finally we can integrate over the magnitude $|\vec{k}|$ to get

$$\frac{d\sigma}{d\Omega} \;=\; \frac{\alpha^2 Q^2}{4q^2}\,(1+\cos^2\theta)$$

where we have used the fine structure constant of QED $\alpha = e^2/4\pi$. Performing the angular integration we obtain

$$\sigma(e^+e^- \to \bar{q}q) \;=\; \frac{4\pi\alpha^2}{3q^2}\, Q^2\,. \qquad (7.3.1)$$

Since experimental uncertainties cancel when ratios of cross sections are measured, we note that the calculation for the cross section for e^+e^- annihilation to muons proceeds just as above, except that $Q = 1$, with the result

$$\sigma(e^+e^- \to \mu^+\mu^-) \;=\; \frac{4\pi\alpha^2}{3q^2}\,. \qquad (7.3.2)$$

Now back to hadronic final states. Let X represent specific final state content, whether there be 2 or more hadrons. Then the invariant amplitude can be written

$$\mathcal{M}_X \;=\; \frac{e^2}{q^2}\langle X|J_h^\mu|0\rangle \bar{v}_e(p_2)\gamma_\mu u_e(p_1)$$

where the hadron current $J_h^\mu = \sum_f Q_f \bar{q}_f \gamma^\mu q_f$ is sandwiched between X and the QCD vacuum. This matrix element cannot be determined perturbatively. The inclusive cross section must now include a sum over all possible final states X, including integrals over the momenta and spins involved:

$$\sigma(e^+e^- \to \text{had}) \;=\; \frac{1}{8p_1^0 p_2^0}\sum_X \frac{1}{4}\sum_{\text{spins},p_X}(2\pi)^4\delta^{(4)}(q-p_X)|\mathcal{M}_X|^2\,.$$

Now we introduce a useful quantity, the hadronic spectral density

$$\rho_h^{\mu\nu}(q) \;=\; (2\pi)^3\sum_{X,p_X}\delta^{(4)}(q-p_X)\langle 0|J_h^\mu|X\rangle\langle X|J_h^\nu|0\rangle\,. \qquad (7.3.3)$$

At the end of this section, we will attempt to give some insight into what this function represents. For the time being, we simply make use of its properties. Since $\rho_h^{\mu\nu}(q^2)$ is symmetric under exchange of indices $\mu \leftrightarrow \nu$, it must be a linear function of $g^{\mu\nu}$ and $q^\mu q^\nu$. Furthermore, the Ward identity which is a consequence of the conserved current implies $q_\mu \rho^{\mu\nu} = q_\nu \rho^{\mu\nu} = 0$, fixing the coefficient between the two terms. Finally, we note that since the states labelled by X have positive energy, $\rho_h^{\mu\nu}(q)$ should vanish for $q^0 < 0$. Thus we have

$$\rho_h^{\mu\nu}(q) \;=\; (-g^{\mu\nu}q^2 \,+\, q^\mu q^\nu)\,\Theta(q^0)\rho_h(q^2)\,. \qquad (7.3.4)$$

Using this in the expression for $|\mathcal{M}_X|^2$, we find

$$\frac{1}{4}\sum_{X,\text{spins}}(2\pi)^4\delta^{(4)}(q-p_X)|\mathcal{M}_X|^2 \;=\; \frac{2\pi e^4}{q^4}\left(q^2\, p_1\cdot p_2 \,+\, 2q\cdot p_1\, q\cdot p_2\right)\rho_h(q^2) \;=\; 2\pi e^4\rho_h(q^2)\,.$$

Therefore

$$\sigma(e^+e^-\to \text{had}) \;=\; \frac{16\pi^3\alpha^2}{q^2}\,\rho_h(q^2)\,. \tag{7.3.5}$$

In general $\rho_h(q^2)$ is a complicated nonperturbative function. However for the cross section $\sigma(e^+e^-\to$ hadrons) we include every hadronic state in the final state. Since the only way to produce hadrons is to produce quark-antiquark pairs, we make the assumption that a sum over all hadronic states can be replaced by a sum over all possible states involving quarks, antiquarks, and gluons

$$\sum_{X\in\text{hadrons}}|X\rangle\langle X| \;=\; \sum_{X\in q,\bar{q},g\text{ states}}|X\rangle\langle X|\,.$$

This assumption leaves out any details involving hadronization, and assumes the dynamics of the virtual photon decay can be separated from the strong dynamics involved in hadronization. Having both hadron and quark-level descriptions of the process is referred to as *quark-hadron duality*.

Making this assumption, then the spectral density can be written

$$\rho_h^{\mu\nu}(q) \;=\; N_c\sum_f Q_f^2\int\frac{d^3k_1}{(2\pi)^3 2k_1^0}\frac{d^3k_2}{(2\pi)^3 2k_2^0}(2\pi)^4\delta^{(4)}(q-k_1-k_2)\text{Tr}\left[(\not{k}_1+m_f)\gamma^\mu(\not{k}_2-m_f)\gamma^\nu\right]\Big|_{k_1^2=k_2^2=m_f^2}$$

To solve the integral, we follow similar steps as those we took to evaluate (6.3.8), the difference being that we presently keep $k_1^2=k_2^2=m_f^2>0$. Writing the integral as $I^{\mu\nu}=Aq^\mu q^\nu+Bg^{\mu\nu}$, contract both sides with $g_{\mu\nu}$ and $q_\mu q_\nu$ to obtain 2 equations which can be solved for A and B. Integrands are simplified by making pre-emptive use of the δ-function: e.g. $q^2=(k_1+k_2)^2=2m_f^2+2k_1\cdot k_2$. Ultimately one should find

$$\begin{aligned} I^{\mu\nu} = \int\frac{d^3k_1}{k_1^0}\frac{d^3k_2}{k_2^0}\,\delta^{(4)}(q-k_1-k_2)k_1^\mu k_2^\nu\Big|_{k_1^2=k_2^2=m_f^2} &= \pi\,\Theta(q^0)\Theta(q^2-4m_f^2) \\ &\times\sqrt{1-\frac{4m_f^2}{q^2}}\left[\frac{q^2+2m_f^2}{3q^2}\left(-g^{\mu\nu}q^2+q^\mu q^\nu\right)+\frac{1}{2}g^{\mu\nu}q^2\right]. \end{aligned}$$

Inserting this into the expression for the spectral density function we find

$$\rho_h(q^2) \;=\; \frac{N_c}{12\pi^2}\sum_f Q_f^2\,\Theta(q^2-4m_f^2)\left(1-\frac{4m_f^2}{q^2}\right)^{\frac{1}{2}}\frac{q^2+2m_f^2}{3q^2}\,.$$

In the limit of massless quarks this reduces to

$$\rho_h(q^2) \;=\; \frac{N_c}{12\pi^2}\sum_f Q_f^2\,.$$

Then we find the leading order cross section is

$$\sigma_{LO}(e^+e^-\to\text{hadrons}) \;=\; N_c\,\frac{4\pi\alpha^2}{3q^2}\sum_f Q_f^2 \tag{7.3.6}$$

as expected. Experimentally, it is useful to measure ratios of cross sections such as

$$R \;=\; \frac{\sigma(e^+e^-\to\text{hadrons})}{\sigma(e^+e^-\to\mu^+\mu^-)} \tag{7.3.7}$$

At leading order we find

$$R_{LO} \;=\; N_c \sum_f Q_f^2 \;=\; \begin{cases} \frac{2}{3} N_c & uds \text{ light} \\ \frac{10}{9} N_c & udsc \text{ light} \\ \frac{11}{9} N_c & udscb \text{ light} \end{cases} \tag{7.3.8}$$

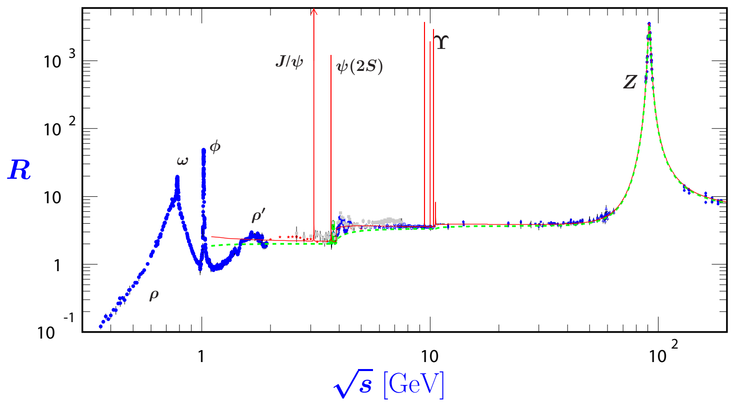

Figure 7.5: Experimental data for $R = \sigma(e^+e^- \to \text{hadrons})/\sigma(e^+e^- \to \mu^+\mu^-)$ vs. $\sqrt{s} = \sqrt{q^2}$. The green dashed line is a naive quark model prediction, which exhibits the same plateaux predicted by (7.3.8). The small jumps as the charm and bottom quarks become active, as well as the general agreement for $N_c = 3$ mark early success for QCD. (Source: Particle Data Group)

Figure 7.5 (made by the Particle Data Group[72]) shows experimental data for R. Naturally our simple calculation does not take into account the strongly interacting resonances which show up as peaks for $\sqrt{q^2} = \sqrt{s} < 10$ GeV, nor the Z peak at $\sqrt{s} = m_Z = 92$ GeV. Nevertheless one can see the green line, which only slightly improves on our calculation, agrees well with the nonresonant contributions to R: for $N_c = 3$, the plateaux described by (7.3.8) follow the data. In particular, notice how the plateau to the left of the charmonium resonances J/ψ and $\psi(2S)$ corresponding to 3 active flavours gives way to the 4-flavour plateau for $\sqrt{s} > 4$ GeV. The jump is smaller as the b quark becomes active, but nevertheless agrees with the data above the bottomonium (Υ) threshold.

[72] J Beringer et al. Review of Particle Physics (RPP). *Phys. Rev.*, D86:010001, 2012

The solid red line in Fig. 7.5 is a 3-loop perturbative QCD prediction. We can write the result as

$$\sigma \;=\; \frac{4\pi\alpha^2}{3q^2}\left[N_c \sum_f Q_f^2\, K\!\left(\alpha_s, \frac{q^2}{\mu^2}\right) \;+\; \left(\sum_f Q_f\right)^2 L\!\left(\alpha_s, \frac{q^2}{\mu^2}\right)\right] \tag{7.3.9}$$

where at the leading order we have $K(0, q^2/\mu^2) = 1$ and $L(0, q^2/\mu^2) = 0$. One-loop diagrams for $e^+e^- \to \bar{q}q$ (Fig. 7.6) are ultraviolet-finite, but diverge in the infrared, where the loop momentum becomes vanishingly small. If one uses dimensional regularization in $4 + 2\epsilon$ dimensions one finds

$$K_{\bar{q}q}^{1-\text{loop}}\left(\alpha_s, \frac{q^2}{\mu^2}\right) \;=\; \frac{C_F \alpha_s(\mu^2)}{2\pi}\left[-\frac{2}{\epsilon^2} - \frac{3}{\epsilon} - 8 + O(\epsilon)\right] H(\epsilon) \tag{7.3.10}$$

where $C_F = \frac{4}{3}$ and $H(\epsilon) = 1 + O(\epsilon)$.

