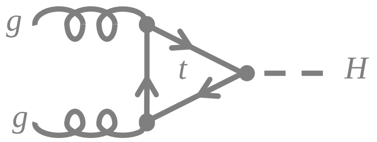
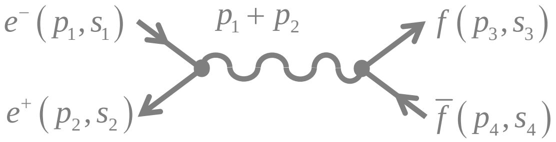
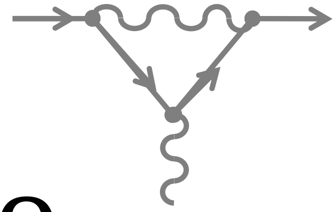
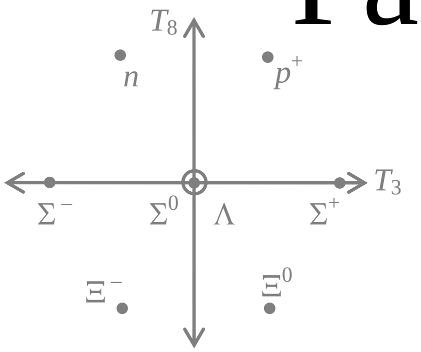
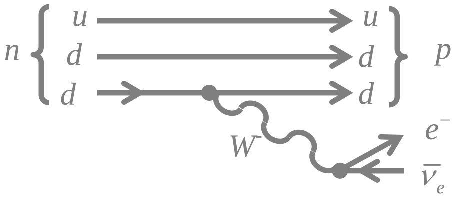




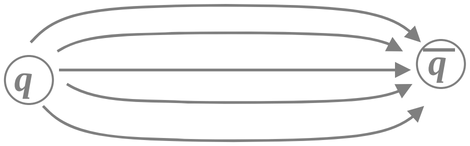


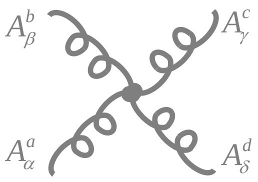
Introduction to Particle Physics

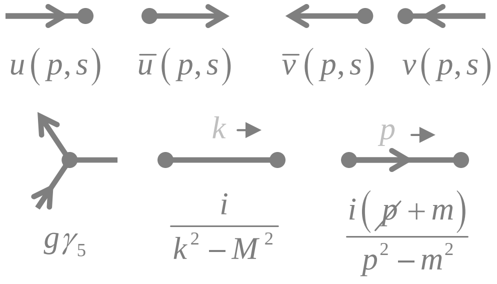




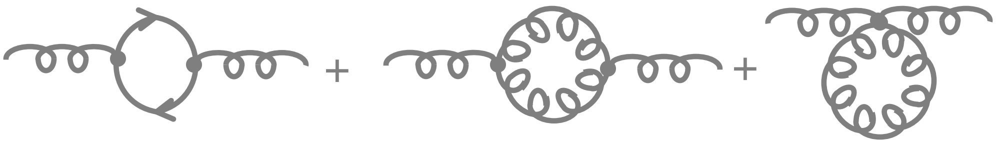
Eric D. Carlson







$$V(\Phi)=\tfrac{1}{2}\lambda\left(\Phi^\dagger\Phi-\tfrac{1}{2}v^2\right)^2$$



Useful Formulas and Identities

Units and Constants

$$c = \hbar = \varepsilon_0 = \mu_0 = 1$$

$$1\text{ s} = 3\times 10^8\text{ m}$$

$$1\text{ kg} = 5.6\times 10^{26}\text{ GeV}$$

$$197\,\text{MeV}\cdot\text{fm} = 1$$

$$1\text{b} = 100\text{ fm}^2$$

$$m_e = 0.51100\text{ MeV}$$

$$m_p = 938.27\text{ MeV}$$

$$\alpha = \frac{e^2}{4\pi} \approx \frac{1}{137.036}$$

$$\alpha_S(M_Z) = \frac{g_S^2}{4\pi} \approx 0.1184$$

$$G_F = 1.166\times 10^{-5}\text{ GeV}^{-2}$$

$$v = 246\text{ GeV}$$

$$\sin^2\theta_W = 0.2312$$

Metric Prefixes

T	10^{12}
G	10^{9}
M	10^{6}
k	10^{3}
m	10^{-3}
μ	10^{-6}
n	10^{-9}
p	10^{-12}
f	10^{-15}

Dirac Matrices

(chiral representation)

$$\gamma^0 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

$$\boldsymbol{\gamma} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}$$

$$\gamma_5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Dirac Properties

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

$$\{\gamma^\mu, \gamma_5\} = 0$$

$$\gamma_5\gamma_5 = 1$$

$$\overline{\Gamma} \equiv \gamma^0\Gamma^\dagger\gamma^0$$

$$\overline{\gamma}^\mu \equiv \gamma^\mu$$

$$\overline{\gamma}_5 = -\gamma_5$$

The Metric

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

More Dirac Identities

$$\gamma^\mu\gamma_\mu = 4$$

$$\gamma^\mu\gamma^\alpha\gamma_\mu = -2\gamma^\alpha$$

$$\gamma^\mu\gamma^\alpha\gamma^\beta\gamma_\mu = 4g^{\alpha\beta}$$

$$\gamma^\mu\gamma^\alpha\gamma^\beta\gamma^\nu\gamma_\mu = -2\gamma^\nu\gamma^\beta\gamma^\alpha$$

☺

Dirac Trace Identities

$$\mathrm{Tr}\left(\gamma^{\mu_1}\gamma^{\mu_2}\cdots\gamma^{\mu_{2N+1}}\right) = \mathrm{Tr}\left(\gamma_5\gamma^{\mu_1}\gamma^{\mu_2}\cdots\gamma^{\mu_{2N+1}}\right) = 0$$

$$\mathrm{Tr}(1) = 4 \qquad \mathrm{Tr}\left(\gamma^\mu\gamma^\nu\right) = 4g^{\mu\nu}$$

$$\mathrm{Tr}\left(\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta\right) = 4\left(g^{\mu\nu}g^{\alpha\beta} + g^{\mu\beta}g^{\nu\alpha} - g^{\mu\alpha}g^{\nu\beta}\right)$$

$$\mathrm{Tr}\left(\gamma_5\right) = \mathrm{Tr}\left(\gamma_5\gamma^\alpha\gamma^\beta\right) = 0$$

$$\mathrm{Tr}\left(\gamma_5\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta\right) = -4i\varepsilon^{\mu\nu\alpha\beta}$$

Cross-Sections and Decay Rates

$$\Gamma = \frac{D}{2M} \qquad \sigma = \frac{D}{4\left|E_2\mathbf{p}_1 - E_1\mathbf{p}_2\right|}$$

$$D(\text{two}) = \frac{p}{16\pi^2 E_{\text{cm}}}\int \left|i\mathcal{M}\right|^2 d\Omega$$

$$D(\text{three}) = \frac{1}{8(2\pi)^5}\int dE_1\, dE_2\, d\Omega_1\, d\phi_{12}\left|i\mathcal{M}\right|^2$$

Spinors and Polarizations

$$\not{p}\,u(p,s) = mu(p,s)$$

$$\not{p}\,v(p,s) = -mv(p,s)$$

$$\sum_s u(p,s)\overline{u}(p,s) = \not{p} + m$$

$$\sum_s v(p,s)\overline{v}(p,s) = \not{p} - m$$

$$q\cdot\varepsilon(q,\lambda) = 0$$

$$\varepsilon^*(q,\lambda)\cdot\varepsilon(q,\tau) = \delta_{\lambda\tau}$$

massless:

$$\sum_\lambda \varepsilon^{*\mu}(q,\lambda)\varepsilon^\nu(q,\lambda) = -g^{\mu\nu}$$

massive:

$$\sum_\lambda \varepsilon^{*\mu}(q,\lambda)\varepsilon^\nu(q,\lambda) = -g^{\mu\nu} + q^\mu q^\nu/M^2$$

Table of Contents

Where can I find …

I. Introduction

Elementary particle physics is the study of the universe at its smallest and most fundamental level. In this course, I will attempt to give you an overview of modern particle physics, including what the fundamental particles are, how they are combined to make other particles, and how we perform probability calculations in particle physics. There will be a particular emphasis on Feynman diagrams as a calculational technique.

A. What is Particle Physics?

Modern particle physics arose as an attempt to combine two of the great concepts of twentieth century physics, namely, quantum mechanics and special relativity. The synthesis of these two proved surprisingly difficult, but when completed, resulted in an exceedingly successful theory. Modern particle physics is a complete theory explaining all known phenomena except gravity and neutrino masses in terms of a handful of particles and parameters.

Einstein's special theory of relativity was an outgrowth of Maxwell's equations, which describe the interactions of electromagnetism. In the special theory of relativity, time is perceived as a fourth dimension on an equal footing with the three conventional space dimensions. Invariance under Lorentz transformations was the main consequence. Another consequence of Einstein's theory is that nothing can travel faster than light. This led to the rejection of action at a distance in favor of local interactions, in which particles interact with each other via intermediate fields. Indeed, in special relativity the introduction of fields is necessary to understand forces. In special relativity, classical particle mechanics gives way to classical field theory.

Quantum mechanics revolutionized our view of the world perhaps more than any other scientific discovery of the twentieth century. In quantum mechanics, momentum is replaced by an operator, and the position of a particle is replaced by a wave function describing the probability amplitude for the particle. This led to surprising results; for example, classically an electron in orbit around a proton can have any energy, but quantum mechanically only certain standing waves were allowed. Quantities which it seemed should be continuous were in fact discrete. Another important consequence of quantum mechanics was wave-particle duality, in which it was understood how an object could behave both like a particle and like a wave.

Quantum field theory, or particle physics, is the union of special relativity and quantum mechanics. If we start with classical field theory, and then quantize it, we end up with quantum field theory. The consequences are many; for example, whereas in conventional electromagnetic theory waves of any strength are possible, in quantum field theory, only discrete values are allowed.

B. Why Particle Physics is Hard

A proper treatment of particle physics would begin with classical particle theory, deriving the equations of motion from a Lagrangian. From there one would proceed to quantize the theory, introducing the Hamiltonian and the momentum operator. Once this is understood, one should then go back and do classical field theory, working once again in the Lagrangian formulation. Then by applying the exact same principles used to quantize the classical particle theory, one would quantize the classical field theory, ending with quantum field theory. Thus, a full course on particle physics should first solve all of the rest of physics, and then do particle physics.

Having done this, the theory still proves incredibly complex. The basic problem with such a theory is that the demands of Lorentz invariance assures that particles of arbitrarily high momentum must be included in the calculations. Quantum perturbation theory leads to sums over an infinite number of intermediate states, and results in infinite contributions. These infinities have to be dealt with before a coherent theory can be built. Theories in which one can consistently eliminate the infinities are called *renormalizable*, and the demands of renormalizability are quite restrictive. For example, in conventional quantum mechanics, describing the Hamiltonian for a single spinless particle in one dimension requires that one give the potential $V(x)$, which has an infinite number of parameters. In quantum field theory, the Hamiltonian for a single spinless field is also described by a potential $V(\phi)$, but this potential has (depending on how you count them) only about three parameters. This restriction is a blessing in disguise, since it restricts us to look at only a relatively small class of theories.

Instead of this approach, we will use a different approach. A seemingly sensible approach is to first build a relativistic theory of one particle, and this is indeed what we will do. There is a reason why, technically, this is not a consistent way to do a relativistic theory. Consider the equation for energy in special relativity, which for small momenta can be expanded as

$$E = \sqrt{m^2c^4 + c^2\mathbf{p}^2} = mc^2 + \frac{\mathbf{p}^2}{2m} - \frac{\left(\mathbf{p}^2\right)^2}{8mc^2} + \cdots. \tag{1.1}$$

If we have a relativistic theory of one particle, we certainly will be including all the terms on the right.

However, in quantum field theory, particle number is not conserved, and particles can appear out of nowhere. In particular, consider the second order perturbation theory energy of a state $|n\rangle$ which looks something like

$$E_n = E_n^0 + \langle n|H_{\text{int}}|n\rangle + \sum_{m\neq n} \frac{\left|\langle m|H_{\text{int}}|n\rangle\right|^2}{E_n^0 - E_m^0} + \cdots. \tag{1.2}$$

States with extra particles in them will have energy greater than those without by an amount of order mc^2, and hence this contribution is vanishingly small in the $c \to \infty$ (non-relativistic) limit. But if we are including the first relativistic correction to E from (1.1), which is proportional to $1/c^2$, how can we justify neglecting this contribution?

This problem is again a blessing in disguise. It tells us that we must build a theory in which particles can be created and destroyed. If we succeed in doing this, then we will be able to describe not only boring processes like electron scattering, but also exciting processes like the creation and destruction of particles never seen before.

However, it means that even the simplest problem secretly have hidden subtleties. The presence of one particle can result in the spontaneous creation, perhaps for only an instant, of an arbitrary number of particles. Indeed, even the “vacuum'” state, which we normally think of as empty, seethes with countless spontaneous particles popping out of nowhere. I have heard the progress in physics described as follows. First there was Newton, and his theory of gravity, and we could exactly solve the two body problem. Then came Einstein and the general theory of relativity, and we could exactly solve the one body problem. Then came particle physics, in which we can't even solve the zero body problem.

Despite these difficulties, we will find that there are problems which, if we can't solve them exactly, at least we can solve them perturbatively, getting successively better approximations to the exact answer. We will focus particularly on the Feynman diagram technique, which efficiently calculates probabilities for various reactions.

C. The Standard Model of Particle Physics

Having told you the road I will travel, let me jump ahead to the conclusion. There is a highly successful theory of particle physics called the *Standard Model* which seems to correctly predict almost all interactions of all known particles. This theory was developed in the 60's and 70’s, and, with only the most minor of modifications, has lasted to the present. In practice, many calculations can be done with little or no precision, but whenever theory and experiment have been compared, the theory has proven accurate. If we could solve the theory exactly, we should be able to use the theory to predict the future behavior of quarks, protons, atoms, molecules, mice, people, and the stock market. In practice, we are lucky if we can explain why protons exist at all (though some other predictions are highly accurate).

The Standard Model consists of seventeen particles interacting via a theory with eighteen parameters. The particles can be categorized in a number of ways, but I have labeled them in the table on the basis of spin, charge (in multiples of the proton charge *e*), color multiplet, and mass (written in energy units, as explained later). Only color multiplet should be an unfamiliar concept to you. The masses (excluding

Particles of the Standard Model

Name	Symbol	Spin	Charge	Color	Mass (GeV)	
Electron	e^-	$\frac{1}{2}$	-1	1	0.000511	leptons
Neutrino 1	ν_1	$\frac{1}{2}$	0	1	$<\sim 10^{-11}$	
Muon	μ^-	$\frac{1}{2}$	-1	1	0.1057	
Neutrino 2	ν_2	$\frac{1}{2}$	0	1	$\sim 10^{-11}$	
Tauon	τ^-	$\frac{1}{2}$	-1	1	1.777	
Neutrino 3	ν_3	$\frac{1}{2}$	0	1	$\sim 10^{-10}$	
Up quark	u	$\frac{1}{2}$	$+\frac{2}{3}$	3	0.0017–0.0031	quarks
Down quark	d	$\frac{1}{2}$	$-\frac{1}{3}$	3	0.0041–0.0057	
Charm quark	c	$\frac{1}{2}$	$+\frac{2}{3}$	3	1.26 ± 0.08	
Strange quark	s	$\frac{1}{2}$	$-\frac{1}{3}$	3	0.100 ± 0.020	
Top quark	t	$\frac{1}{2}$	$+\frac{2}{3}$	3	172.9 ± 1.1	
Bottom quark	b	$\frac{1}{2}$	$-\frac{1}{3}$	3	4.13 ± 0.12	
Photon	γ	1	0	1	0	gauge bosons
Gluon	g	1	0	8	0	
W-boson	$W^\pm$	1	±1	1	80.40 ± 0.02	
Z-boson	Z	1	0	1	91.19	
Higgs boson	H	0	0	1	126	

neutrino masses) can be thought of as twelve of the eighteen parameters, the remaining six describe interactions of the theory. Several of the particles were predicted well before they were discovered, and in some cases, many details of their interactions were approximately or precisely predicted as well. As of this moment (July 2012), all Standard Model particles have been discovered, except for the Higgs boson, for which we have strong but not conclusive evidence at a mass of about 126 GeV.

In addition to the particles listed, there is an *antiparticle* for every particle which has the same mass and spin but opposite electric charge. Some particles are their own anti-particles (γ, g, Z, H) The anti-particles of the quarks and neutrinos are denoted by a bar ($\bar{u}$, $\bar{\nu}_1$, etc.) while the remaining particles are denoted by changing the sign (e^+, μ^+, τ^+, $W^\pm$). The anti-electron is commonly called the *positron*.

The Standard Model is, nonetheless, incomplete. First note that there is no prediction of these eighteen parameters, nor why there are seventeen particles, these are simply arbitrary. But the biggest and most serious flaw is that it does not explain gravity. The gravitational force between two particles is much too weak to actually measure experimentally, and we do not know how to build a renormalizable theory of gravity. The other serious flaw is that we now have experimental evidence that the neutrinos are not, in fact, massless. This is easily fixed in the Standard Model – too easily, in fact. The problem is not that we can't explain it, but we don't know which explanation is likely to be the right one. Even the simplest extensions involving neutrino masses significantly complicate the Standard Model, adding seven or so parameters, but we don't know which one is right. We actually have relatively poor knowledge of these masses and other parameters, so for the most part we will ignore this problem. But it is likely that exciting new discoveries await us in the coming decades in the neutrino sector.

D. High Energy Physics

Why is it that particle physics requires such vast energies? The answer is the de Broglie relation between the momentum of a particle and its wavelength; that is, [1]

$$\lambda p = 2\pi\hbar . \tag{1.3}$$

To see features smaller than some scale, the wavelength of the probing particle must be smaller than the object probed, so we need high momentum, and this momentum means the particle has high energy, since $E^2 = p^2c^2 + m^2c^4$.

How can we possibly take particles and give them such a high momentum? In the early days of particle physics, radioactive sources provided particles of several MeV of energy, and cosmic rays can reach energies much higher than even the largest present-day accelerators. But radioactive decays cannot reach the TeV and higher energies we now expect, and cosmic rays are too sparse to produce sufficient data for most experiments. We therefore require a means of accelerating elementary particles to exceedingly high energy.

Accelerating elementary particles is not performed by tiny catapults, but by electromagnetic fields, which provide a force on an elementary particle given by

---

[1] Note that in particle physics, we essentially never use Planck's constant h, only Planck's reduced constant $\hbar$. This will make even more sense after the discussion about units.

$$\mathbf{F} = q\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right). \tag{1.4}$$

Since the magnetic fields produce only transverse forces, we must use electric fields to accelerate particles to very high velocities. The particles accelerated must be charged, and last long enough to accelerate and collide them before they decay. The only stable charged particles that are elementary or nearly so are the electron and proton and their anti-particles. The muon, with a mean lifetime of a couple of microseconds, lasts just long enough that it could in principle be used, but this has not yet been done. More complex objects *can* be collided; for example, the Large Hadron Collider (LHC) at CERN in Geneva, Switzerland spends some of its time colliding ^{208}Pb ions head-on, but *most* of its work is proton-proton collisions.

The strongest electric fields we can produce have a strength of order 3×10^7 V/m, above which most materials will tend to undergo dielectric breakdown.[1] If we accelerate a particle of charge e with this electric field for a distance of 1 km, the energy will increase by $e\mathbf{E}\cdot\mathbf{s} \sim 30\text{ GeV}$. Indeed, this is likely how future electron-positron colliders will be built. But to reach energies above about 3 TeV would then require an accelerator more than 100 km long, somewhat larger than the largest current accelerators.

There is another way. Magnetic fields provide transverse forces, and using $d\mathbf{p}/dt = \mathbf{F} = q\mathbf{v}\times\mathbf{B}$, it is not hard to show that a steady magnetic field bends a charged particle's orbit into a circle, whose radius R can be shown to be related to the momentum by

$$p = |q|BR = \frac{|q|}{e}\left(\frac{B}{\text{T}}\right)\left(\frac{R}{\text{km}}\right)\left(299.8\text{ GeV}/c\right). \tag{1.5}$$

Realistically, magnets stronger than a few T cannot be created in large quantities, and thus to reach energies of several TeV would require a circular accelerator of radius several km. The LHC, which is not strictly circular, has a circumference of 27 km and a design energy of 7 TeV. As the protons circle the track, small portions of the track are dedicated to accelerating them via electric fields, while the magnetic field is gradually increased (ramped up) so that the orbit remains nearly constant.

The transverse acceleration caused by magnetic fields causes charged particles to radiate energy copiously. For protons, this isn't much of a problem, but for electrons and positrons, because they are so light, they are moving ultra-relativistically, and relativistic effects enhance this effect so much that you have to constantly devote energy to keeping them up to speed. This makes circular accelerators impractical above about 200 GeV of energy. The Large Electron-Positron Collider (LEP), the most powerful such accelerator built to date, operated at energies as high as 104.6 GeV per beam.

Once you have accelerated the particles (protons, electrons, anti-protons, or anti-electrons) you have to decide what to do with them. The simplest thing is to simply dump them into a stationary target, and such *fixed target experiments* are performed by nearly all accelerators. The stationary target can be made of almost any material, and can have a relatively high density of particles, much higher than can ever be created in a beam. This means they will collide at a very high rate, and you get lots of interesting reactions. Not surprisingly, though,

---

[1] Considerably higher electric fields can be achieved with lasers, but the transverse and oscillatory nature of such fields makes them difficult to use for this purpose. Wakefield accelerators are currently being developed that make use of strong lasers, but at present these are not used for particle physics.

colliding particles head-on produces effectively higher energy, just as the collision of two cars moving in opposite directions at 30 mph is much worse than one car hitting a stationary car at 30 mph. It turns out that relativity actually enhances this effect, so that by colliding particles head on you can effectively reach much higher energies. Such *collider* experiments reach the highest energies and provide us with the most useful information. Typical collider experiments involve trillions of particles in each beam, but the vast majority of particles miss each other on any given pass. Fortunately, with circular colliders, the particles will return to the same point and are given another opportunity to collide a fraction of a second later. Counter-rotating beams can circulate for hours before the two beams are sufficiently degraded that they are "dumped" (into a fixed target), the magnets are ramped down again, and the machine is refilled.

E. What Experiments Measure

From a practical standpoint, high energy experiments involve the collision of two particles. It is simply too difficult to combine three beams at a single point and get them to all collide with each other. The collision will be very fast; a pair of relativistic protons are close enough to interact via the short-range strong force, for example, for about a yoctosecond.[1] This interaction produces one or more particles, most of which decay far too quickly to be directly observed. Basically, any particle with a mean lifetime less than about 10^{-12} s cannot be directly observed. This doesn't mean these particles are undetectable, just that their presence must be inferred by their decay products. By the time these primary particles decay, all that remain are photons, electrons and positrons, muons and anti-muons, neutrinos and anti-neutrinos, and a handful of strongly interacting particles called hadrons, which includes protons and neutrons and their anti-particles.

The innermost layer of the detector contains *trackers*, which precisely measure the position of passing charged particles, allowing them to reconstruct the location of the initial interaction and the path taken by the particles as they leave this region. This inner region usually has a strong magnetic field that bends the path of any charged particle, allowing one to measure the sign of the charge of the particle (from the direction of the curvature) and its approximate momentum (from its radius of curvature, assuming we know the charge, usually $\pm e$).

Next come the *calorimeters* which measure the deposition of energy as the particles pass through the detectors. It turns out that electrons, positrons, and photons lose all their energy almost immediately, producing a so-called electromagnetic shower. Next go the hadrons, which can penetrate many meters of the detector before they are stopped. Finally there are the muons, which often make it all the way out of the detector, slowly losing energy all the way. Also, let us not forget the neutrinos, which always escape the detector without leaving the slightest trace.

Analyzing and interpreting the data from such detectors is an art form, and experimentalists perform wonders on a regular basis. However, I will avoid all the details. In this class, whenever we talk about detectors, I will use idealized detectors, in which all particles (except neutrinos) are perfectly identified, and their momentum and energy is measured exactly.

So what exactly is done with all the data that is collected? Typically, the experiment attempts to measure the cross-section and decay rates of the resulting particles. These will require some explanation.

---

[1] 10^{-24} s. Yes, it's a real word. Look it up if you don't believe me. I've never actually seen it used

The *cross-section* tells you how many particles are being produced at any given moment. "Elementary" particles, even composite ones like protons, are far too small to actually aim at specifically. Instead, large numbers of particles are launched at a large number of targets, and we simply hope that at least some of the time, they will hit each other. Imagine yourself firing projectiles at a target of area σ_X , where the *X* denotes a particular goal you have in mind, like hitting the portion of the target where you score (in the analogy), or producing an *X* particle (in particle physics). If your projectile is randomly distributed over a region of size *A*, then the probability that you hit the target is

$$P_X = \frac{\sigma_X}{A} . \tag{1.6}$$

Of course, (1.6) was derived classically, but we may take it as the definition of the *partial cross-section* σ_X , even when doing quantum mechanics. Cross-section has units of area. But now, fire at the target with N_1 projectiles spread over area *A*, and let there be N_2 targets spread over the same area. Then the number of targets that will be hit in a single pass will be $N_1 N_2 \sigma_X / A$. Loop them around, so you get another chance to hit them again, so that they get another chance with frequency *f*. Finally, make *n* bunches of projectiles and *n* bunches of targets, all going around in opposite directions. Then the rate at which some particular outcome *X* will occur, such as the production of *X*-particles, will be

$$\gamma_X = L\sigma_X , \tag{1.7}$$

where the *luminosity L* is

$$L = f \frac{nN_1N_2}{A} . \tag{1.8}$$

If we integrate (1.7) over time we can find the total number of *X*-particles produced, which is

$$N_X = \sigma_X \int L dt , \tag{1.9}$$

where $\int L dt$ is called the *integrated luminosity*.

Equation (1.6) has another form that will be useful when we are calculating cross-sections. Suppose the target is randomly distributed in a volume $V = A\ell$ as viewed by us in the lab frame. Multiply the numerator and denominator of (1.6) by ℓ to yield $P_X = \ell \sigma_X / V$. Because this volume is moving at the target's speed, and the projectiles also are moving, the amount of time *T* it takes for the projectile to move through the volume *V* will be $T = \ell / |\Delta \mathbf{v}|$. Substituting in, this yields the formula

$$P_X = \frac{T|\Delta \mathbf{v}|\sigma_X}{V} . \tag{1.10}$$

It is important to note that $|\Delta \mathbf{v}|$ is simply the magnitude of the difference of the velocities. Hence two photons heading straight towards each other would have $|\Delta \mathbf{v}| = 2c$.

In general, there will be many possible final outcomes when we collide two particles. The *total cross-section* is simply defined as

$$\sigma_{\text{tot}} = \sigma_X + \sigma_Y + \cdots .$$

It is also common that we will want to measure the number of particles scattering into a particular angle, in which case we will want the *differential cross-section* $d\sigma_X / d\Omega$, a notation that will make a little more sense later in this chapter. Cross-section and differential cross-section have units of area.

Now that we have produced some particle, say *X*, it will generally decay. For example, suppose you have a certain number N_X of particles of type *X*. One of the ways it might decay is to some combination of particles *Y*. The number of decays per unit time can be written as

$$\gamma_Y = \Gamma\left(X \to Y\right) N_X , \tag{1.11}$$

where γ_Y is the number of events that occur per second, and $\Gamma\left(X \to Y\right)$ is the *partial decay rate*. Sometimes we will want more detailed information, such as the number going into a specific direction, or *differential decay rate,* which we will write as $d\Gamma\left(X \to Y\right)/d\Omega$. Sometimes there will be multiple decay modes for the *X*, which can be summed to give the *total decay rate*

$$\Gamma_{\text{tot}}\left(X\right) = \Gamma\left(X \to Y\right) + \Gamma\left(X \to Z\right) + \cdots .$$

Of course, any decay of the *X* will decrease the number of *X*-particles by one. If at $t = 0$ we start with N_{X0} *X*-particles, and wait a time *t*, the number remaining will be

$$N_X\left(t\right) = N_{X0} \exp\left(-\Gamma_{\text{tot}} t\right). \tag{1.12}$$

It follows from (1.11) that the number of *Y*'s produced will be

$$\gamma_Y\left(t\right) = N_{X0} \Gamma\left(X \to Y\right) \exp\left(-\Gamma_{\text{tot}} t\right). \tag{1.13}$$

Note in (1.13) that it is the *total* decay rate of *X* that governs the exponential. The *mean lifetime* τ can then be shown to be $\tau = \Gamma^{-1}$. If we wait long enough, the total number of *Y*'s produced from *X*-decay can be found by integrating (1.13) to infinity, and we find

$$N_Y = N_{X0} \frac{\Gamma\left(X \to Y\right)}{\Gamma_{\text{tot}}\left(X\right)} = N_{X0} BR\left(X \to Y\right), \tag{1.14}$$

where the *branching ratio* is defined by

$$BR\left(X \to Y\right) \equiv \frac{\Gamma\left(X \to Y\right)}{\Gamma_{\text{tot}}\left(X\right)} . \tag{1.15}$$

The branching ratio $BR\left(X \to Y\right)$ is simply the fraction of the initial particles *X* that ultimately decay to *Y*. It is often far easier to measure than the decay rates themselves, since many decays are so fast. The total number of events where you produce an *X* particle and it subsequently decays to a *Y* particle will be

$$N\left(X \to Y\right) \equiv \sigma\left(X\right) BR\left(X \to Y\right) \int L dt . \tag{1.16}$$

F. Units in Particle Physics

Metric Prefixes	
T	10^{12}
G	10^{9}
M	10^{6}
k	10^{3}
m	10^{-3}
μ	10^{-6}
n	10^{-9}
p	10^{-12}
f	10^{-15}

Every branch of physics has its preferred system of units, and particle physics is no exception. Because high energies are essentially always achieved by accelerating charged particles (charge is e or a simple multiple thereof) across potential differences (measured in volts), the preferred unit of energy is the *electron volt*, in conventional units,

$$1 \text{ eV} = 1.602 \times 10^{-19} \text{ J}. \tag{1.17}$$

Particle physics is well beyond the eV stage, so more commonly we use the keV, MeV, GeV and TeV.

In particle physics, we are almost always using relativistic particles, and so it is useful to convert time into distances (for example) by letting t be a shorthand notation for ct, so that both time and distances are measured in meters. This effectively means that we are setting $c = 1$. Similarly, when we introduce quantum mechanics, the reduced Planck constant $\hbar$ appears so many places we can make similar choices and set it equal to one as well. When we get into electromagnetism, rather than working with Coulombs, we will work with units such that the permittivity of free space ε_0 is one too. Because $\mu_0 \varepsilon_0 c^2 = 1$, we therefore have

$$c = \hbar = \varepsilon_0 = \mu_0 = 1. \tag{1.18}$$

With these choices, all combinations of length, time, mass, and charge can be converted into a single unit, usually chosen to be the eV or GeV. In terms of GeV, these fundamental scales have units of

$$[\ell] \sim [t] \sim \text{GeV}^{-1}, \quad [m] \sim \text{GeV}, \quad [Q] \sim 1. \tag{1.19}$$

Note that in this system of units, charge is dimensionless, so the fundamental charge e is a pure number (see problem 1.8).

Since $c = \hbar = \varepsilon_0 = 1$, combinations of these can be combined to convert conventional units into GeV (or vice versa), and we have

$$1 = 3.00 \times 10^{8} \text{ m/s} = 0.197 \text{ GeV} \cdot \text{fm} = 5.61 \times 10^{26} \text{ GeV/kg},$$

where $1 \text{ fm} = 10^{-15} \text{ m}$. This allows us to convert almost any quantity into units of powers of GeV; for example, my weight is about

$$\begin{aligned} F &= mg \\ &= (90 \text{ kg})(9.8 \text{ m/s}^2)(5.61 \times 10^{26} \text{ GeV/kg})(0.197 \times 10^{-15} \text{ m} \cdot \text{GeV})(3.00 \times 10^{8} \text{ m/s})^{-2} \\ &= 0.00108 \text{ GeV}^2 . \end{aligned}$$

By comparison, the force between a pair of quarks is about 0.01 GeV^2!

When discussing cross-sections, it is common to use a more metric system-like unit of area. Clearly the square meter is too large for practical use, so instead, the *barn* is used, defined by

$$1 \text{ b} \equiv 100 \text{ fm}^2 . \tag{1.20}$$

The barn is, in most modern contexts, a very large unit of area, so more commonly used are the μb, nb, pb and fb. Cross sections are commonly given in b, and luminosity and integrated luminosity in $\mathrm{b}^{-1}\mathrm{s}^{-1}$ and b^{-1} respectively.

In electromagnetism, it is common to work with materials where there is a dielectric constant, so that $\varepsilon \neq \varepsilon_0$. This is justified because when we look at some object macroscopically, it is too much trouble to think about the shifts of individual charges. But this is particle physics, and in particle physics our particles are much smaller than atoms, and we certainly are interested in individual charges. Effectively, in particle physics, you treat all the particles individually, rather than collectively, and hence every particle is in vacuum, so $\varepsilon = \varepsilon_0 = 1$ and $\mu = \mu_0 = 1$, and therefore $\mathbf{E} = \mathbf{D}$ and $\mathbf{B} = \mathbf{H}$. This will help us understand Maxwell's equations in relativistic notation in the next chapter. But before we get to that, we'd better review some math that we will be using.

G. Math in Three or Four Dimensions

In physics we often perform computations in three dimensions; in particle physics, we often perform them in four. For three dimensions, we will work with vectors, which we will denote by placing them in bold font **x** within these notes, or by placing a vector symbol $\vec{x}$ when writing on the board. When we refer to individual components, we will normally denote those components by putting Latin superscripts on the *x*'s, so that the components of **x** are x^i for $i = 1, 2, 3$, or $\mathbf{x} = \left(x^1, x^2, x^3\right)$. This will be done rarely enough that it will not often be used. Although I may occasionally stumble and call these directions *x*, *y*, and *z*, this notation is rarely used.

We will follow in these notes the *Einstein summation convention*, which means that for any term that has an index that is repeated, it is implied that it is summed over, so for example we could write the dot product between two three-dimensional vectors as

$$\mathbf{x} \cdot \mathbf{y} = x^i y^i \equiv x^1 y^1 + x^2 y^2 + x^3 y^3 . \tag{1.21}$$

When we have an index repeated and don't want to sum it, we will have to explicitly say so, but this will almost never happen. Keep in mind that any such repeated index is a "dummy index," and we can change it at will. The dot product of a vector with itself is denoted $\mathbf{x}^2$, and its square root is the magnitude $|\mathbf{x}|$, which will sometimes confusingly be called *x*, so that

$$x = |\mathbf{x}| \equiv \sqrt{\mathbf{x}^2} = \sqrt{\left(x^1\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2} .$$

It is common that we will prefer to work in spherical coordinates in three dimensions. Our conventions will be those of physics, in which the components of some vector **k** can be written in terms of its magnitude *k* and two angles θ and ϕ as

$$\mathbf{k} = \left(k \sin\theta \cos\phi , k \sin\theta \sin\phi , k \cos\theta\right). \tag{1.22}$$

Note that in math classes, the definition of θ and ϕ are opposite that of physics. It is also common that we will need to perform three-dimensional integrals, which generically will be written as $\int f(\mathbf{k})\, d^3\mathbf{k}$. When performed in spherical coordinates, it takes the form

$$\int f(\mathbf{k})\, d^3\mathbf{k} = \int_0^\infty k^2 dk \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi\ f(\mathbf{k}) = \int_0^\infty k^2 dk \int_{-1}^1 d(\cos\theta) \int_0^{2\pi} d\phi\ f(\mathbf{k}). \qquad (1.23)$$

The latter form is generally easier to use than the former. We will also denote the *differential solid angle* $d\Omega$, defined as

$$d\Omega \equiv d(\cos\theta)\, d\phi. \qquad (1.24)$$

It has units of radians squared, also called *steradians*, which in a sense in dimensionless, since radians are essentially dimensionless. We could write (1.23) in the simpler form

$$\int f(\mathbf{k})\, d^3\mathbf{k} = \int_0^\infty k^2 dk \int d\Omega f(\mathbf{k}). \qquad (1.25)$$

Note that if $f(\mathbf{k})$ is independent of direction, we can use the fact that $\int d\Omega = 4\pi$ to simplify expressions like (1.25).

We will often be working in four dimensions. The additional dimension (time) we will call the *zeroth* dimension, so x^0 will be the time component. A generic four-dimensional vector will be denoted *x*, with no adornments on it, which is a bit confusing, but is standard in this field. The individual components of a vector *x* will be denoted by adding Greek indices, so that x^μ is one of the components, so that

$$x = x^\mu = (x^0, x^1, x^2, x^3) = (x^0, \mathbf{x}).$$

Just like in three dimensions, any index that is repeated will be summed over, but in this case there will be four terms. Hence, for example,

$$g_{\mu\nu} x^\mu y^\nu \equiv \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} g_{\mu\nu} x^\mu y^\nu .$$

As we will discover later, indices that are summed over in four dimensions are almost invariably paired up with down.

H. Taylor Series and Approximations

In mathematics, we usually try to do computations exactly, but in physics this is rarely easy or advisable. After all, theory is ultimately compared with experiment, and experiment always has errors. Sometimes approximations are actually superior to exact formulas in physics, providing insights that are obscured by the exact formulas.

Probably the most powerful tool for approximations is the Taylor Series. Let $f(x)$ be an arbitrary function for which we know the value and derivatives at some point $x = a$, then we can approximate the value at arbitrary x by the series

$$f(x) = f(a) + (x-a) f'(a) + \tfrac{1}{2}(x-a)^2 f''(a) + \tfrac{1}{6}(x-a)^3 f^{(3)}(a) + \cdots, \qquad (1.26)$$

where $f'(a)$, $f''(a)$ and $f^{(3)}(a)$, mean the first three derivatives, and so on. The numerical coefficient is $1/n!$. Taylor expansions are almost always done around $x = 0$. It should be noted that Taylor expansions don't always work; for example, the function $\exp\left(-1/x^2\right)$ and all of its derivatives vanish at the origin, and yet the function is not zero. Very useful is the binomial expansion

$$(1+\varepsilon)^n = 1 + n\varepsilon + \tfrac{1}{2}n(n-1)\varepsilon^2 + \cdots . \tag{1.27}$$

This expansion converges if $|\varepsilon| < 1$, but we will generally only keep a couple of terms. For example, in section 2D we will encounter the formula $E^2 = \mathbf{p}^2 + m^2$. If the momentum is small compared to the mass, then we can approximate

$$E = \sqrt{\mathbf{p}^2 + m^2} = m\left[1 + \left(\frac{\mathbf{p}^2}{m^2}\right)\right]^{1/2} = m\left[1 + \frac{1}{2}\left(\frac{\mathbf{p}^2}{m^2}\right) - \frac{1}{8}\left(\frac{\mathbf{p}^2}{m^2}\right)^2 + \cdots\right] = m + \frac{\mathbf{p}^2}{2m} - \frac{\left(\mathbf{p}^2\right)^2}{8m^3} + \cdots .$$

We will not use this formula; much more common will be the ultrarelativistic limit $|\mathbf{p}| \gg m$.

Exponential and trigonometric functions can similarly be expanded in Taylor Series:

$$\exp(x) \equiv e^x \equiv \sum_{n=0}^{\infty} \frac{1}{n!} x^n = 1 + x + \tfrac{1}{2}x^2 + \tfrac{1}{6}x^3 + \cdots , \tag{1.28a}$$

$$\sin\theta = \theta - \tfrac{1}{6}\theta^3 + \tfrac{1}{120}\theta^5 - \cdots , \tag{1.28b}$$

$$\cos\theta = 1 - \tfrac{1}{2}\theta^2 + \tfrac{1}{24}\theta^4 - \cdots , \tag{1.28c}$$

These converge for all values of x or θ. Indeed, the Taylor Series can be understood if x is anything you know how to multiply and add, such as a square matrix, and we will commonly use eq. (1.28a) when x is a matrix. The *hyperbolic functions* are defined in terms of the exponential:

$$\cosh x \equiv \tfrac{1}{2}\left(e^x + e^{-x}\right) = 1 + \tfrac{1}{2}x^2 + \tfrac{1}{24}x^4 + \cdots , \tag{1.29a}$$

$$\sinh x \equiv \tfrac{1}{2}\left(e^x - e^{-x}\right) = x + \tfrac{1}{6}x^3 + \tfrac{1}{120}x^5 + \cdots , \tag{1.29b}$$

$$\tanh x \equiv \frac{\sinh x}{\cosh x} . \tag{1.29c}$$

By examining the Taylor expansion, it isn't hard to prove Euler's theorem,

$$e^{i\theta} = \cos\theta + i\sin\theta . \tag{1.30}$$

From eqs. (1.28) and (1.29) it is pretty easy to prove a lot of identities; for example

$$\cos\theta = \cosh(i\theta) = \frac{1}{2}\left(e^{i\theta} + e^{-i\theta}\right), \tag{1.31a}$$

$$\sin\theta = \frac{1}{i}\sinh(i\theta) = \frac{1}{2i}\left(e^{i\theta} - e^{-i\theta}\right), \tag{1.31b}$$

$$\cos^2\theta + \sin^2\theta = 1 , \tag{1.31c}$$

$$\cosh^2\phi - \sinh^2\phi = 1 . \tag{1.31d}$$

I. Matrices, and Other Things

We will be working a fair amount with matrices. A matrix A is a set of real or complex numbers arranged in a rectangle, with components $A^i{}_j$, where i designates the row and j designates the column (we will not always be consistent about where we put the indices, but row always comes first). Two matrices A and B can be multiplied only if the number of columns of A matches the number of rows of B, in which case the components of AB are given by

$$\left(AB\right)^i{}_j \equiv A^i{}_k B^k{}_j \,. \tag{1.32}$$

Unlike real or complex numbers, matrices do not always commute; the order in which they are multiplied makes a difference. The same is true of operators in quantum mechanics. For any two matrices A and B, it is useful to define the *commutator* $\left[A,B\right]$ and *anti-commutator* $\left\{A,B\right\}$ as

$$\left[A,B\right] \equiv AB - BA \,, \tag{1.33a}$$

$$\left\{A,B\right\} \equiv AB + BA \,. \tag{1.33b}$$

These definitions are commonly used when we wish to put two matrices or operators in the opposite order, since $AB = \left[A,B\right] + BA = \left\{A,B\right\} - BA$.

One of the more common sets of matrices we will be dealing with is the *Pauli matrices*, given by

$$\sigma^1 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.34}$$

The commutator and anti-commutator of two Pauli matrices are given by

$$\left[\sigma^i, \sigma^j\right] = 2i\varepsilon^{ijk}\sigma^k \,, \tag{1.35a}$$

$$\left\{\sigma^i, \sigma^j\right\} = 2\delta^{ij} \,, \tag{1.35b}$$

where δ^{ij} is the *Kronecker-delta function,* defined by

$$\delta^{ij} \equiv \begin{cases} 1 & \text{if } i = j \,, \\ 0 & \text{if } i \neq j, \end{cases} \tag{1.36}$$

and ε^{ijk} is the three-dimensional *Levi-Civita symbol,* given by

$$\varepsilon^{ijk} \equiv \begin{cases} 1 & \text{if } i,j,k \text{ is an even permutation of } 1,2,3 \,, \\ -1 & \text{if } i,j,k \text{ is an odd permutation of } 1,2,3 \,, \\ 0 & \text{otherwise.} \end{cases} \tag{1.37}$$

From eqs. (1.35), we see that

$$\sigma^i\sigma^j = \tfrac{1}{2}\left\{\sigma^i,\sigma^j\right\} + \tfrac{1}{2}\left[\sigma^i,\sigma^j\right] = \delta^{ij} + i\varepsilon^{ijk}\sigma^k \,. \tag{1.38}$$

Eq. (1.35b) also has an implied 2×2 identity matrix on the right-hand side.

Let's do a simple but useful example of how to work with the Pauli matrices. Let $\hat{\mathbf{n}}$ be any vector of unit length in three-dimensional space, then it is easy to see that

$$\left(\boldsymbol{\sigma}\cdot\hat{\mathbf{n}}\right)^2 = \tfrac{1}{2}\left\{\boldsymbol{\sigma}\cdot\hat{\mathbf{n}},\boldsymbol{\sigma}\cdot\hat{\mathbf{n}}\right\} = \tfrac{1}{2}\left\{\sigma^i\hat{n}^i,\sigma^j\hat{n}^j\right\} = \tfrac{1}{2}\hat{n}^i\hat{n}^j\left\{\sigma^i,\sigma^j\right\} = \hat{n}^i\hat{n}^j\delta^{ij} = \hat{n}^i\hat{n}^i = \hat{\mathbf{n}}^2 = 1 .$$

From this identity, it is not hard to show by looking at the Taylor series expansion that (problem 1.10):

$$\exp\left[-\tfrac{1}{2}i\theta\boldsymbol{\sigma}\cdot\hat{\mathbf{n}}\right] = \cos\left(\tfrac{1}{2}\theta\right) - i\boldsymbol{\sigma}\cdot\hat{\mathbf{n}}\sin\left(\tfrac{1}{2}\theta\right), \tag{1.39a}$$

$$\exp\left[\tfrac{1}{2}\phi\boldsymbol{\sigma}\cdot\hat{\mathbf{n}}\right] = \cosh\left(\tfrac{1}{2}\phi\right) + \boldsymbol{\sigma}\cdot\hat{\mathbf{n}}\sinh\left(\tfrac{1}{2}\phi\right). \tag{1.39b}$$

These formulas will be useful in chapter 3, when we attempt to rotate and Lorentz boost solutions to the Dirac equation.

Some other things we will do with matrices is take their determinants and their trace. Each of these can only be done for square matrices. If we denote the row *i* and column *j* component of a matrix *A* by $A^i{}_j$, then the trace is simply given by

$$\mathrm{Tr}\left(A\right) \equiv A^i{}_i . \tag{1.40}$$

The trace of a sum of two matrices is given by $\mathrm{Tr}\left(A+B\right) = \mathrm{Tr}\left(A\right) + \mathrm{Tr}\left(B\right)$. There is also is a nice identity involving the trace of products:

$$\mathrm{Tr}\left(AB\right) = \left(AB\right)^i{}_i = A^i{}_j B^j{}_i = B^j{}_i A^i{}_j = \left(BA\right)^j{}_j = \mathrm{Tr}\left(BA\right). \tag{1.41}$$

This does not mean that order makes no difference; rather, that objects in a trace can be cyclically permuted, so that, for example, $\mathrm{Tr}\left(ABC\right) = \mathrm{Tr}\left(CAB\right) = \mathrm{Tr}\left(BCA\right)$, but you can't permute them arbitrarily.

For determinants, there is no simple formula for sums, but for products

$$\det\left(AB\right) = \det\left(A\right)\det\left(B\right). \tag{1.42}$$

This rule generalizes to the product of any number of matrices. Clearly, order doesn't matter, since on the right hand side, all factors are real numbers and hence commute.

J. The Dirac Delta Function

We will occasionally encounter the *Heaviside function,* also known as the step-function, defined by

$$\theta\left(x\right) \equiv \begin{cases} 0 & x < 0 , \\ 1 & x > 0 . \end{cases}$$

More commonly used is its derivative, the *Dirac-delta function*, written $\delta\left(x\right)$, having a value of zero everywhere except $x = 0$, where it is infinite, but in such a way that it has integral 1. The product of the delta-function and any function can be easily integrated, in that

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)\,dx = f(a). \tag{1.43}$$

A three- or four-dimensional Dirac delta-function can be defined by multiplying the one-dimensional function, so that

$$\delta^3(\mathbf{x}) \equiv \delta(x^1)\delta(x^2)\delta(x^3), \quad \delta^4(x) \equiv \delta(x^0)\delta(x^1)\delta(x^2)\delta(x^3).$$

Multi-dimensional integrals are similarly trivial; for example,

$$\int f(\mathbf{x})\delta^3(\mathbf{x}-\mathbf{a})\,d^3\mathbf{x} = f(\mathbf{a}).$$

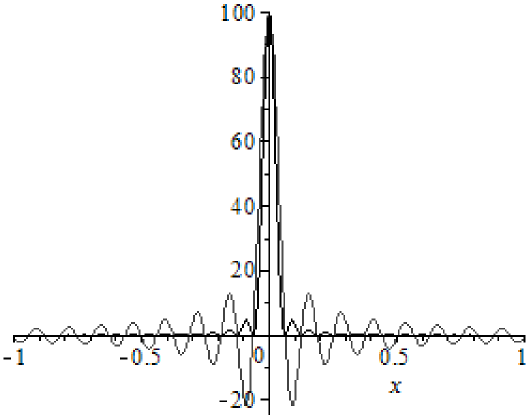


Figure 1-1: The two functions given by Eqs. (1.44a) (gray) and (1.44b) (black) for $L = 100$. They each have maximum height 100 and total integral 2π.

The Dirac delta-function can be thought of as the limit of any tall, narrow function with integral one. For example, consider the two expressions

$$\frac{2}{k}\sin\left(\tfrac{1}{2}kL\right), \tag{1.44a}$$

$$\frac{4}{k^2L}\sin^2\left(\tfrac{1}{2}kL\right). \tag{1.44b}$$

These functions are sketched in Fig. 1-1. For finite *L*, these are two well-behaved functions of *k*. But if we take *L* to be very large, each of these becomes increasingly tall and narrow, having a maximum height of *L* and a width of order $1/L$. With some difficulty one can show that they each have integral 2π. Hence we conclude

$$\lim_{L\to\infty}\frac{2}{k}\sin\left(\tfrac{1}{2}kL\right) = \lim_{L\to\infty}\frac{4}{k^2L}\sin^2\left(\tfrac{1}{2}kL\right) = 2\pi\delta(k), \tag{1.45}$$

We will sometimes have functions inside the Dirac-delta functions. Suppose we are faced with an integral of the form $\int_{-\infty}^{\infty} f(x)\delta\left[g(x)\right]dx$. To make things simple, let's assume that $g(x)$ vanishes only at one place, say $g(x_0)=0$. Then the details of the function *g* away from x_0 are irrelevant, since the delta-function vanishes everywhere else. In the neighborhood of x_0, we Taylor expand $g(x) \approx (x-x_0)g'(x_0)$, so that we have

$$\int_{-\infty}^{\infty} f(x)\delta\left[g(x)\right]dx = \int_{-\infty}^{\infty} f(x)\delta\left[(x-x_0)g'(x_0)\right]dx.$$

We now define $y = (x-x_0)g'(x_0)$, and, paying careful attention to how our limits will change if $g'(x_0) < 0$, we then have

$$\int_{-\infty}^{\infty} f(x)\delta\left[g(x)\right]dx = \frac{1}{\left|g'(x_0)\right|}\int_{-\infty}^{\infty} f\left[x_0 + \frac{y}{g'(x_0)}\right]\delta(y)\,dy,$$

$$\int_{-\infty}^{\infty} f(x)\delta\left[g(x)\right]dx = \frac{f(x_0)}{\left|g'(x_0)\right|}. \tag{1.46}$$

A non-integral version of (1.43) would be

$$\delta\left[g(x)\right] = \frac{1}{\left|g'(x_0)\right|}\delta(x-x_0). \tag{1.47}$$

In particular, if we have a positive constant multiplying everything inside a delta-function, we can bring it out by dividing it, so $\delta\left[cg(x)\right] = \delta\left[g(x)\right]/c$.

It will occasionally be necessary for us to do dimensional analysis on delta functions. It is not hard to convince yourself that $\int dx$ has the same dimensions as *x*, and therefore from Eq. (1.41) we see that $\delta(x)$ must have the same dimensions as $1/x$. Hence a delta-functions dimensionality is always the reciprocal of that of its argument. For three dimensions, $\delta^3(\mathbf{x})$ has the same dimensions as $1/x^3$.

<u>Problems for Chapter 1</u>

1. Find the sum of the charge of all fundamental spin ½ particles in the Standard Model. Multiply by 3 for the quarks, representing three colors. The simplicity of the result is *not* a coincidence.

2. It is possible that the universe has small extra dimensions. If so, we should be able to detect them if we use particles with wavelength shorter than the scale *L* of the extra dimension. Having performed experiments with 4 TeV protons without seeing hints of extra dimensions, estimate the maximum size this extra dimension might be in meters.

3. Look up some of the parameters for the Large Hadron Collider (LHC) for p^+p^+ collisions (2012) in the pdg booklet. Calculate the expected luminosity based on the size of the bunches, the number of particles in the bunches, and the circumference (the protons are moving effectively at the speed of light). Compare with the listed luminosity.

4. At an energy $\sqrt{s} = 8\text{ TeV}$ (next chapter we'll learn what this means), look up the Higgs production cross-section at the LHC for a mass of $M_H = 126\text{ GeV}$. You can find it, for example, at **http://arxiv.org/abs/1012.0530** (you can use the cross section $\sigma(gg \to H)$). Multiply by the luminosity listed in the pdg booklet for the LHC (2012). When operating at this energy, what is the average time between Higgs particles produced?

5. By July 4, 2012, approximately 5 fb^{-1} of integrated lumonisity at $\sqrt{s} = 8\text{ TeV}$ had been analyzed by the CMS and ATLAS detectors. How many Higgs particles were produced? The main signal seen was from the process $H \to \gamma\gamma$. The branching ratio for this decay is about 0.25%. How many $H \to \gamma\gamma$ events should have been seen by each experiment?

6. Look up the decay rate Γ for the *Z*-particle in GeV (summary tables, gauge bosons). What would be its mean lifetime $\tau = \Gamma^{-1}$ in s? Look up the mean lifetime τ for a muon (summary tables, leptons). What would be its decay rate $\Gamma = \tau^{-1}$ in GeV?

7. Look up the total lifetime of the π^+ and K^+ mesons (summary tables, mesons). What would the rate Γ in GeV be? Then look up the branching ratios in each case to decay to $\mu^+\nu_\mu$. Find the partial rates $\Gamma\left(\pi^+ \to \mu^+\nu_\mu\right)$ and $\Gamma\left(K^+ \to \mu^+\nu_\mu\right)$ in each case, and their ratio.

8. Find a dimensionless combination of *e*, ε_0, $\hbar$ and *c*. Then, setting $\varepsilon_0 = \hbar = c = 1$, find the dimensionless value of the fundamental charge *e*.

9. By using a suitable combination of $\hbar$ and *c*, write Newton's constant in the form $G_N = M_P^n$, where M_P has units of mass or energy, and is called the *Planck mass*. Determine the integer *n* and the value of M_P in GeV. If a proton collider were operating at $E = M_P$ and used $B = 10\text{ T}$ magnets, what would be its radius in light-years?

10. What is the total amount of energy in one of the two LHC beams, in ordinary units, like J? In our units, because the protons are ultrarelativistic, $p \approx E$. What is the momentum of one beam in $\text{kg}\cdot\text{m/s}$? If all the energy and momentum of the beam were dumped into your body, would you be thrown across the room? Would you die?

11. Prove eqs. (1.39a) and (1.39b).

12. Show that $\varepsilon^{ijm}\varepsilon^{k\ell m} = \delta^{ik}\delta^{j\ell} - \delta^{i\ell}\delta^{jk}$.

13. Perform traces of the following 2×2 matrices, where 1 stands for the unit 2×2 matrix, and in each case, determine the constants:

$$\operatorname{Tr}(1) = A, \quad \operatorname{Tr}\left(\sigma^i\right) = B, \quad \operatorname{Tr}\left(\sigma^i\sigma^j\right) = C\delta^{ij}, \quad \operatorname{Tr}\left(\sigma^i\sigma^j\sigma^k\right) = D\varepsilon^{ijk},$$
$$\operatorname{Tr}\left(\sigma^i\sigma^j\sigma^k\sigma^\ell\right) = E\left(\delta^{ij}\delta^{k\ell} - \delta^{ik}\delta^{j\ell} + \delta^{jk}\delta^{i\ell}\right).$$

14. Perform the following integrals:
   (a) $\int_0^\infty E^n \delta\left(E^2 - \mathbf{p}^2 - m^2\right) dE$ for arbitrary *n*.
   (b) $\int_0^\infty dE_1 \int_0^\infty dE_2\, \theta\left(\tfrac{1}{2}m - E_1\right)\theta\left(\tfrac{1}{2}m - E_2\right)\theta\left(E_1 + E_2 - \tfrac{1}{2}m\right)\left(\tfrac{1}{2}m^2 E_1 - mE_1^2\right)$
   (c) $\int\left[\left(1 - 2\sin^2\theta_W\right)^2 E^4 + \sin^4\theta_W E^4\left(1 + \cos\theta\right)^2\right] d\Omega$ (note: θ_W is a constant)
   (d) $\int \dfrac{g^4 p\cos^2\theta}{128\pi^2 E\left(E^2 - p^2\cos^2\theta\right)} d\Omega$

II. Relativity and Quantum Mechanics

One of the great advances of 20th century physics was the introduction of special relativity. This introduces the notion that time should be treated as a fourth dimension, on an equal footing with the other three dimensions of space. Combining this with quantum mechanics will prove rather difficult, but we had better make sure we understand special relativity before we start combining it with quantum mechanics.

A. Special Relativity

Special relativity states that the laws of physics must remain invariant under any transformation that leaves the distance formula invariant. The distance formula in four dimensions will be defined in this class[1] using the *mostly minus* convention, in which the distance formula is given by

$$s^2 \equiv \left(\Delta t\right)^2 - \left(\Delta x^1\right)^2 - \left(\Delta x^2\right)^2 - \left(\Delta x^3\right)^2 = \left(\Delta x^0\right)^2 - \left(\Delta \mathbf{x}\right)^2 , \tag{2.1}$$

where we have defined $x^0 = t$ (recall that we are working in units where $c = 1$). The quantity s^2 should be thought of as an entity unto itself, and we normally won't be taking the square root, probably a good idea since it is often negative. To make the notation a bit clearer, we define the *metric tensor* $g_{\mu\nu}$ as

$$g_{\mu\nu} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.2}$$

Then eq. (2.1) can be written a bit more succinctly as

$$S^2 \equiv g_{\mu\nu}\Delta x^\mu \Delta x^\nu . \tag{2.3}$$

What coordinate transformations keep the distance formula (2.3) constant? It is obvious that translations in any space or time direction leave it invariant, *i.e.*, it remains unchanged if we replace x by x', where $x'^\mu = x^\mu - a^\mu$, because the separations all stay the same. It is also true that it will remain invariant under certain linear combinations, called *Lorentz transformations*. Let $\Lambda^\mu{}_\nu$ be a 4×4 matrix, then define

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu . \tag{2.4}$$

Then the distance given by (2.30) in this new coordinate system will be given by

---

[1] This convention is the most common used by particle physicists. For those of us who use it, it is also called the *divinely ordained* convention.

$$S'^2 \equiv g_{\mu\nu}\Delta x'^\mu \Delta x'^\nu = g_{\mu\nu}\Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \Delta x^\alpha \Delta x^\beta \quad ,$$

which will be the same as eq. (2.3) provided

$$g_{\mu\nu}\Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta = g_{\alpha\beta} \; . \tag{2.5}$$

Two examples should illustrate what Lorentz transformations are. A *rotation* is a Lorentz transformation that mixes up two space coordinates while leaving time unchanged. An example is a rotation of the x_1x_2-coordinates by an angle θ, for which the matrix is

$$\Lambda^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.6}$$

The inverse rotation can be achieved by changing θ to $-\theta$. A physical object can achieve a rotation by simply rotating in the correct direction, in this case, about the *z*-axis. Another example of a Lorentz transformation is a *boost*, which mixes time with one direction of space. A boost in the x^3-direction by a *rapidity*[1] ϕ takes the form

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \cosh\phi & 0 & 0 & -\sinh\phi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\phi & 0 & 0 & \cosh\phi \end{pmatrix}. \tag{2.7}$$

The inverse boost can be achieved by changing ϕ to $-\phi$. It is easy to show that an object at rest in the primed system will be moving with velocity $v = \tanh\phi$ as viewed in the unprimed system. Hence boosts can also be achieved (at least in principle) by physical objects simply by accelerating them. Lorentz transformations that can be achieved are called *proper* Lorentz transformations and are simply combinations of rotations and boosts. It is believed that the laws of physics are invariant under translations and all proper Lorentz transformations.

Lorentz boosts have some interesting effects that you should already be familiar with, such as time dilation and Lorentz contraction. These are often written in terms of the *Lorentz factor* $\gamma \equiv \cosh\phi$. I will assume you are familiar with these concepts, and therefore will simply give a quick summary of some formulas involving γ:

$$\gamma \equiv \cosh\phi = \frac{1}{\sqrt{1-v^2}} = \frac{L_0}{L} = \frac{t}{\tau} = \frac{E}{m} \, , \tag{2.8}$$

where *v* is the speed of an object, L_0 is how long it is as viewed in its own rest frame and *L* is the length as viewed by some other observer, τ is how long something takes as viewed in its own rest frame and *t* is how long another observer takes to see it happen, and *E* is the total energy of the object and *m* its rest mass, to be defined in section D. This tells you that when you boost an

---

[1] *Rapidity* is more or less defined by equation (2.7). It is mildly useful in particle physics in its own right, but we won’t really be using it that much.

object, its velocity increases, its energy increases, things like decays take longer to happen, and things get shorter in the direction of motion.

For example, a neutron at rest has a rest mass of m = 939.6 MeV, is roughly 1 fm in size in all directions, and lasts an average of $\tau = 14.7\text{ min}$ in its own rest frame. But if you were to boost it up to E = 4.00 TeV, then $\gamma = 4260$, and it will now be only about 2.34×10^{-4} fm in the direction it is moving (but still 1 fm in the other directions), it will last 43 days and is moving at about 99.99999972% of the speed of light, or a rapidity of about $\phi = 9.05$.

Some Lorentz transformations cannot be performed for a physical object without destroying it. Two examples are *parity* denoted $\mathcal{P}$, and *time reversal,* denoted $\mathcal{T}$, for which the Lorentz transformations take the forms

$$\left(\Lambda_{\mathcal{P}}\right)^{\mu}{}_{\nu} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \text{and} \quad \left(\Lambda_{\mathcal{T}}\right)^{\mu}{}_{\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.9}$$

respectively. Their combination $\mathcal{PT}$ is also unobtainable by physical objects. Parity is reflecting yourself, as in a mirror, while time reversal is going backwards in time. These are examples of *improper* Lorentz transformations. Although it was once thought that $\mathcal{P}$, $\mathcal{T}$, and $\mathcal{PT}$ were all symmetries of nature, we now know that weak interactions violate all three of these.

When we perform a Lorentz transformation, some things (proper distance, electric charge, etc.) remain unchanged, others (momentum, electric current, etc.) will get mixed up with each other. A quantity that is unchanged is called a *scalar*. A scalar s looks the same in both coordinates, so $s' = s$. A *vector u* will have four components, given by $u^{\alpha} = \left(u^0, \mathbf{u}\right)$. When we perform a Lorentz transformation, it will get its components mixed together the same way the coordinates do; that is

$$u'^{\mu} = \Lambda^{\mu}{}_{\nu} u^{\nu} \; . \tag{2.10}$$

Examples of a vector would be the coordinate x^{μ} and the four-momentum p^{μ} (to be defined shortly). There are more complicated objects, called *tensors*, that transform in even more complicated ways. One example is the rank two angular momentum tensor, defined by

$$L^{\mu\nu} \equiv x^{\mu} p^{\nu} - p^{\mu} x^{\nu} \; ,$$

This tensor is anti-symmetric, in that $J^{\mu\nu} \equiv -J^{\nu\mu}$, and hence has six linearly independent components. Under a Lorentz transformation, this tensor transforms as

$$L'^{\mu\nu} = \Lambda^{\mu}{}_{\alpha} \Lambda^{\nu}{}_{\beta} L^{\alpha\beta} \; .$$

This expression is easily generalizable to tensors of higher rank, which simply have more indices on them. The Riemann tensor of general relativity is an example of a rank four tensor.

B. Dot Products, Covectors, Derivatives, and Lorentz Covariant Equations

Let *u* and *v* be any two vectors, then we define the *dot product* as

$$u \cdot v \equiv g_{\alpha\beta} u^\alpha v^\beta = u^0 v^0 - \mathbf{u} \cdot \mathbf{v} . \tag{2.11}$$

Don't forget the minus sign! Suppose you perform a Lorentz transformation, changing these vectors into u' and v', then

$$u' \cdot v' \equiv g_{\alpha\beta} u'^\alpha v'^\beta = g_{\alpha\beta} \Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu u^\mu v^\nu = g_{\mu\nu} u^\mu v^\nu = u \cdot v ,$$

where we used Eq. (2.5) to simplify the expression. Thus the dot product is a scalar quantity, and invariant under Lorentz transformations. You can dot a vector with itself, which we write as

$$u^2 \equiv g_{\alpha\beta} u^\alpha u^\beta = \left(u^0 \right)^2 - \mathbf{u}^2 . \tag{2.12}$$

Note that the square of a vector can be positive, negative, or zero, and just because $u^2 = 0$ doesn't imply $u = 0$.

It is useful for every vector *u* with components u^α to define a *covector* with components u_β defined by

$$u_\beta \equiv g_{\alpha\beta} u^\alpha = \left(u^0 , -\mathbf{u} \right). \tag{2.13}$$

You should think of a covector as a vector waiting to be dotted with another vector. The dot product (2.11) can then be written in any of the equivalent forms

$$u \cdot v \equiv g_{\alpha\beta} u^\alpha v^\beta = u^\alpha v_\alpha = u_\alpha v^\alpha .$$

Given u_β, it is trivial to reverse the process and recover u^α by multiplying by the inverse of the matrix $g_{\alpha\beta}$, which is called the *inverse metric* $g^{\alpha\beta}$, so $g^{\alpha\beta} = \left(g_{\alpha\beta} \right)^{-1}$. Since the metric *g* in (2.2) is its own inverse,[1] clearly the components satisfy $g^{\alpha\beta} = g_{\alpha\beta}$, but as I will explain shortly, this expression is not manifestly Lorentz covariant, and hence physicists don't like writing it. We then reconstruct the vector u^α from u_β as

$$u^\alpha = g^{\alpha\beta} u_\beta . \tag{2.14}$$

Because u^α and u_α contain the same information, we will let *u* stand for either of them, and generally will improperly call both of them vectors. The process of converting between vectors and covectors is called *raising* or *lowering an index.*

Any spacetime index can be raised or lowered. For example, we could raise and/or lower the indices of a Lorentz matrix as

$$\Lambda_\alpha{}^\beta = g_{\alpha\mu} g^{\beta\nu} \Lambda^\mu{}_\nu .$$

---

[1] This statement is only true when using Cartesian coordinates in flat spacetime. We won't be using curved coordinates and will stick to flat spacetime.

Note that when we raise or lower an index, we are careful to leave the horizontal position of the index in place. When we have tensors with multiple indices, it is generally wise to keep them separated, *staggering* the indices like $\Lambda^{\mu}{}_{\nu}$ so we can keep track of their positions. It is then not hard to show that under Lorentz transformations, we have

$$u'_{\alpha} = \Lambda_{\alpha}{}^{\beta} u_{\beta} \, . \tag{2.15}$$

We can even raise the indices on the metric $g_{\alpha\beta}$. Keeping in mind that as matrices, $g_{\alpha\beta}$ and $g^{\alpha\beta}$ are inverses, then it isn't hard to see that

$$g^{\alpha}{}_{\beta} = g_{\beta}{}^{\alpha} = \delta^{\alpha}_{\beta} = \begin{cases} 1 & \text{if } \alpha = \beta \, , \\ 0 & \text{if } \alpha \neq \beta \, . \end{cases}$$

Hence we recognize that when working with spacetime indices, the Kronecker delta-function should always have one index up and one down. If you raise both indices of *g*, it isn't hard to show that $g^{\alpha\mu} g^{\beta\nu} g_{\mu\nu} = g^{\alpha\beta}$, which means our notation is self-consistent.

It is not hard to demonstrate that the metric $g_{\alpha\beta}$, its inverse $g^{\alpha\beta}$, and the Kronecker delta-function are all unchanged if we perform Lorentz transformations on them. It is also true of any combination of products of these things, like $g^{\alpha\beta}\delta^{\nu}_{\mu}$. Are there any other tensors that are Lorentz invariant? There is one other that *almost* works, the four-dimensional Levi-Civita symbol, defined by

$$\varepsilon^{\alpha\beta\mu\nu} \equiv \begin{cases} 1 & \text{if } \alpha, \beta, \mu, \nu \text{ is an even permutation of } 0,1,2,3 \, , \\ -1 & \text{if } \alpha, \beta, \mu, \nu \text{ is an odd permutation of } 0,1,2,3 \, , \\ 0 & \text{otherwise.} \end{cases} \tag{2.16}$$

It can be shown that if you perform a *proper* Lorentz transformation, this tensor is unchanged, but for some improper transformations, such as parity and time reversal ($\mathcal{P}$ and $\mathcal{T}$) it transforms to minus itself. The indices of $\varepsilon^{\alpha\beta\mu\nu}$ can be raised or lowered, just like any other tensor. It is completely anti-symmetric, so if you interchange any pair of indices, you get the same expression back again with a minus sign. It is probably worth noting that if you lower all four indices, you inevitably get a minus sign, so $\varepsilon_{\alpha\beta\mu\nu} = -\varepsilon^{\alpha\beta\mu\nu}$. The only invariant tensors are those that are made from the metric and Kronecker delta function, the Levi-Civita tensor, and their products and sums. Actually, we never have more than one factor of $\varepsilon^{\alpha\beta\mu\nu}$, because it can be shown that the product of two can always be rewritten in terms of the metric. The general expression is too long to include, but when some of the indices are summed over, the expressions are not too bad:

$$\varepsilon^{\alpha\beta\rho\tau} \varepsilon^{\mu\nu}{}_{\rho\tau} = -2g^{\alpha\mu} g^{\beta\nu} + 2g^{\alpha\nu} g^{\beta\mu} \, , \tag{2.17a}$$

$$\varepsilon^{\alpha\beta\rho\tau} \varepsilon^{\mu}{}_{\beta\rho\tau} = -6g^{\alpha\mu} \, , \tag{2.17b}$$

$$\varepsilon^{\alpha\beta\rho\tau} \varepsilon_{\alpha\beta\rho\tau} = -24 \, . \tag{2.17c}$$

We will shortly encounter partial derivatives, which look like $\partial / \partial x^{\alpha}$. Suppose you change coordinates, and have a new coordinate $x'^{\alpha} = \Lambda^{\alpha}{}_{\beta} x^{\beta}$. It is not terribly hard to prove, but not very enlightening, to demonstrate that $\left(\partial/\partial x'^{\alpha}\right) = \Lambda_{\alpha}{}^{\beta}\left(\partial/\partial x^{\beta}\right)$. Comparing with (2.15), we see that partial derivatives act like covectors. To try to make this more obvious (and to shorten our notation) we abbreviate

$$\partial_a \equiv \frac{\partial}{\partial x^{\alpha}} . \tag{2.18}$$

Hence, for example, a time derivative is just ∂_0.

It is helpful to discuss the concept of Lorentz covariant equations. A *Lorentz covariant* equation is any equation that remains the same if we perform a Lorentz transformation on it. An equation is said to be *manifestly Lorentz covariant* if it follows the following rules:

- Any index that appears twice in the same term must appear *exactly* twice, once up and once down, and must be summed over all four values; and
- Any index that appears once in any term must appear once in every term, and must be consistently up or down.

It may be helpful to illustrate these rules with some examples:

$$\text{good:}\quad p^{\mu}p_{\mu} = m^2\ ,\quad F_{\mu\nu} \equiv \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}\ ,\quad J^{\mu\nu} = x^{\mu}p^{\nu} - x^{\nu}p^{\mu}\ ,\quad \tilde{F}_{\alpha\beta} = \tfrac{1}{2}\varepsilon_{\alpha\beta}{}^{\mu\nu}F_{\mu\nu}\ ,$$

$$\text{bad:}\quad p^{\mu}p^{\mu} = m^2\ ,\quad p^{\mu} = q_{\mu}\ ,\quad p^{i} = q^{i}\ ,\quad p^{\mu} = m,\quad g_{\mu\nu} = g^{\mu\nu}\ ,\quad \varepsilon_{\alpha\beta\mu\nu} = -\varepsilon^{\alpha\beta\mu\nu}\ .$$

A spacetime index that is repeated once up and once down is said to be *contracted*. The significance of a manifestly Lorentz covariant expression is that if it is true in one reference frame, it will be true in any reference frame. Of course, if it is false in one reference frame, it will also be false in all reference frames. It should also be noted that equations can be true even when they are not manifestly Lorentz covariant; this is the case for the last two "bad" examples above.

C. Electromagnetism

Let's get some practice writing good Lorentz invariant formulas in the case of electromagnetism. Such equations should be manifestly Lorentz invariant, and they should also be *local*, as I will explain below.

As an easy example, let's start with the concept of conservation of electric charge. Naively, we could write this simply as $dQ/dt = 0$, where Q is simply the total charge of the universe. Though true, this is not very useful experimentally. If we have an experiment and charge were to spontaneously vanish from it, but simultaneously appear in the neighborhood of alpha Centauri, total charge would, in fact, be conserved, but it is difficult to see how this would be discovered experimentally. Such events would also violate special relativity, in that the charge should not instantly move from one place to another. We should actually *see* any charge leaving.

We want instead to imagine that we have a finite volume V with some charge Q_V inside it. This charge may change, but only if we see some current density **J** flowing in or out of the region. Indeed, the rate of change of the total charge Q_V should be given by the amount of current flowing, so we have

$$\frac{d}{dt}Q_V = -\int_S \mathbf{J}\cdot\hat{\mathbf{n}}\,dA\,.$$

Now, the charge in a volume is the integral of the charge density ρ, and we can use Gauss's Law to rewrite the right hand side, so we have

$$0 = \frac{d}{dt}\int_V \rho\, d^3\mathbf{x} + \int_V \nabla\cdot\mathbf{J}\, d^3\mathbf{x} = \int_V\left(\frac{\partial\rho}{\partial t} + \frac{\partial}{\partial x^i}J^i\right)d^3\mathbf{x}\,.$$

Since this is true for any volume, it follows that the integrand must vanish as well, so we have $\partial_0\rho + \partial_i J^i = 0$. If we now define the four-current J^μ as

$$J^\mu \equiv (\rho, \mathbf{J}), \tag{2.19}$$

then conservation of charge becomes as simple as

$$\partial_\mu J^\mu = 0\,. \tag{2.20}$$

This equation is manifestly Lorentz covariant, since the only index is repeated up and down, and it is local, since it only talks about what is happening in the neighborhood of some point *x*. We see from eq. (2.19) that in special relativity, charge density and charge current are simply different components of the same four-vector.

It's time to turn our attention to the much more challenging Maxwell's equations. We already anticipate that we prefer the local version. Keeping in mind that we are working in units where $\varepsilon_0 = \mu_0 = c = 1$, these equations are simply

$$\nabla\cdot\mathbf{E} = \rho\,,\quad \nabla\times\mathbf{E} = -\partial_0\mathbf{B},\quad \nabla\cdot\mathbf{B} = 0,\quad \nabla\times\mathbf{B} = \mathbf{J} + \partial_0\mathbf{E}\,. \tag{2.21}$$

It is not immediately obvious how to proceed. We understand, more or less, that **E** and **B** are somehow related to each other, but it isn't obvious how to put them together into something manifestly Lorentz-covariant.

As a first step, notice that the divergence of **B** vanishes, and it is well known that anything that has no divergence can be written as a curl, so we write

$$\mathbf{B} = \nabla\times\mathbf{A}\,. \tag{2.22}$$

where **A** is called the *vector potential*. Substituting this equation into the second equation of (2.21), we see that $\nabla\times(\mathbf{E} + \partial_0\mathbf{A}) = 0$. Anything whose curl vanishes can be written as the divergence of another quantity, which we call the electrostatic potential Φ, so $\mathbf{E} + \partial_0\mathbf{A} = -\nabla\Phi$, or

$$\mathbf{E} = -\partial_0\mathbf{A} - \nabla\Phi\,, \tag{2.23}$$

If we write out all three components of **B** and **E** explicitly in terms of **A** and Φ, we have

$$\begin{aligned} &E^1 = -\partial_0 A^1 - \partial_1\Phi, \quad E^2 = -\partial_0 A^2 - \partial_2\Phi, \quad E^3 = -\partial_0 A^3 - \partial_3\Phi, \\ &B^1 = \partial_2 A^3 - \partial_3 A^2, \quad\ B^2 = \partial_3 A^1 - \partial_1 A^3, \quad\ B^3 = \partial_1 A^2 - \partial_2 A^1. \end{aligned} \tag{2.24}$$

We now start to make sense of things if we define the four-vector potential as

$$A^{\mu} \equiv (\Phi, \mathbf{A}), \tag{2.25}$$

so that $A_{\mu} \equiv (\Phi, -\mathbf{A})$, and then define the *electromagnetic tensor* $F_{\mu\nu}$ as

$$F_{\mu\nu} \equiv \partial_{\mu} A_{\nu} - \partial_{\nu} A_{\mu} . \tag{2.26}$$

Obviously, $F_{\mu\nu}$ is anti-symmetric, $F_{\mu\nu} = -F_{\nu\mu}$, so it has only six independent components. The components of $F_{\mu\nu}$ (or equivalently, $F^{\mu\nu}$) then are

$$F_{\mu\nu} = \begin{pmatrix} 0 & E^1 & E^2 & E^3 \\ -E^1 & 0 & -B^3 & B^2 \\ -E^2 & B^3 & 0 & -B^1 \\ -E^3 & -B^2 & B^1 & 0 \end{pmatrix}, \quad \text{or} \quad F^{\mu\nu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}. \tag{2.27}$$

The magnetic parts may be a little easier to understand if we simply write them as $F^{ij} = -\varepsilon^{ijk} B^k$, We now work on the inhomogeneous Maxwell's equations. Writing them out explicitly, we have

$$\begin{gathered} J^0 = \rho = \nabla \cdot \mathbf{E} = \partial_i F^{i0} = \partial_i F^{i0} + \partial_0 F^{00} = \partial_{\mu} F^{\mu 0} , \\ J^i = -\partial_0 E^i + (\nabla \times \mathbf{B})^i = \partial_0 F^{0i} + \varepsilon^{ijk} \partial_j B^k = \partial_0 F^{0i} - \partial_j F^{ij} = \partial_{\mu} F^{\mu i} . \end{gathered} \tag{2.28}$$

These equations can be summarized quickly as

$$\partial_{\mu} F^{\mu\nu} = J^{\nu} , \tag{2.29}$$

which covers the two inhomogeneous Maxwell equations. The homogenous equations follow automatically from (2.26), but if we wish to write them in terms of the electromagnetic tensor $F^{\mu\nu}$, they are

$$\varepsilon^{\alpha\beta\mu\nu} \partial_{\beta} F_{\mu\nu} = 0 . \tag{2.30}$$

It turns out that the vector potential A^{μ} is not directly measurable. We can measure the electromagnetic fields $F^{\mu\nu}$, and then try to deduce A^{μ} from Eq. (2.26), but this determination turns out to not be unique. Indeed, let A^{μ} be any vector potential, and let χ be a completely arbitrary function of *x*. Then define

$$A'_{\mu} = A_{\mu} + \partial_{\mu} \chi. \tag{2.31}$$

Then the corresponding electromagnetic tensor will be

$$F'_{\mu\nu} = \partial_{\mu} A'_{\nu} - \partial_{\nu} A'_{\mu} = \partial_{\mu} A_{\nu} + \partial_{\mu} \partial_{\nu} \chi - \partial_{\nu} A_{\mu} - \partial_{\nu} \partial_{\mu} \chi = \partial_{\mu} A_{\nu} - \partial_{\nu} A_{\mu} = F_{\mu\nu} .$$

In other words, the electromagnetic fields remain unchanged if we perform a *gauge transformation* (2.31). Theories that remain unchanged when you perform a gauge transformation are called *gauge invariant*, and they play a special role in particle physics. The three forces of the standard model, the strong, weak, and electromagnetic forces, are all gauge invariant theories. When we make a specific choice about which vector field A^{μ} we will use to describe particular electromagnetic field, we are making a *gauge choice.*

It might seem like a good idea, because of this ambiguity of the vector field, to work directly with the electromagnetic tensor $F^{\mu\nu}$ and avoid A^{μ} entirely. This turns out to be impossible. This is because quantum mechanics works directly with potentials, not forces. When solving the harmonic oscillator, or hydrogen, or any problem in conventional quantum mechanics, the first step is to find the potential V. When you introduce electromagnetic fields into quantum mechanics, their influence on charged particles is described in terms of the electrostatic potential Φ, or if we have magnetic fields, in terms of the vector potential $\mathbf{A}$ as well. We simply don't know how to do quantum mechanics (or quantum field theory) directly in terms of $F^{\mu\nu}$. Hence we are stuck with the difficulties of choosing a gauge whenever we work with such gauge fields.

D. Four-Momentum

Although we use electric and magnetic fields to accelerate particles prior to and after collisions, these fields are much too weak to have any appreciable effect *during* collisions. For the brief moment when a proton collides with an electron, no external forces affect the particles or their products, and no energy is added or subtracted. Therefore both energy and momentum are conserved during collisions and decays. We define the *four-momentum* as

$$p^{\mu} \equiv (E, \mathbf{p}). \tag{2.32}$$

Four-momentum (often just called momentum) is a vector quantity, and Lorentz transforms according to Eq. (2.10). If we dot the four-momentum with itself, we will get a scalar quantity, which we call the mass squared:

$$m^2 \equiv p \cdot p = E^2 - \mathbf{p}^2 . \tag{2.33}$$

The mass squared is an intrinsic property of the particle. Although in principle m^2 could be positive, negative, or zero, in practice no particle has a negative mass squared, so we can always take the square root. The particle of light, or *photon* is believed to be exactly massless, as is a hypothetical particle called the graviton, and the three neutrinos are so small in mass that we will normally treat them as massless. In many collisions, electrons or even quarks may have mass so far below the energy of the experiment that we approximate them as massless as well.

A particle at rest has no three-momentum, and hence $p^{\mu} \equiv (m, 0)$. We can boost it up to rapidity ϕ by observing it in a coordinate system moving with rapidity $-\phi$, and using eq. (2.7), we will find $p^{\mu} \equiv m(\cosh\phi, 0, 0, \sinh\phi)$, and it will be moving at speed $v^3 = \tanh\phi = p^3/E$. Generalizing, we conclude

$$\mathbf{v} = \frac{\mathbf{p}}{E} . \tag{2.34}$$

Eq. (2.34) works even for massless particles, or particles whose masses are negligible. It is easy to see from (2.33) and (2.34) that for massless particles, $|\mathbf{v}| = 1$, so they always go at the speed of light.

In particle physics, it is often desirable to do computations in the *center of mass frame*, a frame where the total initial (and hence final) three-momentum is zero, $\mathbf{p}_{\text{tot}} = 0$. We will then describe the experiment in terms of the total energy in this frame. But how do we figure out what this energy is if we are *not* in the center of mass frame? Define the quantity s as

$$s \equiv \left(p_1 + p_2\right)^2 = \left(E_1 + E_2\right)^2 - \left(\mathbf{p}_1 + \mathbf{p}_2\right)^2 . \tag{2.35}$$

This quantity is a scalar, and hence all observers agree on it. In particular, in the center of mass frame, we have $\mathbf{p}_1 + \mathbf{p}_2 = 0$, and hence in this frame $s = \left(E_1 + E_2\right)^2$, or $E_{\text{cm}} = \sqrt{s}$. Particle physics papers often contain phrases like "Cross-section for proton-proton collisions at $\sqrt{s} = 7.0\ \text{TeV}$", which is just a way of describing the center of mass energy.