The infrared divergence in (7.3.10) is cancelled by contributions from tree-level diagrams for $e^+e^- \to \bar{q}qg$ (Fig. 7.7)

$$K_{\bar{q}qg}^{\text{tree}}\left(\alpha_s, \frac{q^2}{\mu^2}\right) \;=\; \frac{C_F \alpha_s(\mu^2)}{2\pi}\left[\frac{2}{\epsilon^2} + \frac{3}{\epsilon} + \frac{19}{2} + O(\epsilon)\right] H(\epsilon)\,. \tag{7.3.11}$$

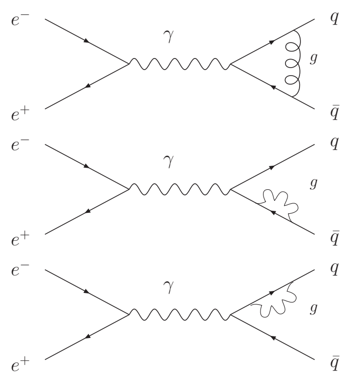

Figure 7.6: One-loop contribution to $e^+e^- \to \bar{q}q$.

Adding (7.3.10) and (7.3.11) to the leading order piece, we find

$$K\left(\alpha_s, \frac{q^2}{\mu^2}\right) = 1 + \frac{\alpha_s(\mu^2)}{\pi}. \qquad (7.3.12)$$

At higher order

$$K\left(\frac{q^2}{\mu^2}, \alpha_s\right) = 1 + \frac{\alpha_s(\mu^2)}{\pi} + \frac{\alpha_s^2(\mu^2)}{\pi^2}\left[1.99 - 0.11 n_f - \frac{\beta_0}{4}\log\frac{q^2}{\mu^2}\right]$$

The function $L(\alpha_s, q^2/\mu^2)$ enters only at $O(\alpha_s^3)$, beginning with Feynman diagrams with 3 gluons in X.

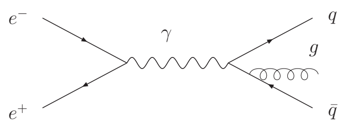

Figure 7.7: Electron-positron annihilation to quark-antiquark-gluon.

Let us conclude this section by revisiting the spectral density function $\rho_h(q^2)$ (7.3.4) and understanding its analytic structure. We first introduce the two-point function

$$\tilde{\Pi}_h^{\mu\nu}(x,y) = i\langle 0|\mathcal{T} J^\mu(x) J^\nu(y)|0\rangle \qquad (7.3.13)$$

where $\mathcal{T}$ stands for time-ordering. We will soon use its Fourier transform

$$\Pi_h^{\mu\nu}(q) = \int d^4(x-y)\, e^{iq\cdot(x-y)}\, \tilde{\Pi}_h^{\mu\nu}(x,y). \qquad (7.3.14)$$

Using Lorentz invariance and the Ward identity that $q_\mu \Pi_h^{\mu\nu} = 0 = q_\nu \Pi_h^{\mu\nu}$,

$$\Pi_h^{\mu\nu}(q) = (-g^{\mu\nu}q^2 + q^\mu q^\nu)\Pi_h(q^2). \qquad (7.3.15)$$

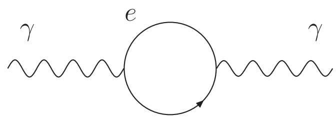

Figure 7.8: Vacuum polarization in QED.

First let us look at the 1-loop vacuum polarization of the photon in QED (Fig. 7.8) which is given in terms of a similar two-point function. Here explicit calculation is possible and sheds some light on the relation between nonanalyticities and physics. A standard but lengthy calculation (e.g. see §7.5 of Peskin and Schroeder[73]) gives

[73] M E Peskin and D V Schroeder. *An Introduction to Quantum Field Theory*. Addison Wesley, 1995. ISBN 0-201-50397-2

$$\Pi(q^2) - \Pi(0) = -\frac{2\alpha}{\pi}\int_0^1 dx\, x(1-x)\, \log\left(\frac{m_e^2}{m_e^2 - x(1-x)q^2}\right).$$

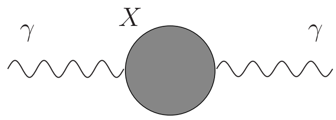

Figure 7.9: Hadronic contribution to photon vacuum polarization.

Since $x(1-x) \leq \frac{1}{4}$, $\Pi(q^2)$ has has a branch cut for real $q^2 > 4m_e^2$. Physically, this branch cut corresponds to the creation of a real electron-positron pair. We will see similar analytic structure in the hadronic contribution $\Pi_h(q^2)$ to the vacuum polarization of the photon (Fig. 7.9).

How are the two-point function $\Pi_h(q^2)$ and spectral density $\rho_h(q^2)$ related? Through the Källén-Lehmann representation. Most textbook derivations are carried out in the context of scalar field theory, looking at $\langle 0|\mathcal{T}\phi(x)\phi(y)|0\rangle$. We have to do a little extra work here for the vector current correlator. Considering the term in $\Pi_h^{\mu\nu}(x,y)$ (7.3.13) where $x^0 > y^0$, we insert a complete set of momentum eigenstates which can be created/annihilated by J_h, and we use the momentum operator $\hat{P}$ as the generator of translations to write $J_h(x) = e^{i\hat{P}x} J_h(0) e^{-i\hat{P}x}$. The result is

$$\begin{aligned}\sum_{X,p_X} i\langle 0|J_h^\mu(x)|X\rangle\langle X|J_h^\nu(y)|0\rangle &= i\sum_{X,p_X} e^{-ip_X\cdot(x-y)}\langle 0|J_h^\mu(0)|X\rangle\langle X|J_h^\nu(0)|0\rangle \\ &= i\int\frac{d^4p}{(2\pi)^3}\, e^{-ip\cdot(x-y)}\rho_h^{\mu\nu}(p) \qquad (7.3.16)\end{aligned}$$

using the spectral density function introduced early in this section, (7.3.3). We find a similar term for the $y^0 > x^0$ term in (7.3.13). Let us now set $y = 0$, without loss of generality. Using these expressions in (7.3.14), multiplying both sides of (7.3.15) by $-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}$ and using (7.3.4), we find

$$\begin{aligned}\Pi_h(q^2) &= \int \frac{d^4p}{(2\pi)^3} \int d^4x\, e^{iq\cdot x}\, i\left[\Theta(x^0)e^{-ip\cdot x} + \Theta(-x^0)e^{ip\cdot x}\right]\Theta(p^0)\rho_h(p^2)\\ &= \int \frac{d^4p}{(2\pi)^3}\int_0^\infty ds\,\delta(s-p^2)\int d^4x\, e^{iq\cdot x}\, i\left[\Theta(x^0)e^{-ip\cdot x} + \Theta(-x^0)e^{ip\cdot x}\right]\Theta(p^0)\,\rho_h(s) \qquad (7.3.17)\end{aligned}$$

In the second step, we introduce the integration variable s along with appropriate δ-function in order to utilize the Feynman propagator and then to carry out the x-integration.

Recall the Feynman propagator for a scalar with mass m is [74]

$$\begin{aligned}i\Delta_F(x;m^2) &= iD(x;m^2)\Theta(x^0) + iD(-x;m^2)\Theta(-x^0)\\ &= i\int \frac{d^3p}{(2\pi)^3 2p^0}\left[\Theta(x^0)e^{-ip\cdot x} + \Theta(-x^0)e^{ip\cdot x}\right]\Big|_{p^0=\sqrt{\vec{p}^2+m^2}}\\ &= i\int \frac{d^4p}{(2\pi)^3}\Theta(p^0)\delta(p^2-m^2)\left[\Theta(x^0)e^{-ip\cdot x} + \Theta(-x^0)e^{ip\cdot x}\right]\\ &= \int \frac{d^4p}{(2\pi)^4}\frac{e^{-ip\cdot x}}{m^2-p^2-i\epsilon} \qquad (7.3.18)\end{aligned}$$

with $\epsilon > 0$. Using this in (7.3.17) and carrying out the x integration, we find

$$\Pi_h(q^2) = \int_0^\infty ds\, \frac{\rho_h(s)}{s-q^2-i\epsilon}\,. \qquad (7.3.19)$$

Just like we saw with the 1-loop QED vacuum polarization $\Pi(q^2)$, $\Pi_h(q^2)$ has a branch cut on the positive part of the real q^2-axis, reaching the origin in the case of massless quarks. Elsewhere in the complex q^2 plane, $\Pi_h(q^2)$ is analytic. One can make use of this analyticity to carry out perturbative QCD calculations with large space-like momenta, $-q^2 \gg 1$, in which case the running coupling becomes small and the quarks and gluons are highly virtual. Then one can analytically continue the result to large time-like momenta for the $e^+e^- \to$ hadrons predictions.

The final ingredient is to solve for $\rho_h(q^2)$ in terms of the quantity computed perturbatively. From (7.3.19) and the analyticity of $\rho_h(s)$ for large s, we can infer that the discontinuity in $\Pi_h(q^2)$ across the branch cut along the real q^2 axis is as for the complex natural logarithm, $2\pi i$ times the residue:

$$\rho_h(\tilde{q}^2) = \frac{1}{2\pi i}\left[\Pi_h(\tilde{q}^2+i\delta) - \Pi_h(\tilde{q}^2-i\delta)\right]$$

with small $\delta > 0$. Taking the contour C_1 as shown in Figure 7.10, we can use the fundamental theorem of calculus to write

$$\rho_h(\tilde{q}^2) = \frac{1}{2\pi i}\int_{C_1} dz \frac{d}{dz}\Pi_h(z)\,.$$

Defining $D(-z) = -z\frac{d}{dz}\Pi_h(z)$,

$$\begin{aligned}\rho_h(\tilde{q}^2) &= \frac{1}{2\pi i}\int_{C_2}\frac{dz}{z}D(-z)\\ &= \frac{1}{2\pi}\int_{-\pi+\delta}^{\pi-\delta} d\theta\, D(\tilde{q}^2 e^{i\theta})\end{aligned}$$

[74] See, e.g. D. Tong's QFT notes. We implicitly define

$$D(x-y;m^2) = \int \frac{d^3p}{(2\pi)^3 2p^0}e^{-ip\cdot(x-y)}$$

with $p^0 = \sqrt{\vec{p}^2+m^2}$.

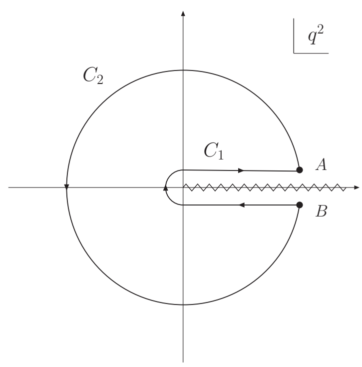

Figure 7.10: Analytic structure of $\Pi_h(q^2)$ and contours of integration. C_2 is a circular contour of radius $\tilde{q}^2$; A and B are the points $\tilde{q}^2 + i\delta$ and $\tilde{q}^2 - i\delta$
.

having used Cauchy's theorem for the integral over the closed contour $C_1 + C_2$. It is $D(-q^2)$ which is calculated in the asymptotic space-like limit $-q^2 \gg 1$ and assumed is valid for all $|q^2| \gg 1$ off of the branch cut.

7.4 *Deep inelastic scattering*

This section is not being lecture this year (2015).

Highly energetic scattering of electrons off of hadronic targets, especially the proton, revealed that hadrons have structure; they are composed of more fundamental particles. The data obtained in early experiments exhibited scaling behaviour which hinted that at high energies the hadronic constituents, at the time called *partons*, were weakly interacting. As the previous sections described, the asymptotic freedom of nonabelian gauge theories can explain this behaviour. In fact, one of the successes of QCD has been to justify the parton model and to reliably calculate corrections to it.