Conservation of four-momentum is a necessary rule we will apply in almost every calculation. Suppose we have a set of particles with momenta $p_1, p_2, \ldots, p_N$ which collide and become a new set of particles with momenta $q_1, q_2, \ldots, q_M$. Conservation of four-momentum tells us that

$$p_1 + p_2 + \ldots + p_N = q_1 + q_2 + \ldots + q_M . \tag{2.36}$$

It is common that we need to find detailed information about the momenta or energy of particles involved in a reaction. A general technique for solving such problems is:

- Label all the momenta and write down equation (2.36).
- Rearrange the equation (this takes some insight).
- Square the resulting equation, i.e., dot each side into itself.
- Replace any momentum squared p^2 by its corresponding mass squared m^2
- Write out explicitly any momenta for which you need to take the remaining dot products. For a particle of energy E and space-momentum p moving in an arbitrary direction, this typically will take the form

  $$p^\mu = \left(E, \mathbf{p}\right) = \left(E, p\sin\theta\cos\phi, p\sin\theta\sin\phi, p\cos\theta\right).$$

  For massless particles, $p = E$.
- Write out the dot products explicitly and solve the problem.

It is also a good idea to set up the problem as conveniently as possible. Usually I take the initial momenta to be along the x^3 axis, since this makes subsequent dot products easier.

For example, suppose a photon of energy E collides with an electron of mass m at rest. After the collision, the photon is moving the opposite direction. What is its final energy E? Following the steps outlined above, we call the initial momenta of the electron and the photon p and k respectively, and the final momenta of the electron and photon p' and k'. Conservation of four-momenta tells us $p + k = p' + k'$, or $p' = p + k - k'$. Squaring this, we have

$$p'^2 = p^2 + k^2 + k'^2 + 2p\cdot k - 2p\cdot k' - 2k\cdot k' .$$

We replace all the momenta squared by masses squared. The photon is massless, so we have

$$m^2 = m^2 + 2p\cdot k - 2p\cdot k' - 2k\cdot k' ,$$
$$p\cdot k' + k\cdot k' = p\cdot k .$$

The initial electron is at rest, so $p^\mu = \left(m, 0\right)$. The initial photon, can be chosen to move in the $+x^3$ direction, so $k^\mu = \left(E, 0, 0, E\right)$. The final photon will be moving in the opposite direction, so

$k'^{\mu} = \left(E', 0, 0, -E'\right)$. We now write out all the dot products, and find $p \cdot k' = mE'$, $p \cdot k = mE$, and $k \cdot k' = EE' - \mathbf{k} \cdot \mathbf{k}' = 2EE'$. Substituting these in yields

$$mE' + 2EE' = mE \,,$$

$$E' = \frac{mE}{m + 2E} \,.$$

E. Quantum Mechanics

Thus far, we have worked hard to develop the formalism of special relativity, but we have done nothing to talk about our other big ingredient, quantum mechanics. In quantum mechanics, particles are not described by giving their position and velocity, but rather by state vectors $\left|\Psi\right\rangle$ in a complex vector space. The Hamiltonian H is then a Hermitian operator that changes vectors into vectors; that is, it has the property that

$$\left\langle\Phi\right|H\left|\Psi\right\rangle = \left\langle\Psi\right|H\left|\Phi\right\rangle^{*} \,. \tag{2.37}$$

Schrödinger's equation is then simply

$$i\partial_0 \left|\Psi\right\rangle = H\left|\Psi\right\rangle \,, \tag{2.38}$$

where we recall that $\hbar = 1$ in our units, and we abbreviated the time derivative as ∂_0. This equation will indeed also apply in quantum field theory.

If our Hamiltonian is independent of time, and is sufficiently simple, we can find *eigenstates* $\left|\phi_n\right\rangle$ of the Hamiltonian, states which satisfy

$$H\left|\phi_n\right\rangle = E_n\left|\phi_n\right\rangle. \tag{2.39}$$

If we were to measure which eigenstate of the Hamiltonian we are in at any given time, the probability will be given by

$$P\left(\left|\phi_n\right\rangle\right) = \left|\left\langle\phi_n\middle|\Psi\right\rangle\right|^2 \,. \tag{2.40}$$

It is pretty easy to show that eigenstates with distinct eigenvalues will be orthogonal; insofar as possible, we often want them to be *orthonormal* as well; that is,

$$\left\langle\phi_n\middle|\phi_m\right\rangle = \delta_{nm} \,. \tag{2.41}$$

They will also be *complete*, which we write as

$$\sum_n \left|\phi_n\right\rangle\left\langle\phi_n\right| = 1 \,. \tag{2.42}$$

Given a complete set of orthonormal eigenstates of the Hamiltonian $\left|\phi_n\right\rangle$, we can, with the help of (2.41) and (2.42). write any state in terms of them, so

$$\left|\Psi\left(t=0\right)\right\rangle = \sum_n c_n\left|\phi_n\right\rangle \,, \quad \text{where} \quad c_n = \left\langle\phi_n\middle|\Psi\left(t=0\right)\right\rangle \,.$$

It is then easy to show that the general solution to Schrödinger's equation (2.38) is just

$$\left|\Psi(t)\right\rangle = \sum_n c_n e^{-iE_n t}\left|\phi_n\right\rangle . \tag{2.43}$$

Everything we have said up to now works fine provided we can find the exact eigenstates of the Hamiltonian. Unfortunately, it is common that this is not the case, in which case we will have to use the technique of time-dependent perturbation theory. Let's divide the Hamiltonian into two pieces,

$$H = H_0 + H_{\text{int}} , \tag{2.44}$$

where H_0 we will assume is large, and we can find its exact eigenstates $\left|\phi_n\right\rangle$, and we hope the interaction piece H_{int} will be small. We will assume both H_0 and H_{int} have no explicit time dependence. Equation (2.39) is now replaced with

$$H_0\left|\phi_n\right\rangle = E_n\left|\phi_n\right\rangle . \tag{2.45}$$

Equation (2.43) will now satisfy (2.38) only if we ignore H_{int}. We can, however, write a general solution in the form (2.43) if we make the c_n's functions of time, so that

$$\left|\Psi(t)\right\rangle = \sum_n c_n(t) e^{-iE_n t}\left|\phi_n\right\rangle . \tag{2.46}$$

Substitute this into Schrödinger's equation (2.38) using our modified Hamiltonian (2.44):

$$\begin{aligned} i\partial_0 \sum_n c_n(t) e^{-iE_n t}\left|\phi_n\right\rangle &= \sum_n \left(H_0 + H_{\text{int}}\right) c_n(t) e^{-iE_n t}\left|\phi_n\right\rangle , \\ \sum_n \left[i\partial_0 c_n(t) + c_n(t) E_n\right] e^{-iE_n t}\left|\phi_n\right\rangle &= \sum_n \left(E_n + H_{\text{int}}\right) c_n(t) e^{-iE_n t}\left|\phi_n\right\rangle , \\ \sum_n \partial_0 c_n(t) e^{-iE_n t}\left|\phi_n\right\rangle &= \sum_n \left(-iH_{\text{int}}\right)\left|\phi_n\right\rangle c_n(t) e^{-iE_n t} . \end{aligned} \tag{2.47}$$

Now act with $\left\langle\phi_m\right|$ on both sides of (2.47). Orthonormality will then collapse the sum on the left to a single term, while on the right we will get

$$\begin{aligned} \partial_0 c_m(t) e^{-iE_m t} &= \sum_n \left\langle\phi_m\right|\left(-iH_{\text{int}}\right)\left|\phi_n\right\rangle c_n(t) e^{-iE_n t} , \\ \partial_0 c_m(t) &= -i\sum_n \left\langle\phi_m\right| H_{\text{int}}\left|\phi_n\right\rangle c_n(t) e^{i(E_m - E_n)t} . \end{aligned} \tag{2.48}$$

Let's assume we know that at $t = 0$, the system is in an eigenstate $\left|\phi_I\right\rangle$, so that $c_n(0) = \delta_{nI}$. Then we can integrate eq. (2.48) to yield

$$c_F(T) = \delta_{FI} - i\sum_n \left\langle\phi_F\right| H_{\text{int}}\left|\phi_n\right\rangle \int_0^T c_n(t) e^{i(E_F - E_n)t} dt , \tag{2.49}$$

where $\left|\phi_F\right\rangle$ is some final state we are interested in. We can now solve (2.49) perturbatively by repeatedly substituting it into itself. For example, to zero'th order in H_{int}, we have $c_n(t) \approx \delta_{nI}$, and therefore to first order we have

$$c_F\left(T\right) \approx \delta_{FI} - i\left\langle \phi_F \left| H_{\text{int}} \right| \phi_I \right\rangle \int_0^T e^{i\left(E_F - E_I\right)t} dt = \delta_{FI} - i\frac{\left\langle \phi_F \left| H_{\text{int}} \right| \phi_I \right\rangle}{i\left(E_F - E_I\right)} \left[e^{i\left(E_F - E_I\right)T} - 1 \right].$$

We can then substitute this expression into eq. (2.49) to get an expression accurate to second order in H_{int}, namely

$$\begin{aligned} c_F\left(T\right) &\approx \delta_{FI} - i\left\langle \phi_F \left| H_{\text{int}} \right| \phi_I \right\rangle \int_0^T e^{i\left(E_F - E_I\right)t} dt + \\ &\left(-i\right)^2 \sum_n \left\langle \phi_F \left| H_{\text{int}} \right| \phi_n \right\rangle \frac{\left\langle \phi_n \left| H_{\text{int}} \right| \phi_I \right\rangle}{i\left(E_n - E_I\right)} \int_0^T e^{i\left(E_F - E_n\right)t} \left[e^{i\left(E_n - E_I\right)t} - 1 \right] dt . \end{aligned} \tag{2.50}$$

Now things get a bit technical, and I don't want to get bogged down in details, but the -1 term in the square brackets is there because we had an abrupt start to our experiment. Basically, a long technical argument would explain why this term is not really very important, and should be ignored. We therefore drop it, and we now have

$$c_F\left(T\right) \approx \delta_{FI} + \int_0^T e^{i\left(E_F - E_I\right)t} dt \left\{ -i\left\langle \phi_F \left| H_{\text{int}} \right| \phi_I \right\rangle + \left(-i\right)^2 \sum_n \frac{\left\langle \phi_F \left| H_{\text{int}} \right| \phi_n \right\rangle \left\langle \phi_n \left| H_{\text{int}} \right| \phi_I \right\rangle}{i\left(E_n - E_I\right)} \right\}.$$

We now define that mess in the curly brackets as the *transition amplitude*,

$$\mathcal{T}_{FI} \equiv -i\left\langle \phi_F \left| H_{\text{int}} \right| \phi_I \right\rangle + \left(-i\right)^2 \sum_n \frac{\left\langle \phi_F \left| H_{\text{int}} \right| \phi_n \right\rangle \left\langle \phi_n \left| H_{\text{int}} \right| \phi_I \right\rangle}{i\left(E_n - E_I\right)} + \cdots, \tag{2.51}$$

where the ellipsis at the end denotes higher order terms, which will not concern us. If we assume the final state is not identical to the initial state, we have

$$c_F\left(T\right) \approx \mathcal{T}_{FI} \int_0^T e^{i\left(E_F - E_I\right)t} dt = \mathcal{T}_{FI} \frac{e^{i\left(E_F - E_I\right)T} - 1}{i\left(E_F - E_I\right)} = \mathcal{T}_{FI} e^{\frac{1}{2}i\left(E_F - E_I\right)T} \frac{2\sin\left[\frac{1}{2}\left(E_F - E_I\right)T\right]}{E_F - E_I}.$$

What we really want to find is the probability that we end up in the state $\left|\phi_F\right\rangle$, which is given by

$$P\left(I \to F\right) = \left|\left\langle \phi_F \middle| \Psi \right\rangle\right|^2 = \left|c_F\left(T\right)\right|^2 = \left|\mathcal{T}_{FI}\right|^2 \frac{4\sin^2\left[\frac{1}{2}\left(E_F - E_I\right)T\right]}{\left(E_F - E_I\right)^2}. \tag{2.52}$$

For large T (and all experiments have large T compared to the rates of particle interactions), the final factor is a sharply peaked function at $E_F = E_I$. Using eq. (1.45), we have

$$P\left(I \to F\right) = 2\pi T \delta\left(E_F - E_I\right)\left|\mathcal{T}_{FI}\right|^2 . \tag{2.53}$$

This formula is known as *Fermi's Golden Rule*. There is a relativistic equivalent, which we will encounter in chapter 4.

F. Plane Waves

For a single particle without spin, a particle is described in quantum mechanics in terms of a wave function $\Psi(\mathbf{x})$. One of the most useful sets of orthogonal states for describing such states is plane wave states, labeled $|\mathbf{k}\rangle$, with wave function

$$\langle \mathbf{x}|\mathbf{k}\rangle = e^{i\mathbf{k}\cdot\mathbf{x}} \,. \tag{2.54}$$

These states are eigenstates of the free Hamiltonian. We immediately run into a problem with (2.54); namely, that this state is not normalizable. The magnitude is everywhere one, and if integrated over all space, it will yield infinity, and hence can't be normalized.

To understand the problem, let's first specialize to one dimension, and let's consider a universe of size L, where $-\tfrac{1}{2}L < x < \tfrac{1}{2}L$. We will work with periodic boundary conditions, in which $\phi\left(-\tfrac{1}{2}L\right) = \phi\left(\tfrac{1}{2}L\right)$, in which case we have $e^{-ikL/2} = e^{ikL/2}$, so

$$1 = e^{ikL} = \cos(kL) + i\sin(kL) \,,$$

$$k = \frac{2\pi}{L} n \,, \tag{2.55}$$

where n is an integer. Now, consider the inner product $\langle k'|k\rangle$, which is given by

$$\langle k'|k\rangle = \int_{-\frac{1}{2}L}^{\frac{1}{2}L} e^{ikx-ik'x}dx = \frac{1}{i(k-k')}\left[e^{i(k-k')L/2} - e^{-i(k-k')L/2}\right] = \frac{2}{k-k'}\sin\left[\tfrac{1}{2}(k-k')L\right]$$

$$= \frac{2}{k-k'}\sin\left[\pi(n-n')\right] = 0 \quad \text{if } k \neq k' \,,$$

$$\langle k|k\rangle = \int_{-\frac{1}{2}L}^{\frac{1}{2}L} e^{ikx-ikx}dx = L \,,$$

$$\langle k'|k\rangle = L\delta_{k,k'} \,. \tag{2.56}$$

Then the states $|k\rangle/\sqrt{L}$, are orthonormal, so that the completeness relation (2.42) is

$$\frac{1}{L}\sum_k |k\rangle\langle k| = 1 \,. \tag{2.57}$$

We next would like to take the infinite universe limit. With the help of (1.45), we see that

$$\langle k'|k\rangle = \lim_{L\to\infty}\frac{2}{k-k'}\sin\left[\tfrac{1}{2}(k-k')L\right] = 2\pi\delta(k-k') \,. \tag{2.58}$$

and what do we make of (2.57)? As $L \to \infty$, the spacing between k values given by (2.55) will become very small, with $\Delta k = 2\pi/L$. If you add up any function $f(k)$ in very small step sizes, you are performing an integral. In particular,

$$\lim_{L\to\infty}\frac{1}{L}\sum_k f(k) = \lim_{\Delta k\to 0}\frac{1}{2\pi}\sum_n f(n\Delta k)\Delta k = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(k)\,dk \,. \tag{2.59}$$

Now that we have everything done for one dimension, we need to generalize to three. We let all three of our space dimensions be restricted to $-\frac{1}{2}L < x^i < \frac{1}{2}L$, and then $\mathbf{k} = \frac{2\pi}{L}\left(n_1, n_2, n_3\right)$. Equations (2.56) and (2.57) become

$$\left\langle \mathbf{k}' \middle| \mathbf{k} \right\rangle = \int d^3\mathbf{x}\, e^{i\mathbf{k}\cdot\mathbf{x} - i\mathbf{k}'\cdot\mathbf{x}} = V\delta_{\mathbf{k},\mathbf{k}'} , \tag{2.60a}$$

$$\frac{1}{V}\sum_{\mathbf{k}} \left|\mathbf{k}\right\rangle\left\langle\mathbf{k}\right| = 1 , \tag{2.60b}$$

where $V = L^3$ is the volume of the universe. Equations (2.58) and (2.59) become a product of three delta functions and a triple integral, with three factors of 2π, so we have

$$\left\langle \mathbf{k}' \middle| \mathbf{k} \right\rangle = \left(2\pi\right)^3 \delta\left(\mathbf{k} - \mathbf{k}'\right), \tag{2.61a}$$

$$\lim_{V\to\infty} \frac{1}{V}\sum_{\mathbf{k}} f\left(\mathbf{k}\right) = \int \frac{d^3\mathbf{k}}{\left(2\pi\right)^3} f\left(\mathbf{k}\right). \tag{2.61b}$$

In general, we will work in finite volume as much as possible, using (2.60), but late in all our calculations we will take the limit of infinite volume. By comparison of (2.60a) and (2.61a), we also see that in the infinite volume limit,

$$\lim_{V\to\infty} V\delta_{\mathbf{k},\mathbf{k}'} = \left(2\pi\right)^3 \delta^3\left(\mathbf{k} - \mathbf{k}'\right). \tag{2.62}$$

G. Spin

For particles with spin, a momentum state $\left|\mathbf{k}\right\rangle$ is an incomplete description of a particle. There is also *spin*, and our state looks more like $\left|\mathbf{k}, s\right\rangle$. Spin comes about in ordinary quantum mechanics because the wave function $\Psi\left(\mathbf{x}\right)$ has multiple components, and when we perform spatial rotations, these components mix up with each other. To make things simple, consider first a particle at rest, so that $\mathbf{k} = 0$, and focus on just the spin state $\left|s\right\rangle$.

Spin is a kind of intrinsic angular momentum of a particle. The intrinsic angular momentum around any direction $\hat{\mathbf{n}}$ is governed by the three Hermitian spin operators **S**, so that the combination $\hat{\mathbf{n}}\cdot\mathbf{S}$ is the operator measuring spin in this direction. The three spin operators do not commute, but instead satisfy the commutation relations

$$\left[S^i, S^j\right] = i\varepsilon^{ijk} S^k . \tag{2.63}$$

Let's review how these equations are dealt with in quantum mechanics. Normally we start by defining $\mathbf{S}^2 = S^i S^i$ and we show that it commutes with all three of the S^i's. Since it commutes with all of them, the different components of $\left|s\right\rangle$ will automatically all have the same value for this operator, so it is an intrinsic property of the particle. We define its eigenvalue as

$$\mathbf{S}^2 = S^2 + S , \tag{2.64}$$

which, for the moment, is just a weird way of writing the eigenvalue in terms of an unknown number *S*. We now pick one of the three S^i's, normally chosen to be S^3, and write the spin states in terms of a set of basis vectors $|M\rangle$, so that

$$S^3|M\rangle = M|M\rangle . \tag{2.65}$$

Next we define raising and lowering operators

$$S^\pm \equiv S^1 \pm iS^2 , \tag{2.66),}$$

which are Hermitian conjugates of each other. Among their properties are

$$\left[S^3, S^\pm\right] = \pm S^\pm , \tag{2.67a}$$

$$\mathbf{S}^2 = S^\pm S^\mp + \left(S^3\right)^2 \pm S^3 . \tag{2.67b}$$

From eq. (2.67a) we can easily show that

$$S^3\left(S^\pm|M\rangle\right) = \left(\left[S^3, S^\pm\right] + S^\pm S^3\right)|M\rangle = \left(\pm S^\pm + MS^\pm\right)|M\rangle = (M \pm 1)\left(S^\pm|M\rangle\right).$$

We see from this that $S^\pm|M\rangle$ has eigenvalue $M \pm 1$ when acted on by S^3, which implies it is proportional to $|M \pm 1\rangle$. Let's call the proportionality constant *c*, so

$$S^\pm|M\rangle = c|M \pm 1\rangle . \tag{2.68}$$

We can determine the constant *c* by taking the inner product of (2.68) with itself, which yields

$$|c|^2 = \left|c|M \pm 1\rangle\right|^2 = \langle M|S^\mp S^\pm|M\rangle = \langle M|\left(\mathbf{S}^2 - \left(S^3\right)^2 \mp S^3\right)|M\rangle = S^2 + S - M^2 \mp M .$$

Taking the square root and substituting into (2.68), we find

$$S^\pm|M\rangle = \sqrt{S^2 + S - M^2 \mp M}\,|M \pm 1\rangle . \tag{2.69}$$

Eq. (2.69) makes no sense if you are taking the square root of a negative number. But if we raise or lower repeatedly using (2.69), we eventually will get values of *M* such that the square root is negative. How do we avoid this? The answer is that there must be a maximum value $M_{\max}$ such that $S^+|M_{\max}\rangle = 0$ and a minimum value $M_{\min}$ such that $S^-|M_{\min}\rangle = 0$. This implies

$$0 = S^2 + S - M_{\max}^2 - M_{\max} \quad \text{and} \quad 0 = S^2 + S - M_{\min}^2 + M_{\min}$$

This has solutions $M_{\max} = S$ and $M_{\min} = -S$, so since *M* increases by steps of size one,

$$M = S, S-1, S-2, \ldots, -S . \tag{2.70}$$

This implies that the difference between the minimum and maximum value, which is 2*S*, must be an integer, so *S* is an integer or half-integer.

There is nothing special about the S^3 operator; we could instead pick any S^i or any linear combination thereof. Let $\hat{\mathbf{n}}$ be any unit vector, then we diagonalize $\hat{\mathbf{n}} \cdot \mathbf{S}$ and find

$$\hat{\mathbf{n}} \cdot \mathbf{S}|\phi_M\rangle = M|\phi_M\rangle , \tag{2.71a}$$

$$M = S, S-1, S-2, \ldots, -S\,. \tag{2.71b}$$

The operators **S** not only measure the angular momentum of a wave function, they also tell you how to rotate it. If you rotate your coordinates by an angle θ about any axis $\hat{\mathbf{n}}$, the state vector gets multiplied by

$$|s\rangle \to \exp\left(-i\mathbf{S}\cdot\hat{\mathbf{n}}\theta\right)|s\rangle\,. \tag{2.72}$$

If we ever need the explicit form of the spin matrices **S**, we start by choosing a basis where S^3 is diagonalized, so that (2.70) gives the diagonal values of the components of the matrix. The explicit form of $S^{\pm}$ can be worked out from (2.69); for example, S^+ will have non-zero matrix elements only just above the diagonal. S^1 and S^2 can then be worked out from (2.66), which tells us

$$S^1 = \tfrac{1}{2}\left(S^+ + S^-\right), \quad S^2 = \tfrac{1}{2}i\left(S^- - S^+\right). \tag{2.73}$$

To make this clear, let's illustrate how this works when $S = \frac{1}{2}$. The spin matrices turn out to just be

$$\mathbf{S} = \tfrac{1}{2}\boldsymbol{\sigma}\,, \tag{2.74}$$

where the $\boldsymbol{\sigma}$'s are the Pauli Matrices, given by Eq. (1.34). It is not hard to show using (1.35a) that **S** then satisfies (2.63), and (1.35b) lets us show that **S** satisfies (2.64) with $S = \frac{1}{2}$. The normalized eigenstates of $\hat{\mathbf{n}}\cdot\mathbf{S}$ will be labeled $\phi_{\pm}$, with eigenvalues $\pm\frac{1}{2}$ and if we write $\hat{\mathbf{n}} = \left(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta\right)$, they are given by

$$\psi_+ = \begin{pmatrix} \cos\left(\frac{1}{2}\theta\right) \\ \sin\left(\frac{1}{2}\theta\right)e^{i\phi} \end{pmatrix}, \quad \psi_- = \begin{pmatrix} \sin\left(\frac{1}{2}\theta\right)e^{i\phi} \\ -\cos\left(\frac{1}{2}\theta\right) \end{pmatrix}. \tag{2.75}$$

A couple of identities involving these two-component eigenstates deserver mention. First is the completeness relation, which in terms of these is

$$\psi_+\psi_+^\dagger + \psi_-\psi_-^\dagger = 1\,. \tag{2.76}$$

Now, these states are eigenstates of $\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}$ with eigenvalues ± 1. If we multiply on the left by $\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}$, the first term in (2.76) is unchanged, and the second changes sign, so

$$\psi_+\psi_+^\dagger - \psi_-\psi_-^\dagger = \hat{\mathbf{n}}\cdot\boldsymbol{\sigma}\,. \tag{2.77}$$

These identities will prove useful later on.

Everything we have done up to now assumes the particle is at rest. And if it isn't? You can simply switch to a frame of reference where the particle *is* at rest, and then the analysis is performed the same way. This works fine for massive particles, but massless particles move at the speed of light, and hence are not at rest in any frame of reference. In such cases, we cannot rotate the state without messing up its direction of motion. The one exception is that you can still rotate *around* the direction of motion. For a massless particle, define the *chirality* of the particle as $\mathbf{v}\cdot\mathbf{S}$ (recall that $v = 1$ for massless particles). As before, we find that our states will be eigenstates of $\mathbf{v}\cdot\mathbf{S}$, so that

$$\mathbf{v}\cdot\mathbf{S}\left|\phi_M\right\rangle = M\left|\phi_M\right\rangle, \tag{2.78}$$

and once again, *M* can be integer or half-integer. But for massless particles, we can't rotate around different axes and get different values of *M*, and the helicity of a particle will be the same as viewed by all observers. Photons, for example, have $M=\pm 1$, but not $M = 0$. Neutrinos, which used to be thought of as massless, were believed to have $M=-\frac{1}{2}$ (left-handed) so they always spin counter-clockwise if you look in their direction of motion, and anti-neutrinos $M=\frac{1}{2}$ (right-handed). Gravitons, hypothetical particles that carry the gravitational force, have $M=\pm 2$.

We should also spend some time mentioning the *spin-statistics* theorem. It turns out that if you exchange two particles of the same type, the state vector always comes out to be either exactly the same or minus itself. The spin-statistics theorem states that if *S* is an integer, interchanging the two particles gives a plus sign, but when it is a half-integer, it yields a minus sign. Integer spin particles are called *bosons*, and half-integer particles are called *fermions*.

$$\begin{aligned} \text{bosons:}\quad & \left|\phi_1,\phi_2\right\rangle = +\left|\phi_2,\phi_1\right\rangle, \\ \text{fermions:}\quad & \left|\psi_1,\psi_2\right\rangle = -\left|\psi_2,\psi_1\right\rangle. \end{aligned} \tag{2.79}$$

Because the only elementary particles that are fermions have spin $\frac{1}{2}$, we sometimes use fermions to mean spin $\frac{1}{2}$. For bosons, spin 0 is referred to as *scalar bosons* and spin 1 is referred to as *vector bosons.*

Now that all the ingredients are in place, we are ready to make our first attempt at a relativistic quantum mechanical theory. The Dirac equation was developed to describe electrons, spin $\frac{1}{2}$ particles, and this is where we turn our attention.

Problems for Chapter 2

1. Find the determinant of each of the Lorentz transformations eqs. (2.6) and (2.7). Assuming the result is the same for all rotations and boosts, show that no combination of rotations and boosts can produce a parity transformation $\mathcal{P}$ nor a time reversal transformation $\mathcal{T}$.

2. Simplify each of the following: (a) $g_{\alpha\beta}g^{\alpha\beta}$

   (b) $\left(p_1^\mu p_2^\nu + p_1^\nu p_2^\mu - p_1\cdot p_2 g^{\mu\nu}\right)\left(p_{3\mu}p_{4\nu} + p_{3\nu}p_{4\nu} - p_3\cdot p_4 g_{\mu\nu}\right)$

   (c) $\varepsilon^{\alpha\mu\beta\nu}\varepsilon^{\rho\ \ \tau}_{\ \ \mu\ \ \nu}p_{1\alpha}p_{2\beta}p_{3\rho}p_{4\tau}$

3. Find formulas for how the components of **E** and **B** are related to $\mathbf{E}'$ and $\mathbf{B}'$ when you perform (a) a rotation around the x^3 axis by an angle θ, and (b) a boost along the x^3 axis by rapidity ϕ. Each part should have six equations.

4. Show that a particle with $E \gg m$ has approximate reciprocal velocity $1/v \approx 1 + m^2/2E^2$. If two neutrinos with different energies E_1 and E_2 arrive at the same time after travelling a distance *d*, find a formula for the difference in time Δt. In 1987, two neutrinos with the

same mass and energies $E_1 = 6\text{ MeV}$ and $E_2 = 20\text{ MeV}$ arrived from SN1987a after travelling a distance $d = 160,000$ light years. Assuming they left at most $\Delta t < 10$ seconds apart, get an approximate limit on the mass *m* of the neutrinos.

5. Suppose two particles of mass *m* and energy *E* collide almost head on, so that $\pi - \theta$ is the angle between them. Find a formula for *s*. Look up the crossing angle θ for the LHC. If they collide protons with $E = 4.00\text{ TeV}$, are they accurate when they publish that the experiments are being performed at $\sqrt{s} = 8.00\text{ TeV}$?

6. Find a formula for *s* if particles of mass *m* and energy *E* collide with a stationary target of mass *m*. If you use *B* = 10 T magnets, how large in Earth radii would you have to make a collider to reach $\sqrt{s} = 8.00\text{ TeV}$?

7. Suppose an electron/positron collider collides beams with energies E_1 and E_2 head on. What is *s*? Treat the electron and positrons as massless. If the BABAR experiment is trying to create the $\Upsilon(4s)$ resonance with mass $M = 10.58\text{ GeV}$ by colliding electrons with energy $E_1 = 9.00\text{ GeV}$ electrons, what energy must the positrons be?

8. A particle of mass *M* decays to two particles. Find a general formula for the magnitude of the final three-momentum:
   (a) If the mass of each final particle is *m*;
   (b) If the mass of one final particle is *m* and the other is 0; and
   (c) If the mass of the final particles are m_1 and m_2, and check that it leads to the correct results for parts (a) and (b).

9. For any process where two particles of momenta p_1 and p_2 collide to make two final particles with momenta p_3 and p_4, define the *Mandelstam variables* by

$$s = (p_1 + p_2)^2 = (p_3 + p_4)^2, \quad t = (p_1 - p_3)^2 = (p_2 - p_4)^2, \quad u = (p_1 - p_4)^2 = (p_2 - p_3)^2.$$

   Show that $s + t + u$ is a constant, and determine it in terms of the masses $m_i^2 = p_i^2$.

10. The neutral Kaon system has two particles $|K_0\rangle$ and $|\bar{K}_0\rangle$. These particles are not mass eigenstates; they are related to mass eigenstates by

$$|K_0\rangle = \tfrac{1}{\sqrt{2}}\left(|K_1\rangle + |K_2\rangle\right), \quad |\bar{K}_0\rangle = \tfrac{1}{\sqrt{2}}\left(|K_1\rangle - |K_2\rangle\right).$$

   These *are* eigenstate of the Hamiltonian, with energies $H|K_1\rangle = M_1|K_1\rangle$ and $H|K_2\rangle = M_2|K_2\rangle$. Suppose at $t = 0$, we have $|\Psi(t=0)\rangle = |K_0\rangle$. What is $|\Psi(t)\rangle$ at all times? At time *t*, the particle is measured to see if it is a $|K_0\rangle$ or $|\bar{K}_0\rangle$. What is the probability of each of these? If $M_1 - M_2 = 3.484\times10^{-6}\text{ eV}$, at what time *t* will the particle first be 100% $|\bar{K}_0\rangle$?

III. The Dirac Equation

We are now ready to start working on an actual relativistic one-particle equation, the Dirac equation. But first, let's remember how we get the free particle Schrödinger equation. We start with the equation for the energy of a free particle in terms of the momentum, namely

$$E = \frac{\mathbf{p}^2}{2m} \tag{3.1}$$

We want to find equations the solve this of the form

$$\Psi = e^{i\mathbf{p}\cdot\mathbf{x} - iEt} = e^{-ip\cdot x} \ . \tag{3.2}$$

We first note that $-i\nabla\Psi = \mathbf{p}\Psi$ and $i\partial_0\Psi = E\Psi$. This suggests multiplying (3.1) on the right by Ψ, and then make the replacements $\mathbf{p} \to -i\nabla$ and $E \to i\partial_0$. This yields the free Schrodinger equation,

$$i\partial_0\Psi = -\frac{1}{2m}\nabla^2\Psi \ .$$

A. The Free Dirac Equation

Will the same technique work to give us a relativistic equation? We still want waves of the form (3.2), but now the relationship between *E* and **p** will be

$$E^2 = \mathbf{p}^2 + m^2 \ . \tag{3.3}$$

We again multiply on the right by Ψ and replace $\mathbf{p} \to -i\nabla$ and $E \to i\partial_0$. We obtain

$$\begin{aligned} -\partial_0\partial_0\Psi &= -\nabla^2\Psi + m^2\Psi \ , \\ -\partial^2\Psi &= m^2\Psi \ . \end{aligned} \tag{3.4}$$

This is the *Klein-Gordon equation*. As a substitute for Schrödinger's equation, however, it leaves a lot to be desired. Normally in quantum mechanics, initial conditions are given by giving the wave function at some initial time, say $\Psi(\mathbf{x}, t=0)$. Because this equation is second order, it is necessary to also know the time derivative, $\dot{\Psi}(\mathbf{x}, t=0)$. There are solutions of (3.4a) where $\Psi(\mathbf{x}, t=0) = 0$, which normally would imply there is nothing, but $\Psi(\mathbf{x}, t) \neq 0$ at other times.

To get around this problem, we would like to, so to speak, take the square root of (3.3). If do this naively, we would end up with an equation like $i\partial_0\Psi = \left(\sqrt{-\nabla^2 + m^2}\right)\Psi$, but interpreting that square root of an operator is problematic at best.

Dirac's brilliant solution to this problem was to write down a *first*-order differential equation that implied the *second*-order Klein-Gordon equation (3.4). His idea was to write

$$i\partial_0\Psi = H\Psi \equiv \left(-i\boldsymbol{\alpha}\cdot\nabla + m\beta\right)\Psi, \tag{3.5}$$

where for the moment, $\boldsymbol{\alpha}$ and β are four unspecified objects. This is the *Dirac Equation*. We recognize $-i\nabla$ as the momentum operator, which is Hermitian, so for H to be Hermitian, we need

$$\boldsymbol{\alpha}^\dagger = \boldsymbol{\alpha} \quad \text{and} \quad \beta^\dagger = \beta\,. \tag{3.6}$$

If you act twice with $i\partial_0$ on Ψ, you get

$$-\partial_0\partial_0\Psi = \left(-i\alpha^i\partial_i + m\beta\right)^2\Psi\,. \tag{3.7}$$

Now, when you square out this operator, you will get diagonal terms like $\left(\alpha^i\right)^2\partial_i\partial_i\Psi$ and $\beta^2\Psi$, and off-diagonal terms like $\left\{\alpha^i,\alpha^j\right\}\partial_i\partial_j\Psi$ and $m\left\{\alpha^i,\beta\right\}\partial_i\Psi$. We want all the off-diagonal terms to vanish, and all the on-diagonal ones to give 1, which can be achieved if

$$\left\{\alpha^i,\alpha^j\right\} = 2\delta^{ij}\,, \quad \left\{\alpha^i,\beta\right\} = 0\,, \quad \beta^2 = 1\,. \tag{3.8}$$

This then turns (3.7) into the Klein-Gordon equation (3.3).

These equations are not satisfied by any sort of ordinary numbers, so Dirac proposed that they should be matrices. Dirac demonstrated that the matrices had to be at least 4×4 matrices. There are several ways to write them, and they are all equivalent, but I will present them in the *chiral* representation, in which these *Dirac matrices* are given by

$$\boldsymbol{\alpha} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & -\boldsymbol{\sigma} \end{pmatrix}, \quad \beta = \begin{pmatrix} 0 & -\mathbf{1} \\ -\mathbf{1} & 0 \end{pmatrix}, \tag{3.9}$$

where $\boldsymbol{\sigma}$'s are the 2×2 Pauli matrices given by eq. (1.34). It is important to understand that each of the matrices in (3.9) are actually 4×4. Hence each 0 represents a 2×2 submatrix of zeros, and **1** represents the 2×2 identity matrix. It is rare that you actually need the full 4×4 form of these matrices; far more often you are working with the representations (3.9), or more commonly, with expressions like (3.8).

Some comments are in order. First note that for the Dirac equation (3.5) to make any sense, Ψ must have four components. We call such an object a *Dirac spinor*. I'd like to note that the $\boldsymbol{\alpha}$'s are all block diagonal; only the matrix β connects the upper half to the lower half. Indeed, if we had a massless particle, then we could divide Ψ into two pieces, each with two components, called the right and left pieces, by writing

$$\Psi = \begin{pmatrix} \Psi_R \\ \Psi_L \end{pmatrix}, \tag{3.10}$$

and then the Dirac equation would split into two independent equations,

$$i\partial_0\Psi_R + i\boldsymbol{\sigma}\cdot\nabla\Psi_R = 0 \quad \text{and} \quad i\partial_0\Psi_L - i\boldsymbol{\sigma}\cdot\nabla\Psi_L = 0\,. \tag{3.11}$$

This will become relevant when we discuss the Higgs field in chapter 11.

It is not obvious that the Dirac equation is Lorentz covariant. To demonstrate it explicitly is more work than I want to go into, but it can be shown that if you have a solution $\Psi\left(x^\mu\right)$ of the

Dirac equation, you can rotate or boost it to get more solutions. When you rotate coordinates about an axis $\hat{\mathbf{n}}$ by an angle θ you must also mix up the components of Ψ according to

$$\Psi \to \exp\left(-i\mathbf{S}\cdot\hat{\mathbf{n}}\theta\right)\Psi \ . \tag{3.12}$$

The spin matrices **S** in the chiral representation are

$$\mathbf{S} = \tfrac{1}{2}\boldsymbol{\Sigma}, \quad \text{where} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}. \tag{3.13}$$

We see that Ψ_R and Ψ_L each are just two separate spin wave functions, each having spin $\tfrac{1}{2}$. Hence the Dirac equation works only for spin $\tfrac{1}{2}$ particles.

Lorentz boosts are like rotations in space-time, so we should not be surprised that when you perform them, they mix up the components of Ψ in a manner very similar to (3.12). It turns out that if you change your coordinates by a boost of rapidity ϕ along the $\hat{\mathbf{n}}$ direction, you also need to perform a mixing of the components given by

$$\Psi \to \exp\left(-\tfrac{1}{2}\boldsymbol{\alpha}\cdot\hat{\mathbf{n}}\phi\right)\Psi \ , \tag{3.14}$$

where $\boldsymbol{\alpha}$ are the usual Dirac matrices given by (3.9). Again note that Ψ_R and Ψ_L are not mixed by Lorentz boosts.

Another comment is in order. Equation (3.11) treats space and time on a different basis, since there is a matrix multiplying the space derivatives but not the time derivative. For this reason, we rewrite the Dirac equation by first multiplying on the left by β to give

$$i\beta\partial_0\Psi = -i\beta\alpha^i\partial_i\Psi + m\Psi.$$

We then define the four matrices γ^μ by

$$\gamma^\mu \equiv \left(\beta, \beta\boldsymbol{\alpha}\right). \tag{3.15}$$

In terms of these, the Dirac equation is

$$i\gamma^\mu\partial_\mu\Psi = m\Psi \ . \tag{3.16}$$

Once we get farther into it, we will always work with the γ^μ's, which I will call the Dirac matrices, rather than $\boldsymbol{\alpha}$ and β. The Dirac matrices are

$$\gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad \gamma^0 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}. \tag{3.17}$$

It is not hard to show that the Dirac matrices anti-commute with each other, yielding the very useful relations

$$\left\{\gamma^\mu, \gamma^\nu\right\} = 2g^{\mu\nu} \ . \tag{3.18}$$

B. Solving the Free Dirac Equation, and the Prediction of Anti-Particles

We are ready to start solving the Dirac equations. As a first attempt, let's solve it for a particle with **p** = 0, so we are looking for solutions that look like $\Psi = ue^{-iEt}$, where *u* is a constant vector. Substituting this into (3.5), we have

$$Eu = m\beta u \,. \tag{3.19}$$

It is easy to find solutions of this equation. The eigenvalues of β are ±1, so we can show that there are two linearly independent solutions with $E = m$ and two with $E = -m$.

Let's focus for the moment on the positive energy solutions. These solutions take the form

$$u(\mathbf{p}=0) = \begin{pmatrix} \psi \\ -\psi \end{pmatrix},$$

where ψ can be any two –component spinor we want. It turns out a sensible way to label things is by picking a direction $\hat{\mathbf{n}}$ and choosing ψ to be an eigenstate of $\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}$. These eigenstates were given in eq. (2.75), and have the property that $\hat{\mathbf{n}} \cdot \boldsymbol{\sigma} \psi_\pm = \pm \psi_\pm$. We can also multiply by an arbitrary constant, which I choose to be $\sqrt{m}$, so *u* works out to

$$u\left(\mathbf{p}=0, \pm\tfrac{1}{2}\right) = \begin{pmatrix} \sqrt{m}\,\psi_\pm \\ -\sqrt{m}\,\psi_\pm \end{pmatrix}. \tag{3.20}$$

It is then easy to see that if you measure the spin of this state in the direction $\hat{\mathbf{n}}$ using the spin operator **S** in (3.13), you will get the result $\pm\tfrac{1}{2}$.

Now, we have found two solutions of the form $\Psi = ue^{-imt}$, which have **p** = 0, but what if we have a moving electron? We can get the result by simply Lorentz boosting this solution. If we boost by an amount $-\phi$ along the $\hat{\mathbf{n}}$ axis, the energy will increase to $E = m\cosh\phi$ and the momentum to $\mathbf{p} = \hat{\mathbf{n}} m \sinh\phi$, so now $\hat{\mathbf{p}} = \hat{\mathbf{n}}$, and our solution will now look like

$$\Psi(x) = u(\mathbf{p}, s)\, e^{-ip\cdot x} \,. \tag{3.21}$$

We can find $u(\mathbf{p}, s)$ by boosting (3.20) using (3.14), and taking advantage of (1.39b) along the way: After a great deal of kind of neat but rather tedious algebra, we find

$$u\left(\mathbf{p}, \pm\tfrac{1}{2}\right) = \begin{pmatrix} \sqrt{E \pm p}\,\psi_\pm \\ -\sqrt{E \mp p}\,\psi_\pm \end{pmatrix}. \tag{3.22}$$

The spin in the direction you are moving, $\hat{\mathbf{p}} \cdot \mathbf{S}$, is called the *helicity* of the particle, and is identical to the chirality defined in eq. (2.78) for particles moving at the speed of light. The waves defined by (3.21) and (3.22) have helicity $\pm\tfrac{1}{2}$. We note that eq. (3.22) makes perfect sense even for massless particles, which is why I included the factor of $\sqrt{m}$ initially.[1] For massless particles, $p = E$, and looking at (3.22), particles which have a positive chirality will live entirely in the upper half (Ψ_R) and those with negative chirality will live entirely in the lower half (Ψ_L) of the Dirac spinor. That's why these pieces are called the right and left pieces.