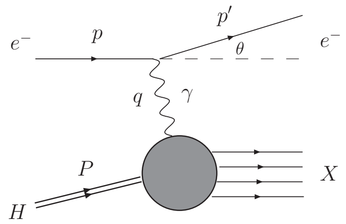

Figure 7.11: Deep inelastic scattering of an electron off of a hadron H.

Let us consider an electron with 4-momentum p scattering off of an initial state hadron H with 4-momentum P and mass M. The electron scatters at an angle θ relative to $\vec{p}$ and has final 4-momentum p'. The hadron breaks up in to a final state X in which we have no detailed interest. The scattering amplitude is given by

$$\mathcal{M} \;=\; (ie)^2 \bar{u}(p')\gamma^\mu u(p) \frac{-g_{\mu\nu}}{q^2} \langle X | J_h^\nu | H(P) \rangle \,. \tag{7.4.1}$$

The differential cross section can be given using (6.2.8). Writing the flux factor in the hadron rest frame as $|\vec{v}_e - \vec{v}_H|/V = 1/V$ we have

$$d\sigma \;=\; \frac{1}{4EM} \frac{d^3p'}{(2\pi)^3 2p'^0} \sum_X (2\pi)^4 \delta^{(4)}(q + P - p_X)\, \tfrac{1}{2} \sum_{\text{spins}} |\mathcal{M}|^2 \,.$$

We can separate the amplitude-squared into lepton and hadron factors

$$\frac{1}{2}\sum_{\text{spins}} |\mathcal{M}|^2 \;=\; \frac{e^4}{2q^4} L_{\mu\nu} \langle H(P) | J_h^\mu | X \rangle \langle X J_h^\nu | H(P) \rangle$$

where, treating the electron as massless,

$$\begin{aligned} L_{\mu\nu} \;&=\; \sum_{\text{spins}} \bar{u}(p)\gamma_\mu u(p')\bar{u}(p')\gamma_\nu u(p) \;=\; \mathrm{Tr}(\not{p}\gamma_\mu \not{p}'\gamma_\nu) \\ &= 4(p_\mu p'_\nu \,+\, p'_\mu p_\nu \,-\, g_{\mu\nu}\, p\cdot p') \end{aligned} \tag{7.4.2}$$

and

$$W_H^{\mu\nu}(q,P) \;=\; \frac{1}{4\pi} \sum_X (2\pi)^4 \delta^{(4)}(q + P - p_X) \langle H(P) | J_h^\mu | X \rangle \langle X J_h^\nu | H(P) \rangle \,. \tag{7.4.3}$$

If the hadron has spin, as it does in most such experiments, then one should also include in (7.4.3) an average over initial state spins. We now have the following expression for the differential cross section

$$E' \frac{d\sigma}{d^3p'} \;=\; \frac{1}{8(2\pi)^2 EM} \frac{e^4}{q^4} L_{\mu\nu} W_H^{\mu\nu} \,.$$

We can ascertain the Lorentz structure as in previous cases, using dimensional analysis, parity, and current conservation to write

$$W_H^{\mu\nu} \;=\; \left(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\right) W_1 + \left(P^\mu - \frac{P\cdot q}{q^2} q^\mu\right)\left(P^\nu - \frac{P\cdot q}{q^2} q^\nu\right) W_2\,.$$

$W_1(\nu, Q^2)$ and $W_2(\nu, Q^2)$ are Lorentz scalars and depend on $Q^2 = -q^2 = 2p\cdot p' = 2EE'(1-\cos\theta) \geq 0$ and $\nu = P\cdot q$. If we treat the final state X as an effective particle with 4-momentum $P_X = (\sqrt{M_X^2 + \vec{P}_X^2}, \vec{P}_X)$, then $P_X = P + q$ implies $M_X^2 = (P+q)^2$ and hence $M_X^2 \geq M^2$. Therefore, we see

$$0 \;\leq\; Q^2 \;\leq\; 2\nu\,.$$

Use $q^\mu L_{\mu\nu} = q^\nu L_{\mu\nu} = 0$ in writing

$$\begin{aligned} L_{\mu\nu} W_H^{\mu\nu} &= 8p\cdot p' W_1 \,+\, 4(2p\cdot P\, p'\cdot P \,-\, M^2 p\cdot p')W_2 \\ &= 4Q^2 W_1 \,+\, 2M^2(4EE' - Q^2)W_2\,. \end{aligned}$$

Now let us focus on very high energies, taking $Q^2 \to \infty$ and $\nu \to \infty$, introducing the following dimensionless variables which stay finite:

$$x = \frac{Q^2}{2\nu}\,, \quad \text{and} \quad y = \frac{\nu}{ME}\,; \tag{7.4.4}$$

these are both bounded to be in the interval $[0,1]$. Then

$$L_{\mu\nu} W_H^{\mu\nu} \;\sim\; 8EM\left(xyW_1 \,+\, \frac{1-y}{y}\,\nu W_2\right)\,.$$

Performing the angular p'-integration

$$d^3p' \;\to\; 2\pi E'^2 d(\cos\theta) dE' \;=\; \pi E' dQ^2 dy \;=\; 2\pi E' \nu dx\, dy$$

and the differential cross section can be written

$$\frac{d\sigma}{dx\,dy} \;=\; \frac{4\pi\alpha^2}{Q^2} 2ME\left[xy^2 F_1(x,Q^2) \,+\, (1-y)F_2(x,Q^2)\right] \tag{7.4.5}$$

where $F_1(x,Q^2) = W_1(\nu,Q^2)$ and $F_2(x,Q^2) = \nu W_2(\nu, Q^2)$ are the dimensionless structure functions for the hadron H.

It is useful to introduce *light cone variables*. For V and U arbitrary 4-vectors, we introduce components along the forward/backward light-cone in the $\vec{e}_3$-direction

$$V^\pm \;=\; V^0 \pm V^3$$

along a 2-vector representing the transverse components

$$\mathbf{V}_\perp \;=\; (V^1, V^2)\,.$$

In terms of the light-cone variables the scalar product is

$$V\cdot U \;=\; \frac{1}{2}(V^+U^- \,+\, V^-U^+) \,-\, \mathbf{V}_\perp\cdot\mathbf{U}_\perp$$

which implies the Minkowski metric tensor has components $g_{+-} = g_{-+} = \frac{1}{2}$, $g_{++} = g_{--} = 0$, and $g_{ij} = -\delta_{ij}$ for $i,j = 1,2$. A Lorentz boost in the $\vec{e}_3$-direction transforms $V^{\pm} \mapsto e^{\pm\theta} V^{\pm}$ and $\mathbf{V}_\perp \mapsto \mathbf{V}_\perp$.

Let us choose a frame where $\mathbf{P}_\perp = \mathbf{q}_\perp = 0$.[75] Then

[75] Choose the rest frame of the hadron, and rotate so that the photon is moving in the $\vec{e}_3$-direction.

$$\begin{aligned} Q^2 &= -q^+ q^- \\ \nu &= \frac{1}{2}\left(q^+ P^- + q^- P^+\right) . \end{aligned}$$

Now we take the deep inelastic limit to be $q^- \to \infty$ with $q^+ = O(P^+)$, so that

$$x \sim -\frac{q^+}{P^+} \quad \text{and} \quad \nu \sim \frac{q^- P^+}{2} .$$

In this frame we have

$$\begin{aligned} W_H^{+-}(q,P) &= -W_1 + \left(P - \frac{P\cdot q}{q^2} q\right)^2 W_2 \\ &= -W_1 + \left(M^2 + \frac{\nu^2}{Q^2}\right)^2 W_2 \\ &\equiv F_L(x, Q^2) . \end{aligned} \tag{7.4.6}$$

In the deep inelastic limit

$$F_L(x,Q^2) \sim -F_1(x,Q^2) + \frac{1}{2x} F_2(x,Q^2) . \tag{7.4.7}$$

The other longitudinal components of W_H are also related to the longitudinal structure function

$$\begin{aligned} W_H^{++}(q,P) &= \frac{(q^+)^2}{Q^2} F_L(x,Q^2) \\ W_H^{--}(q,P) &= \frac{(q^-)^2}{Q^2} F_L(x,Q^2) . \end{aligned}$$

This must be the case in order to satisfy current conservation (via the Ward identities).

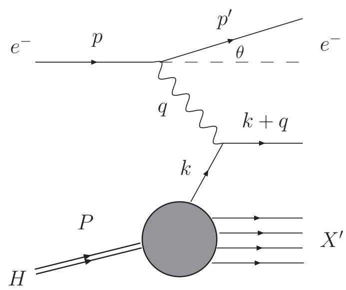

Figure 7.12: Deep inelastic scattering of an electron off of a parton inside hadron H.

Let us use assume that the photon emitted by the electron interacts with a single constituent of the hadron, and that this electromagnetic interaction is unaffected by the strong interactions. This approximation is called *factorization*, and the leading-order model we construct here is the *parton model*. Historically this model preceded the acceptance of QCD as the correct theory of the strong interactions, but now we can associate the partons with quarks[76] and even use QCD to calculate higher-order corrections to the parton model.

[76] The gluons are electrically neutral, so they do not interact directly with the photon.

In the parton model, we assume that the virtual photon strikes a single constituent, carrying momentum k before the interaction and $k+q$ after (Fig. 7.12). Now the sum over final states can be written as the number of ways a parton could have been struck. Assuming the partons (indexed by f) are massless, we have

$$\sum_X = \sum_{X'} \sum_f \frac{1}{(2\pi)^3} \int d^4\tilde{k}\, \Theta(\tilde{k}^0)\, \delta(\tilde{k}^2) \sum_{q\,\text{spins}} .$$

Writing the electromagnetic current as $J_h^\mu = \sum_f Q_f \bar{q}_f \gamma^\mu q_f$ the hadronic contribution to the cross section becomes

$$W_H^{\mu\nu}(q,P) = \sum_f \int d^4k \, \mathrm{Tr}\left[W_f^{\mu\nu}\Gamma_{H,f}(P,k) + \bar{W}_f^{\mu\nu}\bar{\Gamma}_{H,f}(P,k)\right] \qquad (7.4.8)$$

with

$$W_f^{\mu\nu} = \bar{W}_f^{\mu\nu} = \frac{1}{2}Q_f^2\gamma^\mu(\not{k}+\not{q})\gamma^\mu\delta((k+q)^2)\,.$$

$$\Gamma_{H,f}(P,k)_{\beta\alpha} = \sum_{X'}\delta^{(4)}(P-k-p_{X'})\langle H(P)|\bar{q}_{f\alpha}|X'\rangle\langle X'|q_{f\beta}|H(P)\rangle$$
$$\bar{\Gamma}_{H,f}(P,k)_{\beta\alpha} = \sum_{X'}\delta^{(4)}(P-k-p_{X'})\langle H(P)|q_{f\alpha}|X'\rangle\langle X'|\bar{q}_{f\beta}|H(P)\rangle$$

with α,β spin indices.

$$\gamma^\mu\gamma^\lambda\gamma^\nu = s^{\mu\nu\lambda\kappa}\gamma_\kappa + i\epsilon^{\mu\nu\lambda\kappa}\gamma_\kappa\gamma^5$$

with

$$s^{\mu\nu\lambda\kappa} = g^{\mu\lambda}g^{\nu\kappa} + g^{\mu\kappa}g^{\nu\lambda} - g^{\mu\nu}g^{\lambda\kappa}\,.$$

Also noting in the DIS limit $\not{k}+\not{q} \sim \frac{1}{2}q^-\gamma^+$ then

$$\gamma^j\gamma^+\gamma^i = \gamma^+(\delta^{ji} + i\epsilon^{ji}\gamma^5)\,.$$

Define functions for the integrals

$$\frac{1}{2}\int d^4k\,\delta\left(\frac{k^+}{P^+} - x\right)\mathrm{Tr}\left(\gamma^+\Gamma_{H,f}(P,k)\right) = P^+q_f(x)$$
$$\frac{1}{2}\int d^4k\,\delta\left(\frac{k^+}{P^+} - x\right)\mathrm{Tr}\left(\gamma^+\bar{\Gamma}_{H,f}(P,k)\right) = P^+\bar{q}_f(x)\,. \qquad (7.4.9)$$

Putting the pieces together we arrive at

$$F_1(x,Q^2) \sim \frac{1}{2}\sum_f Q_f^2\left[q_f(x) + \bar{q}_f(x)\right]\,. \qquad (7.4.10)$$

A similar calculation for $F_L(x,Q^2)$ shows that it vanishes in the deep inelastic limit, which implies that

$$F_2(x,Q^2) \sim 2xF_1(x,Q^2) \qquad (7.4.11)$$

a relation first derived by Callan and Gross.