And now it's time to face the music: We noticed, but did not deal with the fact that for particles at rest, eq. (3.19) has negative energy solutions. We can find them, and we can boost them to find new solutions. They look like

$$\Psi(\mathbf{x},t) = v(\mathbf{p},s)e^{ip\cdot x}, \qquad (3.23)$$

where $E = \sqrt{\mathbf{p}^2 + m^2}$, and

$$v\left(\mathbf{p},\pm\tfrac{1}{2}\right) = \begin{pmatrix} \sqrt{E \mp p}\,\psi_{\mp} \\ \sqrt{E \pm p}\,\psi_{\mp} \end{pmatrix}. \qquad (3.24)$$

What do these solutions mean? It isn't hard to show that they have momentum $-\mathbf{p}$ (not a problem), energy $-E$ (a big problem) and spin in the direction of $\hat{\mathbf{p}}\cdot\mathbf{S} = \mp\frac{1}{2}$.This would imply that there exist negative energy states, and hence, for example, an electron in a hydrogen atom can not only fall from the 2p to 1s state and emit a few eV of energy, it can also fall into one of these negative energy states, emitting at least an MeV of energy. Indeed, by then increasing its momentum, it could further lower the energy, ultimately reaching negatively infinite energy and emitting enough energy to destroy the universe. But there is strong experimental evidence that the universe exists, so this must not be happening.

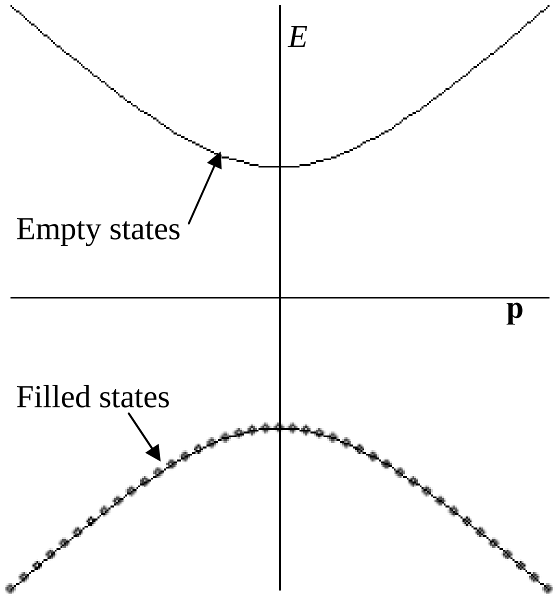


Figure 3-1: There are an infinite number of possible positive and negative energy states for the electrons, but the negative energy ones are all already filled up, and by the Pauli exclusion principle, no more electrons can go into them.

How can we reconcile this apparent problem? Dirac's brilliant solution was as follows: Since the universe tends to be in the lowest energy state, let's assume that what we call empty space, or the vacuum, isn't empty at all, but rather is filled with electrons of negative energy already, forming a *Dirac sea* of filled states. By the Pauli exclusion principle, any additional electrons cannot fall into these already occupied states, as illustrated in Fig. 3-1.

Just because these states are already filled does not mean they are irrelevant. Suppose we added energy to one of the negative energy states and promoted it to a positive energy state, as illustrated in Fig. 3-2. This positive energy state would suddenly become evident, appearing as a physical electron. But the "hole" left behind would also be evident, just as a bubble underwater is detectable by the absence of water. Because it would represent an absence of energy *E*, momentum **p**, helicity $\mp\frac{1}{2}$, and the absence of charge $-e$, such a hole would look like a particle

---

[1] Not all authors follow this normalization convention. My choice works for massless particles as well.

with energy E, momentum $\mathbf{p}$, helicity $\pm\frac{1}{2}$ and charge $+e$. At the time, the only elementary particles known were the proton, neutron, electron, and photon, and Dirac thought that it must be the proton, and was puzzled why the proton had a different mass than the electron. We now recognize that it is its own particle, often called a *positron*, but probably better called the anti-electron. Note that if an electron ever encounters an anti-electron, it can fall back into the empty "hole" and emit energy, usually in the form of two or three photons.

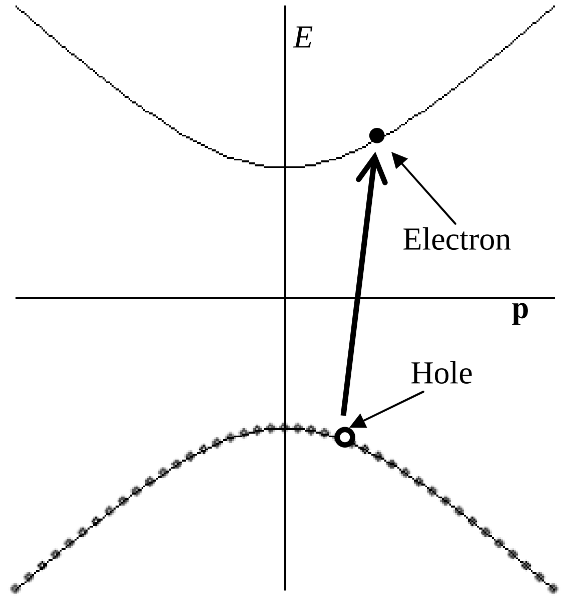


Figure 3-2: One of the negative energy states is promoted to a positive energy state. An electron appears as if by magic. The "hole" that is left behind would also be evident, and would look like a positively charged particle with positive energy.

C. Discrete Symmetries: $\mathcal{C}$, $\mathcal{P}$, and $\mathcal{T}$

We have already argued (without really proving it) that the Dirac equation is unchanged when we perform rotations and boosts. Hence it is unchanged under all *proper* Lorentz transformations. But what about the improper ones? Also note that the anti-electron has the same mass as the electron, and many similar properties. Is there some symmetry of the Dirac equation that relates particles and anti-particles? The answer is in every case yes.

Let's start with the *parity*, denoted $\mathcal{P}$, the symmetry that reverses space but leaves time unchanged. You might think that this would be as simple as taking $\mathcal{P}\Psi(\mathbf{x},t)=\Psi(-\mathbf{x},t)$, but just as proper Lorentz transformations also mix up the components of Ψ, we anticipate that parity will do so as well, so

$$\mathcal{P}\Psi(\mathbf{x},t)=P\Psi(-\mathbf{x},t), \tag{3.25}$$

where P is some matrix we have yet to determine. The goal is to make sure that if $\Psi(\mathbf{x},t)$ is a solution of the Dirac equation (3.16), so also will be $\mathcal{P}\Psi(\mathbf{x},t)$. We want

$$i\gamma^\mu\partial_\mu\left[P\Psi(-\mathbf{x},t)\right]=mP\Psi(-\mathbf{x},t).$$

Now, we can rename $\mathbf{x}\to-\mathbf{x}$ in this formula, but in this case this will cause the space derivatives to change sign, so the equation we want to be true will be

$$i\gamma^o P\partial_0\Psi(\mathbf{x},t)-i\gamma^i P\partial_i\Psi(\mathbf{x},t)=mP\Psi(\mathbf{x},t).$$

Now, we know that $\Psi(\mathbf{x},t)$ satisfies (3.16), so if we multiply this on the left by P, we get

$$iP\gamma^o\partial_0\Psi(\mathbf{x},t)+iP\gamma^i\partial_i\Psi(\mathbf{x},t)=mP\Psi(\mathbf{x},t).$$

These two equations will be the same if $P\gamma^{\iota} = -\gamma^{i} P$ and $P\gamma^{0} = \gamma^{0} P$. Is there such a matrix? From eq. (3.18) it is obvious that $P = \gamma^{0}$ satisfies these, so

$$\mathcal{P}\Psi\left(\mathbf{x},t\right) = \gamma^{0}\Psi\left(-\mathbf{x},t\right), \tag{3.26}$$

Let's tackle *charge conjugation*, or $\mathcal{C}$ next. We want to take the solutions that look like $e^{-ip\cdot x}$ and turn them into solutions that look like $e^{ip\cdot x}$. It is pretty obvious that we are going to want some sort of complex conjugation going on. Let's assume that

$$\mathcal{C}\Psi\left(\mathbf{x},t\right) = C\Psi\left(\mathbf{x},t\right)^{*}$$

for some matrix *C*. We want $\mathcal{C}\Psi\left(\mathbf{x},t\right)$ to satisfy (3.16), so we want

$$i\gamma^{\mu}\partial_{\mu}\left[C\Psi\left(\mathbf{x},t\right)^{*}\right] = mC\Psi\left(\mathbf{x},t\right)^{*} .$$

Taking the complex conjugate of this expression,, we have

$$-i\gamma^{*\mu}\, C^{*}\partial_{\mu}\left[\Psi\left(\mathbf{x},t\right)\,\right] = mC^{*}\Psi\left(\mathbf{x},t\right)^{*} .$$

This will be equivalent to (3.16) if $\gamma^{*\mu}\, C^{*} = -C^{*}\gamma^{\mu}$. It is pretty easy to see from (3.17) that all the γ^{μ}'s are real except for γ^{2}, which implies we want C^{*} to anti-commute with the other γ^{μ}'s and commute with γ^{2}. From (3.18) we note that γ^{2} has this property, as would any multiple of γ^{2}. Presumably to keep things real the convention is $C^{*} = C = i\gamma^{2}$, so

$$\mathcal{C}\Psi\left(\mathbf{x},t\right) = i\gamma^{2}\Psi\left(\mathbf{x},t\right)^{*} . \tag{3.27}$$

This leaves only time reversal $\mathcal{T}$. The obvious thing to do with time reversal is to simply change $t \to -t$, but this would change $e^{i\mathbf{p}\cdot\mathbf{x}-iEt}$ to $e^{i\mathbf{p}\cdot\mathbf{x}+iEt}$, which changes the sign of the energy but not the momentum, which is doubly wrong. We intuitively understand that if you reverse time, momentum should reverse, but energy should stay the same. This can be corrected by simultaneously changing $t \to -t$ and *also* taking the complex conjugate. We therefore expect that time reversal will look something like

$$\mathcal{T}\Psi\left(\mathbf{x},t\right) = T\Psi\left(\mathbf{x},-t\right)^{*} .$$

If you substitute this in the Dirac equation, take the complex conjugate, and change $t \to -t$, you will find that we need T^{*} to commute with γ^{0} and γ^{2} while anti-commuting with γ^{1} and γ^{3}. It isn't hard to show that $T^{*} = \Sigma^{2}$ has this property, where $\mathbf{\Sigma}$ is defined in eq. (3.13). I don't know how the phase is chosen, but this means $T = -\Sigma^{2}$, so we have

$$\mathcal{T}\Psi\left(\mathbf{x},t\right) = -\Sigma^{2}\Psi\left(\mathbf{x},-t\right)^{*} . \tag{3.28}$$

Thus parity, charge conjugation, and time reversal are all symmetries of the free Dirac equation. It turns out they are all symmetries of the Dirac equation when you include electromagnetism as well. At one time, it was thought that $\mathcal{C}$, $\mathcal{P}$, and $\mathcal{T}$ are all symmetries of nature, and all theories of particles would naturally respect these symmetries and their various combinations. It is now known that this is false. However, one combinations, the product of all three or $\mathcal{CPT}$ is believed to definitely be a symmetry of nature. Basically, this conclusion is derived simply from assuming that the Hamiltonian is Hermitian plus Lorentz invariance under proper Lorentz transformations. It is also well demonstrated experimentally.

D. Building Lorentz Scalars and Vectors from Dirac Spinors

In ordinary quantum mechanics, when we perform a rotation or a Galilean boost, the magnitude of the wave function $|\Psi|^2$ does not change. But it is easy to see using (3.21) and (3.22) that

$$\Psi^\dagger\Psi = u^\dagger u = 2E \, . \tag{3.29}$$

The same is true for the anti-particle solutions (3.23) and (3.24). Why did this happen? It turns out this is not surprising. Think of $\Psi^\dagger\Psi$ as the probability density of a particle. For a particle at rest, this is $\Psi^\dagger\Psi = 2m$. Let's say this is spread out over a volume *V*. When we perform a Lorentz boost of this wave function, the region where the particle is located will shrink in the direction of motion by a factor of the Lorentz factor $\gamma = E/m$. Hence the probability *density* must increase by this amount, $\Psi^\dagger\Psi \to (E/m)2m = 2E$.

There is another reason to see that this was inevitable. When we perform a rotation or Lorentz boost we know from (3.12) and (3.14) that Ψ transforms as

$$\begin{aligned} \text{rotation:} \quad & \Psi \to \exp\left(-\tfrac{1}{2} i\boldsymbol{\Sigma}\cdot\hat{\mathbf{n}}\theta\right)\Psi \, , \\ \text{boost:} \quad & \Psi \to \exp\left(\tfrac{1}{2}\boldsymbol{\alpha}\cdot\hat{\mathbf{n}}\phi\right)\Psi \, . \end{aligned} \tag{3.30}$$

Taking the Hermitian conjugate of these equations, and keeping in mind that $\boldsymbol{\Sigma}$ and $\boldsymbol{\alpha}$ are Hermitian matrices, we see that

$$\begin{aligned} \text{rotation:} \quad & \Psi^\dagger \to \Psi^\dagger \exp\left(\tfrac{1}{2} i\boldsymbol{\Sigma}\cdot\hat{\mathbf{n}}\theta\right) , \\ \text{boost:} \quad & \Psi^\dagger \to \Psi^\dagger \exp\left(\tfrac{1}{2}\boldsymbol{\alpha}\cdot\hat{\mathbf{n}}\phi\right) . \end{aligned} \tag{3.31}$$

Putting (3.30) and (3.31) together, we see that under a rotation or boots, $\Psi^\dagger\Psi$ transforms as

$$\begin{aligned} \text{rotation:} \quad & \Psi^\dagger\Psi \to \Psi^\dagger\Psi \, , \\ \text{boost:} \quad & \Psi^\dagger\Psi \to \Psi^\dagger \exp\left(\boldsymbol{\alpha}\cdot\hat{\mathbf{n}}\phi\right)\Psi \, . \end{aligned}$$

So it is unchanged under rotations but not boosts. This is not the way a scalar behaves, but rather how the time component of a vector behaves; for example, energy is like this.

Can we make a scalar quantity out of $\Psi^\dagger$ and Ψ? We will need to do so in chapter 6 when we come back to fermions. The idea is to come up with some quantity *S* of the form

$$S = \Psi^\dagger M \Psi \, ,$$

where *M* is a matrix yet to be determined. Under rotations and boosts, this changes to:

$$\text{rotation:}\quad \Psi^\dagger M\Psi \to \Psi^\dagger \exp\left(\tfrac{1}{2}i\mathbf{\Sigma}\cdot\hat{\mathbf{n}}\theta\right) M \exp\left(-\tfrac{1}{2}i\mathbf{\Sigma}\cdot\hat{\mathbf{n}}\theta\right)\Psi\,,$$

$$\text{boost:}\quad \Psi^\dagger M\Psi \to \Psi^\dagger \exp\left(\tfrac{1}{2}\boldsymbol{\alpha}\cdot\hat{\mathbf{n}}\phi\right) M \exp\left(\tfrac{1}{2}\boldsymbol{\alpha}\cdot\hat{\mathbf{n}}\phi\right)\Psi\,.$$

It isn't hard to see that this will simplify to $\Psi^\dagger M\Psi$ provided we can find a matrix M that commutes with $\mathbf{\Sigma}$ and anti-commutes with $\boldsymbol{\alpha}$. We already know that γ^0 anti-commutes with $\boldsymbol{\alpha}$, and it isn't hard to show it commutes with $\mathbf{\Sigma}$, so the quantity

$$S \equiv \Psi^\dagger\gamma^0\Psi \tag{3.32}$$

is, in fact, a scalar, at least under proper Lorentz transformations.

The quantity $\Psi^\dagger\gamma^0$ comes up so often that it is given its own symbol. For any Dirac spinor, we define

$$\bar{\Psi} \equiv \Psi^\dagger. \tag{3.33}$$

We will be using bars more often than daggers when working with fermions. Under a Lorentz transformation, $\bar{\Psi}$ transforms according to

$$\text{rotation:}\quad \bar{\Psi} \to \bar{\Psi}\exp\left(\tfrac{1}{2}i\mathbf{\Sigma}\cdot\hat{\mathbf{n}}\theta\right), \qquad \text{boost:}\quad \bar{\Psi} \to \bar{\Psi}\exp\left(-\tfrac{1}{2}\boldsymbol{\alpha}\cdot\hat{\mathbf{n}}\phi\right) \tag{3.34}$$

The scalar quantity S can then be written as $S = \bar{\Psi}\Psi$, and is a scalar under Lorentz transformations. Are there any other combinations that work? Consider the *chirality matrix*

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.35}$$

This chirality matrix has eigenvalue +1 on Ψ_R and -1 on Ψ_L. Because it has identity matrices in a block-diagonal form, and both $\mathbf{\Sigma}$ and $\boldsymbol{\alpha}$ are block diagonal, it commutes with both of them. It therefore isn't hard to show that the combination

$$P \equiv \bar{\Psi}\gamma_5\Psi \tag{3.36}$$

also behaves as a scalar under proper Lorentz transformations.

What about improper ones? Let's focus on parity. Ignoring the space dependence, parity transforms $\mathcal{P}\Psi = \gamma^0\Psi$, and taking the bar of this quantity,

$$\overline{\mathcal{P}\Psi} = \left(\gamma^0\Psi\right)^\dagger\gamma^0 = \Psi^\dagger\gamma^0\gamma^0 = \bar{\Psi}\gamma^0\,.$$

Under parity, then,

$$\Psi \to \gamma^0\Psi \quad \text{and} \quad \bar{\Psi} \to \bar{\Psi}\gamma^0\,.$$

We therefore see that under parity,

$$S = \bar{\Psi}\Psi \to \bar{\Psi}\gamma^0\gamma^0\Psi = \bar{\Psi}\Psi = S\,.$$

A scalar that also stays the same under parity is called a *true scalar*. It is left as an exercise (problem 3.3) to determine if P is a true scalar or not.

Can we get anything more interesting than scalars, like, for example, vectors? We already commented that $\Psi^\dagger\Psi$ transforms like the time component of a vector If we rewrite this as $\Psi^\dagger\gamma^0\gamma^0\Psi = \bar{\Psi}\gamma^0\Psi$ we get a good hint of how to make a vector quantity. Define

$$V^\mu \equiv \bar{\Psi}\gamma^\mu\Psi\,. \tag{3.37}$$

Then it can be shown that this quantity transforms as a vector under proper Lorentz transformations. Finally, since γ_5 commutes with the spin and boost matrices, we define the quantity

$$A^\mu \equiv \bar{\Psi}\gamma^\mu\gamma_5\Psi\,, \tag{3.38}$$

which is also a vector under proper Lorentz transformations. The properties of V^μ and A^μ under parity are also left as an exercise (problem 3.3).

E. Working with Dirac Matrices

We will later be doing a lot of work with Dirac matrices, so let's summarize what we know about them. First, the four Dirac matrices satisfy anti-commutation relations (3.18), and similar relations for γ_5:

$$\left\{\gamma^\mu,\gamma^\nu\right\} = 2g^{\mu\nu}\,, \quad \left\{\gamma_5,\gamma^\mu\right\} = 0\,, \quad \gamma_5^2 = 1\,. \tag{3.39}$$

It is easy to see that only some of the Dirac matrices are Hermitian, in fact,

$$\gamma^{0\dagger} = \gamma^0\,, \quad \gamma^{i\dagger} = -\gamma^i\,, \quad \gamma_5^\dagger = \gamma_5\,. \tag{3.40}$$

We will often need to take the complex conjugate of expressions like

$$\bar{\Psi}_A\Gamma_1\Gamma_2\cdots\Gamma_n\Psi_B$$

where Ψ_A and Ψ_B are Dirac spinors with four components, and $\Gamma_1,\Gamma_2,\ldots,\Gamma_n$ are some collection of matrices. Note that this is a number, not a matrix, but we can take the complex conjugate for a number by taking the Hermitian conjugate. We first reinsert $\bar{\Psi}_A = \Psi_A^\dagger\gamma^0$, then we have

$$\left(\bar{\Psi}_A\Gamma_1\Gamma_2\cdots\Gamma_n\Psi_B\right)^* = \left(\Psi_A^\dagger\gamma^0\Gamma_1\Gamma_2\cdots\Gamma_n\Psi_B\right)^\dagger = \Psi_B^\dagger\Gamma_n^\dagger\cdots\Gamma_2^\dagger\Gamma_1^\dagger\gamma^{0\dagger}\Psi_A\,.$$

We now take advantage of the fact that $\gamma^{0\dagger} = \gamma^0$, and use $\gamma^0\gamma^0 = 1$ to insert a pair of γ^0's between every pair of matrices above to turn it into

$$\left(\bar{\Psi}_A\Gamma_1\Gamma_2\cdots\Gamma_n\Psi_B\right)^* = \Psi_B^\dagger\gamma^0\gamma^0\Gamma_n^\dagger\gamma^0\cdots\gamma^0\Gamma_2^\dagger\gamma^0\gamma^0\Gamma_1^\dagger\gamma^0\Psi_A\,. \tag{3.41}$$

We replace $\Psi_B^\dagger\gamma^0 = \bar{\Psi}_B$, and for any matrix Γ we define

$$\bar{\Gamma} \equiv \gamma^0\Gamma\gamma^0\,. \tag{3.42}$$

We then use this in (3.41) to write

$$\left(\bar{\Psi}_A\Gamma_1\Gamma_2\cdots\Gamma_n\Psi_B\right)^* = \bar{\Psi}_B\bar{\Gamma}_n\cdots\bar{\Gamma}_2\bar{\Gamma}_1\Psi_A \,. \tag{3.43}$$

It isn't hard to show that

$$\bar{\gamma}^\mu = \gamma^\mu \quad \text{and} \quad \bar{\gamma}_5 = -\gamma_5 \,. \tag{3.44}$$

In summary, to take the complex conjugate of some complicated expression with matrices and Dirac spinors, we simply bar it, which is done by

- Reverse the order of everything;
- Change any barred spinors to unbarred, of unbarred to barred;
- Bar any matrices; and
- Take the complex conjugate of any complex numbers.

For example, suppose we were asked to simplify the expression

$$A = i\bar{\Psi}\gamma^\mu\gamma^\nu\gamma_5\Psi + \left(i\bar{\Psi}\gamma^\mu\gamma^\nu\gamma_5\Psi\right)^* .$$

We would first rewrite the complex conjugate to give

$$A = i\bar{\Psi}\gamma^\mu\gamma^\nu\gamma_5\Psi - i\bar{\Psi}\bar{\gamma}_5\bar{\gamma}^\nu\bar{\gamma}^\mu\Psi \,,$$

which with the help of (3.44) simplifies to

$$A = i\bar{\Psi}\gamma^\mu\gamma^\nu\gamma_5\Psi + i\bar{\Psi}\gamma_5\gamma^\nu\gamma^\mu\Psi \,.$$

We then use the fact that γ_5 anti-commutes with the Dirac matrices to rewrite

$$\gamma_5\gamma^\nu\gamma^\mu = -\gamma^\nu\gamma_5\gamma^\mu = \gamma^\nu\gamma^\mu\gamma_5$$

so that

$$A = i\bar{\Psi}\left(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu\right)\gamma_5\Psi \,.$$

We then use (3.39) to simplify this to

$$A = 2ig^{\mu\nu}\bar{\Psi}\gamma_5\Psi \,.$$

Let's get a couple of other expressions that will be useful later on. We know that $\Psi = u(\mathbf{p},s)e^{-ip\cdot x}$ is a solution of the Dirac equation, $i\gamma^\mu\partial_\mu\Psi = m\Psi$. Substituting this in, we have

$$p_\mu\gamma^\mu u(\mathbf{p},s) = mu(\mathbf{p},s) \,. \tag{3.45}$$

We can also take the bar of this equation, which yields

$$\bar{u}(\mathbf{p},s)\gamma_\mu p^\mu = \bar{u}(\mathbf{p},s)m \,. \tag{3.46}$$

Combinations of the Dirac matrices are common enough that it is very helpful to define the *Feynman slash notation,*

$$\not{p} \equiv p_\mu\gamma^\mu \,. \tag{3.47}$$

In this notation, (3.45) and (3.46) would be written

$$\not{p}u(\mathbf{p},s) = mu(\mathbf{p},s) \quad \text{and} \quad \bar{u}(\mathbf{p},s)\,\not{p} = \bar{u}(\mathbf{p},s)m\,. \tag{3.48}$$

In a similar way, we can use the fact that $\Psi = v(\mathbf{p},s)e^{ip\cdot x}$ is a solution of the Dirac equation to show that

$$\not{p}v(\mathbf{p},s) = -mv(\mathbf{p},s) \quad \text{and} \quad \bar{v}(\mathbf{p},s)\,\not{p} = -\bar{v}(\mathbf{p},s)m\,. \tag{3.49}$$

Two other identities will prove useful when we return to fermions in chapter 6. Consider the expression $u(\mathbf{p},s)\bar{u}(\mathbf{p},s)$, which is a 4×4 matrix. We write this out explicitly:

$$\begin{aligned} u\left(\mathbf{p},\pm\tfrac{1}{2}\right)\bar{u}\left(\mathbf{p},\pm\tfrac{1}{2}\right) &= \begin{pmatrix} \sqrt{E\pm p}\psi_\pm \\ -\sqrt{E\mp p}\psi_\pm \end{pmatrix} \begin{pmatrix} \sqrt{E\pm p}\psi_\pm^\dagger & -\sqrt{E\mp p}\psi_\pm^\dagger \end{pmatrix} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{E^2-p^2}\psi_\pm\psi_\pm^\dagger & -(E\pm p)\psi_\pm\psi_\pm^\dagger \\ (E\mp p)\psi_\pm\psi_\pm^\dagger & \sqrt{E^2-p^2}\psi_\pm\psi_\pm^\dagger \end{pmatrix} = \begin{pmatrix} m\psi_\pm\psi_\pm^\dagger & (-E\mp p)\psi_\pm\psi_\pm^\dagger \\ (-E\pm p)\psi_\pm\psi_\pm^\dagger & m\psi_\pm\psi_\pm^\dagger \end{pmatrix}. \end{aligned}$$

We will often want to sum this over spins. With the help of (2.76) and (2,77), we can then show that

$$\begin{aligned} \sum_s u(\mathbf{p},s)\bar{u}(\mathbf{p},s) &= \begin{pmatrix} m\left(\psi_+\psi_+^\dagger + \psi_-\psi_-^\dagger\right) & -E\left(\psi_+\psi_+^\dagger + \psi_-\psi_-^\dagger\right) + p\left(\psi_+\psi_+^\dagger - \psi_-\psi_-^\dagger\right) \\ -E\left(\psi_+\psi_+^\dagger + \psi_-\psi_-^\dagger\right) + p\left(\psi_+\psi_+^\dagger - \psi_-\psi_-^\dagger\right) & m\left(\psi_+\psi_+^\dagger + \psi_-\psi_-^\dagger\right) \end{pmatrix} \\ &= \begin{pmatrix} m & -E - p\hat{\mathbf{p}}\cdot\boldsymbol{\sigma} \\ -E + p\hat{\mathbf{p}}\cdot\boldsymbol{\sigma} & m \end{pmatrix} = m\mathbf{1} + E\gamma^0 - p\hat{\mathbf{p}}\cdot\boldsymbol{\gamma} = p_0\gamma^0 + p_i\gamma^i + m = \not{p} + m \end{aligned}$$

A similar relation can be found for the *v*'s. Putting these together, we have

$$\sum_s u(\mathbf{p},s)\bar{u}(\mathbf{p},s) = \not{p} + m\,, \quad \sum_s v(\mathbf{p},s)\bar{v}(\mathbf{p},s) = \not{p} - m\,. \tag{3.50}$$

For now we have exhausted the information we want to know about the free Dirac equation. But to give us an idea of what is coming up, let's couple it to an electromagnetic field.

F. The Dirac Equation with Electromagnetic Fields

We would now like to include electromagnetic fields interacting with an electron. We start with the original Dirac equation (3.5). In the presence of an electrostatic potential $\Phi(x)$, an electron at point *x* would acquire a potential energy $-e\Phi$. It makes sense to modify (3.5) to yield

$$i\partial_0\Psi = \left[-i\boldsymbol{\alpha}\cdot\nabla + m\beta - e\Phi\right]\Psi.$$

As before, we now multiply this by $\beta = \gamma^0$ on the left, and use $\Phi = A^0 = A_0$ together with the definitions (3.15) to rewrite this as

$$i\gamma^\mu\partial_\mu\Psi + e\gamma^0 A_0\Psi = m\Psi\,.$$

This equation is obviously not Lorentz covariant, but this is hardly surprising, since we only included the effects of the scalar potential Φ and ignored the vector potential **A**. The obvious extension of this equation is then

$$\gamma^{\mu}\left(i\partial_{\mu}+eA_{\mu}\right)\Psi=m\Psi\,. \tag{3.51}$$

This is, indeed, the correct equation for an electron. If we were working with a different fermion with charge Qe, we would simply replace $e\rightarrow -Qe$ in (3.51).

It is helpful to define the *covariant derivative* for a charged particle as

$$D_{\mu}\equiv\partial_{\mu}+ieQA_{\mu}\,. \tag{3.52}$$

The Dirac equation can then be written at its most compact in terms of this derivative, namely,

$$\left(i\not{D}-m\right)\Psi=0\,. \tag{3.53}$$

Covariant derivatives have some remarkable properties. The first thing I want to notice about them is that they do not commute. Specifically,

$$\begin{aligned}\left[D_{\mu},D_{\nu}\right]\Psi&=\left(\partial_{\mu}+ieQA_{\mu}\right)\left(\partial_{\nu}+ieQA_{\nu}\right)\Psi-\left(\partial_{\nu}+ieQA_{\nu}\right)\left(\partial_{\mu}+ieQA_{\mu}\right)\Psi\\&=ieQ\partial_{\mu}\left(A_{\nu}\Psi\right)+ieQA_{\mu}\partial_{\nu}\Psi-ieQ\partial_{\nu}\left(A_{\mu}\Psi\right)-ieQA_{\nu}\partial_{\mu}\Psi\\&=ieQ\left(\partial_{\mu}A_{\nu}\right)\Psi-ieQ\left(\partial_{\nu}A_{\mu}\right)\Psi\,,\\\left[D_{\mu},D_{\nu}\right]\Psi&=ieQF_{\mu\nu}\Psi\,.\end{aligned} \tag{3.54}$$

We will use this property later when we consider other gauge theories, such as QCD in chapter 9.

The other remarkable property of the covariant derivative has to do with gauge invariance. Recall from eq. (2.31) that electromagnetic fields are unchanged when you perform a gauge transformation $A_{\mu}\rightarrow A'_{\mu}=A_{\mu}+\partial_{\mu}\chi$. Keep in mind that χ is a completely arbitrary function of *x*. It should be obvious that we now have $\left(i\not{D}'-m\right)\Psi\neq 0$.. However, let us simultaneously change the phase of Ψ by

$$\Psi\rightarrow\Psi'=e^{-iQe\chi}\Psi\,. \tag{3.55}$$

Then the covariant derivative of Ψ' will be

$$\begin{aligned}D'_{\mu}\Psi'&=\left(\partial_{\mu}+ieQA'_{\mu}\right)\left[e^{-iQe\chi}\Psi\right]=e^{-iQe\chi}\left(\partial_{\mu}+ieQA'_{\mu}\right)\Psi-ieQe^{-iQe\chi}\left(\partial_{\mu}\chi\right)\Psi\\&=e^{-iQe\chi}\left(\partial_{\mu}+ieQA_{\mu}\right)\Psi\,.\end{aligned}$$

It follows that

$$i\not{D}'\Psi'=\exp\left(-iQe\chi\right)\not{D}\Psi=m\exp\left(-iQe\chi\right)\Psi=m\Psi'\,,$$

i.e., Ψ' satisfies the Dirac equation in the new gauge. Hence a gauge transformation transforms not just the vector potential, but also the wave function as well:

$$A_{\mu}\rightarrow A'_{\mu}=A_{\mu}+\partial_{\mu}\chi\,, \tag{3.56a}$$

$$\Psi\rightarrow\Psi'=e^{-iQe\chi}\Psi\,. \tag{3.56b}$$

Let's explore one interesting consequence of the Dirac equation, magnetic dipole moments. Act on the left of eq. (3.53) with $i\not{D}+m$. The result is

$$0=\left(i\not{D}+m\right)\left(i\not{D}-m\right)\Psi=-\gamma^{\mu}\gamma^{\nu}D_{\mu}D_{\nu}\Psi-m^{2}\Psi . \tag{3.57}$$

Now, rewrite the product of the two Dirac matrices as

$$\gamma^{\mu}\gamma^{\nu}=\tfrac{1}{2}\left\{\gamma^{\mu},\gamma^{\nu}\right\}+\tfrac{1}{2}\left[\gamma^{\mu},\gamma^{\nu}\right]=g^{\mu\nu}+\tfrac{1}{2}\left[\gamma^{\mu},\gamma^{\nu}\right]$$

Substitute this into eq. (3.57) to yield

$$-g^{\mu\nu}D_{\mu}D_{\nu}\Psi-\tfrac{1}{2}\left(\gamma^{\mu}\gamma^{\nu}-\gamma^{\nu}\gamma^{\mu}\right)D_{\mu}D_{\nu}\Psi=m^{2}\Psi .$$

Since the indices μ and ν are dummy indices that are being summed over, we can switch them $\mu\leftrightarrow\nu$ on one term to rewrite this as

$$-D^{2}\Psi=\tfrac{1}{2}\left(\gamma^{\mu}\gamma^{\nu}D_{\mu}D_{\nu}-\gamma^{\mu}\gamma^{\nu}D_{\nu}D_{\mu}\right)\Psi+m^{2}\Psi=\tfrac{1}{2}\gamma^{\mu}\gamma^{\nu}\left[D_{\mu},D_{\nu}\right]\Psi+m^{2}\Psi ,$$

$$-D^{2}\Psi=\tfrac{1}{2}ieQ\gamma^{\mu}\gamma^{\nu}F_{\mu\nu}\Psi+m^{2}\Psi . \tag{3.58}$$

If we ignore the first term on the right of eq. (3.58), this is just the covariant version of the Klein-Gordon equation (3.4). What is that new term? To understand it, let's imagine we have a weak magnetic field so that $F_{ij}=F^{ij}=-\varepsilon^{ijk}B^{k}$. Let's imagine a particle at rest, so $\Psi\propto e^{-iEt}$ and assume we have no scalar potential, and assume weak fields, so we can ignore effects proportional to $\mathbf{A}^2$. Then eq. (3.58) becomes

$$E^{2}\Psi=-\tfrac{1}{2}ieQ\gamma^{i}\gamma^{j}\varepsilon^{ijk}B^{k}\Psi+m^{2}\Psi . \tag{3.59}$$

With a bit of work, one can then show $\tfrac{1}{2}i\gamma^{i}\gamma^{j}\varepsilon^{ijk}=\Sigma^{k}$, defined by eq. (3.13), so this simplifies to

$$E^{2}\Psi=-eQ\mathbf{B}\cdot\mathbf{\Sigma}\Psi+m^{2}\Psi=m^{2}\Psi-2eQ\mathbf{B}\cdot\mathbf{S}\Psi . \tag{3.60}$$

It is clear from eq. (3.60) that the magnetic field causes a shift in the energy based on the spin state. If we replace $\mathbf{B}\cdot\mathbf{S}$ by its eigenvalue, which we give the same name, we cancel Ψ and obtain

$$E=\sqrt{m^{2}-2eQ\mathbf{B}\cdot\mathbf{S}}\approx m-\frac{eQ}{m}\mathbf{B}\cdot\mathbf{S} , \tag{3.61}$$

where we used the weak field limit to Taylor expand our expression to first order in the magnetic field. The first term in eq. (3.61) is just the rest energy, the second is the *magnetic dipole* interaction. It simply says that a spinning charged particle is like a little magnet. In general, for a particle, we define the *g*-factor by stating that the magnetic shift in energy is

$$\Delta E=-\frac{geQ}{2m}\mathbf{B}\cdot\mathbf{S} . \tag{3.62}$$

The classical non-relativistic prediction is g =1, the Dirac equation predicts g = 2. Some experimental values are g = 2.00232 for the electron, g = 2.00230 for the muon, and g = 5.5857 for the proton. Clearly, the prediction is good but not perfect for the electron and the muon, and totally worthless for the proton. Ultimately, this is a clue that the proton is not elementary.

G. Scattering From Electromagnetic Fields

I'd now like to calculate the probability of scattering an electron in the presence of an electromagnetic field. Rather than working with an arbitrary electromagnetic field, I will work with a field of the form

$$A^\mu(x) = a^\mu e^{-ik\cdot x} , \tag{3.63}$$

where a^μ is an arbitrary constant vector. We wish to think of a^μ as small, so we can use perturbation theory. The vector k should be thought of as completely arbitrary and unrestricted, because this isn't necessarily a wave, but rather an electromagnetic field produced by some unknown source. It is known that an arbitrary vector field can be written as sums or integrals of expressions like (3.63), so this is in fact pretty general. Substituting this into (3.51) yields[1]

$$\gamma^\mu \left(i\partial_\mu + e a_\mu e^{-ik\cdot x} \right) \Psi = m\Psi ,$$

or, multiplying by γ^0 on the left and rearranging a bit,

$$i\partial_0 \Psi = -i\boldsymbol{\alpha}\cdot\nabla\Psi + \beta m\Psi - e^{-ik\cdot x} e\gamma^0 \not{a}\Psi .$$

The right hand side must be $H\Psi$. We split up the Hamiltonian into a main piece and an interaction piece, namely

$$H = H_0 + H_{\text{int}} , \quad H_0 = -i\boldsymbol{\alpha}\cdot\nabla + \beta m , \quad H_{\text{int}} = -e^{-ik\cdot x} e\gamma^0 \not{a}\Psi \ .$$

The unperturbed Hamiltonian H_0 just corresponds to the free Dirac equation. For now we'll choose our initial state $|I\rangle$ to be a positive energy solution with momentum **p** and spin s, so

$$\Psi_I = \frac{1}{\sqrt{2EV}} u(p,s) e^{-ip\cdot x} ,$$

where V is the volume of the universe. The factors in the square root in the denominator are present to make sure that the state is properly normalized, $\langle \Psi_I | \Psi_I \rangle = 1$. Similarly, the final state will be chosen to be

$$\Psi_F = \frac{1}{\sqrt{2E'V}} u(p',s') e^{-ip'\cdot x} .$$

We now wish to use our formulas from time-dependent perturbation theory, section 2E. We start by finding the transition amplitude (2.51), which we will calculate only to first order:

$$\begin{aligned} \mathcal{T}_{FI} &\equiv -i \langle \Psi_F | H_{\text{int}} | \Psi_I \rangle = -i(-e) \int d^3\mathbf{x}\, e^{-ik\cdot x} \Psi_F^\dagger(x) \gamma^0 \not{a} \Psi_I(x) \\ &= \frac{ie}{V\sqrt{2E}\sqrt{2E'}} \int d^3\mathbf{x}\, e^{-ik\cdot x} e^{ip'\cdot x} \bar{u}(\mathbf{p}',s') \not{a} u(\mathbf{p},s) e^{-ip\cdot x} , \end{aligned}$$

---

[1] Note that e is used here both as the fundamental charge and the base of the natural logarithm. You can distinguish the latter from the former because it is always raised to some interesting power.

$$\mathcal{T}_{FI} = \frac{ie}{V\sqrt{2E}\sqrt{2E'}} e^{-i\left(E+k^0-E'\right)t}\left[\bar{u}\left(\mathbf{p}',s'\right)\not{A}u\left(\mathbf{p},s\right)\right]\int d^3\mathbf{x}\, e^{i\mathbf{p}\cdot\mathbf{x}+i\mathbf{k}\cdot\mathbf{x}-i\mathbf{p}'\cdot\mathbf{x}} ,$$

$$\mathcal{T}_{FI} = \frac{ie}{\sqrt{2E}\sqrt{2E'}} e^{-i\left(E+k^0-E'\right)t}\left[\bar{u}\left(\mathbf{p}',s'\right)\not{A}u\left(\mathbf{p},s\right)\right]\delta_{\mathbf{p}',\mathbf{p}+\mathbf{k}} , \tag{3.64}$$

where we used eq. (2.60a) for the last step. We then want to substitute this expression into (2.53) to get the final answer, but we have a slight problem because that equation was derived assuming that Hamiltonian had no explicit time dependence. Hence it assumed that (3.58) would contain a factor like $e^{-i(E-E')t}$ instead of the additional factor of e^{-ik^0t}. It is obvious, however, that this will only change the final energy delta function into $\delta\left(E+k^0-E'\right)$, so we have

$$P\left(I\to F\right) = 2\pi T\delta\left(E'-E-k^0\right)\left|\mathcal{T}_{FI}\right|^2 = \frac{2\pi T}{4EE'}\delta\left(E'-E-k^0\right)\left|\bar{u}\left(\mathbf{p}',s'\right)\not{A}u\left(p,s\right)\right|^2\delta_{\mathbf{p}',\mathbf{p}+\mathbf{k}} .$$

Note that the Kronecker delta-function is unaffected by squaring, since it is always zero or one. We can go to the infinite volume limit by using eq. (2.62) to give us

$$P\left(I\to F\right) = \left(2\pi\right)^4\frac{T}{4EE'V}\delta^4\left(p'-p-k\right)\left|\bar{u}\left(\mathbf{p}',s'\right)\not{A}u\left(p,s\right)\right|^2 . \tag{3.65}$$

The factors of time T and volume of the universe V may look a little puzzling in this formula, but we will deal with them in chapter 4. For now, I want you to notice that the factor $\delta^4\left(p'-p-k\right)$ simply tells us that the external electromagnetic field is adding momentum **k** and energy k^0 to the electron. In Fig. 3-3, I show what is going on, keeping track of the Dirac sea of negative energy electrons. The negative energy electrons simply stay put as the positive energy electron scatters to a new positive energy state.