We see in (7.4.10) and (7.4.11) that the parton model predicts the structure functions F_1 and F_2 are independent of Q^2: they do not depend on the absolute center-of-momentum energy of the collision, but only on the ratio $x = Q^2/2P\cdot q$. This prediction is referred to as Bjorken scaling, and x is often called Bjorken-x. Figure 7.13 shows a recent compilation of representative data for the proton's structure function $F_2(x,Q^2)$.[77] Indeed for a wide range of x, the data are relatively independent of Q^2 over several orders of magnitude.

[77] J Beringer et al. Review of Particle Physics (RPP). *Phys. Rev.*, D86:010001, 2012

The parton model is justified within QCD by asymptotic freedom. The power of QCD comes in using the full theory to calculate corrections to the leading order parton model description. We will simply give a schematic outline of what happens in such calculations.

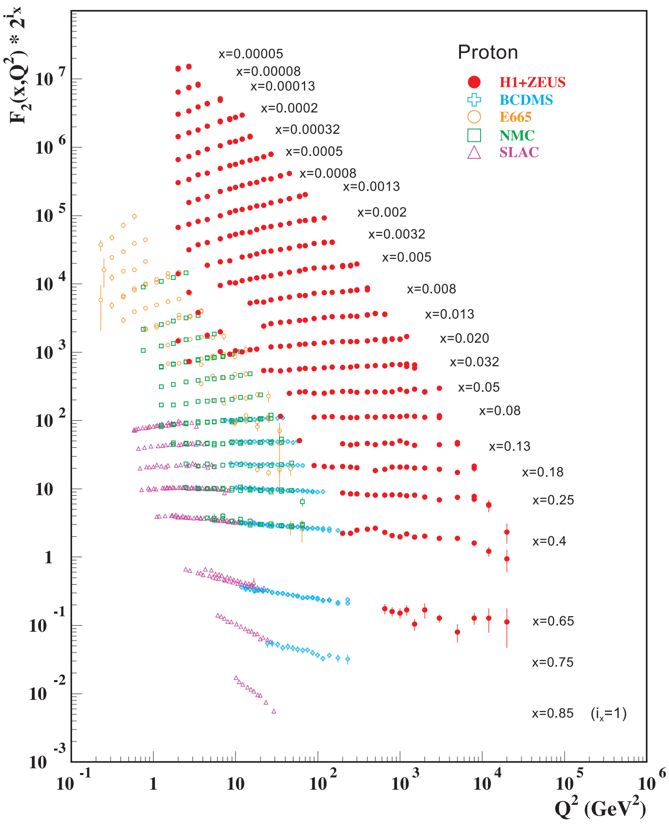

Figure 7.13: The proton structure function $F_2(x, Q^2)$ scaled by an x-dependent factor 2^{i_x} (for the purposes of plotting). For moderate values of x, F_2 is nearly independent of Q^2. (Source: Particle Data Group)

Consider a structure function $F(x, Q^2)$ (the story is the same for each of them). At next-to-leading order in QCD, we must account for the fact that the parton will have had an interaction proportional to α_s before being struck by the photon. For example, it could have radiated a gluon (Fig. 7.14) or it could have been a gluon which pair-creates a quark-antiquark pair (Fig. 7.15).

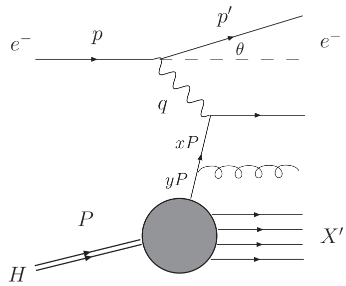

Figure 7.14: Deep inelastic scattering of an electron off of a parton inside hadron H.

$$F(x, Q^2) \sim \sum_{i \in \{q_f, \bar{q}_f, G\}} \int_x^1 \frac{dy}{y} C_i\left(\frac{x}{y}, \frac{Q^2}{\mu_F^2}; \alpha_s\right) f_i(y, \mu_F^2) \qquad (7.4.12)$$

where $f_i(y, \mu_F^2)$ is one of $\{q_f(y, \mu_F^2), \bar{q}_f(y, \mu_F^2), G(y, \mu_F^2)\}$.

$F(x, Q^2)$ must be independent of the unphysical factorization scale μ_F. Thus, $\mu_F \frac{d}{d\mu_F} F = 0$ implies[78]

$$\begin{aligned} \mu_F \frac{d}{d\mu_F} C_i\left(x, \frac{Q^2}{\mu_F^2}; \alpha_s\right) &= -\sum_j \int_x^1 \frac{dy}{y} C_j\left(y, \frac{Q^2}{\mu_F^2}; \alpha_s\right) P_{ji}\left(\frac{x}{y}; \alpha_s\right) \\ \mu_F \frac{d}{d\mu_F} f_i(y, \mu_F^2) &= \sum_j \int_y^1 \frac{dz}{z} P_{ij}\left(\frac{y}{z}; \alpha_s\right) f_j(z, \mu_F^2) \,. \end{aligned} \qquad (7.4.13)$$

These equations are called the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP), or sometimes just Altarelli-Parisi, equations.

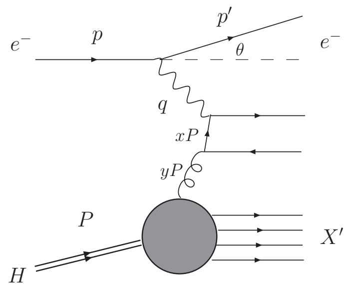

Figure 7.15: Deep inelastic scattering of an electron off of a parton inside hadron H.

[78] We used, for general A, B, the fact that if $\mu \frac{d}{d\mu} A_i = -A_j P_{ji}$ and $\mu \frac{d}{d\mu}(A_i B_i) = 0$ then $\mu \frac{d}{d\mu} B_i = P_{ij} B_j$.

Effective field theory

As we can infer from the preceding chapter, calculations in QCD are difficult and contain subtleties, even in the perturbative limit. Solving QCD in the nonperturbative regime is even more difficult. While numerical methods using lattice field theory yield first principles results in some cases, insight and accurate predictions can be made using a framework called effective field theory.

Effective field theory exploits large separations in energy scales in order to construct a simpler description of low energy physics. We already saw a nice example in Fermi's weak theory, an effective description of the full electroweak theory, which exploits the large mass of the W boson to describe weak decays at the scale of a few GeV and below by local 4-fermion interactions. In that case it was safe to expand the W propagator

$$\frac{1}{p^2 - m_W^2} \approx -\frac{1}{m_W^2} - \frac{p^2}{m_W^4} + \dots$$

since the external momenta involved were small enough that the nonanalytic structure of the W propagator played a negligible role.

Here we wish to build effective low-energy Lagrangians: a series of local operators involving only light degrees-of-freedom. There are several useful reviews of effective field theory (EFT) in the literature. The one by Georgi[79] explains the framework very well. The next 2 sections follow lectures by Kaplan.[80] Both of those works cite other useful references.

[79] H Georgi. Effective field theory. *Ann. Rev. Nucl. Part. Sci.*, 43:209–252, 1993

[80] D B Kaplan. Five lectures on effective field theory. 2005. arXiv:nucl-th/0501023

8.1 Scaling dimensions of local operators

Since we accept that our effective field theory is valid only up to some mass scale Λ we cannot use renormalizability as a constraint to determine which operators can and cannot enter the Lagrangian. In principle, there are an infinite number of terms which enter. Let us write the Lagrangian, separating the kinetic energy terms from the interactions

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{kin}} + \sum_n \mathcal{L}_{\text{int}}^{(n)} .$$

Georgi refers to the last term as an infinite "tower of interactions." We will see that the infinite height of the tower does not trouble us if we are simply interested in the view from finite heights.

Recalling that observables may be computed using path integrals, we can observe from the normalization factor (the partition function of QFT)

$$Z = \int \mathcal{D}(\text{fields})\, e^{i\int d^4x\,\mathcal{L}^{\text{eff}}}$$

that the Lagrangian must have mass dimension 4 in order to cancel the dimensions of the integration measure.[81] Thus we write $\mathcal{L}_{\text{int}}^{(n+4)}$ as a sum of dimension-$n+4$ operators times dimensionless coefficients, with the dimensions made correct with factors of Λ

$$\mathcal{L}_{\text{int}}^{(n+4)} = \sum_i \frac{c_i^{(n)}}{\Lambda^n}\, O_i^{(n+4)}\,.$$

[81] The dimensional analysis throughout this chapter relies on using $\hbar = c = 1$ units. Minor changes are necessary for other systems of natural units such as nonrelativisitic natural units $\hbar = 2m = 1$.

We need to make a couple assumptions in order to proceed:

1. There are a finite number of parameters for each $\mathcal{L}_{\text{int}}^{(n+4)}$, correspondingly a finite number of independent operators.

2. The coefficients of operators can be written as $\frac{c^{(n)}}{\Lambda^n}$ where the $c^{(n)}$ are dimensionless coefficients, at most of order 1, and Λ represents some heavy mass scale which is independent of n.

If these assumptions are valid, then we can truncate the tower of interactions depending on how far and accurately we want to see. For a given dynamical energy E, contributions to observables from $\mathcal{L}_{\text{int}}^{(n+4)}$ are corrections of order $(E/\Lambda)^n$. If we desire accuracy of order ε then we must find the power n_ε large enough so that

$$\left(\frac{E}{\Lambda}\right)^{n_\varepsilon} \approx \varepsilon\,.$$

that is,

$$n_\varepsilon \approx \frac{\log(1/\varepsilon)}{\log(\Lambda/E)}\,.$$

We see from this that we must increase n_ε if we seek greater accuracy (decreasing ε) or if we wish to describe higher energy behaviour. We shall see later the important role symmetries play in further constraining the types of operators which appear in $\mathcal{L}_{\text{eff}}$.

Let us consider a real scalar field in 4 dimensions. Let us also work in Euclidean spacetime where the path integral is well-behaved. The effective Lagrangian which will describe physics up to a cutoff scale Λ is generally represented by an infinite number of terms which can be written[82]

$$\begin{aligned}\mathcal{L} =& \frac{1}{2}(\partial\phi)^2 + \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4!}\phi^4 \\ &+ \sum_{n=1}^{\infty}\left[\frac{c_n}{\Lambda^{2n}}\,\phi^{4+2n} + \frac{d_n}{\Lambda^{2n}}\,(\partial\phi)^2\phi^{2n} + \dots\right].\end{aligned} \tag{8.1.1}$$

[82] We also assume the theory should be invariant under $\phi \mapsto -\phi$.

Given that the mass dimension of $\mathcal{L}$ is 4 and that of m and ∂_μ is 1, the field must carry mass dimension 1. Therefore, as written the couplings λ, c_n, and d_n are dimensionless. Let us assume here that the theory is perturbative, so that all of these dimensionless couplings are small.

Correlations functions in terms of path integrals are given by expressions of the form

$$\langle \phi_1 \cdots \phi_n \rangle = \frac{1}{Z} \int \mathcal{D}\phi \, \phi_1 \cdots \phi_n \, e^{-S}$$

where $S = \int d^4x \, \mathcal{L}$ is the Euclidean action and $Z = \int \mathcal{D}\phi \, e^{-S}$. Below we will give 2 arguments justifying an expansion which treats lower dimension opertors as the most important operators in the effective field theory expansion, with higher dimension operators representing small corrections.

Consider a specific field configuration $\tilde{\phi}(x)$ which is localized in a volume L^4 where $L \approx 2\pi/k$, and k is a wavenumber (or momentum). Take the amplitude of the wavelet to be ϕ_k, and let us define the dimensionless amplitude which is the ratio $\hat{\phi}_k = \phi_k/k$. For fun, an example wavelet is given in Figure 8.1.

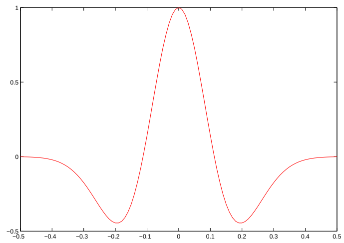

Figure 8.1: Radial part of the 4-d wavelet

$$\tilde{\phi}(x) = \phi_k(1 - 2k^2r^2)e^{-k^2r^2} .$$

For such wavelet field configurations we can crudely approximate the integrals of the terms in the Lagrangian (8.1.1) which sum to give the Euclidean action. Since the wavelet only has support in a volume $L^4 = (2\pi/k)^4$ we can estimate

$$\begin{aligned}
\int d^4x \, m^2 \tilde{\phi}^2 &\approx L^4 m^2 \phi_k^2 = \left(\frac{2\pi}{k}\right)^4 m^2 k^2 \hat{\phi}_k^2 \\
\int d^4x \, (\partial\tilde{\phi})^2 &\approx L^4 k^2 \phi_k^2 = (2\pi)^4 \hat{\phi}_k^2 \\
\left(\frac{1}{\Lambda}\right)^{2p+q-4} \int d^4x \, (\partial\tilde{\phi})^p \tilde{\phi}^q &\approx (2\pi)^4 \left(\frac{k}{\Lambda}\right)^{2p+q-4} \hat{\phi}_k^{p+q} .
\end{aligned} \tag{8.1.2}$$

Then for this field configuration, the Euclidean action is given approximately by

$$\begin{aligned}
S \approx (2\pi)^4 \Bigg\{ & \frac{\hat{\phi}_k^2}{2} + \frac{m^2 \hat{\phi}_k^2}{2k^2} + \frac{\lambda}{4!} \hat{\phi}_k^4 \\
& + \sum_n \left[c_n \left(\frac{k}{\Lambda}\right)^{2n} \hat{\phi}_k^{4+2n} + d_n \left(\frac{k}{\Lambda}\right)^{2n} \hat{\phi}_k^{2+2n} + \ldots \right] \Bigg\} .
\end{aligned} \tag{8.1.3}$$

For this field configuration, its contribution to path integrals becomes an ordinary integral over $\hat{\phi}_k$. Due to the exponential factor $\exp(-S)$, integrals will be dominated by values of $\hat{\phi}_k$ which minimize S. In the perturbative regime, the quadratic terms in (8.1.3) are the most important. In the relativistic theories which concern us here, $k \gg m$ and the kinetic energy term dominates. Thus S is minimized for $\hat{\phi}_k \approx 1/(2\pi)^2$ or smaller. Note that this condition is independent of k.

As k is reduced, the terms in the square bracket of (8.1.3) are reduced as $(k/\Lambda)^{2n}$, hence are called *irrelevant*, in the sense of the renormalization group (which is the proper context in which to view our arguments in this section). Of course some these terms are likely not to be irrelevant for low energy physics, but we now see how to rank them from least to most irrelevant. The mass term is seen here to be *relevant*, since it becomes larger in the infrared. The $\lambda\phi^4/4!$ term is termed *marginal* since at leading order we cannot tell whether it is relevant or irrelevant; quantum loop corrections will ultimately determine its behaviour.

A more general but less visual way to understand the relative importance of terms in the tower of interactions is to consider scale transformations. Let $\phi(x)$ now be an arbitrary field configuration, with action $S(\phi(x); m^2, \lambda, c_n, d_n, \ldots)$. Now consider the family of configurations related by a scale transformation $\phi_\xi(x) = \phi(\xi x)$. Defining $x' = \xi x$ and $\phi'(x') = \xi^{-1}\phi(\xi x)$, we have the action for the scaled field as

$$\begin{aligned} S(\phi_\xi(x); m^2, \lambda, c_n, d_n, \ldots) &= \int d^4x \Bigg\{ \frac{1}{2}(\partial\phi(\xi x))^2 + \frac{1}{2}m^2\phi^2(\xi x) + \frac{\lambda}{4!}\phi^4(\xi x) \\ &\qquad + \sum_n \left[c_n \frac{\phi^{4+2n}(\xi x)}{\Lambda^{2n}} + d_n \frac{(\partial\phi(\xi x))^2\phi^{2n}(\xi x)}{\Lambda^{2n}} + \ldots \right] \Bigg\} \\ &= \int d^4x' \Bigg\{ \frac{1}{2}(\partial'\phi'(x'))^2 + \frac{1}{2}m^2\xi^{-2}(\phi'(x'))^2 + \frac{\lambda}{4!}(\phi'(x'))^4 \\ &\qquad + \sum_n \left[c_n\xi^{2n} \frac{(\phi'(x'))^{4+2n}}{\Lambda^{2n}} + d_n\xi^{2n} \frac{(\partial'\phi'(x'))^2(\phi'(x'))^{2n}}{\Lambda^{2n}} + \ldots \right] \Bigg\} . \end{aligned} \tag{8.1.4}$$

Since x' is just an integration variable, we can identify

$$S(\phi(\xi x); m^2, \lambda, c_n, d_n, \ldots) = S(\xi^{-1}\phi(x); \xi^{-2}m^2, \lambda, \xi^{2n}c_n, \xi^{2n}d_n, \ldots) . \tag{8.1.5}$$

Under scaling, we see how each term in $\mathcal{L}_{\text{eff}}$ behaves. As we take $\xi \to 0$ we expose the infrared "flow" of the couplings in (8.1.1): $\phi(x) = e^{ik\cdot x}$ is mapped to $\phi_\xi(x) = e^{i(\xi k)\cdot x}$.

We define the *scaling dimension* of fields and couplings by their behaviour under scale transformations $x' \mapsto \xi x$. We use square brackets as a symbol for scaling dimension, so if we say the scaling dimension of some object y is equal to Δ, that implies:

$$[y] = \Delta \quad \Rightarrow \quad y \mapsto \xi^{-\Delta} y . \tag{8.1.6}$$

In the scalar field example we found $[\phi] = 1$, $[m^2] = 2$, $[\lambda] = 0$, $[c_n] = [d_n] = -2n$. What we have seen from the above derivation is that the scaling dimension is equivalent to the mass dimension (assuming all factors of Λ have been absorbed back into the couplings). This is a consequence of using $\hbar = c = 1$. In nonrelativistic units, the same ideas apply, but the details differ.

8.2 Rayleigh scattering

As a simple example with which we apply these ideas of effective field theory, let us consider the low energy, elastic scattering of light off of atoms or molecules. This *Rayleigh scattering* is usually treated in classical electrodynamics texts. Here we show that we can get the main result rather straightforwardly using EFT.

We assume, as is the case with sunlight scattering off of air molecules in the Earth's atmosphere, that the energy of the photons is much less than the smallest atomic or molecular excitation energy ΔE:

$$E_\gamma \ll \Delta E \ll r_0^{-1} \ll M_{\text{atom}} \,. \tag{8.2.1}$$

Here we also note that the energy scale needed to probe the size of the atom[83] is much higher than the excitation energy.[84] The largest scale in the problem is the atomic mass M_{atom}, hence the atom is nonrelativistic. In fact we will treat the atom as *static*, i.e. its velocity is unchanged, to a very good approximation, by the scattering of the photon.

[83] I'm going to stop writing "or molecule."

[84] Typically $\Delta E = O(\alpha^2 m_e)$ while $1/r_0 = O(\alpha m_e)$, where $\alpha = e^2/4\pi$ is the fine structure constant.

In keeping with this, let $\phi_v^\dagger$ be a field which creates an atom with 4-velocity v; in the atom's rest frame $v = (1, 0, 0, 0)$. Since the atom is neutral[85] ϕ_v does not couple to the photon field A_μ directly, but to the field strength tensor $F_{\mu\nu}$.

[85] ϕ_v transforms trivially under $U(1)_{\text{EM}}$.

Now we are nearly ready to begin constructing the lower floors of the tower of interactions. There are some simplifying factors we should consider first, however. We know from the Maxwell equation(s)-of-motion that $\partial_\mu F^{\mu\nu} = 0$, so we need not include such terms in our effective Lagrangian. Furthermore, if we insist the atomic ground state have zero energy in the atom's rest frame, then there $\partial_t \phi_v = 0$ or $v^\mu \partial_\mu \phi_v = 0$. Similarly, using the nonrelativistic kinetic energy operator, we can infer $\partial_\mu \partial^\mu \phi_v = 0$.

The last ingredient we need are the scaling dimensions of our building blocks. We still work in relativistic units, and in the static limit the atomic mass does not enter anywhere, so the scaling dimension is again equal to the mass dimension. Thus $[\partial_\mu] = 1$. From the Maxwell Lagrangian $\mathcal{L}_M = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$ and $[\mathcal{L}_M] = 4$ we see $[F_{\mu\nu}] = 2$. Finally $[\phi_v] = \frac{3}{2}$ from considering the atomic wavefunction resulting from the creation operator $\phi_v^\dagger(x)|0\rangle = \Psi_A(x)|A\rangle$. Taking nonrelativistic normalization for the states, $\langle 0|0\rangle = \langle A|A\rangle = 1$ implies $\int d^3x \, |\Psi_A|^2 = 1$, which confirms the dimensions work out as claimed.

We now can write out the first few terms of the effective Lagrangian for Rayleigh scattering:

$$\begin{aligned} \mathcal{L}_{\text{eff}} \;=\; \mathcal{L}_M \;+\; g_1 \, \phi_v^\dagger \phi_v \, F_{\mu\nu} F^{\mu\nu} + g_2 \, \phi_v^\dagger \phi_v \, v^\alpha F_{\alpha\mu} \, v_\beta F^{\beta\mu} \\ + g_3 \, \phi_v^\dagger \phi_v (v^\alpha \partial_\alpha) F_{\mu\nu} F^{\mu\nu} \;+\; \ldots \end{aligned} \tag{8.2.2}$$

Higher powers of $v \cdot \partial$ yield higher dimension operators, which are more irrelevant.