In anticipation of future work, I would like to introduce a diagrammatic representation of eq. (3.59), very similar to what we will later call Feynman diagrams. Let's denote an electron by an arrow going to the right and a positron by an arrow going to the left. Time will increase from left to right. We'll use a squiggly line to denote the electromagnetic field, and a dot to denote the interaction. I also marked the four-momentum of all the particles and the field, and drew an arrow to remind us that the electromagnetic momentum is flowing in. The process would be sketched as shown in Fig. 3-4.

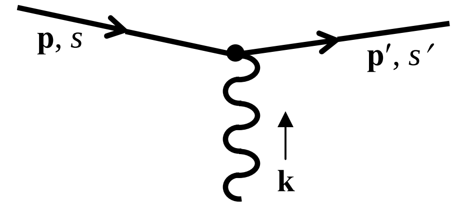


Figure 3-4: A Feynman-like diagrammatic depiction of the process described by eq. (3.65)

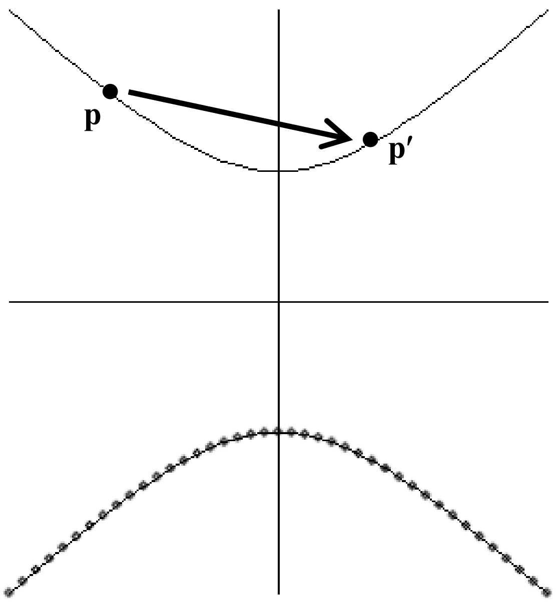


Figure 3-3: An electron scatters to a new momentum, as described by eq. (3.65).

There is no reason to restrict ourselves to situation where both the initial and final states have positive energy. For example, let's let the initial state be a negative energy state,

$$\Psi_I = \frac{1}{\sqrt{2EV}} v(p,s) e^{ip\cdot x} .$$

In terms of the computation, it is clear we can immediately get the probability by taking $u \to v$ and $p \to -p$, so (3.65) becomes

$$P(I \to F) = (2\pi)^4 \frac{T}{4EE'V} \delta^4(p' + p - k) \left| \bar{u}(\mathbf{p}', s') \not{A} v(\mathbf{p}, s) \right|^2 . \tag{3.66}$$

We have just calculated a probability, but the probability of what? We have converted a negative energy electron to a positive one, much as we sketched back in Fig. 3-2. Experimentally, we would see initially an empty universe (save for the electromagnetic fields), and in the end we would see an electron of momentum **p**′ and a positron of momentum **p**. Diagrammatically, this would lead to Fig. 3-5.

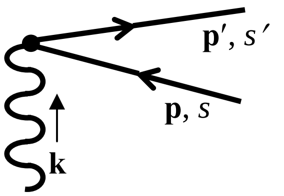


Figure 3-5: A Feynman-like diagrammatic depiction of the process described by eq. (3.66)

There is no reason we can't do the same thing with the final state particle as well. If we make it negative energy, eq. (3.65) becomes

$$P(I \to F) = \frac{(2\pi)^4 T}{4EE'V} \delta^4(-p' - p - k) \left| \bar{v}(\mathbf{p}', s') \not{A} u(\mathbf{p}, s) \right|^2 . \tag{3.67}$$

In this case, the initial state is an electron, but the final state fills a negative energy state that must have been empty beforehand. Hence experimentally, it started as a state with both an electron and a positron, and ends with nothing, with all the energy and momentum being absorbed by the electromagnetic field. A sketch in terms of the sea of electrons is shown in Fig. 3-6, and a Feynman diagram style sketch is in Fig. 3-7.

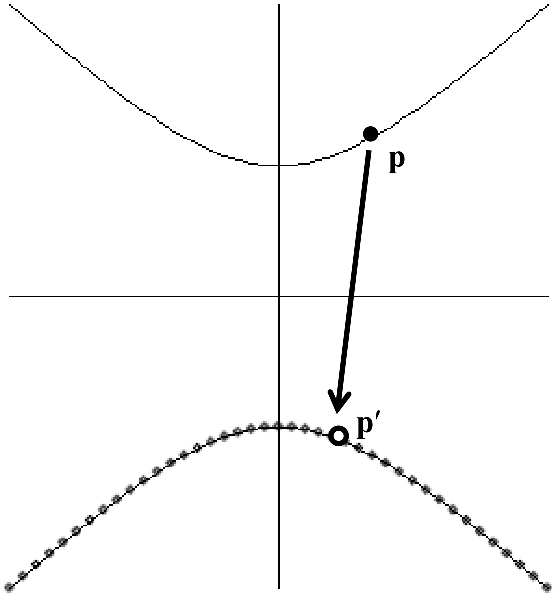


Figure 3-6: An electron annihilates a positron, as described by eq. (3.67)

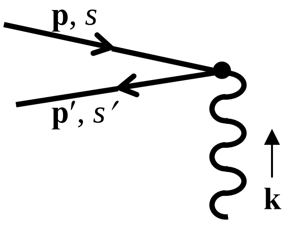


Figure 3-7: A Feynman-like diagrammatic depiction of the process described by eq. (3.67).

In many ways, the Dirac equation is quite successful. It is a quantum mechanical description of spin $\frac{1}{2}$ particles, and predicts the existence of anti-particles with the opposite charge. It approximately predicts the magnetic dipole moment of certain particles, such as the electron and muon. It is clear that the process of creating an electron is closely related to the process of eliminating a positron, and vice versa. From a diagrammatic standpoint, it is clear these are related simply by twisting the lines representing electrons or positrons around to make one change into the other.

On careful reflection, however, the Dirac theory must be incomplete. When an electron scatters from one momentum state to another, it absorbed energy and momentum from the electromagnetic field. But we have not included the effects on the field anywhere in our calculations. Furthermore, the number of electrons was fixed in this theory. At the moment we have no idea how to do a process as simple as electron-electron scattering. Also, the Dirac theory works well for fermions, but we cannot use the Pauli exclusion principle to save us from the negative energy states for bosons.

We therefore abandon the Dirac theory and attempt to build up a complete theory of an arbitrary number of particles of any type. But we will use the Dirac theory as our inspiration. In particular, we will assume that anti-particles exist for *all* particles, and that their interactions are related in some simple manner to the interactions of particles. Furthermore, when we develop Feynman rules for electromagnetic interactions in Chapter 6, we will get most of the rules we need right here, in the Dirac theory.

H. The Dirac Equation with scalar fields

Because of its importance to the development of QED, we have focused exclusively on interactions with the electromagnetic field, which is a vector field. For the development of the Higgs field in chapter 11, we must also consider how to include a scalar field. This is easy to do. If $\phi(x)$ is a scalar field, we can include its interaction in the Dirac equation simply by modifying eq. (3.16) to

$$i\not\partial\Psi - g\phi\Psi - m\Psi = 0 . \tag{3.68}$$

Since the scalar field is unaffected by Lorentz transformations (proper or improper) the new term transforms exactly the same as the mass term, so this expression is still Lorentz covariant. In particular, if the field $\phi(x)$ is non-zero, it would change the mass of the Ψ. It could cause the Ψ to act like it has a mass, even if $m = 0$. This will be relevant when we discuss the Higgs field in chapter 11.

The coupling in eq. (3.68) is appropriate if $\phi(x)$ is a true scalar field. If it is a pseudoscalar, it can be shown that the appropriate modification is

$$i\not\partial\Psi - ig\phi\gamma_5\Psi - m\Psi = 0 . \tag{3.69}$$

Giving $\phi(x)$ a non-zero value still has the effect of modifying the mass, though in a less obvious way.

Problems for Chapter 3

1. Show explicitly that eq. (3.21) with (3.22) satisfies the free Dirac equation, (3.5). Then repeat with eq. (3.23) with (3.24).

2. The Dirac Matrices satisfy $\left\{\gamma^\mu,\gamma^\nu\right\}=2g^{\mu\nu}$, $\gamma^{0\dagger}=\gamma^0$ and $\gamma^{i\dagger}=-\gamma^i$. These equations are not sufficient to define them, however. We can take the chiral representation γ_C^μ defined by eqs. (3.17) and produce some other representation γ_O^μ related by $\gamma_O^\mu=U\gamma_C^\mu U^\dagger$, where U is an arbitrary 4×4 *unitary* matrix (a unitary matrix satisfies $U^\dagger U=UU^\dagger=1$).
   (a) Show that the new Dirac matrices satisfy the same anti-commutation relations as the old ones, and have the same properties under Hermitian conjugation.
   (b) Dirac originally wrote his equations in the *Dirac* representation, related to ours by
   $$\gamma_D^\mu=U\gamma_C^\mu U^\dagger \quad \text{where} \quad U=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix},$$
   where it is understood that 1 inside U stands for the 2×2 identity matrix. Find explicitly the Dirac representation.
   (c) Solve the Dirac equation for particles at rest in the Dirac representation. Why do you think people who do non-relativistic physics prefer this representation?

3. Let Ψ be any four-component spinor. Consider the four quantities
   $$S=\bar\Psi\Psi\,,\quad P=\bar\Psi\gamma_5\Psi\,,\quad V^\mu=\bar\Psi\gamma^\mu\Psi\,,\quad A^\mu=\bar\Psi\gamma^\mu\gamma_5\Psi\,.$$
   (a) Which of these are real, and which are pure imaginary?
   (b) A *true scalar* is a quantity that is unchanged under parity. A *pseudoscalar* is a quantity that goes to minus itself under parity. A *true vector*, under parity, has its time component stay the same, while its vector part changes sign. An *axial vector*, under parity, has its time component change sign, while its space component stays the same. Classify these four quantities into the corresponding categories.
   (c) Suppose that a neutrino only has components in its lower half, so that $\Psi_R=0$ and $\Psi_L\neq0$. Which of these four quantities would be non-zero for a neutrino?

4. Suppose that a neutrino only has components in its lower half, so that $\Psi_R=0$ and $\Psi_L\neq0$. Which of the following seven symmetries carry the lower half to the lower half? Those that do not cannot be symmetries of neutrinos: $\mathcal{C}$, $\mathcal{P}$, $\mathcal{T}$, $\mathcal{CP}$, $\mathcal{PT}$, $\mathcal{CT}$, $\mathcal{CPT}$.

5. Define the current density four-vector $J^\mu=\bar\Psi\gamma^\mu\Psi$, where Ψ is a solution to the Dirac equation in the presence of electromagnetic fields, eq. (3.51). Show that $\partial_\mu J^\mu=0$.

6. Show that for an arbitrary electromagnetic field $A^\mu(x)$, if $\Psi(x)$ is a solution of eq. (3.51), so also is $\mathcal{C}\Psi(x)$ if we simultaneously let $A^\mu(x)\to-A^\mu(x)$.

7. Show that for an arbitrary electromagnetic field $A^\mu(x)$, if $\Psi(x)$ is a solution of eq. (3.51), so also is $\mathcal{P}\Psi(x)$ if we simultaneously let $A(x) \to \Lambda_P A(\Lambda_P x)$, that is, $A^0(x^0, \mathbf{x}) \to A^0(x^0, -\mathbf{x})$ and $\mathbf{A}(x^0, \mathbf{x}) \to -\mathbf{A}(x^0, -\mathbf{x})$.

8. Suppose an electron coming from $x^3 = -\infty$ is impacting an electromagnetic potential given by $\mathbf{A}(x) = 0$, $A^0(x) = V\theta(x^3)$, *i.e.*, it has zero potential that jumps to V at $x^3 = 0$. Assume the initial, reflected, and transmitted waves take the form

$$\Psi_I(x) = \begin{pmatrix} \alpha_I \\ 0 \\ \beta_I \\ 0 \end{pmatrix} e^{ipx^3 - iEt}, \quad \Psi_R(x) = \begin{pmatrix} \alpha_R \\ 0 \\ \beta_R \\ 0 \end{pmatrix} e^{-ipx^3 - iEt}, \quad \Psi_T(x) = \begin{pmatrix} \alpha_T \\ 0 \\ \beta_T \\ 0 \end{pmatrix} e^{ip'x^3 - iEt}$$

The initial and reflected wave exist only for $x^3 < 0$, the transmitted wave only for $x^3 > 0$.

(a) Write out explicitly the Dirac equation for all three waves. For each wave, this should give you two equation relating α and β.

(b) For the incident and transmitted reflected waves, one can show $p^2 = E^2 - m^2$. What is the corresponding expression for p'?

(c) Find an expression for α/β for each of the three waves. For definiteness, use $E = \frac{13}{12}m$ and $eV = \frac{1}{6}m$.

(d) Assume that $\beta_I = 1$, find the other five components by assuming Ψ is continuous across the boundary.

9. In the lecture notes, we discussed what "scattering" really means in a variety of situations, such as positive energy → positive energy (scattering), negative → positive (pair creation), and positive → negative (pair annihilation). The only case we haven't discussed is what it means when negative energy → negative energy. Let $-p$ be the momentum of an electron that scatters into a state of momentum $-p'$. How do we interpret this in the hole picture? What do we see in the initial state, and in the final state? Draw both kinds of pictures (filled/empty state diagrams and Feynman-like diagrams) to show what is going on. What is the probability after time *T* that such a transition will have occurred, to linear order in the field?

10. Alice and Bob are doing electromagnetic scattering from an electromagnetic field given by eq. (3.63); however, they are working in different gauges, so $A_\mu^A(x) = a_\mu^A e^{-ik\cdot x}$ and $A_\mu^B(x) = a_\mu^B e^{-ik\cdot x}$.

(a) Given that they are related by a gauge transformation, eq. (3.56a), what can we say about the relation between a_μ^A and a_μ^B?

(b) Show that if they both use eq. (3.65), they will nonetheless get the same answer.

(c) Show that they also get the same result if they are doing pair production (eq. (3.66)) or pair annihilation (eq. (3.67)).

IV. A Relativistic Hamiltonian

In the previous chapter we developed the Dirac theory as far as it would go. We found several faults with this theory, principally that it was not designed to handle multiple particles. In this chapter, I will develop a general theory that deals with arbitrary numbers of particles. We will then set up a simple toy model, called the $\psi^*\psi\phi$ model, and use it to learn how to calculate amplitudes, probabilities, and cross sections using a technique called Feynman diagrams.

A. Fock Space

The first step in our program is to develop a set of basis states that describe an arbitrary number of particles. This space will be called *Fock space.* We start with the zero particle (vacuum) state, which we will call $\left|0\right\rangle$. We assume this state is properly normalized, so $\left\langle 0|0\right\rangle = 1$.

Now we want to consider one particle states, which I will write as $\left|t, p, s\right\rangle$, where *t* describes what type of particle we are dealing with, *p* is its four momentum, and *s* describes its spin. For example, an electron with wave function $\Psi(x) = u\left(\mathbf{p}, +\tfrac{1}{2}\right)e^{-ip\cdot x}$ will be simply written as $\left|e^-, p, +\tfrac{1}{2}\right\rangle$. This state will be normalized in a finite volume universe as

$$\left\langle t, p, s|t, p, s\right\rangle = 2EV \, .$$

The factor of 2*E* is included so that when we Lorentz transform the state, the probability density will increase due to Lorentz contraction, as discussed in section 3D.

Note that in relativistic quantum mechanics, our state vectors are normally assumed to be functions of both space and time, whereas in ordinary quantum mechanics, a state $\left|\mathbf{p}, s\right\rangle$ would have only space dependence, $e^{i\mathbf{p}\cdot\mathbf{x}}$. But note that when we did time-dependent perturbation theory in section 2E, we explicitly put the time dependence back in in equation (2.46). Hence, though our notation has changed, all the work of this section is still valid; in particular, equations (2.51) and (2.53) still work, but we have to remove the time dependence of the state vectors to use them. The easiest way to do this is to simply modify (2.51) by setting the time to zero, so we have

$$\mathcal{T}_{FI} \equiv -i\left\langle \phi_F \left| H_{\text{int}}(0) \right| \phi_I \right\rangle + (-i)^2 \sum_n \frac{\left\langle \phi_F \left| H_{\text{int}}(0) \right| \phi_n \right\rangle \left\langle \phi_n \left| H_{\text{int}}(0) \right| \phi_I \right\rangle}{i\left(E_n - E_I\right)} + \cdots . \tag{4.1}$$

But note we must be careful to work with properly normalized states in this formula.

To describe multiple particles, we simply make a list of them, so for *N* particles, a typical state will look like

$$\left|\Psi\right\rangle = \left|t_1, p_1, s_1 \, ; t_2, p_2, s_2 \, ; \cdots ; t_N, p_N, s_N\right\rangle \tag{4.2}$$

This state will be assumed to have total four-momentum

$$p = p_1 + p_2 + \cdots + p_N \, . \tag{4.3}$$

It will be normalized so that

$$\langle\Psi|\Psi\rangle = V^N\left(2E_1\right)\left(2E_2\right)\cdots\left(2E_N\right). \tag{4.4}$$

The order of the particles in (4.2) doesn't make a whole lot of difference. The one exception will be when the two particles are fermions, for which interchanging two particles will introduce a minus sign, in accordance with eq. (2.79). We therefore have

$$\left|t_1,p_1,s_1\,;t_2,p_2,s_2\right\rangle = \begin{cases} -\left|t_2,p_2,s_2\,;t_1,p_1,s_1\right\rangle & \text{if two fermions ,} \\ \left|t_2,p_2,s_2\,;t_1,p_1,s_1\right\rangle & \text{otherwise .} \end{cases} \tag{4.5}$$

The only case where this will be relevant is when the two particles are the same type, since otherwise we could just adopt some sort of convention and avoid this problem, like list electrons before muons. A similar rule applies when exchanging any pair of particles if there are $N > 2$ particles.

B. The Hamiltonian

It's now time to write our Hamiltonian. We first write our Hamiltonian in the form

$$H = H_0 + H_{\text{int}}\ . \tag{4.6}$$

The unperturbed Hamiltonian H_0 will be chosen such that our Fock space states given by (4.2) are eigenstates of it with eigenvalues

$$H_0|\Psi\rangle = \left(E_1+E_2+\cdots+E_N\right)|\Psi\rangle\,.$$

These states are *not* eigenstates of the full Hamiltonian, since if they were, particles could never collide, decay, interact, etc. We now move on to the interactions. What sort of effects might H_{int} include?

The first thing to realize about interactions is that they must be *local*. In ordinary quantum mechanics, it is fine to have terms in the Hamiltonian like $e^2\big/\left[4\pi\left|\mathbf{x}^1-\mathbf{x}^2\right|\right]$, but in relativity, it is impossible for one particle to instantaneously know another's location. It follows that the Hamiltonian must be the integral of some sort of *Hamiltonian density,* $\mathcal{H}_{\text{int}}(x)$, so that

$$H_{\text{int}}\left(x^0\right) = \int d^3\mathbf{x}\,\mathcal{H}_{\text{int}}\left(x^0,\mathbf{x}\right). \tag{4.7}$$

We will need to calculate the matrix elements of this operator between various states, so we will need

$$\langle\phi_F|H_{\text{int}}(0)|\phi_I\rangle = \int d^3\mathbf{x}\langle\phi_F|\mathcal{H}_{\text{int}}(0,\mathbf{x})|\phi_I\rangle\,. \tag{4.8}$$

Although $\mathcal{H}_{\text{int}}(x)$ is a function of *x*, the interaction should be the same everywhere, since the laws of physics are independent of its location in space and time. They should care only about what the state vector is doing at the point *x*. The initial state is assumed to have four-momentum p_I and the final state p_F. It follows that the only difference between the matrix element

$\langle\phi_F|\mathcal{H}_{\text{int}}(x)|\phi_I\rangle$ and $\langle\phi_F|\mathcal{H}_{\text{int}}(0)|\phi_I\rangle$ should be a factor of $e^{ip_F\cdot x-ip_I\cdot x}$. Because $\mathcal{H}_{\text{int}}(0)$ is going to be featured prominently in subsequent discussion, we define

$$\mathcal{H}\equiv\mathcal{H}_{\text{int}}(0)\ . \tag{4.9}$$

Then we have

$$\langle\phi_F|\mathcal{H}_{\text{int}}(x)|\phi_I\rangle=\langle\phi_F|\mathcal{H}|\phi_I\rangle e^{ip_F\cdot x-ip_I\cdot x}=e^{i(E_F-E_I)t}\langle\phi_F|\mathcal{H}|\phi_I\rangle e^{i\mathbf{p}_I\cdot\mathbf{x}-i\mathbf{p}_F\cdot\mathbf{x}}\ . \tag{4.10}$$

Substituting this in eq. (4.8), we therefore have

$$\langle\phi_F|H_{\text{int}}(0)|\phi_I\rangle=\int d^3\mathbf{x}\langle\phi_F|\mathcal{H}|\phi_I\rangle e^{i\mathbf{p}_I\cdot\mathbf{x}-i\mathbf{p}_F\cdot\mathbf{x}}=\langle\phi_F|\mathcal{H}|\phi_I\rangle V\delta_{\mathbf{p}_F,\mathbf{p}_I}\ , \tag{4.11}$$

where we used (2.60a) to perform the space integral. Note that (4.11) assures that the interaction Hamiltonian automatically conserves three-momentum. So you see that three-momentum conservation comes from the assumption that the interaction Hamiltonian is the same everywhere in space, and energy conservation comes from the assumption that the Hamiltonian is not time-dependent.

C. The Interaction Hamiltonian Density $\mathcal{H}$

It remains to discuss the matrix elements $\langle F|\mathcal{H}|I\rangle$. We will rely a bit on intuition and a bit on experience with the Dirac equation in chapter 3. The first property I'd like it to have is that the Hamiltonian must be Hermitian. This will be assured if $\mathcal{H}$ is Hermitian as well. In other words,

$$\begin{aligned}&\langle t_1,p_1,s_1;\ldots;t_N,p_N,s_N|\mathcal{H}|t_1',p_1',s_1';\ldots;t_M',p_M',s_M'\rangle\\&\quad=\langle t_1',p_1',s_1';\ldots;t_M',p_M',s_M'|\mathcal{H}|t_1,p_1,s_1;\ldots;t_N,p_N,s_N\rangle^*\ .\end{aligned} \tag{4.12}$$

The second property I want is for $\mathcal{H}$ to be Lorenz invariant.[1] This is most easily illustrated in terms of spinless states. Let's say we want to calculate a matrix element that turns two particles into one, say

$$\langle t_1,p_1|\mathcal{H}|t_2,p_2;t_3,p_3\rangle=f(p_1,p_2,p_3)\,.$$

Then $f(p_1,p_2,p_3)$ must be a Lorentz invariant function of the four-momenta. Expressions like $p_1\cdot p_2$ or p_1^2 are fine, but $\mathbf{p}_1\cdot\mathbf{p}_2$ or E_1E_2 are not.

The third property is what I call the *spectator property*. Suppose you have any two Fock states $|X\rangle$ and $|Y\rangle$ that are connected by the Hamiltonian density, so that $\langle Y|\mathcal{H}|X\rangle\neq 0$. Then the presence of one or more other particles *Z* should not prevent this process from occurring. Furthermore, the resulting matrix element should be related, so we can "factor out" the *Z* particle(s), so that

---

[1] Okay, I'm lying to you, it's actually the Lagrangian density that is Lorentz invariant. Usually they are the same thing. When they aren't, I'll simply talk fast enough that you won't know the difference.

$$\langle Z;Y|\mathcal{H}|Z;X\rangle = \langle Z|Z\rangle\langle Y|\mathcal{H}|X\rangle . \tag{4.13}$$

Basically, the *Z* is simply acting as a *spectator*, standing by while *X* and *Y* do their thing. In some cases, eq. (4.13) will require additional terms on the right, in that there may be additional interactions where the *Z* participates as well.

In situations where there are spectators, like in (4.13), we normally will immediately factor them out. The matrix elements that remain I will call *basic* matrix elements. There will be an additional constraint called *renormalizability* that will apply to these basic matrix elements, which we will discuss later.

The fourth property I want all Hamiltonians to have is what I call the anti-particle property. We learned in studying the Dirac equation that processes where we create an electron are closely related to processes where we eliminate a positron, and vice versa. To show how this works, suppose we have a process where two particles are being converted to one, so we have a matrix element of the form

$$\langle t_3, p_3, s_3|\mathcal{H}|t_1, p_1, s_1; t_2, p_2, s_2\rangle = f\left(p_1, s_1, p_2, s_2, p_3, s_3\right). \tag{4.14}$$

Then it will also connect states of the form

$$\langle \bar{t}_1, p_1, s_1; t_3, p_3, s_3|\mathcal{H}|t_2, p_2, s_2\rangle = g\left(p_1, s_1, p_2, s_2, p_3, s_3\right), \tag{4.15}$$

where $\bar{t}_1$ is the anti-particle of t_1 Furthermore, the matrix elements (4.14) and (4.15) will be related, namely

$$f\left(p_1, s_1, p_2, s_2, p_3, s_3\right) = g\left(-p_1, \mathcal{P}s_1, p_2, s_2, p_3, s_3\right). \tag{4.16}$$

where $\mathcal{P}s_1$ is the parity-reversed spin state. The details of this won't come up too much.

It is even possible to move all the particles to one side of the interaction, so we have matrix elements that look like

$$\langle 0|\mathcal{H}|t_1, p_1, s_1; t_2, p_2, s_2; t_3, p_3, s_3\rangle .$$

This may look odd, since it clearly cannot conserve energy. Don't let this bother you. We are calculating the Hamiltonian at one instant of time (and for that matter, at one point in space) and by the uncertainty principle, you are allowed to violate conservation of energy. When we ultimately integrate over all times, conservation of energy will be restored, though intermediate states that violate conservation of energy are allowed.

A fifth property is a bit beyond us at the moment, because it involves something called *internal symmetries*, something we haven't discussed yet. Internal symmetries are transformations of the state vector $|\Psi\rangle$ that do not involve spacetime, and which should leave matrix elements unchanged. Consider electric charge. Let *Q* be the operator that measures the charge on a state, so for example, $Q|e^-\rangle = -|e^-\rangle$. Now, for an arbitrary state vector, let $|X\rangle \to e^{-iQ\theta}|X\rangle$. We would like the matrix elements of these new states to be unchanged, so that we have

$$\langle Y|e^{iQ\theta}\mathcal{H}e^{-iQ\theta}|X\rangle = \langle Y|\mathcal{H}|X\rangle . \tag{4.17}$$

Clearly, this simply demands that

$$[Q, \mathcal{H}] = 0 \,. \tag{4.18}$$

This is accomplished simply by demanding that $\mathcal{H}$ only connect states with the same charge.

In summary, the Hamiltonian density $\mathcal{H}$ must:

(1) be Hermitian;
(2) be Lorentz invariant;
(3) allow spectators;
(4) respect the anti-particle property,
(5) respect internal symmetries.

All relativistic particle theories must obey these four properties. There is one additional property that requires discussion. It is satisfied by the standard model, and by any fully-self-consistent theory that we know of:

(6) The theory must be *renormalizable*.

Renormalizability is a complicated concept, and we won't be able to get into it in too much detail here. Basically, if you go to high enough order in perturbation theory, you end up with intermediate states including all possible particles in the theory with all possible momenta. Because of the demands of Lorentz invariance, the interactions of these high-momentum states are related to the low-momentum states we actually experiment with. We cannot arbitrarily ignore or eliminate them without ruining Lorentz invariance. But there are many more high-momentum states than low-momentum states, and these high momentum states tend to completely dominate the high-order perturbative expansions. Indeed, they often yield infinity.

Fortunately, a technique called *renormalization* has been developed that, in some cases, allows us to eliminate these infinite contributions. One can add new terms to the interaction Hamiltonian that cancel out these infinities. In renormalizable theories, one need only add a finite number of these terms, and you end up with a theory with a finite number or parameters, one where one is able to make theoretical predictions and compare them with experiment. To put it in the simplest possible terms: renormalizable good, non-renormalizable bad.

A caveat is appropriate here, however. Even though non-renormalizable theories often have an infinite number of parameters, at low energies it is often the case that most of them are very small, and hence you can ignore all but a finite number of terms. This allows one to make at least approximate predictions at low energy. However, as you increase the energy, new interactions will inevitable become important, until the entire perturbative expansion breaks down, and you have to abandon your non-renormalizable theory. This, in fact, is exactly what happened with weak interactions, where the non-renormalizable Fermi interaction ultimately was abandoned in favor of the Glashow-Weinberg-Salam model.

And what distinguishes a renormalizable theory? Renormalizable theories:

- Have only spin 0, $\frac{1}{2}$, and 1 particles in them;
- The spin-1 interactions are always gauge interactions (which we will discuss later, but we already encountered one example with electromagnetism);
- The basic matrix elements of $\mathcal{H}$ are always polynomials in the momenta;
- The coefficients of said polynomials have dimensions of non-negative powers of the mass;

These restrictions severely limit the types of interactions that are possible. Often, such a theory has only a finite and very limited number and types of interactions.

D. The $\psi^*\psi\phi$ Model

To illustrate how all this works, I will use a simple example, which I call the $\psi^*\psi\phi$ model. This model consists of only spin-0 particles, and has only three types of particles, the ϕ, which is its own anti-particle, and the ψ and its anti-article, the ψ^*. The ψ is assumed to carry some sort of conserved charge, say +1, and its anti-particle ψ^* has charge -1, while the ϕ is neutral. This charge isn't associated with any sort of force, however, it is just conserved. The theory will be assumed to be renormalizable. The mass of the ψ we will call m and the mass of the ϕ will be called M.

We need to get all the basic matrix elements of $\mathcal{H}$. By the anti-particle property, any matrix element of the form $\langle Y|\mathcal{H}|X\rangle$ can be rewritten by bringing the Y to the other side, so we need only focus on matrix elements like $\langle 0|\mathcal{H}|X\rangle$. We want matrix elements of the form

$$\langle 0|\mathcal{H}|t_1, p_1; t_2, p_2; \cdots; t_N, p_N\rangle, \tag{4.19}$$

where we have dropped the unnecessary spin labels.

Now it's time to do dimensional analysis. For a single particle in four space-time dimensions, $\langle t, p|t, p\rangle = 2EV$, and since E has dimensions of M and V has dimensions of M^{-3}, $\langle t, p|t, p\rangle \sim M^{-2}$. It is a little arbitrary how one defines the dimensions of a bra or ket, but the simplest way is to define them to have the same dimensions, so $\langle t, p| \sim |t, p\rangle \sim M^{-1}$. For many particles,

$$|t_1, p_1; t_2, p_2; \cdots; t_N, p_N\rangle \sim M^{-N} \tag{4.20}$$

The vacuum state $|0\rangle$ and its corresponding bra $\langle 0|$ are dimensionless. The Hamiltonian H has the dimensions of mass, and since H is the space integral of $\mathcal{H}$, the Hamiltonian density has dimensions of $\mathcal{H} \sim M/V \sim M^4$. Putting this all together, we have

$$\langle 0|\mathcal{H}|t_1, p_1; t_2, p_2; \cdots; t_N, p_N\rangle \sim M^{N-4}. \tag{4.21}$$

Now we start using the power of renormalizability. This tells us these matrix elements must be polynomials in the momentum with coefficients that are non-negative powers of the mass. The constant term of this polynomial will have dimensions M^{N-4}, which tells us that it will only be non-zero if $N \le 4$. Because momentum has dimensions $p \sim M$, linear terms exist only if $N \le 3$, and so on. In conclusion, any matrix elements of the form (4.21) exist only for $N \le 4$.

A matrix element of the form $\langle 0|\mathcal{H}|0\rangle$ is really just a constant addition to the Hamiltonian, which is irrelevant, so we ignore it. It turns out that matrix elements with one particle, like $\langle 0|\mathcal{H}|\phi\rangle$ can be eliminated by shifting fields, as we discuss in section 5D, and therefore are normally assumed not to be present. Matrix elements for N = 2 particles contribute to the mass, and will also be discussed in section 5D. Hence we need only consider $N \ge 3$ in general, or in four space-time dimensions, we only need to consider N = 3 or 4.

We now take advantage of the fact that we have a conserved charge. This means that if $\left\langle 0\middle|\mathcal{H}\middle|X\right\rangle \neq 0$, X must contain equal numbers of ψ 's and ψ^* 's It is easy to see that the only types of matrix elements we might have for $N = 3$ are then

$$\left\langle 0\middle|\mathcal{H}\middle|\psi\psi^*\phi\right\rangle = g \quad \text{and} \quad \left\langle 0\middle|\mathcal{H}\middle|\phi\phi\phi\right\rangle = h\,, \tag{4.22}$$

where, for the moment, g and h are arbitrary complex unknown functions of the momenta. I have temporarily dropped the momenta labels, for good reason as we will show soon. Similarly, if we have $N = 4$ particles in the initial state, the only non-vanishing matrix elements would be

$$\left\langle 0\middle|\mathcal{H}\middle|\psi\psi\psi^*\psi^*\right\rangle = \lambda_1\,, \quad \left\langle 0\middle|\mathcal{H}\middle|\psi\psi^*\phi\phi\right\rangle = \lambda_2\,, \quad \text{and} \quad \left\langle 0\middle|\mathcal{H}\middle|\phi\phi\phi\phi\right\rangle = \lambda_3\,. \tag{4.23}$$

But we aren't done using renormalizability! Recall that all of these functions must be polynomials in the momenta. Let's just focus on the first of these five interactions. We write, in general,

$$\left\langle 0\middle|\mathcal{H}\middle|\psi, p; \psi^*, p'; \phi, k\right\rangle = g\left(p, p', k\right).$$

The function g has dimensions of mass to the first. Since the coefficients of g must be non-negative powers of the momenta, g has at most a constant term and a term linear in the momenta. But there is no way to make a scalar quantity that is linear in the momenta. Hence the linear term must vanish. From this, it can be concluded that g is a constant, which at the moment is an arbitrary complex number. Similar arguments (see problem 4.5) can show that $h, \lambda_1, \lambda_2,$ and λ_3 are also constants.

Now, the matrix element $\left\langle 0\middle|\mathcal{H}\middle|\psi\psi^*\phi\right\rangle$ is related, by the anti-particle property, to seven other matrix elements:

$$\begin{aligned} g &= \left\langle 0\middle|\mathcal{H}\middle|\psi\psi^*\phi\right\rangle = \left\langle \psi\middle|\mathcal{H}\middle|\psi\phi\right\rangle = \left\langle \psi^*\middle|\mathcal{H}\middle|\psi^*\phi\right\rangle = \left\langle \psi\psi^*\middle|\mathcal{H}\middle|\phi\right\rangle, \\ &= \left\langle \psi\psi^*\phi\middle|\mathcal{H}\middle|0\right\rangle = \left\langle \psi\phi\middle|\mathcal{H}\middle|\psi\right\rangle = \left\langle \psi^*\phi\middle|\mathcal{H}\middle|\psi^*\right\rangle = \left\langle \phi\middle|\mathcal{H}\middle|\psi\psi^*\right\rangle. \end{aligned} \tag{4.24}$$

By the Hermitian property, the four objects in the first line of (4.24) are the complex conjugates of those in the second row, so $g = g^*$ and g is real. Similar arguments (see problem 4.5) can show that $h, \lambda_1, \lambda_2,$ and λ_3 are also real.

We began the $\psi^*\psi\phi$ theory merely by specifying a list of particles and their spins and charges. This completely specified the theory up to seven real parameters: the masses m and M, and the five couplings g, $h, \lambda_1, \lambda_2,$ and λ_3. Before proceeding, I'd like to simplify things a bit. I will for the moment assume only the interaction g is non-zero. Effectively, this reduces us to a three parameter theory, the two masses and g. I'll call this the restricted $\psi^*\psi\phi$ theory, in contrast to the general $\psi^*\psi\phi$ theory.

I'd also like to introduce some simple diagrammatics. I will denote a ψ particle by an arrow going to the right, a ψ^* by an arrow going to the left, and a ϕ by a line with no arrow on it, as illustrated in Fig. 4-1. Interactions will be denoted by a dot. Keeping in mind that time increases to the right, the eight related interactions described by the constant coupling g appearing in eq. (4.24) are sketched in Fig. 4-2. Note

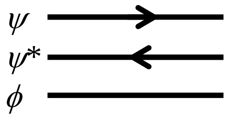


Figure 4-1: Symbols used in the $\psi^*\psi\phi$ theory.

that the anti-particle property is simply reflected in the topological nature of these diagrams. One is allowed to shift any line from the right to the left. The arrows keep track of if you have ψ particles or ψ^* anti-particles. Note that however you draw it, the interaction always has one arrow going in, one coming out, and one line.

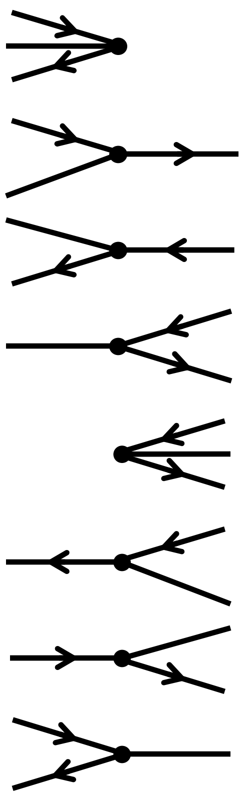


Figure 4-2: The eight interactions described by Eq. (4.24).

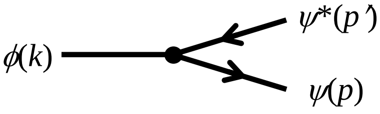


Figure 4-3: The Feynman diagram for the decay of the ϕ particle.

E. A Decay Rate Computation

Let's calculate the rate for the ϕ particle decay via the process $\phi \to \psi^* \psi$ to first order in perturbation theory (which is as far as we will do the computation). We first draw the appropriate diagram, as shown in Fig. 4-3, labeling the particle types and their momenta as indicated. When we get to Feynman diagrams in chapter 5, this will in fact be the Feynman diagram for this decay.

We now want to calculate the decay rate. We need to properly normalize the initial and final states, which are

$$\left|\Psi_I\right\rangle = \frac{1}{\sqrt{2E_k V}}\left|\phi,k\right\rangle, \quad \left|\Psi_F\right\rangle = \frac{1}{\sqrt{2EV}}\frac{1}{\sqrt{2E'V}}\left|\psi,p;\psi^*,p'\right\rangle .$$

We now calculate the first order transition matrix given by eq. (4.1), which we calculate using eq. (4.11)

$$\begin{aligned}\mathcal{T}_{FI} &= -i\left\langle\Psi_F\right|H_{\text{int}}(0)\left|\Psi_I\right\rangle = -i\left\langle\Psi_F\right|\mathcal{H}\left|\Psi_I\right\rangle V\delta_{\mathbf{p}_F,\mathbf{p}_I} \\ &= -i\frac{1}{\sqrt{2EV}}\frac{1}{\sqrt{2E'V}}\frac{1}{\sqrt{2E_k V}}\left\langle\psi,p;\psi^*,p'\right|\mathcal{H}\left|\phi,k\right\rangle V\delta_{\mathbf{p}+\mathbf{p}',\mathbf{k}} ,\end{aligned}$$

$$\mathcal{T}_{FI} = \frac{-ig}{\sqrt{(2E_k V)(2EV)(2E'V)}}V\delta_{\mathbf{p}+\mathbf{p}',\mathbf{k}} . \tag{4.25}$$

Substitute this into eq. (2.53) to give

$$P = \frac{1}{2E_k V}\frac{1}{2EV}\frac{1}{2E'V}2\pi V^2 T\delta_{\mathbf{p}+\mathbf{p}',\mathbf{k}}\delta\left(E+E'-E_k\right)\left|-ig\right|^2 .$$

We now take the infinite volume limit (2.62) to eliminate the Kronecker delta function, yielding

$$P = \frac{1}{2E_k V}\frac{1}{2EV}\frac{1}{2E'V}(2\pi)^4 VT\delta^4\left(p+p'-k\right)\left|-ig\right|^2 . \tag{4.26}$$

The formula still looks odd, because it has so many left over factors of the volume of the universe V and even the time for the experiment T. The reason all those factors of V are in the denominator is because the probability of going into any state with *exactly* momentum **p** is zero. The key to finishing the problem is to realize that we must sum over all possible final state momenta. We therefore have

$$P = \frac{1}{2E_k}(2\pi)^4 T \frac{1}{V^2}\sum_{\mathbf{p}}\sum_{\mathbf{p}'}\frac{1}{2E}\frac{1}{2E'}\delta^4\left(p + p' - k\right)\left|-ig\right|^2 .$$

We now take the infinite volume limit using eq. (2.61b), which gives us

$$\frac{P}{T} = \frac{1}{2E_k}(2\pi)^4 \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{d^3\mathbf{p}'}{(2\pi)^3}\frac{1}{2E}\frac{1}{2E'}\delta^4\left(p + p' - k\right)\left|-ig\right|^2 \qquad (4.27)$$

We note that the decay rate Γ is the probability per unit time that the particle decays, so $\Gamma = P/T$. In the next section we will learn how to complete this computation, but for now I will simply give you the answer in the center of mass frame:

$$\Gamma = \frac{g^2}{16\pi M^2}\sqrt{M^2 - 4m^2} . \qquad (4.28)$$

Note that this answer only makes sense if $M > 2m$; if $M \le 2m$, the ϕ particle cannot decay and is stable.

F. The Feynman Invariant Amplitude, Decay Rates, and Cross-Sections

Before proceeding further, let's look at equation (4.26). There are many features that will occur repeatedly as we perform more complex computations. First note that there is a Dirac delta function that makes sure that both energy and momentum are preserved. Next note that we get a factor of $1/2E_iV$ from each of the initial and final particle states. There is also a factor of $(2\pi)^4 VT$. These will appear in every calculation we make. There is also something inside absolute value symbols that is always squared. This is the only part that will change.

We therefore make the following general statement of the form of the probability. The probability is always given by

$$P = \prod_{i\in\text{in}}\left(\frac{1}{2E_iV}\right)\prod_{i\in\text{out}}\left(\frac{1}{2E_iV}\right)(2\pi)^4 VT\delta^4\left(p_F - p_I\right)\left|i\mathcal{M}\right|^2 , \qquad (4.29)$$

where $i\mathcal{M}$ is called the *Feynman invariant amplitude*. In the case just completed, $i\mathcal{M} = -ig$. Equation (4.29) is the relativistic equivalent of Fermi's golden rule, eq. (2.53). The Feynman invariant amplitude gets its name from the fact that it is an amplitude that needs to be squared to give probabilities, that it is Lorentz invariant, and that it was first introduced by Richard Feynman. Fortunately, Feynman not only introduced the Feynman invariant amplitude, he also introduced *Feynman diagrams* which allow us to efficiently calculate the Feynman amplitude. We will learn the rules for generating Feynman diagrams and calculating Feynman amplitudes in chapter 5.

For now I would like to focus on what we do with the Feynman amplitude once we have it. We note that we have a lot of factors of the volume of the universe in the denominator. For the final states, we already know what to do with them: sum over final state momenta and change the sums into integrals, which yields

$$\frac{P}{T} = V\prod_{j\,\text{in}}\left(\frac{1}{2E_jV}\right)\prod_{i\,\text{out}}\left[\int\frac{d^3\mathbf{p}_i}{(2\pi)^3\,2E_i}\right](2\pi)^4\,\delta^4\left(p_F - p_I\right)\left|i\mathcal{M}\right|^2 , \tag{4.30}$$

It is common to split up (4.30) into two parts. Let's define by far the hardest part as

$$D \equiv \prod_{i\,\text{out}}\left[\int\frac{d^3\mathbf{p}_i}{(2\pi)^3\,2E_i}\right](2\pi)^4\,\delta^4\left(p_F - p_I\right)\left|i\mathcal{M}\right|^2 . \tag{4.31}$$

Then eq. (4.30) becomes

$$\frac{P}{T} = V\prod_{j\,\text{in}}\left(\frac{1}{2E_jV}\right)D . \tag{4.32}$$

Actually performing the integrals in (4.31) tends to be a mess. The first thing to note is that the expression actually is Lorentz invariant. This is not obvious. I already mentioned that the Feynman amplitudes are Lorentz invariant. The expression $\delta^4\left(p_F - p_I\right)$ looks Lorentz invariant as well. But what about all those momentum integrals? Not only do they have only integrals over three-momentum, they also have explicit factors of the energy!

Let m be the mass of a particle in the final state. Consider the integral:

$$\int d^4p_i\;\delta\left(p_i^2 - m_i^2\right)\theta\left(p_i^0\right) .$$

The integral is Lorentz invariant, as is the argument of the delta function, which ensures that $E_i^2 = \mathbf{p}_i^2 + m^2$. It is not obvious, but true, that Lorentz transformations for momenta with positive energy leads to momenta with positive energy, so the combination $\delta\left(p_i^2 - m_i^2\right)\theta\left(p_i^0\right)$ is in fact Lorenz invariant as well. We now perform the p^0 integral, which yields

$$\int d^4p_i\;\delta\left(p_i^2 - m_i^2\right)\theta\left(p_i^0\right) = \int d^3\mathbf{p}_i\int_0^\infty \delta\left(E_i^2 - \mathbf{p}_i^2 - m_i^2\right)dE_i = \int\frac{d^3\mathbf{p}_i}{2E_i} . \tag{4.33}$$

Since the left hand side of (4.33) is Lorentz invariant, so also is the right hand side, even though it is not obvious. It follows that the expression (4.31) is Lorentz invariant.

Why is this important? It means that we can calculate D in any frame we want. Usually the easiest frame to work in is the center of mass frame, the frame in which the three-momentum is zero. In this frame, the initial three-momentum is $\mathbf{p}_{\text{in}} = 0$, and the energy is given by E_{cm}.

Let's start by calculating D for two particles in the final state. We first use the trick of equation (4.33) to rewrite one of the integrals as a four-dimensional integral, so we have

$$\begin{aligned} D_2 &= \int\frac{(2\pi)^4\,d^3\mathbf{p}_1 d^3\mathbf{p}_2}{(2\pi)^6\,2E_1 2E_2}\delta^4\left(p_1 + p_2 - p_I\right)\left|i\mathcal{M}\right|^2 \\ &= \int\frac{d^3\mathbf{p}_1 d^4p_2}{8\pi^2E_1}\delta\left(p_2^2 - m_2^2\right)\theta\left(p_2^0\right)\delta^4\left(p_1 + p_2 - p_I\right)\left|i\mathcal{M}\right|^2 , \\ D_2 &= \int\frac{d^3\mathbf{p}_1}{8\pi^2E_1}\delta\left[\left(p_I - p_1\right)^2 - m_2^2\right]\theta\left(E_{\text{cm}} - E_1\right)\left|i\mathcal{M}\right|^2 . \end{aligned} \tag{4.34}$$

The function $\theta\left(E_{\text{cm}} - E_1\right)$ simply restricts the integral to energies that make sense (you can't have $E_1 > E_{\text{cm}}$), so we drop this. We note that $p_I^2 = E_{\text{cm}}^2$, $p_I \cdot p_1 = E_{\text{cm}} E_1$, and $p_1^2 = m_1^2$, so (4.34) simplifies to

$$D_2 = \int \frac{d^3\mathbf{p}_1}{8\pi^2 E_1} \delta\left[E_{\text{cm}}^2 - 2E_{\text{cm}}E_1 + m_1^2 - m_2^2\right] \left|i\mathcal{M}\right|^2 .$$

We now write out the $\mathbf{p}_1$ integral into a radial and angular part, and we have[1]

$$D_2 = \int \frac{p_1^2 dp_1 \, d\Omega_1}{8\pi^2 E_1} \delta\left[E_{\text{cm}}^2 - 2E_{\text{cm}}E_1 + m_1^2 - m_2^2\right] \left|i\mathcal{M}\right|^2 .$$

Using the fact that $E_1^2 = p_1^2 + m_1^2$, we see that $E_1 dE_1 = p_1 dp_1$, and therefore

$$D_2 = \int \frac{p_1 E_1 dE_1 d\Omega_1}{8\pi^2 E_1} \delta\left[E_{\text{cm}}^2 - 2E_{\text{cm}}E_1 + m_1^2 - m_2^2\right] \left|i\mathcal{M}\right|^2 ,$$

$$D\left(\text{two particles}\right) = \frac{p}{16\pi^2 E_{\text{cm}}} \int d\Omega \left|i\mathcal{M}\right|^2 , \tag{4.35}$$

where p is the momentum of either of the final state particles in the center of mass system, and E_{cm} is the center of mass energy. And how do we find this energy? For decays, it is simply the mass of the decaying particle. For two particles in the initial state, it is given by $\sqrt{s}$, where $s = \left(p_1 + p_2\right)^2$, defined in eq. (2.35).

Now let's work out a (relatively) simple expression for D when there are three particles in the final state. We have

$$D_3 = \int \frac{\left(2\pi\right)^4 d^3\mathbf{p}_1 d^3\mathbf{p}_2 d^3\mathbf{p}_3}{\left(2\pi\right)^9 2E_1 2E_2 2E_3} \delta^4\left(p - p_1 - p_2 - p_3\right) \left|i\mathcal{M}\right|^2 .$$

We use the usual trick to turn the last momentum integral into a $d^4 p_3$, and then perform this integral:

$$\begin{aligned}
D_3 &= \int \frac{d^3\mathbf{p}_1 d^3\mathbf{p}_2 d^4 p_3}{\left(2\pi\right)^5 2E_1 2E_2} \delta\left(p_3^2 - m_3^2\right) \delta^4\left(p_I - p_1 - p_2 - p_3\right) \left|i\mathcal{M}\right|^2 \\
&= \int \frac{d^3\mathbf{p}_1 d^3\mathbf{p}_2}{\left(2\pi\right)^5 2E_1 2E_2} \delta\left[\left(p_I - p_1 - p_2\right)^2 - m_3^2\right] \left|i\mathcal{M}\right|^2 , \\
&= \int \frac{d^3\mathbf{p}_1 d^3\mathbf{p}_2}{\left(2\pi\right)^5 2E_1 2E_2} \delta\left[p_I^2 + m_1^2 + m_2^2 - 2p_I \cdot p_1 - 2p_I \cdot p_2 + 2p_1 \cdot p_2 - m_3^2\right] \left|i\mathcal{M}\right|^2 ,
\end{aligned}$$

---

[1] Note that p_1 has changed from meaning the four momentum to meaning the magnitude of the three-momentum. I know this is confusing, but this confusing notation is pretty standard.

$$D_3 = \int \frac{d^3\mathbf{p}_1 d^3\mathbf{p}_2}{(2\pi)^5\, 2E_1 2E_2}\,\delta\left[\begin{array}{c} E_{\rm cm}^2 + m_1^2 + m_2^2 - 2E_{\rm cm}E_1 - 2E_{\rm cm}E_2 \\ +2E_1E_2 - 2\mathbf{p}_1\cdot\mathbf{p}_2 - m_3^2 \end{array}\right]\left|i\mathcal{M}\right|^2 . \tag{4.36}$$

We now write out the remaining integrals more explicitly. For $\mathbf{p}_1$, we rewrite the integral as $d^3\mathbf{p}_1 = p_1^2 dp_1\, d\Omega_1$, but for $\mathbf{p}_2$ we write out the angles relative to the direction of $\mathbf{p}_1$, so we have $d^3\mathbf{p}_2 = p_2^2 dp_2\, d\Omega_{12} = p_2^2 dp_2\, d\phi_{12}\, d\cos\theta_{12}$. So we have

$$\begin{aligned} D_3 &= \int \frac{p_1^2 dp_1\, d\Omega_1\, p_2^2 dp_2\, d\phi_{12}\, d\cos\theta_{12}}{(2\pi)^5\, 2E_1 2E_2}\left|i\mathcal{M}\right|^2 \delta\left[\begin{array}{c} E_{\rm cm}^2 + m_1^2 + m_2^2 - 2E_{\rm cm}E_1 - 2E_{\rm cm}E_2 \\ +2E_1E_2 - 2p_1p_2\cos\theta_{12} - m_3^2 \end{array}\right] \\ &= \int \frac{p_1^2 dp_1\, d\Omega_1\, p_2^2 dp_2\, d\phi_{12}}{(2\pi)^5\, 2E_1 2E_2}\left|i\mathcal{M}\right|^2 \frac{1}{2p_1p_2} . \end{aligned}$$

We now rewrite, as before, $p_1 dp_1 = E_1 dE_1$ and $p_2 dp_2 = E_2 dE_2$ to simplify this to

$$D(\text{three particles}) = \frac{1}{8(2\pi)^5}\int dE_1\, dE_2\, d\Omega_1\, d\phi_{12} \left|i\mathcal{M}\right|^2 . \tag{4.37}$$

In this form, the equation doesn't look so bad, but in fact, for arbitrary mass particles, the limits on the integrals tend to be so messy that the integrals are almost impossible to do in closed form. We will only do a couple of problems with three particles, and we will never go as high as four particles in the final state. You're welcome.

Equations (4.35) and (4.37) allow us to quickly calculate the final state integrals and derive D for two or three particles in the final state. We now need to figure out how to finish the problem; that is, how to actually get physically measurable quantities like cross sections and decay rates.

Decay rates are easy. They are used when we have one particle in the initial state. Going back to eq. (4.32), and recalling that $\Gamma = P/T$, we have $\Gamma = D/2E$. Remembering that D is Lorentz invariant, this simply tells you that the decay rate is inversely proportional to the energy. The decay is fastest if the particle is at rest, in which case $\Gamma = D/2M$. If the particle is moving, this is suppressed by a factor of M/E, which simply represents the time dilation of relativity. If you look up the decay rate in the particle data book, instead of giving you a table of several values for different energies, they will give you the rate in the rest frame, or

$$\Gamma = \frac{D}{2M} , \tag{4.38}$$

or more commonly the mean lifetime $\tau = \Gamma^{-1}$, and assume you know special relativity. But it is nice to see that time dilation is built into our equations; we don't need to put it in by hand afterwards.

Cross sections are a little more work. Returning once more to eq. (4.32), we see that

$$\frac{P}{T} = \frac{D}{4VE_1E_2} .$$

What's that factor of volume doing in there? If we have one particle in the universe trying to collide with one other particle in the universe, then the odds of them colliding are pretty small. Comparing with eq. (1.10), we see that

$$\sigma = \frac{PV}{T\left|\Delta\mathbf{v}\right|} = \frac{D}{4E_1E_2\left|\Delta\mathbf{v}\right|}\,.$$

Keeping in mind that $\mathbf{p} = E\mathbf{v}$, this can be written in the form

$$\sigma = \frac{D}{4E_1E_2\left|\Delta\mathbf{v}\right|} = \frac{D}{4\left|E_2\mathbf{p}_1 - E_1\mathbf{p}_2\right|}\,. \tag{4.39}$$

It is worth commenting that the cross section is invariant in any frame where the two particles are moving along the same axis. We already know that *D* is coordinate invariant, and it isn't hard to show that $\left|E_1\mathbf{p}_2 - E_2\mathbf{p}_1\right|$ is unchanged when you perform a Lorentz transformation, provided $\mathbf{p}_1$ and $\mathbf{p}_2$ are parallel (or anti-parallel) and we perform a boost along this direction (problem 4.8). In particular, the cross-section is the same in the center of mass frame and the frame where the target is at rest. For this reason I usually compute it in the center of mass frame.

To make sure we understand how this all works, let's do the decay rate of the $\phi \to \psi^*\psi$ in the theory with just the one interaction *g*. We note that $-i\mathcal{M} = -ig$. In the center of mass frame, half the energy goes to each particle, which means the momentum will be $p = \sqrt{\tfrac{1}{4}M^2 - m^2}$. Eqs. (4.38) and (4.35) between them give us the decay rate, which in the center of mass frame is

$$\Gamma = \frac{D}{2M} = \frac{1}{2M}\frac{p}{16\pi^2 E_{\text{cm}}}\int d\Omega\left|-i\mathcal{M}\right|^2 = \frac{1}{2M}\frac{4\pi g^2}{16\pi^2 M}\sqrt{\tfrac{1}{4}M^2 - m^2} = \frac{g^2}{16\pi M^2}\sqrt{M^2 - 4m^2}\,.$$

The general procedure is now clear. To calculate a decay rate or cross-section, first calculate the Feynman invariant amplitude $-i\mathcal{M}$. Then you calculate *D*, using (4.31) in general, or more likely (4.35) or (4.37) for two- or three-particle final states, working in the center of mass frame. Then you get the decay rate or cross-section from (4.38) or (4.39). The center of mass energy for decays is $E_{\text{cm}} = m$, while for cross-sections it is $E_{\text{cm}} = \sqrt{s}$, where $s = \left(p_1 + p_2\right)^2$. The only missing ingredient at this point is the Feynman amplitude, and we now turn our attention there.

Problems for Chapter 4

1. Three fermions would be described as $\left|a,b,c\right\rangle$, where we have combined the type, momentum and spin indices into a single letter. Show how all six orderings of the three states are related to each other.

2. The Δ^- baryon is made of three down quarks and has spin 3/2. If we ignore momentum, then a Δ^- baryon with spin $S_z = +\tfrac{3}{2}$ would be described as $\left|\Delta^-\right\rangle = k\left|d,+\tfrac{1}{2};d,+\tfrac{1}{2};d,+\tfrac{1}{2}\right\rangle$. Show that this state is actually zero. The problem will only be resolved in section 9B.

3. Write down all nine (eight plus the original one) matrix elements that correspond to the coupling λ_1 related by the anti-particle property. Also, draw all nine corresponding diagrams akin to Fig. 4-2 for this interaction.

4. Consider the matrix elements of the form $\langle q_i | \mathcal{H} | q_j, W^- \rangle$, where W^- is the *W*-boson, and q_i and q_j are any of the six quarks. Use the internal symmetry of electric charge to argue that of the 36 possible matrix elements, only nine of them can be non-zero. Which ones? In fact, all nine of them are non-zero.

5. Argue that $h, \lambda_1, \lambda_2,$ and λ_3 in eqs. (4.22) and (4.23) are constants, and they are all real.

6. Consider a renormalizable theory with two charged spin 0 particles, ψ_1 with charge Q_1 and ψ_2 with charge Q_2. They are not equivalent to their anti-particles ψ_1^* and ψ_2^*.
   (a) Write down all possible renormalizable matrix elements of the form $\langle 0 | \mathcal{H} | X \rangle$, where *X* has more than two particles, if $Q_2 = Q_1$, and figure out which ones must be real.
   (b) Suppose instead that $Q_2 = -Q_1$. Show that you don't have to redo the work of part (a) (hint: it is arbitrary what we call a particle and what we call an anti-particle)
   (c) Repeat part (a) if $Q_2 = 2Q_1$.
   (d) Repeat part (a) if $Q_2 = 5Q_1$. Argue that the number of particles (minus anti-particles) for ψ_1 and ψ_2 are separately conserved. Such a conservation law that is demanded by renormalizability is called an *accidental* symmetry. Baryon number in the Standard Model is an accidental symmetry.

7. Consider particle physics in some arbitrary number *D* dimensions ($D-1$ space dimensions). Imagine a theory with one scalar particle ϕ that is its own anti-particle.
   (a) Using the normalization condition $\langle p | p \rangle = 2EV$, and assuming $| p \rangle$ and $\langle p |$ have the same dimensions, what is the dimensionality of $| p \rangle$? Keep in mind that *V* is a $D-1$ dimensional volume.
   (b) What is the dimensionality of $\mathcal{H}$, given that $H = \int d^{D-1}\mathbf{x}\, \mathcal{H}$ has dimensionality of mass?
   (c) If we have an interaction $\langle 0 | \mathcal{H} | p_1, p_2, \ldots, p_n \rangle = g$, what is the dimensionality of *g*?
   (d) Find an inequality that *n* and *d* must satisfy so that *g* is a non-negative power of mass. Check that in $D = 4$ space-time dimensions, the restriction is $n \le 4$.
   (e) Work out explicitly the limit on *n* for *D* = 2, *D* = 3, and *D* = 6.

8. Suppose two particles are moving along the x^3 axis, so that $p_1^\mu = (E_1, 0, 0, p_1)$ and $p_2^\mu = (E_2, 0, 0, p_2)$. Note that one particle might be at rest. Show that the combination $|E_2\mathbf{p}_1 - E_1\mathbf{p}_2|$ is unaffected by a boost along the x^3-axis.

9. A charged pion or kaon decays via $\pi^+ \to \mu^+\nu_\mu$ or $K^+ \to \mu^+\nu_\mu$. The Feynman invariant amplitude takes the form $|i\mathcal{M}|^2 = a_\pi^2 \left(p \cdot p'\right)$ or $|i\mathcal{M}|^2 = a_K^2 \left(p \cdot p'\right)$, where a_π and a_K are constants, and p and p' are the four-momenta of rhe final state particles. Calculate the decay rate in terms of a and the pion or kaon mass m_π or m_K and muon mass m_μ (a formula for momentum of the final particles was found in problem 2.9b). Treat the anti-neutrino mass as zero. Using the decay rates you found (or I found) from problem 1.7, find the ratio a_K / a_π.

10. The Feynman amplitude for $e^+e^- \to \mu^+\mu^-$, in the high energy limit (energy much larger than the electron or muon mass), is $|i\mathcal{M}|^2 = e^4\left(1+\cos^2\theta\right)$, where θ is the angle between the initial electron and final muon. Find the total cross section $\sigma\left(e^+e^- \to \mu^+\mu^-\right)$ when they collide head on with energy E each. How would your answer change if one had energy E and the other energy E'?

11. Muons decay via $\mu^- \to e^-\bar{\nu}_e\nu_\mu$, where all the particles on the right are light enough to be treated as massless. The Feynman invariant amplitude for this process is:

$$|i\mathcal{M}|^2 = 64G_F^2\left(p_\mu \cdot p_2\right)\left(p_1 \cdot p_3\right)$$

where p_μ, p_1, p_2, and p_3 are the momenta of μ^-, e^-, $\bar{\nu}_e$ and ν_μ respectively, and $G_F = 1.166\times10^{-5}\ \text{GeV}^{-2}$.

(a) Show that in the rest frame of the muon, the amplitude squared can be written in terms of just *one* of the final energies.

(b) Argue that the energy of no one particle can be more than $\frac{1}{2}m_\mu$. From this, derive exactly *three* inequalities on the energies E_1 and E_2. Draw a plot of E_1 vs. E_2, and shade in the allowed region.

(c) Find a formula for $\Gamma\left(\mu^- \to e^-\bar{\nu}_e\nu_\mu\right)$ in terms of m_μ and G_F. Work it out numerically, and compare the resulting mean lifetime $\tau = \Gamma^{-1}$ to the experimental value.

12. Suppose we were doing a cross-section or decay to two particles, so that the final state is $\left|\Psi_F\right\rangle = \left|t, p_1; t, p_2\right\rangle$. This is identical to $\left|\Psi_F\right\rangle = \left|t, p_2; t, p_1\right\rangle$. Hence when you integrate over all final states using eq. (4.35), you will encounter this final state twice. How do you think you need to modify eq. (4.35) in this case? If you had three identical final state particles, how do you think the final state integral eq. (4.37) would need to be modified?

13. When we have spin, the initial spin for a decay or cross-section might be random. Suppose we were performing a decay with an initial state $\left|\Psi_I\right\rangle = \left|t, p_1, \pm\frac{1}{2}\right\rangle$, where it is equally likely to be spin $\pm\frac{1}{2}$. How would the decay formula eq. (4.38) need to be modified? Suppose we were doing a cross-section with $\left|\Psi_I\right\rangle = \left|t_1, p_1, \pm\frac{1}{2}; t_2, p_2, \pm\frac{1}{2}\right\rangle$, where both spins are random and uncorrelated. How would the cross-section formula eq. (4.39) need to be modified?

V. Feynman Diagrams

In the previous chapter we explained how to calculate cross-sections and decay rates if we know the Feynman invariant amplitude, but we never explained how to get this, other than finding that it is $i\mathcal{M} = -ig$ for the decay process sketched in Fig. 4-3. The Feynman invariant amplitude is basically defined by eq. (4.29). Let's calculate it for one more process, the scattering process $\psi\psi \to \psi\psi$. We will continue (for the moment) to include only the one interaction labeled *g*, given by (4.24). Our initial and final states will be

$$\left|\Psi_I\right\rangle = \frac{1}{\sqrt{2E_1V}\sqrt{2E_2V}}\left|\psi, p_1; \psi, p_2\right\rangle, \quad \left|\Psi_F\right\rangle = \frac{1}{\sqrt{2E_3V}\sqrt{2E_4V}}\left|\psi, p_3; \psi, p_4\right\rangle. \tag{5.1}$$

It is obvious that because the interactions (4.24) always create or destroy a ϕ-particle, the leading order contribution to the transition amplitude (4.1) will vanish. We therefore need to go to second order, and we need to calculate

$$\mathcal{T}_{FI} = (-i)^2 \sum_n \frac{\left\langle\Psi_F\right|H_{\text{int}}(0)\left|\Psi_n\right\rangle\left\langle\Psi_n\right|H_{\text{int}}(0)\left|\Psi_I\right\rangle}{i\left(E_n - E_I\right)}. \tag{5.2}$$

We now carefully consider what sort of states $\left|\Psi_n\right\rangle$ might contribute to this sum. It will be useful to use diagrammatic notation to keep track of things.

A. Not Quite Feynman Diagrams

We want to know what sort of intermediate states $\left|\Psi_n\right\rangle$ can have non-vanishing matrix elements for both $\left\langle\Psi_F\right|H_{\text{int}}(0)\left|\Psi_n\right\rangle$ and $\left\langle\Psi_n\right|H_{\text{int}}(0)\left|\Psi_I\right\rangle$. Since $\mathcal{H}$ always creates or annihilates a ϕ, there must be a ϕ particle in this state. It is not hard to convince yourself that $\left|\Psi_n\right\rangle$ must be of the form $\left|\Psi_n\right\rangle \propto \left|\phi\psi\psi\right\rangle$. We have to make it properly normalized, and if we allow the intermediate momenta to be completely arbitrary, this state will be

$$\left|\Psi_n\right\rangle = \frac{1}{\sqrt{2E_5V}\sqrt{2E_6V}\sqrt{2E_kV}}\left|\phi, k; \psi, p_5; \psi, p_6\right\rangle.$$

We have to sum up this expression over all possible intermediate states, which would look something like[1]

$$\mathcal{T}_{FI} = (-i)^2 \sum_{\mathbf{k}}\sum_{\mathbf{p}_5}\sum_{\mathbf{p}_6} \frac{\left\langle\Psi_F\right|H_{\text{int}}(0)\left|\phi, k; \psi, p_5; \psi, p_6\right\rangle\left\langle\phi, k; \psi, p_5; \psi, p_6\right|H_{\text{int}}(0)\left|\Psi_I\right\rangle}{i\left(E_n - E_I\right)\left[2E_kV\right]\left[2E_5V\right]\left[2E_6V\right]}. \tag{5.3}$$

---

[1] Actually, expression (5.3) is a little subtle. Because interchanging the two momenta p_5 and p_6 produce the same intermediate state, we should restrict the sum to avoid double counting. This apparent error will go away in a moment, so we won't worry about it.

This is a mess, but it will get better soon.

The problem with this complicated expression is that, in fact, almost every term in this sum is zero. The reason is that matrix elements like $\langle \phi,k;\psi,p_5;\psi,p_6|\mathcal{H}|\Psi_I\rangle$ are not *basic* matrix elements, since our basic operations only include things like $\langle \phi\psi|\mathcal{H}|\psi\rangle$. This implies that one of the particles must be a spectator in this process. This means one of the two intermediate ψ momenta must match one of the initial momenta, say $\mathbf{p}_2$. Similarly, one of the intermediate momenta must match one of the final momenta, say $\mathbf{p}_3$. This will produce one of four contributions to $\mathcal{T}_{FI}$, which is given by

$$\mathcal{T}_{FI}^{(1)} = (-i)^2 \sum_{\mathbf{k}} \frac{\langle \Psi_F | H_{\text{int}}(0) | \phi,k;\psi,p_3;\psi,p_2\rangle\langle\phi,k;\psi,p_3;\psi,p_2|H_{\text{int}}(0)|\Psi_I\rangle}{i(E_3+E_2+E_k-E_1-E_2)[2E_kV][2E_2V][2E_3V]}. \tag{5.4}$$

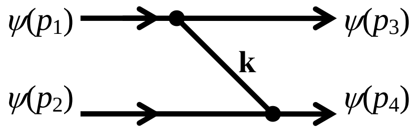


Figure 5-1: A not-quite Feynman diagram representation of eq. (5.4).

It will help to have a diagrammatic approach for this, which I will call a *not quite Feynman diagram*, as sketched in Fig. 5-1. Recalling that time increases as we move to the right, we see that we start with two ψ's with four-momenta p_1 and p_2, but at some time p_1 changes to p_3 while producing a ϕ, with three-momentum $\mathbf{k}$, while the spectator with momentum p_2 looks on. Then the ϕ is absorbed by the ψ with momentum p_2 and changed to momentum p_4, and this time p_3 is the spectator.

Let's start calculating (5.4). Writing out the explicit initial and final states, we have

$$\mathcal{T}_{FI}^{(1)} = (-i)^2 \sum_{\mathbf{k}} \frac{\langle \psi,p_3;\psi,p_4 | H_{\text{int}}(0) | \phi,k;\psi,p_3;\psi,p_2\rangle\langle\phi,k;\psi,p_3;\psi,p_2|H_{\text{int}}(0)|\psi,p_1;\psi,p_2\rangle}{i(E_3+E_k-E_1)[2E_kV][2E_2V][2E_3V]\sqrt{2E_1V}\sqrt{2E_2V}\sqrt{2E_3V}\sqrt{2E_4V}}.$$

By the spectator property, we can remove the spectators, from each interaction, which will yield a factor of $\langle \psi,p_2|\psi,p_2\rangle = 2VE_2$ and $\langle \psi,p_3|\psi,p_3\rangle = 2VE_3$, which simplifies things to

$$\mathcal{T}_{FI}^{(1)} = (-i)^2 \prod_{i=1}^{4}\frac{1}{\sqrt{2E_iV}} \sum_{\mathbf{k}} \frac{\langle \psi,p_4 | H_{\text{int}}(0) | \phi,k;\psi,p_2\rangle\langle\phi,k;\psi,p_3|H_{\text{int}}(0)|\psi,p_1\rangle}{i(E_3+E_k-E_1)[2E_kV]}.$$

We now use eq. (4.11) to simplify this to

$$\mathcal{T}_{FI}^{(1)} = (-i)^2 V \prod_{i=1}^{4}\frac{1}{\sqrt{2E_iV}} \sum_{\mathbf{k}} \frac{\langle \psi,p_4 | \mathcal{H} | \phi,k;\psi,p_2\rangle\langle\phi,k;\psi,p_3|\mathcal{H}|\psi,p_1\rangle}{i(E_3+E_k-E_1)2E_k}\delta_{\mathbf{p}_4,\mathbf{p}_2+\mathbf{k}}\delta_{\mathbf{p}_3+\mathbf{k},\mathbf{p}_1}.$$

One of the two Kronecker delta functions can then be used to do the sum on $\mathbf{k}$. The remaining matrix elements each yield a factor of g, so this simplifies further to

$$\mathcal{T}_{FI}^{(1)} = V\delta_{\mathbf{p}_4,\mathbf{p}_2+\mathbf{p}_1-\mathbf{p}_3}\prod_{i=1}^{4}\frac{1}{\sqrt{2E_iV}}\left.\frac{(-ig)^2}{i(E_3+E_k-E_1)2E_k}\right|_{\mathbf{k}=\mathbf{p}_1-\mathbf{p}_3}. \tag{5.5}$$

It's time to pause and catch our breath. The Kronecker delta function out front is really just enforcing conservation of momentum. Let's compare this result, for example, with eq. (4.25), the equation we got for when we were doing decays. There we also had a factor of V, a momentum preserving delta function, and factors of $1/\sqrt{2EV}$ for each initial and final state particle. The remaining factor of $-ig$ in (4.24) became the Feynman invariant amplitude. By analogy, we conclude that in this case we have

$$i\mathcal{M}^{(1)} = \left.\frac{(-ig)^2}{i\left(E_3 + E_k - E_1\right)2E_k}\right|_{\mathbf{k}=\mathbf{p}_1-\mathbf{p}_3} . \tag{5.6}$$

Let's look at eq. (5.6) and see if we can understand where all the factors are coming from. Because we have two factors of $-i\mathcal{H}$ in our computation, corresponding to the two dots in Fig. 5-1, we ended up with two factors of $-ig$. It's clear that the rest of it has something to do with the intermediate ϕ particle. Why is its momentum $\mathbf{k} = \mathbf{p}_1 - \mathbf{p}_3$? The answer is clear from Fig. 3-1. We have momentum $\mathbf{p}_1$ flowing into the vertex and momentum $\mathbf{p}_3$ flowing out to the right. To balance momentum, the ϕ must carry off the difference in momentum, much as current must be balanced at vertices in a Kirchhoff's Laws problem. We could have used the other vertex and written $\mathbf{k} = \mathbf{p}_4 - \mathbf{p}_2$ instead. For these not quite Feynman diagrams, we know which way the momentum is flowing: it is always to the right.

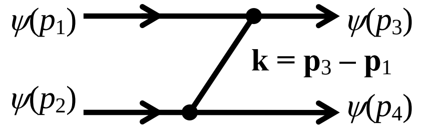


Figure 5-2: A not-quite Feynman diagram representation of eq. (5.7).

I'd now like to calculate a second contribution to the Feynman amplitude, as sketched in Fig. 5-2. The computation is similar to before. Using conservation of momentum at the right vertex, we see that $\mathbf{k} = \mathbf{p}_3 - \mathbf{p}_1$. The energy denominator $E_m - E_I$ can be written as $E_m - E_F = E_1 + E_k - E_3$. So we have

$$i\mathcal{M}^{(2)} = \left.\frac{(-ig)^2}{i\left(E_1 + E_k - E_3\right)2E_k}\right|_{\mathbf{k}=\mathbf{p}_3-\mathbf{p}_1} . \tag{5.7}$$

Note that the intermediate momentum **k** is the opposite of that for $i\mathcal{M}^{(1)}$, but the corresponding energy $E_k = \sqrt{\mathbf{k}^2 + M^2}$ is identical.

Now let's add these two contributions.

$$i\mathcal{M}^{(1+2)} = \left.\left[\frac{(-ig)^2}{i\left(E_3 + E_k - E_1\right)2E_k} + \frac{(-ig)^2}{i\left(E_1 + E_k - E_3\right)2E_k}\right]\right|_{\mathbf{k}=\mathbf{p}_1-\mathbf{p}_3}$$

$$= \left.\frac{(-ig)^2 i}{\left(E_1 - E_3 - E_k\right)\left(E_1 - E_3 + E_k\right)}\right|_{\mathbf{k}=\mathbf{p}_1-\mathbf{p}_3} = \frac{(-ig)^2 i}{\left(E_1 - E_3\right)^2 - \left(\mathbf{p}_1 - \mathbf{p}_3\right)^2 - M^2} ,$$

$$i\mathcal{M}^{(1+2)} = \frac{i(-ig)^2}{\left(p_1 - p_3\right)^2 - M^2} . \tag{5.8}$$

We note that when we added these two terms, not only was there a great simplification, but the result is manifestly Lorentz invariant, while both of the pieces $i\mathcal{M}^{(1)}$ and $i\mathcal{M}^{(2)}$ were not.

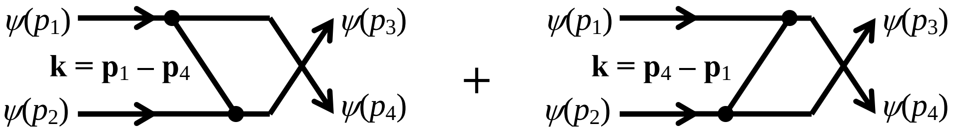


Figure 5-3: A not-quite Feynman diagram representation of eq. (5.9).

There are two other diagrams that need to be included, which I have sketched in Fig. 5-3. You can see from the diagrams that the only difference between these diagrams and the previous two is that the role of p_3 and p_4 has been changed. We therefore immediately write

$$i\mathcal{M}^{(3+4)} = \frac{i(-ig)^2}{(p_1 - p_4)^2 - M^2}. \tag{5.9}$$

The total Feynman amplitude, therefore, is

$$i\mathcal{M} = \frac{i(-ig)^2}{(p_1 - p_3)^2 - M^2} + \frac{i(-ig)^2}{(p_1 - p_4)^2 - M^2}. \tag{5.10}$$

Now that we have the Feynman amplitude, it is a relatively straightforward process to calculate the final state integrals and proceed to the cross-section. We'll do so in section C, but for now let's pause and think about how we did the computation. We first computed two not-quite Feynman diagrams shown in Fig. 5-1 and 5-2. Neither of them separately was Lorentz invariant, but their sum was. The diagrams look very similar; they differ only in which of the two interaction vertices came first. But in special relativity, which of two events occurred first is ambiguous. It is therefore not surprising that we get a Lorentz invariant combination only when we combined them. Similarly, the two not-quite Feynman diagrams in Fig. 5-3 only yield a Lorentz invariant amplitude when we add them.

It would be nice to have a diagrammatic approach in which Figs. 5-1 and 5-2 are not two separate diagrams but one, and where the rules for computing them are manifestly Lorentz invariant. This is what Feynman diagrams are.

B. Feynman Diagrams

Feynman diagrams are a quick and efficient way of computing the Feynman amplitude. They differ from our previous not-quite Feynman diagrams in that the time-ordering of all intermediate vertices is ambiguous. Even if we draw one vertex clearly to the right of the other, it is understood that we don't know which one occurred first.

Fig. 5-4 shows the two Feynman diagrams for the scattering process we have been working with. In Feynman diagrams, we label interior lines with their four-momentum, which is computed by conserving four-momentum at vertices. Because it is not obvious

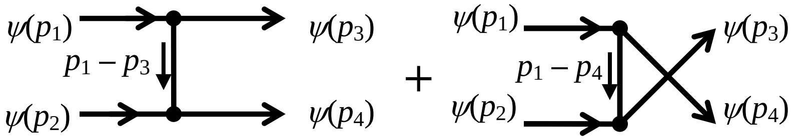


Figure 5-4: Feynman diagrams representing eq. (5.10).

which way the momentum is flowing, I have added arrows indicating this, but this is rarely done, and in this case is irrelevant. Note that in these diagrams, we normally "anchor" the initial and final state particles, so that, for example, I am consistent to always put $\psi(p_3)$ on the upper right, and just rearrange the lines inside.

If we look at eq. (5.10), it is clear that each of the factors of $-ig$ is coming from the interaction, the g from the basic matrix element of $\mathcal{H}$, and the $-i$ from the way we do the time-dependent perturbation theory. The remaining factor coming from either diagram in Fig. (5-4) is called the *propagator*, and is given by $i/(k^2 - M^2)$, where k is the four-momentum associated with the interior line. Particles that appear as internal lines in Feynman diagrams are called *virtual particles*, and unlike external lines, they generally do not satisfy $k^2 = M^2$. When they do, it creates special problems, which we will discuss in section F.

I'm now prepared to give the Feynman rules, which tell you how to compute Feynman amplitudes for an arbitrary process:

(1) For each incoming or outgoing ψ, draw an arrow coming in on the left, or an arrow coming out on the right. For ψ^* draw the arrow pointing the other way. For ϕ, draw an exterior line with no arrows

(2) Draw all topologically distinct diagrams you can think of that connect the initial and final particles. The only vertex allowed (in this theory) is one that has one arrow going in, one going out, and one line.

(3) Label all interior lines by their four-momenta, conserving four-momentum at each vertex.

(4) For each vertex, there is one factor of $-ig$.

(5) For each interior ϕ line with momentum k, there is a factor of $i/(k^2 - M^2)$.

(6) For each interior ψ line with momentum p, there is a factor of $i/(p^2 - m^2)$.

(7) Add up the contributions from each diagram.

That's it! These are all the Feynman rules we will be dealing with for this theory. Rules 4, 5, and 6 are illustrated in Fig. 5-5. The only rule we haven't encountered in the specific examples we have done so far is rule 6, but it is obviously analogous to rule 5, with the ϕ mass M replaced with the ψ mass m.

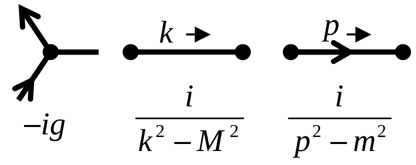


Figure 5-5: The Feynman rules for the restricted $\psi^*\psi\phi$ theory

Note that for exterior lines, the momentum *always* is assumed to be flowing left to right. For interior lines, you can orient the momentum whichever direction you prefer, though it is conventional that for the ψ particles, you orient it in the direction of the arrow. But because the momentum is squared, it doesn't matter for bosons. When we get into fermions in the next chapter, it is crucial that you keep track of which way the momentum is flowing.

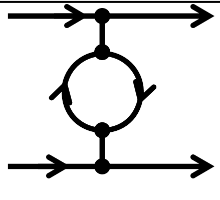


Figure 5-6: One of many loop diagrams contributing to $\psi\psi$ scattering.

Though these are all the rules *we* will deal with, there are additional rules required when computing *loops*, diagrams where it is possible to complete a full circuit and return to the same point. ***Please note that the following is <u>not</u> relevant for this class, it is included only for your cultural edification***. When you have a Feynman diagram involving a closed loop, as illustrated in Fig. 5-6, there is one ambiguous momentum. It turns out the correct thing to do is to sum over all possible momenta, which leads to the additional rule

(8) For each loop, include a factor of $\int d^4k\big/(2\pi)^4$.

Unfortunately, when you do so, you often have places where the propagators given by rules (5) and (6) have vanishing denominators. We need a prescription for dealing with these. It turns out these rules need to be slightly modified:

(5′) For each interior ϕ line with momentum *k*, there is a factor of $i\big/\left(k^2-M^2+i\varepsilon\right)$.

(6′) For each interior ψ line with momentum *p*, there is a factor of $i\big/\left(p^2-m^2+i\varepsilon\right)$.

The quantity ε should be thought of as an arbitrarily small positive real number. This allows you to perform the integrals specified by rule (8). However, even when you do so, you often find integrals that diverge. These are dealt with when you

(9) Regulate and renormalize the theory.