Since $[\mathcal{L}_{\text{eff}}] = 4$, $[g_1] = [g_2] = -3$ and $[g_3] = -4$. Leading order scattering is governed by the g_1 and g_2 terms. We have multiple high energy scales (8.2.1) which we could associate with the Λ of § 8.1. In the case of leading order scattering, we expect purely classical scattering, and so the high energy scale making up the dimensions would more likely be the inverse atomic size $1/r_0$ rather than the excitation energy of the valence electron(s). Therefore we assume g_1 and g_2 are $O(r_0^3)$, hence we define dimensionless coupling constants $a_1 = g_1/r_0^3$ and $a_2 = g_2/r_0^3$, so that

$$\mathcal{L}_{\text{eff}} \;=\; \mathcal{L}_M \;+\; r_0^3 (a_1 \, \phi_v^\dagger \phi_v \, F_{\mu\nu} F^{\mu\nu} \;+\; a_2 \, \phi_v^\dagger \phi_v \, v^\alpha F_{\alpha\mu} \, v_\beta F^{\beta\mu}) \,. \tag{8.2.3}$$

The cross section is proportional to the scattering amplitude squared, i.e. proportional to $|\langle\gamma(k'), A|\mathcal{L}_{\text{int}}|\gamma(k), A\rangle|^2$, so $\sigma = O(r_0^6)$. Given that the cross section has dimensions of area, the mismatch in dimensions must be made up by the only relevant dynamical energy scale E_γ, so we deduce that

$$\sigma \propto E_\gamma^4 r_0^6 . \tag{8.2.4}$$

We see that blue light scatters more than red, and with a power law that agrees with the full derivation in classical electrodynamics.

For higher energy photons, the leading order effective Lagrangian is not sufficient. Effects due to the excitation energies of the air molecules start to become non-negligible for large enough E_γ,

$$\sigma \propto E_\gamma^4 r_0^6 \left(1 + O\left(\frac{E_\gamma}{\Delta E}\right)\right) .$$

These arguments give a result which agrees with the full classical calculation, up to constants of proportionality which need to be determined by knowing the full theory (or perhaps by performing experiments). In the case of Rayleigh scattering, the effective field theory calculation has the virtue of being simpler. However, the real power of EFT comes in being able to make progress when calculations in the full theory are not possible.

8.3 Chiral Lagrangian

References for this section include [86] as well as the classic papers.[87]

In QCD, it appears that $SU(3)_F$ is an approximate global symmetry which allows us to classify the light and strange hadrons. The octet of pseudoscalar mesons have much smaller masses than the rest of the hadrons. We can understand these particles to be the Goldstone bosons which arise from the spontaneous breaking of a larger, chiral symmetry $SU(3)_L \times SU(3)_R$.[88]

Recalling the discussion of § 4.2, we can interpret the Goldstone bosons as excitations of a specific vacuum configuration ϕ_0 where the field in a localized volume is transformed away from this vacuum to another. In the $O(N)$ model with a spontaneously broken vacuum $\phi_0 = (0, 0, \dots, v)^T$, the excitations were those of the form

$$\phi(x) = \begin{pmatrix} \pi_1(x) \\ \pi_2(x) \\ \dots \\ v + \sigma(x) \end{pmatrix} .$$

We can think of the massless field $\pi_j(x)$ as being localized transformations away from the vacuum ϕ_0. Below we develop a more general method for describing Goldstone excitations, using group transformations $\tilde{g} \in G$ with $\tilde{g} \notin H$.

We proceed following the steps:

1. Goldstone fields should represent physical excitations – we want them to create/annihilate asymptotic out/in states. Since the

[86] J F Donoghue, E Golowich, and B R Holstein. *Dynamics of the Standard Model*. Cambridge University Press, 1992. ISBN 0-521-47652-6; H Georgi. *Weak interactions and modern particle theory*. Benjamin/Cummings, 1984. ISBN 0-8053-3163-8; and S Weinberg. *The Quantum Theory of Fields, Volume II*. Cambridge University Press, 1996. ISBN 0-521-55002-5

[87] S R Coleman, J Wess, and B Zumino. Structure of phenomenological Lagrangians. 1. *Phys. Rev.*, 177:2239–2247, 1969; and C G Callan, S R Coleman, J Wess, and B Zumino. Structure of phenomenological Lagrangians. 2. *Phys. Rev.*, 177:2247–2250, 1969. DOI: 10.1103/PhysRev.177.2247

[88] The fact that these mesons have small rather than zero mass arises from the small masses of the u, d, and s quarks, and can be treated as a perturbation.

Goldstone excitations correspond to local fluctuations from one vacuum to a different vacuum, the fields should correspond to coordinates in the coset space G/H.

2. The remaining symmetry corresponding to H should be manifest.

3. The effective Lagrangian should be invariant under G.

Our notation below foreshadows what happens in QCD, which has Lagrangian with chiral symmetry – under separate left- and right-handed transformations – but has axial-vector combinations broken by a nonzero expectation value for the chiral condensate $\langle 0|\bar{q}q|0\rangle$, leaving a remnant symmetry under vector-like transformations. (This will be discussed more precisely later.) Let us denote the generators of the unbroken subgroup H as V^a and the remaining generators of G as A^b ($a = 1, \ldots, \dim G$ and $b = 1, \ldots, \dim G - \dim H$).

Let us write an element of G, $g \in G$ as [89]

$$g = e^{i\alpha\cdot A} e^{i\beta\cdot V} \tag{8.3.1}$$

[89] Note that this is equivalent to $\exp(i\gamma \cdot A + i\delta \cdot V)$, where $\gamma \neq \alpha$ and $\delta \neq \beta$ because the generators generally do not commute. Nevertheless we know there do exist γ and δ so that (8.3.1) holds due to closure of the Lie algebra of G. In other words, both (α, β) and (γ, δ) are valid choices for coordinates on the group manifold.

Since the Goldstone excitations are local misalignments in group space (away from the coset eH, where e is the identity of G), they can be represented by coordinates $\xi(x)$ on the coset space G/H. At each point we can imagine the excitation as a group transformation on the vacuum of

$$e^{i\xi(x)\cdot A}\,. \tag{8.3.2}$$

The $\xi(x)$ fields will be related (up to a dimensional factor) to the fields that create/annihilate Goldstone bosons. However it is easier to make the symmetry properties manifest working with elements of the group G than of its corresponding algebra.

If we perform a general group transformation $g_0 \in G$ on our field

$$g_0 e^{i\xi\cdot A} = e^{i\xi'\cdot A} e^{iu'\cdot V} \tag{8.3.3}$$

the result being a field with elements in both G/H and H ($\xi'(x)$ and u' depend on $\xi(x)$ and g_0). Let us check the behaviour under group multiplication. If we similarly write

$$g_1 e^{i\xi'\cdot A} = e^{i\xi''\cdot A} e^{iu''\cdot V} \tag{8.3.4}$$

then

$$g_1 g_0\, e^{i\xi\cdot A} = e^{i\xi''\cdot A} e^{iu'''\cdot V} \tag{8.3.5}$$

where $e^{iu'''\cdot V} = e^{iu''\cdot V} e^{iu'\cdot V}$. We see that choosing to write group elements as products of elements of G/H and H (8.3.1) allows us to factorize transformations.

As usual, we assume we are working with orthogonal generators so that $\mathrm{Tr} t^a t^b \propto \delta^{ab}$. Then the commutator $[V^i, A^j]$ is in the span of the generators A^j:

$$[V^i, A^j] \in \mathrm{span}(A)\,. \tag{8.3.6}$$

This can be checked by showing $\mathrm{Tr}([V^i, A^b]V^{i'}) = 0$ using the closure of the subalgebra $[V^i, V^{i'}] = if^{ii'i''}V^{i''}$. Therefore for $h = e^{iu\cdot V} \in H$,

$$he^{i\xi(x)\cdot A} = e^{i\xi'(x)\cdot A}h\,. \tag{8.3.7}$$

In words, commuting $e^{i\xi(x)\cdot A}$ with an element h of the invariant subgroup only alters the transformation in the coset space, not on the subgroup manifold due to (8.3.6). Right-multiplying (8.3.7) by h^{-1} we get

$$e^{i\xi'(x)\cdot A} = he^{i\xi(x)\cdot A}h^{-1} = e^{iu\cdot V}e^{i\xi(x)\cdot A}e^{-iu\cdot V}\,. \tag{8.3.8}$$

Now we need to bring in a further symmetry property of QCD. Here $G = SU(3)_L \times SU(3)_R$ and $H = SU(3)_F$. Sometimes this is also written $H = SU(3)_V$ because the vector symmetries are preserved while the axial-vector symmetries are broken. Writing the generators of $SU(3)_L \times SU(3)_R$ as L^a and R^a, then the generators of H are $V^a = \frac{1}{2}(L^a + R^a)$ with the generators $A^a = \frac{1}{2}(L^a - R^a)$ broken.[90]

[90] We could use more explicit notation to emphasize on which part of the direct product group the generators act: $L^a = T^a \otimes 1$ and $R^a = 1 \otimes T^a$, where T^a are generators of $SU(3)$. With this notation it is clear that $[L^a, L^b] = if^{abc}L^c$, $[R^a, R^b] = if^{abc}R^c$, $[L^a, R^b] = 0$.

Consider an automorphism $g \mapsto \mathcal{R}(g)$ which takes $L^a \mapsto R^a$ and $R^a \mapsto L^a$, and hence $V^a \mapsto V^a$ and $A^a \mapsto -A^a$. This is a simple relabeling of left and right, which should leave the QCD physics invariant. We will make use of this below.

Let us return to the result of left- multiplying by $g \in G$. From (8.3.1)

$$\begin{aligned} ge^{i\xi(x)\cdot A} &= e^{i\alpha\cdot A}e^{i\beta\cdot V}e^{i\xi(x)\cdot A} \\ &= e^{i\xi'(x)\cdot A}e^{i\beta\cdot V} \end{aligned} \tag{8.3.9}$$

where in this instance we write $\exp(i\xi'(x)\cdot A)$ as the product of $\exp(i\alpha\cdot A)$ and the result of applying (8.3.8). Applying the automorphism (flipping the sign of A but not V) and in the next line inverting the result

$$\begin{aligned} \mathcal{R}(g)e^{-i\xi\cdot A} &= e^{-i\xi'\cdot A}e^{i\beta\cdot V} \\ e^{i\xi\cdot A}\mathcal{R}(g^{-1}) &= e^{-i\beta\cdot V}e^{i\xi'\cdot A} \end{aligned} \tag{8.3.10}$$

using $\mathcal{R}^{-1}(g) = \mathcal{R}(g^{-1})$. Multiplying the latter lines of (8.3.9) and (8.3.10) we arrive at

$$ge^{2i\xi\cdot A}\mathcal{R}(g^{-1}) = e^{2i\xi'\cdot A}\,. \tag{8.3.11}$$

Finally, writing using $g = e^{i\alpha\cdot A}e^{i\beta\cdot V}$ and $\mathcal{R}(g^{-1}) = e^{-i\beta\cdot V}e^{i\alpha\cdot A}$ we arrive at

$$e^{2i\xi'(x)\cdot A} = e^{i\alpha\cdot A}e^{i\beta\cdot V}e^{2i\xi(x)\cdot A}e^{-i\beta\cdot V}e^{i\alpha\cdot A}\,. \tag{8.3.12}$$

This important results tells us how the field of coset elements transforms under transformations $g \in G$. Note the important plus sign in the last exponent. If it were a minus sign then we would have shown $\exp(2i\xi\cdot A)$ transformed linearly under G (we would have had a similarity transformation $k' = gkg^\dagger$.) Of course we want to describe a theory where this is *not* the case, so (8.3.12) is appropriate.