For some theories, this process fails; those theories where it does not are called *renormalizable* theories. Since we won't be doing any loop computations, we won't need modified rules 5 and 6, nor will we need rules 8 and 9. ***We now return to our regularly scheduled programming.*** We will consider only diagrams without loops, which are called *tree diagrams.*

C. Calculating with Feynman Diagrams

Let's do a couple of computations to show how we calculate things using Feynman diagrams. I won't redo the decay of the ϕ, since we already did that in sections 4E and 4F. First, let's finish the work of finding the cross-section for $\psi\psi\to\psi\psi$. The first step would be to draw the Feynman diagrams in Fig. 5-4. In a single step, we would then get the Feynman invariant amplitude of eq. (5.10). To proceed further, we need to write these amplitudes out more explicitly. It is usually easiest to work in the center of mass frame. In this frame, the three-momenta must have equal magnitude, which we will call *p*, and since they have the same mass, they will have the same energy as well, which we will call *E*. The final particles have the same mass and also must have canceling three-momentum, and therefore will again have equal energy. The total energy is 2*E*, so they will have energy *E* each, and hence must also have momentum *p*. Let the angle between $\mathbf{p}_1$ and $\mathbf{p}_3$ be θ, then the angle between $\mathbf{p}_1$ and $\mathbf{p}_4$ will be $\pi-\theta$, as you can see from Fig. 5-7. We therefore have

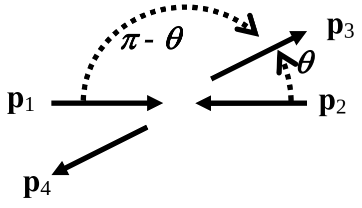


Figure 5-7: The two incoming particles with momenta $\mathbf{p}_1$ and $\mathbf{p}_2$ are back to back, as are the final particles. with momenta $\mathbf{p}_3$ and $\mathbf{p}_4$.

$$\left(p_1-p_3\right)^2=\left(E_1-E_3\right)^2-\left(\mathbf{p}_1-\mathbf{p}_3\right)^2=0-\mathbf{p}_1^2-\mathbf{p}_3^2+2\mathbf{p}_1\cdot\mathbf{p}_3=-2p^2+2p^2\cos\theta . \qquad (5.11)$$

Taking advantage of the fact that $\cos\left(\pi-\theta\right)=-\cos\theta$, we can similarly show that

$$\left(p_1-p_4\right)^2=-2p^2-2p^2\cos\theta . \qquad (5.12)$$

Substituting (5.11) and (5.12) into (5.10), we have

$$i\mathcal{M}=\frac{ig^2}{2p^2-2p^2\cos\theta+M^2}+\frac{ig^2}{2p^2+2p^2\cos\theta+M^2},$$
$$i\mathcal{M}=\frac{2ig^2\left(2p^2+M^2\right)}{\left(2p^2+M^2\right)^2-4p^4\cos^2\theta}. \tag{5.13}$$

We now use eqs (4.35) and (4.39) to get the cross-section:

$$\sigma=\frac{D}{4\left|E_1\mathbf{p}_2-E_2\mathbf{p}_1\right|}=\frac{p}{64\pi^2E_{\text{cm}}\left|E_1\mathbf{p}_2-E_2\mathbf{p}_1\right|}\int d\Omega\left|i\mathcal{M}\right|^2. \tag{5.14}$$

The center of mass energy is $E_{\text{cm}} = 2E$, and since the incoming particles are back-to-back, $\left|E_1\mathbf{p}_2-E_2\mathbf{p}_1\right|=2Ep$. Substituting these expressions and (5.13) into (5.14), we have

$$\sigma=\frac{g^4}{64\pi^2E^2}\int\frac{\left(2p^2+M^2\right)^2d\Omega}{\left[\left(2p^2+M^2\right)^2-4p^4\cos^2\theta\right]^2}. \tag{5.15}$$

If we want the differential cross-section, it is easy to do, because we simply don't perform the angular integrals. We have

$$\frac{d\sigma}{d\Omega}=\frac{g^4\left(2p^2+M^2\right)^2}{64\pi^2E^2\left[\left(2p^2+M^2\right)^2-4p^4\cos^2\theta\right]^2}.$$

If we want the total cross section, we need to perform $\int d\Omega=\int d\phi\, d\left(\cos\theta\right)$, but there is one surprising subtlety. Recall that the two states $\left|\psi,p_3;\psi,p_4\right\rangle$ and $\left|\psi,p_4;\psi,p_3\right\rangle$ are actually the same state. If we integrate over all angles, we will be double counting. It is easiest to get the answer by multiplying the resulting integral by $\frac{1}{2}$. The ϕ integral yields 2π, so

$$\sigma=\frac{g^4}{64\pi E^2}\int_{-1}^{1}\frac{\left(2p^2+M^2\right)^2d\cos\theta}{\left[\left(2p^2+M^2\right)^2-4p^4\cos^2\theta\right]^2}=\frac{g^4}{64\pi E^2}\left[\frac{\ln\left(1+4p^2/M^2\right)}{4M^2p^2+8p^4}+\frac{1}{M^4+4p^2M^2}\right].$$

The final integral was performed by a great deal of tedious calculus. This is the cross-section in the center of mass frame. To get it in some other frame, where at least the particles are collinear, we note that $s=4E^2=4p^2+4m^2$, so we can rewrite this formula in terms of s to obtain

$$\sigma=\frac{g^4}{16\pi s}\left[\frac{1}{M^2\left(s-4m^2\right)+\frac{1}{2}\left(s-4m^2\right)^2}\ln\left(1+\frac{s-4m^2}{M^2}\right)+\frac{1}{M^4+sM^2-4m^2M^2}\right].$$

This formula works equally well if one of our particles is at rest, for example.

D. Other Interactions

Up until now, the only coupling we have allowed in the $\psi^*\psi\phi$ theory is g, defined in eq. (4.22). But there is another interaction in (4.22), and three more in (4.23), for a total of four more:

$$\langle 0|\mathcal{H}|\phi\phi\phi\rangle = h\,, \quad \langle 0|\mathcal{H}|\psi\psi\psi^*\psi^*\rangle = \lambda_1\,, \quad \langle 0|\mathcal{H}|\psi\psi^*\phi\phi\rangle = \lambda_2\,, \quad \langle 0|\mathcal{H}|\phi\phi\phi\phi\rangle = \lambda_3\,.$$

We assumed these were all zero, but they might not be, so we should include Feynman rules for them as well. This isn't hard to figure out: we simply draw a vertex for each, with an arrow in for each ψ, an arrow out for each ψ^*, and a line for each ϕ. Just as the interaction g was multiplied by $-i$, so will be each of these interactions. The corresponding Feynman rules are listed diagrammatically in Fig. 5-8.

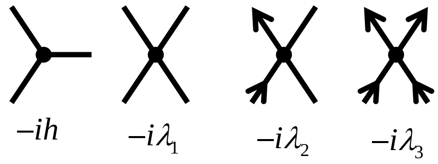


Figure 5-8: The extra Feynman rules for the complete $\psi^*\psi\phi$ theory

Let's do a computation to make sure we understand how this works. Let's calculate in the complete theory the Feynman amplitude for pair annihilation, $\psi^*\psi \to \phi\phi$. We'll call the initial momenta p and p', and the final momenta k and k'. There will be four Feynman diagrams, as shown in Fig. 5-9.

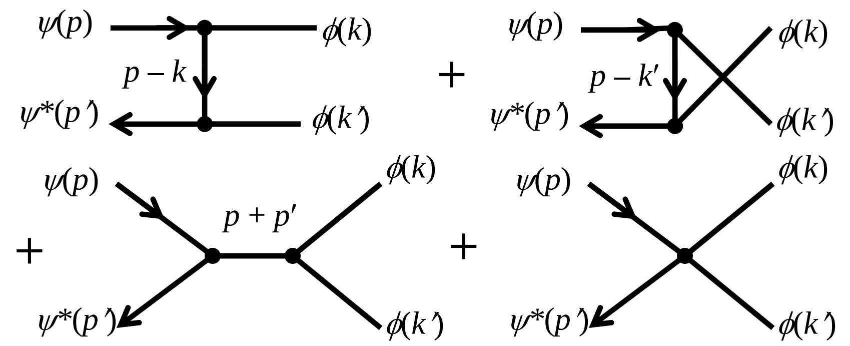


Figure 5-9: The Feynman diagrams contributing to $\psi^*\psi \to \phi\phi$.

The Feynman amplitude is then simply

$$i\mathcal{M} = \frac{(-ig)^2 i}{(p-k)^2 - m^2} + \frac{(-ig)^2 i}{(p-k')^2 - m^2} + \frac{(-ig)(-ih)i}{(p+p')^2 - M^2} - i\lambda_2\,. \tag{5.16}$$

We could then proceed forward to calculate the differential cross-section, or whatever else we want. Because the particles in the final state are identical type, there will be the same factor of $\frac{1}{2}$ that we encountered in section C.

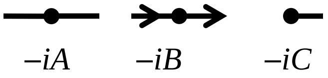


Figure 5-10: The Feynman diagrams that would come from the basic matrix elements in eq. (5.17).

When we first started talking about interactions, we mentioned that we were most interested in interactions involving at least three particles. Why is that? What about interactions with only one or two particles? Suppose we have the following three non-zero basic matrix elements:

$$\langle 0|\mathcal{H}|\phi\phi\rangle = A\,, \quad \langle 0|\mathcal{H}|\psi\psi^*\rangle = B\,, \quad \langle 0|\mathcal{H}|\phi\rangle = C\,. \tag{5.17}$$

These would result in the three additional Feynman rules sketched in Fig. 5-10. Let's first consider the one coming from A. Suppose we were doing some sort of computation with a ϕ propagator, such as the $\psi\psi$ scattering we did in sections A and B. Let the momentum for this propagator be *k*. Now, let us add a series of interactions of type *A*, so we get a whole lot of similar looking terms, as illustrated in Fig. 5-11. Because each of these *A* vertices has only two lines connected to it, conservation of momentum tells us the momentum of each line piece is *k*. Therefore, the propagator, including all these other terms, becomes

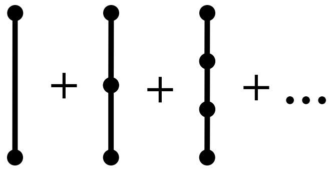


Figure 5-11: In the presence of the coupling *A*, a propagator turns into a sum of an infinite number of terms.

$$\Pi(k) = \frac{i}{k^2-M^2} + \frac{i}{k^2-M^2}(-iA)\frac{i}{k^2-M^2} + \frac{i}{k^2-M^2}(-iA)\frac{i}{k^2-M^2}(-iA)\frac{i}{k^2-M^2} + \cdots .$$

This is a geometric series, which we can add up to yield

$$\Pi(k) = \frac{i}{k^2-M^2}\left[1-\frac{i(-iA)}{k^2-M^2}\right]^{-1} = \frac{i}{k^2-M^2-A}. \tag{5.18}$$

And what does this mean? Effectively, this means that the ϕ particle has a mass squared given by M^2+A, not M^2. Basically, when we divided our Hamiltonian up in eq. (4.6) into $H = H_0 + H_{\text{int}}$, we inappropriately left some of the mass term in H_{int}. By rearranging, we can get rid of this interaction simply by including it in H_0. A similar discussion applies to the term *B*, which adds to the ψ mass. Hence we won't be including the *A* or *B* interactions.

It is nice, however, to sometimes think of masses as just another type of interaction. For one thing, this puts the masses on the same basis as all the other interactions; they are all just different terms in the Hamiltonian. It is also handy when we are sometimes a little unsure of what our particle states are. For example, suppose we were working in a theory with a whole bunch of particle types ψ_i. Our mass terms would all look like

$$\langle 0|\mathcal{H}|\psi_i^*\psi_j\rangle = \langle\psi_i|\mathcal{H}|\psi_j\rangle = m_{ij}^2 . \tag{5.19}$$

By the Hermitian property of $\mathcal{H}$, m^2 is a Hermitian matrix. We could then diagonalize it and find *mass eigenstates* $|\psi'_a\rangle$, linear combinations of the $|\psi_i\rangle$'s such that

$$\langle\psi'_a|\mathcal{H}|\psi'_b\rangle = m_a^2\delta_{ab} . \tag{5.20}$$

Our current situation with neutrinos is in this rather confused state. For years we have worked in the so called flavor eigenstates $\{|\nu_e\rangle,|\nu_\mu\rangle,|\nu_\tau\rangle\}$, and most texts still list these as three of the elementary particles of the standard model, but we now have convincing proof that these are not identical with the mass eigenstates $\{|\nu_1\rangle,|\nu_2\rangle,|\nu_3\rangle\}$. But more on this in chapter 12.

This still leaves the vertex of type *C*. Because this interaction has only one line on it, conservation of four-momentum at this vertex will assure that the momentum of this line is $k = 0$. So, for example, suppose we have an interaction of this type that then went and interacted with, say, a passing ψ particle, as sketched in Fig. 5-12. The combination of the *C* vertex, the ϕ propagator, and the $\psi^*\psi\phi$ vertex will contribute to the overall Feynman amplitude a factor of

$$\frac{(-ig)(-iC)i}{0-M^2} = i\frac{gC}{M^2}$$

We could have achieved the same effect by simply introducing a *B*-type vertex with $B = -gC/M^2$. One can similarly show that the effect of a *C* interaction is just to modify other interactions.

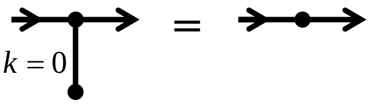


Figure 5-12: **Left:** A passing ψ particle interacts with a scalar field that is created from the vacuum state by the *C* interaction. **Right:** The diagram can be replaced by an equivalent *B*-type interaction with $B = -gC/M^2$.

It is helpful to gain a little insight into what is going on here. Though we haven't really discussed it in this class, it turns out that for every particle there is a corresponding field, so for the ϕ particles there is a corresponding field also called ϕ. There is a *potential*, for this field, and all the interactions of the field ϕ correspond to terms in this potential. Keeping track of only the mass term and the *C* term, the potential is given by

$$V(\phi) = \tfrac{1}{2}M^2\phi^2 + C\phi. \tag{5.21}$$

We have been doing all our perturbation theory about $\phi = 0$, but it is clear that the minimum of the potential is not at $\phi = 0$ but rather at

$$\phi_{\min} = v \equiv -C/M^2. \tag{5.22}$$

We can then define a new field $\phi' \equiv \phi - v$, in terms of which the potential (5.21) has its minimum at $\phi' = 0$. We then work out all our Feynman rules for the field ϕ', which we rename ϕ. Basically, the new rules imply that for any basic matrix element of the form

$$\langle 0|\mathcal{H}|\phi, X\rangle = f \tag{5.23}$$

there will appear an additional new basic matrix element of the form

$$\langle 0|\mathcal{H}|X\rangle = fv, \tag{5.24}$$

with *v* given by eq. (5.22).

In summary, it is never necessary to consider new interactions of the types *A*, *B*, and *C* as sketched in Fig. 5-10. The terms *A* and *B* should be included in the masses M^2 and m^2 respectively, while the presence of *C* simply modifies the other interactions according to eqs. (5.23) and (5.24). Hence in general we will ignore these interactions. But the interactions governed by the parameters *h*, λ_1, λ_2, and λ_3, as illustrated with the Feynman rules in Fig. 5-8, do need to generally be included.

E. Effective Theories

Let us momentarily return to the theory with only the $\psi^*\psi\phi$ interaction with Feynman rule $-ig$. Let us imagine for the moment that our colliders do not yet have sufficient energy to actually create ϕ particles. We discover from $\psi\psi$ scattering that these particles seem to be interacting, but we don't necessarily attribute it to an as yet undiscovered ϕ particle. We think there is just some fundamental new $\psi\psi$ coupling.

Let's look at the energy denominators in eq. (5.10). Because the energy is presumably low compared to the ϕ mass *M*, we can Taylor expand the energy denominators as

$$\frac{1}{M^2-k^2} \approx \frac{1}{M^2}\left(1-k^2/M^2\right)^{-1} = \frac{1}{M^2}+\frac{k^2}{M^4}+\cdots .$$

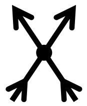
Substituting this into eq. (5.10), we find that the amplitude would be

$$i\mathcal{M} = \frac{2ig^2}{M^2}+\frac{ig^2}{M^4}\left[\left(p_1-p_3\right)^2+\left(p_1-p_4\right)^2\right]+\cdots . \tag{5.25}$$

At low energy, the first term dominates, but we would have no reason to suspect the interaction is of this form. We would probably guess that there is just a fundamental interaction of the form $\left\langle 0\middle|\mathcal{H}\middle|\psi\psi\psi^*\psi^*\right\rangle = -\lambda$. As we increase our collider energies further, we might eventually realize our experiments are deviating from the predictions of this simple theory, and we would discover that the matrix elements have some apparent momentum dependence, eventually deducing the form of the second term in eq.(5.25). Our Feynman rule would look something like that given in Fig. 5-13. Secretly, the two parameters λ and *f* are given by

$$\lambda = \frac{2g^2}{M^2}, \quad f = \frac{g^2}{M^4},$$

but we don't know this. We are now working in an *effective theory*, a low energy theory that gives correct predictions. This may, in fact, be the most efficient way to do calculations at low energies.



$$i\lambda + if\left[\left(p_1-p_3\right)^2+\left(p_1-p_4\right)^2\right]$$

Figure 5-13: The new Feynman rule for the effective theory as defined by eq. (5.25).

A sufficiently clever theorist, however, might suspect something is wrong. The interaction parameter λ is dimensionless, which is fine, but *f* has dimensions of mass to the negative two power. According to our rules for renormalizability, all the parameters describing interactions are supposed to have positive powers of mass. That means the effective theory is not renormalizable. In general, non-renormalizable theories tend to break down at high energies. In this case, for example, note that as the incoming increase in energy, eventually $f\left[\left(p_1-p_3\right)^2+\left(p_2-p_4\right)^2\right]$ will get bigger than λ. We will see a rising cross-section. There are, in fact, theoretical limits on how large the cross-section can get, basically because probabilities can never exceed one, and we know the theory breaks down. Indeed, at $E^2 \sim M^2 \sim \lambda/f$, the *f* and λ terms become comparable, and it is at this approximate scale that the effective theory breaks down and the ϕ particle starts to appear. We could, in fact, predict the breakdown of the theory at the energy scale $E^2 \sim \lambda / f$, and a sufficiently clever theories could predict the ϕ before it was discovered.

The scenario just outlined is, in fact, approximately the story of the discovery of the *W* and *Z* particles. Non-renormalizable interactions were measured that were being mediated by these massive particles when experiments had insufficient energy to produce the particles themselves. This will be explored further in chapter 10.

And that leads to a nice segue. Suppose theorists postulated the existence of the ϕ particle. What would be the most effective way to discover it? If you have sufficient energy, you could produce it through pair annihilation $\psi\psi^* \to \phi\phi$. But at even lower energies, it would be possible to detect its clear presence through the process $\psi\psi^* \to \psi\psi^*$.

F. Resonances

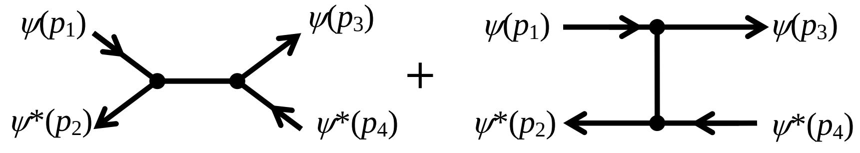


Figure 5-14: Feynman diagrams contributing to $\psi^*\psi \to \psi^*\psi$

Consider the process $\psi\psi^* \to \psi\psi^*$. In the restricted $\psi^*\psi\phi$ theory, there are only two Feynman diagrams that contribute, sketched in Fig. 5-14. The Feynman amplitude would be given by

$$i\mathcal{M} = \frac{i(-ig)^2}{(p_1+p_2)^2 - M^2} + \frac{i(-ig)^2}{(p_1-p_3)^2 - M^2}. \tag{5.26}$$

It isn't hard to show that $(p_1 - p_3)^2 < 0$, and hence the second denominator never vanishes. Consider, however, the first term. Assuming $M > 2m$, it is possible to have $(p_1+p_2)^2 = s = M^2$, making this denominator vanish. What are we to make of this? Does the cross-section become infinite at this point?

The correct way to deal with this problem is to sum an infinite number of loop diagrams, but that sounds like a lot of work, so let's find an easier way. To understand what happens, let's first think about the mass of the ϕ. We defined it to be *M*, which would be the energy of a ϕ at rest. Naively, this means a ϕ at rest would have a wave function that looks like $\Psi(t) \sim e^{-iMt}$. But if this were *actually* the wave function, then the probability of finding a ϕ particle at time *t* would be given by $\Psi^*\Psi = 1$. We want this probability to decrease with time, so that $\Psi^*\Psi \propto e^{-\Gamma t}$. This suggests we should instead have

$$\Psi(t) \sim e^{-iMt} e^{-\frac{1}{2}\Gamma t} = e^{-i\left(M - \frac{1}{2}i\Gamma\right)t} .$$

This suggests that the correct procedure is to replace $M \to M - \frac{1}{2}i\Gamma$ for unstable intermediate particles. Since the mass always appears as M^2, we would replace

$$M^2 \to M^2 - \tfrac{1}{4}\Gamma^2 - i\Gamma M . \tag{5.27}$$

The appearance of the imaginary part then prevents the corresponding denominators from ever vanishing.

Let's see how this works for the process we are considering. To simplify, let's make the substitution (5.27) into the troublesome first term of eq. (5.26), and ignore the second term, to try to see what's going on. We have

$$i\mathcal{M} = \frac{-ig^2}{s - M^2 + \frac{1}{4}\Gamma^2 + i\Gamma M}.$$

We then substitute this into (4.35) and (4.39) to yield

$$\sigma = \frac{D}{4\left|E_2\mathbf{p}_1 - E_1\mathbf{p}_2\right|} = \frac{1}{8Ep}\frac{p}{16\pi^2 E_{cm}}\int d\Omega \left|i\mathcal{M}\right|^2 = \frac{1}{256\pi^2 E^2}\int \frac{g^4 d\Omega}{\left(s - M^2 + \frac{1}{4}\Gamma^2\right)^2 + \Gamma^2 M^2},$$

$$\sigma = \frac{g^4}{16\pi s\left[\left(s - M^2 + \frac{1}{4}\Gamma^2\right)^2 + \Gamma^2 M^2\right]}. \quad (5.28)$$

The decay rate Γ appears in eq. (4.28).

Examining eq. (5.28), we see that if Γ is small compared to *M*, then the cross section will be large but finite near $\sqrt{s} = M$. Such *resonances* in the cross-section are often the best way to look for a new particle. Of course, in the minimal $\psi^*\psi\phi$ theory we have been discussing, one would have to include this effect on *both* terms in the amplitude (5.26), and then square the entire complex amplitude to get the cross-section. But the dramatic increase in cross-section persists, as sketched in Fig. 5-15.

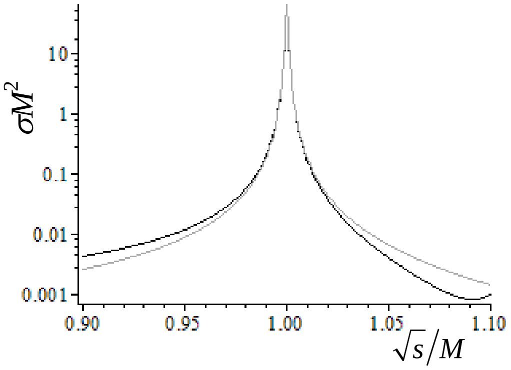


Figure 5-15: The dramatic enhancement of the cross section near the resonance coming from the ϕ particle, evaluated for $m = g = \frac{1}{4}M$. The gray line is the result from eq. (5.28), the black includes both terms from eq. (5.26).

It is worth mentioning that in many cases, it is impossible to directly measure the decay rate in terms of time. Most particles decay far too quickly to measure the time directly. You can do much better if the particle is moving quickly, not just because time dilation enhances its lifetime by a factor of $\gamma = E/m$, but mostly because we can then deduce how long the particle lasted from the distance it traveled. Still, for particles that last less than about 10^{-14} s, this technique doesn't work very well.

For much shorter times, it is much easier to measure the cross-section for the production. One can show that if Γ is small compared to *M*, then the peak will be largest near to $\sqrt{s} = M$, and the cross-section will fall to half of this value near $\sqrt{s} = M \pm \frac{1}{2}\Gamma$. It is common, therefore, to define the *full width at half maximum*, or FWHM, as the difference in $\sqrt{s}$ between these two points, and to a very good approximation $\text{FWHM} = \Gamma$. For this reason, Γ is also called the *width*, and often listed (for short lived particles) in units like MeV. For lifetimes $\Gamma^{-1} < 10^{-18}$ s, this produces widths larger than a keV, which you have at least a chance of measuring.

It should be noted that, like lifetimes, the width Γ, and hence the cross-section as a function of energy, is sensitive to the *total* decay rate of the intermediate particles. For example, when the Large Electron-Positron collider (LEP) was measuring the production of hadrons (to be defined later) near the Z resonance, they could easily measure the total decay rate of the Z, even though some of those decays (neutrinos) could not directly be detected. It was graphs like the one in Fig. 5-16 that convinced physicists that there were exactly three neutrinos.

We could explore many other theories of spin 0 particles, or scalar bosons, but the actual standard model contains only one scalar boson, the Higgs particle. It is therefore time to move on and consider something more realistic, a theory that includes fermions.

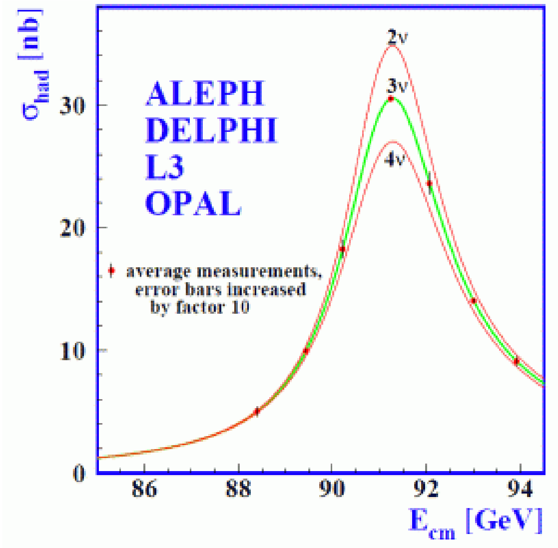


Figure 5-16: The cross-section for $e^+e^- \to \text{hadrons}$ in the neighborhood of the Z – resonance. The dots are the data, and the three curves are the predictions for the standard model including two, three, or four neutrinos. The data clearly indicates there are three neutrinos. Figure stolen from **http://t2k-experiment.org/neutrinos /oscillations-today/zwidth/**

Problems for Chapter 5

1. Fig. 5-6 has a one loop diagram contributing to the scattering $\psi\psi \to \psi\psi$. Draw at least five more.

2. In the restricted theory (g only), there is no tree-level diagram contributing to the process $\phi\phi \to \phi\phi$. Draw the one loop diagrams that *do* contribute to this process. There are six total.

3. In the restricted theory (g only), draw all six tree level diagrams contributing to $\psi\psi^* \to \phi\phi\phi$. Then draw all six diagrams contributing to $\psi\phi \to \psi\phi\phi$. You don't have to do anything else with them.

4. Starting from eq. (5.26), find the differential and total cross-section for $\psi\psi^* \to \psi\psi^*$. Work in the center of mass frame, let E be the energy of each of the initial particles, and assume you are away from resonance, so you don't have to make the modification eq. (5.27).

5. Calculate in the center of mass frame the cross-section for $\psi^*\psi^* \to \psi^*\psi^*$. Let p_1 and p_2 be the initial momenta, and p_3 and p_4 the final momenta. Once you get the Feynman invariant amplitude, I strongly recommend you compare with eq. (5.10) before finishing the problem.

6. In the restricted theory (g coupling only), find the annihilation differential cross-section and total cross-section for $\psi\psi^* \to \phi\phi$, where the energies of the initial particles be E in the center of mass frame. Note that the Feynman amplitude can be found with trivial modification of eq. (5.16).

7. In the restricted theory (g coupling only), draw the two relevant Feynman diagrams for $\psi(p)\phi(k)\to\psi(p')\phi(k')$, labeling the intermediate momenta. Find the Feynman invariant amplitude. Calculate the differential cross-section for this process in the center of mass frame. You can let p stand for the common magnitude of their momenta, E_p and E_k their initial energies, and θ can be the angle they end up at compared to where they started.

8. In the general theory with all couplings, find the Feynman amplitude (four diagrams) for $\phi\phi\to\phi\phi$, working in the cm frame where both initial particles have energy E. Now calculate the cross-section if $h=0$.

9. In the general theory with all couplings, find the Feynman amplitude (four diagrams) for $\phi\phi\to\psi\psi^*$. To simplify, let $g = 0$, and now calculate the cross-section.

10. It was claimed in section E that for M large, we could pretend that there is just the four-particle interaction sketched in Fig. 5-13, with Feynman rule $-i\lambda=-2ig^2/M^2$. Using this approach, find the cross section for $\psi\psi\to\psi\psi$. Check that it yields the same formula as we got with the exact formula in the limit $M\to\infty$. It may help to recall that $\ln(1+\varepsilon)\approx\varepsilon$.

11. In problem 4.6c, you found all relevant matrix elements for a theory with two particles, one of which had twice the charge of the other. Make up a notation for the two particles and their corresponding anti-particles and give me a complete list of Feynman rules: propagators and couplings. Let m_1 be the mass of ψ_1 and m_2 be the mass of ψ_2. Name the Hamiltonian matrix elements as:

$$\langle 0|\mathcal{H}|\psi_2^*\psi_1\psi_1\rangle=g\,,\quad \langle 0|\mathcal{H}|\psi_2\psi_1^*\psi_1^*\rangle=g^*\,,$$
$$\langle 0|\mathcal{H}|\psi_1\psi_1\psi_1^*\psi_1^*\rangle=\lambda_1\,,\quad \langle 0|\mathcal{H}|\psi_1\psi_2\psi_1^*\psi_2^*\rangle=\lambda_2\,,\quad \langle 0|\mathcal{H}|\psi_2\psi_2\psi_2^*\psi_2^*\rangle=\lambda_3\,.$$

12. Using the Feynman rules from problem 11, calculate the decay rate for $\psi_2\to\psi_1\psi_1$ in this theory. What inequality must be true for this decay to occur? Can the ψ_1 particle be unstable in this theory?

13. Using the Feynman rules from problem 11, calculate the cross-section $\psi_2\psi_1^*\to\psi_2\psi_1^*$. Is there any chance that there will be resonance?

14. Using the Feynman rules from problem 11, calculate the cross-section $\psi_1\psi_1\to\psi_1\psi_1$. Include the possibility that you are working very near resonance, $2E\approx m_2$.

15. In general, we can always add an arbitrary phase to the definition of a one particle state. Suppose we let $|\psi_2'\rangle=e^{i\theta}|\psi_2\rangle$. If we work with $|\psi_2'\rangle$ instead of $|\psi_2\rangle$, how would the three couplings in problem 11 change? Argue that by such a choice, they can all be chosen real.

VI. Fermions

It is now time to reintroduce some particles with spin, namely, spin $\frac{1}{2}$ particles, which we will call fermions. The majority of the particles in the standard model have spin $\frac{1}{2}$, and the vast majority of the particles things are made of have this spin. Electrons, protons, and neutrons, as well as the quarks from which neutrons are made, are all fermions. For fermions, to describe the particle we must not only name the type and momentum, but also the spin, $|t, p, s\rangle$. Although it is possible to choose spin in an arbitrary direction, we will normally choose our states to be helicity eigenstates, so $\hat{\mathbf{p}} \cdot \mathbf{S} |t, p, \pm\frac{1}{2}\rangle = \pm\frac{1}{2} |t, p, \pm\frac{1}{2}\rangle$.. However, we will almost always be summing or averaging over spins, so in fact, explicit forms will rarely be used. But, before we get started, we should put some math equations we will need together.

A. Math with Dirac Spinors

This section is fairly heavy with math, and we will need to get comfortable with a lot of notation. I therefore recall a number of formulas previously derived or defined, which will prove very useful:

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}, \quad \{\gamma_5, \gamma^\mu\} = 0, \quad \gamma_5^2 = 1. \tag{3.39}$$

$$\bar{\Gamma} \equiv \gamma^0 \Gamma \gamma^0 . \tag{3.42}$$

$$\left(\bar{\Psi}_A \Gamma_1 \Gamma_2 \cdots \Gamma_n \Psi_B\right)^* = \bar{\Psi}_B \bar{\Gamma}_n \cdots \bar{\Gamma}_2 \bar{\Gamma}_1 \Psi_A . \tag{3.43}$$

$$\bar{\gamma}^\mu = \gamma^\mu , \quad \bar{\gamma}_5 = -\gamma_5 . \tag{3.44}$$

$$\not{p} \equiv p_\mu \gamma^\mu . \tag{3.47}$$

$$\not{p} u(\mathbf{p}, s) = m u(\mathbf{p}, s) \quad \text{and} \quad \bar{u}(\mathbf{p}, s) \not{p} = \bar{u}(\mathbf{p}, s) m . \tag{3.48}$$

$$\not{p} v(\mathbf{p}, s) = -m v(\mathbf{p}, s) \quad \text{and} \quad \bar{v}(\mathbf{p}, s) \not{p} = -\bar{v}(\mathbf{p}, s) m . \tag{3.49}$$

$$\sum_s u(\mathbf{p}, s) \bar{u}(\mathbf{p}, s) = \not{p} + m , \quad \sum_s v(\mathbf{p}, s) \bar{v}(\mathbf{p}, s) = \not{p} - m . \tag{3.50}$$

More often than not, these will be used in situations where one or more of the Dirac matrices is contracted with momentum vectors. For example, suppose we have an expression like $\not{p}\not{k}$ and we wish we had them in the other order. Then we could rewrite this as

$$\not{p}\not{k} = \{\not{p}, \not{k}\} - \not{k}\not{p} = 2p \cdot k - \not{k}\not{p} .$$

This sort of trick will be used all the time to simplify things. Note if we set $k = p$ we can then show $\not{p}^2 = p^2$.

In addition to all these identities, we will often be working with traces of products of Dirac matrices. First recall the cyclical nature of the trace, so you can have $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. Note that this is true even if *A* and *B* are not square matrices, just so long as the number of rows and columns in *A* matches the number of columns and rows in *B* respectively.

We will often need to take the trace of the products of lots of Dirac matrices. Because of the block off-diagonal nature of the Dirac matrices in the chiral representation, it is clear from eq. (3.17) that the trace of an odd number of Dirac matrices is automatically zero. Since γ_5 is block diagonal, as we can see from eq. (3.35), multiplying by it doesn't change this result, so we have

$$\mathrm{Tr}\left(\gamma^{\mu_1}\gamma^{\mu_2}\cdots\gamma^{\mu_{2N+1}}\right)=\mathrm{Tr}\left(\gamma_5\gamma^{\mu_1}\gamma^{\mu_2}\cdots\gamma^{\mu_{2N+1}}\right)=0\,. \tag{6.1}$$

For an even numbers of γ^ν's, the situation is more complicated. If we have zero factors of γ^ν, then we have $\mathrm{Tr}\left(1\right)=4$, because 1 in this case represents a 4×4 identity matrix. For two Dirac matrices, we can use the anti-commutation relation eq. (3.39) to rewrite $\gamma^\mu\gamma^\nu=2g^{\mu\nu}-\gamma^\nu\gamma^\mu$, and then use the cyclic property of the trace to derive

$$\begin{aligned}
&\mathrm{Tr}\left(\gamma^\mu\gamma^\nu\right)=\mathrm{Tr}\left(2g^{\mu\nu}-\gamma^\nu\gamma^\nu\right)=2g^{\mu\nu}\mathrm{Tr}\left(1\right)-\mathrm{Tr}\left(\gamma^\nu\gamma^\mu\right)=8g^{\mu\nu}-\mathrm{Tr}\left(\gamma^\mu\gamma^\nu\right),\\
&\mathrm{Tr}\left(\gamma^\mu\gamma^\nu\right)=4g^{\mu\nu}\ .
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\mathrm{Tr}\left(\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta\right)&=\mathrm{Tr}\left(2g^{\mu\nu}\gamma^\alpha\gamma^\beta-\gamma^\nu\gamma^\mu\gamma^\alpha\gamma^\beta\right)=8g^{\mu\nu}\mathrm{g}^{\alpha\beta}+\mathrm{Tr}\left(-2g^{\mu\alpha}\gamma^\nu\gamma^\beta+\gamma^\nu\gamma^\alpha\gamma^\mu\gamma^\beta\right)\\
&=8g^{\mu\nu}\mathrm{g}^{\alpha\beta}-8g^{\mu\alpha}\mathrm{g}^{\beta\nu}+\mathrm{Tr}\left(2g^{\mu\beta}\gamma^\nu\gamma^\alpha-\gamma^\nu\gamma^\alpha\gamma^\beta\gamma^\mu\right)\\
&=8g^{\mu\nu}\mathrm{g}^{\alpha\beta}-8g^{\mu\alpha}\mathrm{g}^{\beta\nu}+8g^{\mu\beta}g^{\nu\alpha}-\mathrm{Tr}\left(\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta\right),\\
\mathrm{Tr}\left(\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta\right)&=4g^{\mu\nu}\mathrm{g}^{\alpha\beta}-4g^{\mu\alpha}\mathrm{g}^{\beta\nu}+4g^{\mu\beta}g^{\nu\alpha}\ .
\end{aligned}$$

This technique can be continued indefinitely. Summarizing, we have

$$\mathrm{Tr}\left(1\right)=4, \tag{6.2a}$$

$$\mathrm{Tr}\left(\gamma^\mu\gamma^\nu\right)=4g^{\mu\nu}\,, \tag{6.2b}$$

$$\mathrm{Tr}\left(\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta\right)=4\left(g^{\mu\nu}g^{\alpha\beta}+g^{\mu\beta}g^{\nu\alpha}-g^{\mu\alpha}g^{\nu\beta}\right). \tag{6.2c}$$

In general, you can only get a non-zero result if any Dirac matrices γ^μ that appear are present an even number of times. It is then not surprising that if you include $\gamma_5=i\gamma^0\gamma^1\gamma^2\gamma^3$, as defined by eq. (3.35), you must have at least four other Dirac matrices, so

$$\mathrm{Tr}\left(\gamma_5\right)=\mathrm{Tr}\left(\gamma_5\gamma^\alpha\gamma^\beta\right)=0\,, \tag{6.3a}$$

$$\mathrm{Tr}\left(\gamma_5\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta\right)=-4i\varepsilon^{\mu\nu\alpha\beta}. \tag{6.3b}$$

These equations are important enough that they have been summarized on the inside front cover. We will be using them regularly when we square the Feynman amplitude and sum over spin states, as we will in section 6C and beyond.

It's time to start working out a specific model to understand how things work. In the real world, nuclei are made of protons and neutrons, which are collectively called *nucleons*, and they are held together by the interchange of three particles called *pions*, which are a specific example of a group or particles called *mesons*. We will briefly explore this more realistic model in chapter 8. But for now, let's build a model with just one type of nucleon and one meson.

B. The $\bar{\psi}\psi\phi$ Model

We'd like to start with a simple model that is actually fairly realistic, the nucleon-meson or $\bar{\psi}\psi\phi$ model. This theory consists of a scalar particle ϕ (the meson) with mass *M* which is its own anti-particle, and a fermion particle ψ (the nucleon) that has mass *m* and is not its own anti-particle. There will be a conserved "charge" associated with the nucleon. A single meson is described entirely in terms of its momentum, $\left|\phi,k\right\rangle$, while the nucleon or anti-nucleon also has spin, $\left|\psi,p,s\right\rangle$.

What about interactions? The smallest number of particles we can have in interactions is three. One possible type of interaction we could consider would be of the form

$$\left\langle\psi,p',s'\middle|\mathcal{H}\middle|\psi,p,s;\phi,k\right\rangle = g\left(p,s;p',s';k\right), \tag{6.4}$$

where *g* is some sort of Lorentz invariant combination of the spins and momenta. As we learned in chapter 3, a fermion with momentum *p* and spin *s* has a Dirac spinor $u\left(p,s\right)$ associated with it, and we'd expect a factor of this appearing on the right side of (6.4). Similarly, for the final state particle, we'd expect there to be factor of $u^\dagger\left(p',s'\right)$ in eq. (6.4). Roughly speaking, we'd expect (6.4) to take the form

$$\left\langle\psi,p',s'\middle|\mathcal{H}\middle|\psi,p,s;\phi,k\right\rangle = u^\dagger\left(p',s'\right)\Gamma u\left(p,s\right), \tag{6.5}$$

where Γ is some horrible matrix that may depend on the various momenta in a complicated way.

But wait – we know that whatever Γ is, it should be some sort of polynomial in the momenta, with coefficients that have dimensions of non-negative powers of mass. What are the dimensions of Γ? In the exact way we did in chapter 5, we realize that $\mathcal{H}$ has dimensions of M^4 and each particle contributes M^{-1}, so the total dimension on the left is M^1. On the right, we see from eqs. (3.22) and (3.24) that Dirac spinors have dimension $M^{1/2}$, so Γ is, in fact, dimensionless. This tells us Γ in fact has no momentum dependence, but instead is a constant matrix.

But we also need eq. (6.5) to be Lorentz invariant. It turns out we know two ways of putting Dirac spinors together to make something Lorentz invariant, as given by eqs. (3.32) and (3.36). To be as general as possible, we put these together to make the combination

$$\left\langle\psi,p',s'\middle|\mathcal{H}\middle|\psi,p,s;\phi,k\right\rangle = \bar{u}\left(p',s'\right)\left(g_1+ig_2\gamma_5\right)u\left(p,s\right), \tag{6.6}$$

where g_1 and g_2 are arbitrary constants. The factor of *i* is chosen for later convenience.

According to the anti-particle property, we should be able to switch the ϕ to the other side without changing the expression, so

$$\left\langle\psi,p',s'\middle|\mathcal{H}\middle|\psi,p,s;\phi,k\right\rangle = \left\langle\psi,p',s';\phi,-k\middle|\mathcal{H}\middle|\psi,p,s\right\rangle = \bar{u}\left(p',s'\right)\left(g_1+ig_2\gamma_5\right)u\left(p,s\right).$$

The two expressions on the left side should be complex conjugates of each other if we exchange primes and non-primes, and change the sign of *k* (which doesn't appear on the right anyway) because $\mathcal{H}$ is Hermitian, so we have

$$\bar{u}(p,s)(g_1 + ig_2\gamma_5)u(p',s') = \left[\bar{u}(p',s')(g_1 + ig_2\gamma_5)u(p,s)\right]^* = \bar{u}(p,s)\left(g_1^* - ig_2^*\bar{\gamma}_5\right)u(p',s')$$
$$= \bar{u}(p,s)(g_1 + ig_2\gamma_5)u(p',s') \ ,$$
$$g_1^* + ig_2^*\gamma_5 = g_1 + ig_2\gamma_5 \ ,$$

and hence we see that g_1 and g_2 must both be real.

In general, there is nothing wrong with having a theory with both couplings g_1 and g_2. But let's assume now that parity is a symmetry that is respected by the Hamiltonian; that is, that the universe works the same if you take the mirror image of it. We already know from problem (3.3) that when you perform a parity transformation on expressions like $\bar{\psi}\psi$ it is unchanged, but $\bar{\psi}\gamma_5\psi$ gets a sign change. But what effect does parity have on a scalar particle ϕ? If you think about it, if you take a mirror image twice, you end up back where you started, so $\mathcal{P}^2 = 1$, so $\mathcal{P}^2|\phi\rangle = |\phi\rangle$. But this allows two possible outcomes, either $\mathcal{P}|\phi\rangle = +|\phi\rangle$ or $\mathcal{P}|\phi\rangle = -|\phi\rangle$. The first case is a *true scalar* particle and the latter a *pseudoscalar* particle. Either is acceptable, in other words, our coupling is going to be something like

$$\langle\psi, p', s'|\mathcal{H}|\psi, p, s; \phi, k\rangle = \begin{cases} \bar{u}(p',s')gu(p,s) & \text{for scalar particles,} \\ \bar{u}(p',s')ig\gamma_5 u(p,s) & \text{for pseudoscalar particles.} \end{cases} \tag{6.7}$$

Only experiment can resolve which of the two possibilities are realized. In the real world, pions are pseudoscalars, and we will normally work exclusively with the second case, and our coupling will be $ig\gamma_5$.