It is conventional to work with dimensionful fields, so let us introduce a constant F which carries mass dimension 1, so that

$$\xi(x) \cdot A = \xi^a(x) A^a = \frac{1}{F}\Pi^a(x) t^a \equiv \frac{1}{F}\Pi(x)\,. \tag{8.3.13}$$

In the case of 3-flavour QCD,[91] with quarks transforming in the fundamental representation of $SU(3)_F$, the Π field which can be identified with the $SU(3)_F$ octet of pseudoscalar mesons (by supposition) as

$$\Pi = \begin{pmatrix} \frac{\pi^0}{\sqrt{2}} + \frac{\eta}{\sqrt{6}} & \pi^+ & K^+ \\ \pi^- & -\frac{\pi^0}{\sqrt{2}} + \frac{\eta}{\sqrt{6}} & K^0 \\ K^- & \bar{K}^0 & -\frac{2\eta}{\sqrt{6}}\,, \end{pmatrix} \tag{8.3.14}$$

where we assumed the Gell-Mann basis for the generators, $t^a = T^a = \lambda^a/2$ (see Note 67). From this, we can determine the coefficients Π^a, for example $\Pi^1 = \pi^+ + \pi^-$ and $\Pi^2 = i(\pi^+ - \pi^-)$.

This Π field satisfies the requirement that it transform linearly under $SU(3)_F$. Setting $\alpha = 0$ in (8.3.12)

$$\begin{aligned} e^{i\beta\cdot V} e^{2i\Pi/F} e^{-i\beta\cdot V} &= \sum_{n=0}^{\infty} \frac{1}{n!} e^{i\beta\cdot V} \left(\frac{2i\Pi}{F}\right)^n e^{-i\beta\cdot V} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(e^{i\beta\cdot V} \frac{2i\Pi}{F} e^{-i\beta\cdot V}\right)^n \\ &= e^{2i\Pi'/F} \end{aligned} \tag{8.3.15}$$

where $\Pi' = e^{i\beta\cdot V}\Pi e^{-i\beta\cdot V}$ as required.

Let us write $\Sigma = e^{2i\Pi/F}$, as well as $U_L = e^{i\alpha\cdot A} e^{i\beta\cdot V}$ and $U_R = e^{i\alpha\cdot A} e^{i\beta\cdot V}$, then (8.3.12) becomes

$$\Sigma' = U_L \Sigma U_R^\dagger\,. \tag{8.3.16}$$

In the case that $\alpha = 0$, $U_L = U_R = U_V$, and (8.3.16) is just a similarity transformation leaving the vacuum invariant, reflecting the unbroken symmetry of H, in this case $SU(3)_F$. For $\alpha \neq 0$, the transformation takes us from one vacuum to another.

The effective Lagrangian still must be invariant under the full symmetry group of the full Lagrangian, $SU(3)_L \times SU(3)_R$ in QCD. Given (8.3.16), we can infer operators in the effective Lagrangian $\mathcal{L}_\chi$ must have Σ and $\Sigma^\dagger$ appearing in pairs. Since the operators should be scalars, terms should be traced over $SU(3)$ indices.

The final observation we should make before writing down the Lagrangian is that the scale at which chiral perturbation theory breaks down, generically labelled Λ_χ is not known *a priori*. One might use the experimental observation that the vector mesons, e.g. the ρ meson, cannot be Goldstone bosons,[92] to suggest $\Lambda_\chi \approx m_\rho = 770$ MeV. Another argument might be that $\Lambda_\chi \approx 4\pi F$. In fact this is up for debate. One must examine the convergence of chiral perturbation theory empirically.

[91] If one in interested in energies where strange mesons play no role, one might work only with 2-flavour QCD and develop the chiral perturbation theory of $SU(2)$-isospin. Mesons with heavier quarks (e.g. c, b) are too massive to be treated as approximate Goldstone bosons.

[92] Recall Goldstone bosons must have spin 0.

Now we write down the leading order chiral Lagrangian. The lowest dimension term we can consider is dimension 0, but it is too trivial: $\mathrm{Tr}\Sigma^\dagger\Sigma = 3$. The next highest is dimension 2

$$\mathcal{L}_\chi^{\mathrm{LO}} \;=\; \frac{F^2}{4}\mathrm{Tr}\,\partial_\mu\Sigma\,\partial^\mu\Sigma^\dagger\,. \qquad (8.3.17)$$

The normalization here yields the canonical normalization for the Π^a kinetic energy term, i.e. $\frac{1}{2}\partial_\mu\Pi^a\partial^\mu\Pi^a$, when the $\Sigma(x)$ field is expanded about small $\Pi^a(x)/F$. One can also examine the leading-order strong interactions between Goldstone bosons after expanding out $\Sigma(x)$.

The whole point of using the chiral Lagrangian, indeed of using any effective Lagrangian, is to replace the more complicated full Lagrangian with a tower of simpler operators. As long as we look at low energy processes, the results of calculating Greens functions should be nearly equal. A convenient method for carrying out this matching between full and effective theories makes use of external sources. Therefore, let us introduce the following external sources

$$\begin{aligned} \ell_\mu(x) &= \ell^0_\mu(x) + \ell^a_\mu(x)T^a\,, \quad r_\mu(x) = r^0_\mu(x) + r^a_\mu(x)T^a\,, \\ s(x) &= s^0(x) + s^a(x)T^a\,, \quad p(x) = p^0(x) + p^a(x)T^a\,, \end{aligned} \qquad (8.3.18)$$

corresponding to left-handed, right-handed, scalar, and pseudoscalar sources, with T^a being generators of $SU(3)_F$ in the fundamental representation [93]. These enter the QCD Lagrangian as

[93] E.g. the Gell-Mann matrices λ^a (Note 67).

$$\begin{aligned} \mathcal{L}^{\mathrm{src}}_{\mathrm{QCD}} \;=\; & -\frac{1}{4}F^a_{\mu\nu}F^{a,\mu\nu} \;+\; \bar{q}_L\gamma^\mu(iD_\mu - \ell_\mu)q_L \;+\; \bar{q}_R\gamma^\mu(iD_\mu - r_\mu)q_R \\ & -\; \bar{q}_L(s+ip)q_R \;-\; \bar{q}_R(s-ip)q_L\,. \end{aligned} \qquad (8.3.19)$$

Take care to note the different spaces in which terms operate: e.g. in the second term iD^μ has $SU(3)_c$ indices but is diagonal in flavour while ℓ^μ is diagonal in colour but is generally nondiagonal in flavour. These sources represent probes external to the QCD sector of the standard model. We will see shortly that the insertion of a weak current or an electromagnetic current can be represented by using ℓ_μ or $v_\mu = \ell_\mu + r_\mu$, respectively.

When we say we match the effective theory to the full theory, we mean that we require the generating functionals of the two theories to be equal (in the low energy limit, up to some finite precision). Usually we work with the generating functional $W(\ell_\mu, r_\mu, p, s)$ which gives the connected Green's functions, and is related to the generating functional for all Green's functions via $iW(\ell_\mu, r_\mu, p, s) = \log Z(\ell_\mu, r_\mu, p, s)$. In the full theory,

$$e^{iW(\ell_\mu,r_\mu,p,s)} = \int \mathcal{D}q\,\mathcal{D}\bar{q}\,\mathcal{D}A^a_\mu\, e^{i\int d^4x\,\mathcal{L}^{\mathrm{src}}_{\mathrm{QCD}}(q,\bar{q},A_\mu,\ell_\mu,r_\mu,p,s)}\,, \qquad (8.3.20)$$

while in the effective theory

$$e^{iW(\ell_\mu,r_\mu,p,s)} = \int \mathcal{D}\Sigma\, e^{i\int d^4x\,\mathcal{L}^{\mathrm{src}}_\chi(\Sigma,\ell_\mu,r_\mu,p,s)}\,. \qquad (8.3.21)$$

The next step is to determine how to introduce the external sources into the effective Lagrangian. We do this by observing that the QCD Lagrangian with external sources (8.3.19) can be invariant under *local* $SU(3)_L \times SU(3)_R$ transformations

$$
\begin{aligned}
q_L(x) &\mapsto U_L(x) q_L(x)\,, \quad \bar{q}_L(x) \mapsto \bar{q}_L(x) U_L^\dagger(x) \\
q_R(x) &\mapsto U_R(x) q_R(x)\,, \quad \bar{q}_R(x) \mapsto \bar{q}_R(x) U_R^\dagger(x)
\end{aligned}
\tag{8.3.22}
$$

provided that the external fields transform according to

$$
\begin{aligned}
\ell_\mu(x) &\mapsto U_L(x)\ell_\mu(x)U_L^\dagger(x) + i(\partial_\mu U_L(x))U_L^\dagger(x) \\
r_\mu(x) &\mapsto U_R(x)r_\mu(x)U_R^\dagger(x) + i(\partial_\mu U_R(x))U_R^\dagger(x) \\
(s+ip)(x) &\mapsto U_L(x)(s+ip)(x)U_R^\dagger(x)\,.
\end{aligned}
\tag{8.3.23}
$$

In particular, note that $\ell_\mu(x)$ and $r_\mu(x)$ transform just as $SU(3)_{L,R}$ gauge fields and enter the covariant derivative accordingly.

The chiral Lagrangian can be expanded to include the external sources, where under $SU(3)_L \times SU(3)_R$ gauge transformations (8.3.16) becomes

$$
\Sigma(x) \mapsto U_L(x)\Sigma(x)U_R^\dagger(x)\,.
$$

The external field transformations (8.3.23) restrict the terms which can appear in the gauge invariant chiral Lagrangian to be

$$
\mathcal{L}_\chi^{\mathrm{src}} = \frac{F^2}{4}\mathrm{Tr} D_\mu\Sigma\, D_\mu\Sigma^\dagger + \frac{F^2}{4}\mathrm{Tr}(\chi\Sigma^\dagger + \Sigma\chi^\dagger) \tag{8.3.24}
$$

where

$$
\chi = 2B_0(s+ip)
$$

for some constant B_0 which has mass dimension 1, and

$$
\begin{aligned}
D_\mu\Sigma &= \partial_\mu\Sigma + i\ell_\mu\Sigma - i\Sigma r_\mu \\
D_\mu\Sigma^\dagger &= \partial_\mu\Sigma^\dagger + ir_\mu\Sigma^\dagger - i\Sigma^\dagger\ell_\mu\,.
\end{aligned}
\tag{8.3.25}
$$

This procedure of enforcing $SU(3)_L \times SU(3)_R$ *gauge* invariance is equivalent (in the absence of anomalies) to satisfying Ward identities and can also be modified to deal with anomalous Ward identities.[94]

[94] H Leutwyler. On the foundations of chiral perturbation theory. *Annals Phys.*, 235:165–203, 1994

Just to give an idea of what lies beyond leading order, here is the next-to-leading-order (NLO) term in the chiral Lagrangian:

$$
\begin{aligned}
\mathcal{L}_\chi^{\mathrm{NLO}} &= \alpha_1 \mathrm{Tr}(D_\mu\Sigma\, D^\mu\Sigma^\dagger)^2 + \alpha_2\mathrm{Tr}(D_\mu\Sigma\, D_\nu\Sigma^\dagger)\mathrm{Tr}(D^\mu\Sigma\, D^\nu\Sigma^\dagger) + \alpha_3\mathrm{Tr}(D_\mu\Sigma\, D^\mu\Sigma^\dagger D_\nu\Sigma\, D^\nu\Sigma^\dagger) \\
&+ \alpha_4\mathrm{Tr}(D_\mu\Sigma\, D^\mu\Sigma^\dagger)\mathrm{Tr}(\chi\Sigma^\dagger + \Sigma\chi^\dagger) + \alpha_5\mathrm{Tr}[D_\mu\Sigma\, D^\mu\Sigma^\dagger(\chi\Sigma^\dagger + \Sigma\chi^\dagger)] \\
&+ \alpha_6\left[\mathrm{Tr}(\chi^\dagger\Sigma + \Sigma\chi^\dagger)\right]^2 + \alpha_7\left[\mathrm{Tr}(\chi^\dagger\Sigma - \Sigma\chi^\dagger)\right]^2 + \alpha_8\mathrm{Tr}\left[(\chi^\dagger\Sigma)^2 + (\Sigma^\dagger\chi)^2\right] \\
&+ i\alpha_9\mathrm{Tr}(L_{\mu\nu}D^\mu\Sigma\, D^\nu\Sigma^\dagger + R_{\mu\nu}D^\mu\Sigma\, D^\nu\Sigma^\dagger) + \alpha_{10}\mathrm{Tr}(L_{\mu\nu}\Sigma\, R_{\mu\nu}\Sigma^\dagger)
\end{aligned}
\tag{8.3.26}
$$

where $L_{\mu\nu}$ and $R_{\mu\nu}$ are the field strength tensors for the external left- and right-handed sources

$$
\begin{aligned}
L_{\mu\nu} &= \partial_\mu\ell_\nu - \partial_\nu\ell_\mu + i[\ell_\mu, \ell_\nu] \\
R_{\mu\nu} &= \partial_\mu r_\nu - \partial_\nu r_\mu + i[r_\mu, r_\nu]\,.
\end{aligned}
\tag{8.3.27}
$$

Determining these coefficients α_j (traditionally denoted with ℓ_j or L_j) is an area of active research, both experimentally and using lattice QCD.

8.4 A few uses of χPT

If we set $\ell_\mu = r_\mu = p = 0$ in (8.3.19), then we have

$$\mathcal{L}^{\rm src}_{\rm QCD} \;=\; -\frac{1}{4}F^a_{\mu\nu}F^{a,\mu\nu} \;+\; \bar{q}i\not{D}q \;-\; \bar{q}sq\,. \tag{8.4.1}$$

This is equivalent to the Lagrangian for massive QCD (7.1.1) when the scalar source $s(x)$ is set equal to the mass matrix

$$M = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_d & 0 \\ 0 & 0 & m_s \end{pmatrix}.$$

Fixing $s(x)$ in this way explicitly breaks $SU(3)_L \times SU(3)_R$ down to $SU(3)_F$. We say this is a "soft" symmetry breaking, because the quark masses are small compared to Λ_χ. The effect of this is to make the Nambu-Goldstone bosons become *pseudo*-Nambu-Goldstone bosons with finite mass. The story of light mesons remains approximately true, with small corrections due to the small quark masses.

The masses for the pseudo-Nambu-Goldstone bosons can be read off from the $s(x) = M$ term in (8.3.24):

$$\frac{F^2}{4}{\rm Tr}(\chi\Sigma^\dagger + \Sigma\chi^\dagger) \;=\; \frac{B_0F^2}{2}{\rm Tr}[M(\Sigma+\Sigma^\dagger)]\,.$$

Expanding out $\Sigma = \exp(2i\Pi/F)$ and focusing on the quadratic terms yields

$$\frac{1}{2}(m_u+m_d)B_0(|\pi^+|^2+|\pi^-|^2) + \frac{1}{2}(m_u+m_s)B_0(|K^+|^2+|K^-|^2) + \ldots\,.$$

Thus, leading order chiral perturbation theory predicts the Nambu-Goldstone boson masses depend on the quark masses as

$$\begin{aligned} m^2_{\pi^\pm} &= B_0(m_u+m_d)\,, \quad m^2_\eta \;=\; \frac{B_0}{3}(m_u+m_d+4m_s) \\ m^2_{K^\pm} &= B_0(m_u+m_s)\,, \quad m^2_{K^0} \;=\; B_0(m_d+m_s)\,. \end{aligned} \tag{8.4.2}$$

Using experimental data for the meson masses, then gives a leading order prediction for the ratios of quark masses. Since the "isospin" splittings, i.e. those due to swapping up and down quarks, are small compared to splittings due to swapping either up or down with strange, we can make a decent approximation by assuming $m_u \approx m_d$. Writing the average up/down quark mass as $\hat{m}$, we obtain

$$\frac{\hat{m}}{m_s} \;=\; \frac{m_\pi^2}{2m_K^2 - m_\pi^2} \;\approx\; \frac{1}{26}\,.$$

Also in this approximation we can obtain the Gell-Mann–Okubo relation

$$m^2_\eta \;=\; \frac{1}{3}(4m_K^2 - m_\pi^2)\,.$$

In QCD, the scalar quantity which gets a nonzero vacuum expectation value due to spontaneous symmetry breaking is the *chiral* *condensate*, $\langle 0|\bar{q}q|0\rangle$. Substituting u, d, or s, may lead to different values for the corresponding condensates. Looking at (8.3.20) and (8.4.1), we see that we can obtain the chiral condensates from differentiating the generating functional

$$\left.\frac{\delta W}{\delta s_{ij}(x)}\right|_{s=0} = -\langle 0|\bar{q}_i(x)q_j(x)|0\rangle$$

where $i, j = 1, 2, 3$ are flavour indices.[95] If we are interested in studying the spontaneous symmetry breaking itself, we wish to find the chiral condensate for $s \to 0$. We may also be interested in the physical case where we instead set $s = M$. Since we are able to approximate (8.3.20) by (8.3.21), we find that at leading order, expanding $\Sigma(x) \approx \mathbb{1} + \ldots$,

$$\langle 0|\bar{q}_i(x)q_j(x)|0\rangle = -F^2 B_0\,\delta_{ij}\,. \tag{8.4.3}$$

[95] Instead of differentiating with respect to the matrix elements of $s = s^0 + s^a T^a$, one could differentiate with respect to the coefficients of each term as appropriate. For example, noting that we can use the Gell-Mann matrices $T^a = \lambda^a/2$ (Note 67) to write

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix} = \frac{1}{2}\left(\frac{2}{3}\mathbb{1} + \lambda^3 + \frac{1}{\sqrt{3}}\lambda^8\right)$$

we find

$$\langle 0|\bar{u}(x)u(x)|0\rangle = -\left[\frac{1}{3}\frac{\delta W}{\delta s^0(x)} + \frac{\delta W}{\delta s^3(x)} + \frac{1}{\sqrt{3}}\frac{\delta W}{\delta s^8(x)}\right].$$

We can use the chiral Lagrangian (8.4.1) to determine the leading order hadronic matrix element governing weak decays, e.g. $\pi \to e\bar{\nu}_e$ as in § 6.4. Notice that we can obtain left-handed currents with any flavour structure using

$$j^a_\mu(x) = -\frac{\partial \mathcal{L}^{\text{src}}_{\text{QCD}}}{\partial \ell^{a\mu}(x)} = \bar{q}_L \gamma_\mu T^a q_L = \bar{q}\gamma_\mu \tfrac{1}{2}(1-\gamma^5)T^a\,q\,. \tag{8.4.4}$$

For example the Standard Model weak current (6.4.1)[96] which is indeed external to QCD,

[96] Here we consider only the 3 quark flavours relevant for chiral perturbation theory.

$$J_\mu = \bar{u}\gamma_\mu(V_{ud}d + V_{us}s)$$

can be written

$$J_\mu = V_{ud}(j^1_\mu + i\,j^2_\mu) + V_{us}(j^4_\mu + i\,j^5_\mu)\,.$$

Now to derive the weak current within chiral perturbation theory, we differentiate (8.3.24) with respect to the left-handed sources

$$\begin{aligned} j^{\text{eff},a}_\mu &= -\frac{\partial \mathcal{L}^{\text{src}}_\chi}{\partial \ell^{a\mu}(x)} \\ &= -\frac{iF^2}{2}\,\text{Tr}(T^a\Sigma\,\partial_\mu\Sigma^\dagger) \\ &= -\frac{F}{2}\partial_\mu\Pi^a + \ldots \end{aligned} \tag{8.4.5}$$

having used $\partial_\mu(\Sigma\Sigma^\dagger) = 0$ to rearrange and combine terms, and then, after expanding Σ and keeping only the leading nontrivial term, $\text{Tr}(T^a T^b) = \frac{1}{2}\delta^{ab}$. Finally we conclude the leading order current governing π^- leptonic decay is

$$J_\mu = -V_{ud}\frac{F}{2}\partial_\mu(\Pi^1 + i\Pi^2) = -V_{ud}\,F\,\partial_\mu\pi^-\,, \tag{8.4.6}$$

using (8.3.14). Acting with J_μ on a π^- state with definite momentum p yields the matrix element

$$\langle 0|J_\mu|\pi^-(p)\rangle = -V_{ud}\,F\,p_\mu\,. \tag{8.4.7}$$

Thus we can identify, at leading order in the chiral expansion, the F appearing in the chiral Lagrangian with the pion decay constant introduced in (6.4.3).[97]

[97] I am still checking factors of $\sqrt{2}$ and i to be sure they are consistent throughout the notes.

Bibliography

J Beringer et al. Review of Particle Physics (RPP). *Phys. Rev.*, D86: 010001, 2012.

C G Callan, S R Coleman, J Wess, and B Zumino. Structure of phenomenological Lagrangians. 2. *Phys. Rev.*, 177:2247–2250, 1969. DOI: 10.1103/PhysRev.177.2247.

F Close. *The Infinity Puzzle*. Oxford University Press, 2011. ISBN 978-0199593507.

S R Coleman, J Wess, and B Zumino. Structure of phenomenological Lagrangians. 1. *Phys. Rev.*, 177:2239–2247, 1969.

J F Donoghue, E Golowich, and B R Holstein. *Dynamics of the Standard Model*. Cambridge University Press, 1992. ISBN 0-521-47652-6.

H Georgi. *Weak interactions and modern particle theory*. Benjamin/Cummings, 1984. ISBN 0-8053-3163-8.

H Georgi. Effective field theory. *Ann. Rev. Nucl. Part. Sci.*, 43: 209–252, 1993.

F Halzen and A D Martin. *Quarks and Leptons*. Wiley & Sons, 1984. ISBN 0-471-88741-2.

D B Kaplan. Five lectures on effective field theory. 2005. arXiv:nucl-th/0501023.

H Leutwyler. On the foundations of chiral perturbation theory. *Annals Phys.*, 235:165–203, 1994.

M E Peskin and D V Schroeder. *An Introduction to Quantum Field Theory*. Addison Wesley, 1995. ISBN 0-201-50397-2.

J C Romão and J P Silva. A resource for signs and Feynman diagrams of the Standard Model. *Int. J. Mod. Phys.*, A27:1230025, 2012. arXiv:1209.6213.

R.F. Streater and A.S. Wightman. *PCT, Spin and Statistics, and All That*. Addison-Wesley, 1989. ISBN 0-201-09410-X.

S Weinberg. *The Quantum Theory of Fields, Volume I*. Cambridge University Press, 1995. ISBN 0-521-55001-7.

S Weinberg. *The Quantum Theory of Fields, Volume II*. Cambridge University Press, 1996. ISBN 0-521-55002-5.