By the anti-particle property, we should also be able to move the ψ's to the other side. We already know from chapter 3 that the effect of this is to replace u's by v's. Hence there are a total of eight non-zero matrix elements, which are given by

$$\langle\psi, p', s'|\mathcal{H}|\psi, p, s; \phi, k\rangle = \langle\psi, p', s'; \phi, k|\mathcal{H}|\psi, p, s\rangle = \bar{u}(p',s')ig\gamma_5 u(p,s), \tag{6.8a}$$
$$\langle 0|\mathcal{H}|\bar{\psi}, p', s'; \psi, p, s; \phi, k\rangle = \langle\phi, k|\mathcal{H}|\bar{\psi}, p', s'; \psi, p, s\rangle = \bar{v}(p',s')ig\gamma_5 u(p,s), \tag{6.8b}$$
$$\langle\bar{\psi}, p, s; \psi, p', s'|\mathcal{H}|\phi, k\rangle = \langle\bar{\psi}, p, s; \psi, p', s'; \phi, k|\mathcal{H}|0\rangle = \bar{u}(p',s')ig\gamma_5 u(p,s), \tag{6.8c}$$
$$\langle\bar{\psi}, p, s|\mathcal{H}|\bar{\psi}, p', s'; \phi, k\rangle = \langle\bar{\psi}, p, s; \phi, k|\mathcal{H}|\bar{\psi}, p', s'\rangle = -\bar{v}(p',s')ig\gamma_5 v(p,s). \tag{6.8d}$$

Don't worry about the minus signs; this has to do with the fact that we are dealing with fermions. The rule about minus signs will be discussed in the context of Feynman diagrams later.

The complications of eqs. (6.8) might lead you to believe there will be something like four different rules for the fermion coupling, but this is not actually the case. We will think of the *coupling* as just the factor of $ig\gamma_5$. All the Dirac spinors will be thought of as being associated with the fermion lines, internal or external. It *will* lead to some slightly more complicated Feynman diagram rules. And with that note, we proceed to considering Feynman diagrams.

C. Feynman Rules for Fermions

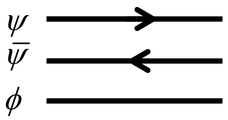


Figure 6-1: Symbols used in the $\bar{\psi}\psi\phi$ theory.

We want to develop Feynman rules for the $\bar{\psi}\psi\phi$ theory. In a manner similar to before, we will denote ϕ by a line with no adornments, a ψ by an arrow going to the right, and a $\bar{\psi}$ by an arrow going to the left, as illustrated in Fig. 6-1. There will be only one interaction, which involves one arrow going in, one going out, and a line. The general rule for going from matrix elements like eqs. (6.8) into Feynman rules is to multiply by a factor of $-i$. The only part we'll associate with the vertex itself, however, will be the matrix part, so the rule for this vertex is just $g\gamma_5$. All the factors of *u*, *v*, $\bar{u}$ and $\bar{v}$ will be associated with their corresponding fermion lines.

This leads us to our first big departure from the previous chapter. Any time we have external lines for fermions, there will be an additional factor involved. The way I remember the rule is: if it's a particle, it gets a *u*, and if it's an anti-particle, it's a *v*. If it is an arrow leaving the diagram, it has a bar on it, otherwise not.

This leads us to the propagators, the factors associated with internal lines. Recall in the previous chapter where the propagator came from. We got it from second order perturbation theory, namely, from a factor of $1/i\left(E_n - E_I\right)$ coming from eq. (2.51), together with a factor$1/2E$ from the normalization. Furthermore, a propagator represents the sum of two separate not-quite Feynman diagrams. Suppose we have some internal fermion line in a Feynman diagram, which represents the sum of two not-quite Feynman diagrams, as illustrated in Fig. 6-2. We will label as *p* the four-momentum flowing along this arrow assuming four-momentum is conserved at each vertex. Technically, we should, for the moment label these intermediate lines not only by their momentum, but also by the spin *s*.

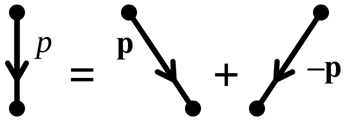


Figure 6-2: The Feynman propagator on the left is the sum from the two not-quite Feynman diagrams on the right.

Now, the propagator comes from summing the two not-quite Feynman diagrams on the right in Fig. 6-2. Consider the first one. Since it is creating and then annihilating a particle, it should contribute a factor of $u(\mathbf{p},s)\bar{u}(\mathbf{p},s)$. Similarly, for the second one, there will be an additional factor of $v(-\mathbf{p},s)\bar{v}(-\mathbf{p},s)$. There is an additional minus sign that is a little difficult to explain, but basically, it has to do with the fact that when you interchange fermions in a state vector, a minus sign is introduced. Following an argument analogous to what we did before with scalars, but including these additional factors, we would logically think the propagator should be

$$\Pi(p,s) = \frac{u(\mathbf{p},s)\bar{u}(\mathbf{p},s)}{i\left(E_p - E\right)2E_p} - \frac{v(-\mathbf{p},s)\bar{v}(-\mathbf{p},s)}{i\left(E_p + E\right)2E_p}, \tag{6.9}$$

where $E_p = \sqrt{\mathbf{p}^2 + m^2}$. We are supposed to include *all* intermediate states when computing amplitudes, however, and therefore we need to sum this on spin. We can then use eqs. (3.50) to simplify the resulting expression:

$$
\begin{aligned}
\Pi(p) = \sum_s \Pi(p,s) &= \sum_s \left[\frac{u(\mathbf{p},s)\bar{u}(\mathbf{p},s)}{i\left(E_p - E\right)2E_p} - \frac{v(-\mathbf{p},s)\bar{v}(-\mathbf{p},s)}{i\left(E_p + E\right)2E_p} \right] \\
&= \frac{E_p\gamma^0 - \mathbf{p}\cdot\boldsymbol{\gamma} + m}{i\left(E_p - E\right)2E_p} - \frac{E_p\gamma^0 + \mathbf{p}\cdot\boldsymbol{\gamma} - m}{i\left(E_p + E\right)2E_p} \\
&= \frac{i}{2}\gamma^0\left(\frac{1}{E - E_p} + \frac{1}{E + E_p} \right) + \frac{i}{2E_p}\left(\mathbf{p}\cdot\boldsymbol{\gamma} - m\right)\left(-\frac{1}{E - E_p} + \frac{1}{E + E_p} \right) \\
&= \frac{i}{E^2 - E_p^2}\left(E\gamma^0 - \mathbf{p}\cdot\boldsymbol{\gamma} + m\right) = \frac{i\left(\not{p} + m\right)}{E^2 - \mathbf{p}^2 - m^2},
\end{aligned}
$$

$$
\Pi(p) = \frac{i\left(\not{p} + m\right)}{p^2 - m^2}. \tag{6.10}
$$

This leads to the Feynman rules. Let's try to summarize them.

(1) For each incoming or outgoing ψ, draw an arrow coming in on the left, or an arrow coming out on the right. For $\bar{\psi}$ draw the arrow pointing the other way. For ϕ, draw an exterior line with no arrows.

(2) Draw all topologically distinct diagrams you can think of that connects the initial and final particles. The only vertex allowed (in this theory) is one that has one arrow going in, one going out, and one line.

(3) Label all interior lines by their four-momenta, conserving four-momentum at each vertex.

(4) For each vertex, there is one factor of $g\gamma_5$.

(5) Include a factor of u for any incoming ψ, a $\bar{u}$ for any outgoing ψ, a $\bar{v}$ for each incoming $\bar{\psi}$, and a v for each outgoing $\bar{\psi}$.

(6) For each interior ϕ line with momentum k, there is a factor of $i/\left(k^2 - M^2\right)$.

(7) For each interior ψ line with momentum p, oriented in the direction of the arrow, there is a factor of $i\left(\not{p} + m\right)/\left(p^2 - m^2\right)$.

(8) Add up the contributions from diagrams that differ by switching a pair of external boson lines. Subtract diagrams that differ by switching a pair of external fermion lines.

These rules are illustrated in Fig. 6-3. That last rule will take a specific example before we can explain it, but basically it comes from eq. (2.79). Just as was the case in the previous chapter, there are additional rules if you are going to be dealing with loops, which we will avoid.

The rules for the propagators and interactions are a little confusing as well, as it may not be immediately obvious how you put all the factors

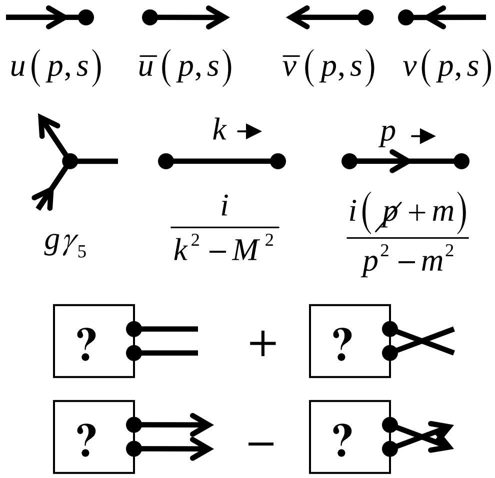


Figure 6-3: The Feynman rules for the restricted $\bar{\psi}\psi\phi$ theory.

together. Since the fermion propagator, coupling, and all the external line rules involve matrices, we must carefully specify what order we multiply things, since order matters when multiplying matrices. The way to deal with the matrices is to find an arrow pointing out of the diagram and then follow it backwards to the tail. I generally recommend doing this once for each set of connected fermion lines before including the other factors. I will give ample examples of how this computation is done in the next section.

The fermion propagator is sometimes expressed in an alternate form. Keeping in mind that $\not{p}^2 = p^2$, we can rewrite this propagator as

$$\frac{i\left(\not{p}+m\right)}{p^2-m^2}=\frac{i\left(\not{p}+m\right)}{\not{p}^2-m^2}=\frac{i\left(\not{p}+m\right)}{\left(\not{p}+m\right)\left(\not{p}-m\right)}=\frac{i}{\not{p}-m}. \tag{6.11}$$

This looks simpler, but is not, because (6.11) involves matrix division, which is not easy. My rule of thumb is that the standard formula for the propagator, $i\left(\not{p}+m\right)\big/\left(p^2-m^2\right)$, is easier to work with when doing calculations, but $i\big/\left(\not{p}-m\right)$ is easier to use for proofs.

D. Summing over Final Spins

Let's now do several examples to make clear how this works. Assume $M > 2m$, so that the decay process $\phi \to \psi\bar{\psi}$ is possible. At tree level, there is only one diagram, illustrated in Fig. 6-4. We start with the ψ particle, because that is an arrow pointing outwards. According to rule (5), this contributes a factor of $\bar{u}\left(p,s\right)$. Going backwards along the arrows, this leads us to the vertex, where we will pick up a factor of $g\gamma_5$. We then proceed back to the $\bar{\psi}$ particle, which contributes a factor of $v\left(p',s'\right)$. So far, our Feynman amplitude is then

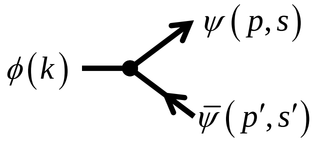


Figure 6-4: The Feynman diagram for the decay $\phi \to \psi\bar{\psi}$.

$$i\mathcal{M} = g\left(\bar{u}\gamma_5 v'\right), \tag{6.12}$$

where I have abbreviated $\bar{u} = \bar{u}\left(p,s\right)$ and $v' = v\left(p',s'\right)$. And, in fact, there is no factor for the external boson line, so this is the amplitude. Note how I put parentheses around the combination $\left(\bar{u}\gamma_5 v'\right)$ to remind me that the factors inside are matrices, and I should be cautious about putting anything in the middle of it. These expressions will always start with the bar of a Dirac spinor and end with an unbarred Dirac spinor.

Let's proceed forward and see if we can't find the decay rate. We first need

$$\left|i\mathcal{M}\right|^2 = \left(i\mathcal{M}\right)\left(i\mathcal{M}\right)^* = g^2\left(\bar{u}\gamma_5 v'\right)\left(\bar{u}\gamma_5 v'\right)^*$$

Using eqs. (3.43) and (3.44), we know how to take the complex conjugate of something like this. We therefore have

$$\left|i\mathcal{M}\right|^2 = -g^2\left(\bar{u}\gamma_5 v'\right)\left(\bar{v}'\gamma_5 u\right). \tag{6.13}$$

We could easily proceed forward and calculate the decay rate to a particular combination of momentum and spin states (p, s) and (p', s'). The rate would go to zero in the infinite volume limit.

The solution is to sum over all possible final states, which means summing over all the final momenta states *and* sum over the final spin states. This is something we have not encountered before. In general when calculating a cross-section or decay rate, ***always sum over the final state spins***. Therefore, what we really want is not eq. (6.13), but instead

$$\sum_{s,s'} |i\mathcal{M}|^2 = -g^2 \sum_{s,s'} \left(\bar{u}\gamma_5 v' \bar{v}' \gamma_5 u\right). \tag{6.14}$$

I deliberately smushed the two factors in parentheses together to point out that we have v' right next to $\bar{v}'$, and therefore we can use eq. (3.50) to get rid of the sum over spins and these Dirac spinors in one fell swoop. Notice also that I could have written the factors in (6.13) in the other order, as $(\bar{v}'\gamma_5 u)(\bar{u}\gamma_5 v') = (\bar{v}'\gamma_5 u\bar{u}\gamma_5 v')$, in which case I could have gotten rid of the *u*'s instead. If only there were some way to do both…

But wait! We suddenly remember that (6.14), though a number, can be thought of as a 1×1 matrix. We can take the trace of this matrix without having any effect, and then we can use the cyclic property of the trace to bring the *u* on the right to the left, as follows:

$$\sum_{s,s'} |i\mathcal{M}|^2 = -g^2 \sum_{s,s'} \mathrm{Tr}\left(\bar{u}\gamma_5 v' \bar{v}' \gamma_5 u\right) = -g^2 \sum_{s,s'} \mathrm{Tr}\left(u\bar{u}\gamma_5 v' \bar{v}' \gamma_5\right).$$

Using eq. (3.50), this becomes

$$\sum_{s,s'} |i\mathcal{M}|^2 = -g^2 \mathrm{Tr}\left[\left(\not{p} + m\right)\gamma_5 \left(\not{p}' - m\right)\gamma_5\right]. \tag{6.15}$$

We need to simplify these traces. We'd like to bring the γ_5's together, so we can use the fact that $\gamma_5^2 = 1$. We note that γ_5 obviously commutes with *m* (which is just a number), but since it anti-commutes with γ^μ we have $\gamma_5\left(\not{p} - m\right) = \left(-\not{p} - m\right)\gamma_5$, so we have

$$\sum_{s,s'} |i\mathcal{M}|^2 = g^2 \mathrm{Tr}\left[\left(\not{p} + m\right)\left(\not{p}' + m\right)\gamma_5\gamma_5\right] = g^2 \mathrm{Tr}\left[\left(\not{p} + m\right)\left(\not{p}' + m\right)\right],$$

$$\sum_{s,s'} |i\mathcal{M}|^2 = g^2 \mathrm{Tr}\left(\not{p}\not{p}' + m\not{p} + m\not{p}' + m^2\right) = g^2\left(4p\cdot p' + 4m^2\right). \tag{6.16}$$

We are almost done computing the decay rate, but will need the dot product in (6.16). This can be most easily obtained using conservation of momentum, $k = p + p'$, or squaring this,

$$M^2 = k^2 = \left(p + p'\right)^2 = p^2 + p'^2 + 2p\cdot p' = 2m^2 + 2p\cdot p'.$$

Solving for $p\cdot p'$ and substituting in (6.16), we obtain

$$\sum_{s,s'} |i\mathcal{M}|^2 = 2g^2 M^2. \tag{6.17}$$

We can then use this in eqs. (4.35) and (4.38) to obtain

$$\Gamma = \frac{D}{2M} = \frac{1}{2M}\frac{p}{16\pi^2 M}\int d\Omega \sum_{s,s'} \left| i\mathcal{M} \right|^2 = \frac{p}{32\pi^2 M^2} 2g^2 M^2 \int d\Omega = \frac{g^2 p}{4\pi} .$$

Since the two final state particles have equal mass, they will split the energy evenly, and will hence each have energy $\frac{1}{2}M$, and hence their momentum is $p = \sqrt{\frac{1}{4}M^2 - m^2}$, so

$$\Gamma = \frac{g^2}{8\pi}\sqrt{M^2 - 4m^2} . \qquad (6.18)$$

Keeping in mind that g is dimensionless, this has units of mass, as it must.

E. Averaging Over Initial Spins

Let's now calculate the cross section $\psi\bar{\psi} \to \phi\phi$ There are two relevant Feynman diagrams, as sketched in Fig. 6-5. As always, we conserve four-momenta at the relevant vertices. Note that we have to get the direction of the momentum right – it is flowing to the right for all external lines, but for the interior fermion line, it flows in the direction of the arrow. It is not obvious which way we want to write this intermediate momentum, since the first one is $p - k = k' - p'$ and the second is $p - k' = k - p'$. I have chosen an expression that will lead to later simplification. If we had chosen another choice, we would ultimately get the same result, but only after more work. Note also from Fig. 6-5 that I put a big plus sign between the two diagrams. If you imagine taking the first diagram, and switch the external lines for k and k', you can turn the first diagram into the second. Since they differ only by the switching of a boson, they get a relative plus sign.

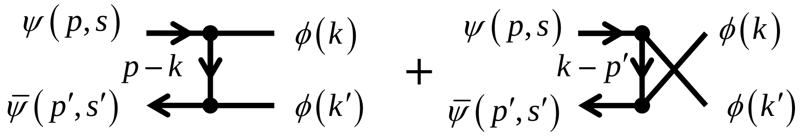


Figure 6-5: The two Feynman diagrams contributing to $\psi\bar{\psi} \to \phi\phi$.

Now that we have them all carefully drawn, we put in the relevant factors. For example, for the first diagram, we start at the arrow out for a factor of $\bar{v}'$, then a coupling $g\gamma_5$, a fermion propagator, another $g\gamma_5$, and finally a factor of u from the incoming fermion. The other diagram is identical except that the propagator has a different momentum. The Feynman amplitude, therefore, is

$$i\mathcal{M} = g^2 \bar{v}' \gamma_5 \frac{i\left(\not{p} - \not{k} + m\right)}{\left(p-k\right)^2 - m^2}\gamma_5 u + g^2 \bar{v}' \gamma_5 \frac{i\left(\not{k} - \not{p}' + m\right)}{\left(k-p'\right)^2 - m^2}\gamma_5 u . \qquad (6.19)$$

There is a straightforward way to proceed from here to a final answer, but if you follow it you will go mad. The expressions are already pretty complicated, and all those Dirac matrices and Dirac spinors will proliferate when you square it. We therefore very much would like to simplify it first. The denominators can be simplified pretty straightforwardly, for what it's worth:

$$\left(p-k\right)^2 - m^2 = p^2 + k^2 - 2p\cdot k - m^2 = M^2 - 2p\cdot k\,,$$
$$\left(k-p'\right)^2 - m^2 = k^2 + p'^2 - 2p'\cdot k - m^2 = M^2 - 2p'\cdot k\,.$$

What would really simplify things is if we could get rid of those γ_5's since they square to one, but they aren't next to each other. No problem! They anti-commute with the other Dirac matrices and commute with the *m*'s, so a bit of work and (6.19) becomes

$$i\mathcal{M} = ig^2 \frac{\bar{v}'\left(\not{k} - \not{p} + m\right)u}{M^2 - 2p\cdot k} + ig^2 \frac{\bar{v}'\left(\not{p}' - \not{k} + m\right)u}{M^2 - 2p'\cdot k}\,. \tag{6.20}$$

And now we have a bit of luck. From eqs. (3.48) and (3.49), $\not{p}u = mu$ and $\bar{v}'\not{p}' = -\bar{v}'m$, which simplifies eq. (6.20) considerably:

$$i\mathcal{M} = ig^2\left(\bar{v}'\not{k}\,u\right)\left(\frac{1}{M^2 - 2p\cdot k} - \frac{1}{M^2 - 2p'\cdot k}\right). \tag{6.21}$$

The matrix part of (6.21) has simplified to the point where it looks tractable.

We need to square eq. (6.21). We know we are supposed to sum over final state spins, but what about the initial state spins? For the initial states, *most* beams are *unpolarized*, which means that the initial state spin is just as likely to be spin up as spin down. Hence we ***average over the initial spins***, i.e. sum them and divide by two, or in this case four since there are two spins. We therefore need to calculate

$$\frac{1}{4}\sum_{s,s'}\left|i\mathcal{M}\right|^2 = g^4\left(\frac{1}{M^2 - 2p\cdot k} - \frac{1}{M^2 - 2p'\cdot k}\right)^2 \frac{1}{4}\sum_{s,s'}\left(\bar{u}\not{k}v'\right)\left(\bar{v}'\not{k}u\right). \tag{6.22}$$

The sums on spins turn into traces, which then are expanded. Keeping in mind that only an even number of Dirac matrices has a non-zero trace, we have

$$\begin{aligned}\tfrac{1}{4}\sum_{s,s'}\left(\bar{u}\not{k}v'\right)\left(\bar{v}'\not{k}u\right) &= \tfrac{1}{4}\mathrm{Tr}\left[\left(\not{p}+m\right)\not{k}\left(\not{p}'-m\right)\not{k}\right] = \tfrac{1}{4}\mathrm{Tr}\left(\not{p}\not{k}\not{p}'\not{k} - m^2\not{k}\not{k}\right)\\ &= 2\left(p\cdot k\right)\left(p'\cdot k\right) - \left(p\cdot p'\right)k^2 - m^2k^2\,.\end{aligned} \tag{6.23}$$

We will need to substitute explicit expressions for the dot products, so let's start working them out in the center of mass frame. We'll let the initial ψ come in along the *z*-axis, then $p^\mu = \left(E,0,0,p\right)$, and since we are in the center of mass frame, the $\bar{\psi}$ must have cancelling momentum $p'^\mu = \left(E,0,0,-p\right)$. The final state particles must share the same total energy, and therefore they must have energy *E* as well, but their momentum *k* will be different, and their direction will be arbitrary, though they will be moving in opposite directions. The four momenta will therefore be

$$\begin{aligned}p^\mu &= \left(E,0,0,p\right), \quad k^\mu = \left(E, k\sin\theta\cos\phi, k\sin\theta\sin\phi, k\cos\theta\right),\\ p'^\mu &= \left(E,0,0,-p\right), \quad k'^\mu = \left(E, -k\sin\theta\cos\phi, -k\sin\theta\sin\phi, -k\cos\theta\right).\end{aligned}$$

From these we obtain all the dot products we will need:

$$p\cdot k = E^2 - pk\cos\theta\,, \quad p'\cdot k = E^2 - pk\cos\theta\,, \quad p'\cdot p = E^2 + p^2\,, \tag{6.24}$$

and, of course, $k^2 = M^2$. Substituting these into (6.23) and (6.24), we have

$$\begin{aligned}\frac{1}{4}\sum_{s,s'}\left|i\mathcal{M}\right|^2 &= g^4\left(\frac{1}{M^2-2E^2+2pk\cos\theta}-\frac{1}{M^2-2E^2-2pk\cos\theta}\right)^2\times\\ &\quad\times\left[2\left(E^2-pk\cos\theta\right)\left(E^2+pk\cos\theta\right)-M^2\left(E^2+p^2+m^2\right)\right]\\ &=\frac{32g^4p^2k^2\cos^2\theta}{\left[\left(2E^2-M^2\right)^2-4p^2k^2\cos^2\theta\right]^2}\left[E^4-p^2k^2\cos^2\theta-M^2E^2\right],\\ \frac{1}{4}\sum_{s,s'}\left|i\mathcal{M}\right|^2 &= \frac{32g^4p^2k^4\cos^2\theta\left(E^2-p^2\cos^2\theta\right)}{\left[\left(2E^2-M^2\right)^2-4p^2k^2\cos^2\theta\right]^2}.\end{aligned} \tag{6.25}$$

It is now a short but messy step to go from here to the cross-section, or at least the differential cross-section:

$$\sigma=\frac{D}{4\left|Ep+Ep\right|}=\frac{1}{8Ep}\frac{k}{16\pi^2\left(2E\right)}\int\tfrac{1}{4}\sum_{s,s'}\left|i\mathcal{M}\right|^2d\Omega=\frac{32g^4pk^5}{256\pi^2E^2}\int\frac{\cos^2\theta\left(E^2-p^2\cos^2\theta\right)d\Omega}{\left[\left(2E^2-M^2\right)^2-4p^2k^2\cos^2\theta\right]^2},$$

$$\frac{d\sigma}{d\Omega}=\frac{g^4k^5p\cos^2\theta\left(E^2-p^2\cos^2\theta\right)}{8\pi^2E^2\left[\left(2E^2-M^2\right)^2-4p^2k^2\cos^2\theta\right]^2}.$$

To get the total cross-section, we must be a little careful, since we have identical particles in the final states. As describes in section 5C, we need to include a factor of $\tfrac{1}{2}$. The $\int d\phi$ just gives a factor of 2π, but the $\cos\theta$ integral just gives a mess, which I will not bother displaying in general

To simplify, let's treat the scalar as massless compared to the fermion. This is fairly realistic if we are thinking of nucleons and mesons, since $M^2/m^2 \approx 0.02$. Then we have $M^2 = 0$ and $k^2 = E^2 - M^2 = E^2$, and the differential cross-section simplifies to

$$\frac{d\sigma}{d\Omega}=\frac{g^4p\cos^2\theta}{128\pi^2E\left(E^2-p^2\cos^2\theta\right)}.$$

The remaining integral isn't too bad and yields

$$\sigma=\frac{g^4}{64\pi p^2}\left[\tanh^{-1}\left(\frac{p}{E}\right)-\frac{p}{E}\right]=\frac{g^4}{16\pi\left(s-4m^2\right)}\left[\ln\left(\frac{\sqrt{s}+\sqrt{s-4m^2}}{2m}\right)-\sqrt{1-\frac{4m^2}{s}}\right].$$

We also made the answer Lorentz invariant, using the relationships $\tfrac{1}{4}s = E^2 = p^2 + m^2$.

F. Swapping Fermion Lines

We have not yet done any examples involving the exchanging of fermion lines. Consider, for example, the scattering process $\psi\bar{\psi} \to \psi\bar{\psi}$. There are two Feynman diagrams, as illustrated in Fig. 6-6. And why is there a relative minus sign between them? Keeping in mind that Feynman diagrams are topological in nature, we could grab the two outgoing arrows p_2 and p_3 and exchange them, and this would turn one Feynman diagram into the other.

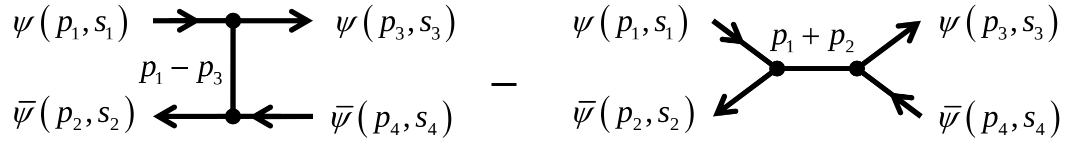


Figure 6-6: The two Feynman diagrams contributing to $\psi\bar{\psi} \to \psi\bar{\psi}$.

We now want to start actually calculating these diagrams. I'll start with the diagram on the left, which has two continual fermion lines. We work our way along the upper line from the arrow out back to the arrow in, which yields a factor of $g\left(\bar{u}_3\gamma_5 u_1\right)$. The lower line yields a factor of $g\left(\bar{v}_2\gamma_5 v_4\right)$. Finally, there is the boson propagator, which you can multiply anywhere it is convenient. For the second diagram, you have a factor of $g\left(\bar{v}_2\gamma_5 u_1\right)$ and $g\left(\bar{u}_3\gamma_5 v_4\right)$, and another boson propagator. Putting everything together, the Feynman amplitude is

$$i\mathcal{M} = \frac{ig^2\left(\bar{u}_3\gamma_5 u_1\right)\left(\bar{v}_2\gamma_5 v_4\right)}{\left(p_1 - p_3\right)^2 - M^2} - \frac{ig^2\left(\bar{v}_2\gamma_5 u_1\right)\left(\bar{u}_3\gamma_5 v_4\right)}{\left(p_1 + p_2\right)^2 - M^2}. \tag{6.26}$$

This expression is about to get very messy, so let's go ahead and take a limit to try to make it a bit more tractable. Let's arbitrarily decide that *M* is very heavy, much higher than any other masses or momenta in the problem, and *m* is very light, much smaller than the momenta in the problem. Then we always ignore *m* terms and we always neglect momenta compared to *M*. Then eq. (6.26) simplifies to

$$i\mathcal{M} = -\frac{ig^2}{M^2}\left[\left(\bar{u}_3\gamma_5 u_1\right)\left(\bar{v}_2\gamma_5 v_4\right) - \left(\bar{v}_2\gamma_5 u_1\right)\left(\bar{u}_3\gamma_5 v_4\right)\right]. \tag{6.27}$$

We will need to multiply this by its complex conjugate, which is

$$\left(i\mathcal{M}\right)^* = \frac{ig^2}{M^2}\left[\left(\bar{u}_1\gamma_5 u_3\right)\left(\bar{v}_4\gamma_5 v_2\right) - \left(\bar{u}_1\gamma_5 v_2\right)\left(\bar{v}_4\gamma_5 u_3\right)\right] , \tag{6.28}$$

where the process of barring caused us to get two minus signs in each term, one from each γ_5. We now need to multiply eqs. (6.27) and (6.28), which yields

$$\left|i\mathcal{M}\right|^2 = \frac{g^4}{M^4}\left[\left(\bar{u}_3\gamma_5 u_1\right)\left(\bar{v}_2\gamma_5 v_4\right) - \left(\bar{v}_2\gamma_5 u_1\right)\left(\bar{u}_3\gamma_5 v_4\right)\right]\left[\left(\bar{u}_1\gamma_5 u_3\right)\left(\bar{v}_4\gamma_5 v_2\right) - \left(\bar{u}_1\gamma_5 v_2\right)\left(\bar{v}_4\gamma_5 u_3\right)\right].$$

Unfortunately, there isn't really much we can do other than multiply this all out, yielding four very complicated terms. I will, however, use a bit of discretion when doing so, since any quantity in parentheses is a pure number, so we can write the combinations in any order we want. We have

$$\left|i\mathcal{M}\right|^2 = \frac{g^4}{M^4}\left[\begin{array}{l}(\bar{u}_3\gamma_5 u_1)(\bar{u}_1\gamma_5 u_3)(\bar{v}_2\gamma_5 v_4)(\bar{v}_4\gamma_5 v_2)-(\bar{u}_3\gamma_5 u_1)(\bar{u}_1\gamma_5 v_2)(\bar{v}_2\gamma_5 v_4)(\bar{v}_4\gamma_5 u_3)\\ -(\bar{v}_2\gamma_5 u_1)(\bar{u}_1\gamma_5 u_3)(\bar{u}_3\gamma_5 v_4)(\bar{v}_4\gamma_5 v_2)+(\bar{v}_2\gamma_5 u_1)(\bar{u}_1\gamma_5 v_2)(\bar{u}_3\gamma_5 v_4)(\bar{v}_4\gamma_5 u_3)\end{array}\right].$$

I arranged them so that, wherever possible, each grouping ends with the same Dirac spinor that the next grouping begins with. We can now introduce $\mathrm{Tr}(\ )$ wherever we want, as long as it includes complete groupings. The goal is to choose it so it starts and ends with the same Dirac spinor. We therefore have

$$\left|i\mathcal{M}\right|^2 = \frac{g^4}{M^4}\left[\begin{array}{l}\mathrm{Tr}(\bar{u}_3\gamma_5 u_1\bar{u}_1\gamma_5 u_3)\mathrm{Tr}(\bar{v}_2\gamma_5 v_4\bar{v}_4\gamma_5 v_2)-\mathrm{Tr}(\bar{u}_3\gamma_5 u_1\bar{u}_1\gamma_5 v_2\bar{v}_2\gamma_5 v_4\bar{v}_4\gamma_5 u_3)\\ -\mathrm{Tr}(\bar{v}_2\gamma_5 u_1\bar{u}_1\gamma_5 u_3\bar{u}_3\gamma_5 v_4\bar{v}_4\gamma_5 v_2)+\mathrm{Tr}(\bar{v}_2\gamma_5 u_1\bar{u}_1\gamma_5 v_2)\mathrm{Tr}(\bar{u}_3\gamma_5 v_4\bar{v}_4\gamma_5 u_3)\end{array}\right].$$

It's time to deal with the sums on spins. Because we have two spins in the initial states, we will want to average over these spins, so we want a factor of $\frac{1}{4}$ as well. Keeping in mind that we are setting *m* = 0, we have

$$\frac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 = \frac{g^4}{4M^4}\left[\begin{array}{l}\mathrm{Tr}(\not{p}_3\gamma_5\not{p}_1\gamma_5)\mathrm{Tr}(\not{p}_2\gamma_5\not{p}_4\gamma_5)-\mathrm{Tr}(\not{p}_3\gamma_5\not{p}_1\gamma_5\not{p}_2\gamma_5\not{p}_4\gamma_5)\\ -\mathrm{Tr}(\not{p}_2\gamma_5\not{p}_1\gamma_5\not{p}_3\gamma_5\not{p}_4\gamma_5)+\mathrm{Tr}(\not{p}_2\gamma_5\not{p}_1\gamma_5)\mathrm{Tr}(\not{p}_3\gamma_5\not{p}_4\gamma_5)\end{array}\right].$$

We anti-commute the γ_5's together and make them go away, and then we take the traces, and we obtain

$$\begin{aligned}\frac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 &= \frac{g^4}{4M^4}\left[\begin{array}{l}\mathrm{Tr}(\not{p}_3\not{p}_1)\mathrm{Tr}(\not{p}_2\not{p}_4)-\mathrm{Tr}(\not{p}_3\not{p}_1\not{p}_2\not{p}_4)\\ -\mathrm{Tr}(\not{p}_2\not{p}_1\not{p}_3\not{p}_4)+\mathrm{Tr}(\not{p}_2\not{p}_1)\mathrm{Tr}(\not{p}_3\not{p}_4)\end{array}\right]\\ &= \frac{g^4}{M^4}\left\{\begin{array}{c}4(p_1\cdot p_3)(p_2\cdot p_4)+4(p_1\cdot p_2)(p_3\cdot p_4)\\ -2\left[(p_1\cdot p_3)(p_2\cdot p_4)+(p_1\cdot p_2)(p_3\cdot p_4)-(p_1\cdot p_4)(p_3\cdot p_2)\right]\end{array}\right\},\end{aligned}$$

$$\frac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 = \frac{2g^4}{M^4}\left[(p_1\cdot p_3)(p_2\cdot p_4)+(p_1\cdot p_2)(p_3\cdot p_4)+(p_1\cdot p_4)(p_3\cdot p_2)\right]. \qquad (6.29)$$

Working in the center of mass frame, the energy *E* of each particle and the space momenta are all the same. If the two incoming particles are coming in on the x^3-axis, then we can choose our momenta to be

$$\begin{aligned}p_1^\mu &= (E,0,0,E), \quad p_3^\mu = (E, E\sin\theta\cos\phi, E\sin\theta\sin\phi, E\cos\theta),\\ p_2^\mu &= (E,0,0,-E), \quad p_4^\mu = (E,-E\sin\theta\cos\phi,-E\sin\theta\sin\phi,-E\cos\theta).\end{aligned}$$

The dot products we need are then

$$p_1\cdot p_2 = p_3\cdot p_4 = 2E^2, \quad p_1\cdot p_3 = p_2\cdot p_4 = E^2 - E^2\cos\theta, \quad p_1\cdot p_4 = p_3\cdot p_2 = E^2 + E^2\cos\theta. \qquad (6.30)$$

Substituting eqs. (6.30) into (6.29), we therefore have

$$\frac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 = \frac{2g^4E^4}{M^4}\left[\left(1-\cos\theta\right)^2+4+\left(1+\cos\theta\right)^2\right] = \frac{4g^4E^4}{M^4}\left(3+\cos^2\theta\right).$$

We now proceed in the usual way towards the cross section. We have

$$\sigma = \frac{D}{4\left(E^2+E^2\right)} = \frac{1}{8E^2}\frac{E}{16\pi^2\left(2E\right)}\int \tfrac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 d\Omega$$

$$= \frac{1}{256\pi^2E^2}\int d\Omega\frac{4g^4E^4}{M^4}\left(3+\cos^2\theta\right) = \frac{g^4E^2}{64\pi^2M^4}\int d\Omega\left(3+\cos^2\theta\right),$$

$$\frac{d\sigma}{d\Omega} = \frac{g^4E^2}{64\pi^2M^4}\left(3+\cos^2\theta\right), \quad \text{and} \quad \sigma = \frac{5g^4E^2}{24\pi M^4} = \frac{5g^4s}{96\pi M^4}.$$

G. Other Interactions

Up to now, we have been assuming there is only one type of interaction, of the type given by the coupling *g*, with corresponding Feynman rule $g\gamma_5$ for the pseudoscalar theory. Are there other renormalizable interactions in the $\bar{\psi}\psi\phi$ model? What if we had chosen the scalar interaction instead?

We already discovered that matrix elements of the form $\langle\psi|\mathcal{H}|\psi\phi\rangle$ are characterized by a single dimensionless parameter *g*. The reason is that we needed to also put in two Dirac spinors (collectively dimension *M*) so that $\langle\psi|\mathcal{H}|\psi\phi\rangle$, which is dimension *M* as well, and hence the remaining factor *g* came out dimensionless. If we got greedy and put even one more particle in, say $\langle\psi|\mathcal{H}|\psi\phi\phi\rangle$, we know that $|\phi\rangle$ has dimensions M^{-1}, and hence the resulting parameter describing this coupling would also be dimension M^{-1}, and hence not allowed. So you can't add anything to $\langle\psi|\mathcal{H}|\psi\phi\rangle$. You also can't have only a single ψ or $\bar{\psi}$, since that would not only violate charge conservation, but also conservation of angular momentum (you always need an even number of fermions).

It follows that there are no more interactions involving fermions in this theory. There could be other interactions involving only scalars, like

$$\langle 0|\mathcal{H}|\phi\phi\phi\rangle = h \quad \text{or} \quad \langle 0|\mathcal{H}|\phi\phi\phi\phi\rangle = \lambda. \qquad (6.31)$$

The former violates parity if ϕ is a pseudoscalar, since letting parity act on it produces three minus signs, so it isn't allowed. Hence there is only one other possible interaction in the pseudoscalar theory. The two interactions that are possible in this theory are sketched in Fig. 6-7. Since the only simple computation where this new interaction is relevant would be $\phi\phi\to\phi\phi$ scattering, and includes no fermions, it is mostly irrelevant for our purposes.

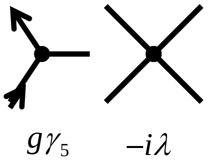


Figure 6-7: The couplings in the full $\bar{\psi}\psi\phi$ pseudoscalar theory.

What if we prefer a scalar interaction to a pseudoscalar? Then the parameter *g* has no matrix attached to it, and the resulting Feynman rule would lead to a factor of $-ig$. The boson and fermion propagators (and the rest of the rules) would all stay the same. The additional interactions as described in (6.31) would both be acceptable. The three Feynman rules for a scalar boson interaction are given in Fig. 6-8.

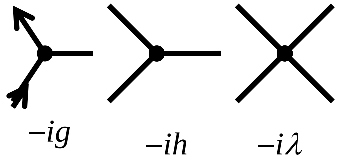


Figure 6-8: The couplings in the full $\bar{\psi}\psi\phi$ scalar theory.

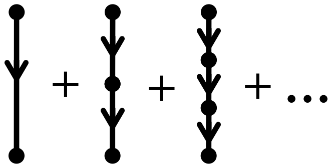


Figure 6-9: In the presence of the coupling *B*, a propagator turns into a sum of an infinite number of terms.

In either theory, we could also consider interactions with fewer than three particles. Two of these, interactions of the form $\langle 0|\mathcal{H}|\phi\phi\rangle$ and $\langle 0|\mathcal{H}|\phi\rangle$, will act exactly as they did in section 5D, the first adding a constant to M^2 and the latter corresponding to a redefinition of the ϕ field. But consider a basic interaction of the form $\langle 0|\mathcal{H}|\bar{\psi}\psi\rangle$ or one of the related diagrams. For example, suppose we have

$$\langle \psi, p', s'|\mathcal{H}|\psi, p, s\rangle = B\bar{u}(p', s')u(p, s) \qquad (6.32)$$

This will lead to a new interaction which will have only one arrow in and one out, with a corresponding factor of simply $-iB$.

Consider some internal fermion propagator carrying four-momentum *p*. If we add in the *B* interaction, it will result in an infinite sum of diagrams, as sketching in Fig. 6-9. As a result, the propagator will be modified to

$$\Pi(p) = \frac{i}{\not{p}-m} + \frac{i}{\not{p}-m}(-iB)\frac{i}{\not{p}-m} + \frac{i}{\not{p}-m}(-iB)\frac{i}{\not{p}-m}(-iB)\frac{i}{\not{p}-m} + \cdots.$$

This can again be seen to be a geometric series, which we can sum to yield

$$\Pi(p) = \frac{i}{\not{p}-m}\left(1-\frac{B}{\not{p}-m}\right)^{-1} = \frac{i}{\not{p}-m}\frac{\not{p}-m}{\not{p}-m-B} = \frac{i}{\not{p}-m-B}.$$

In other words, the interaction (6.32) simply contributes to the mass. This is very similar to the situation with bosons, the most important difference being that for fermions the Hamiltonian matrix elements give rise to factors of mass to the first power. When thinking about masses, we'll ignore the Dirac spinors and write expressions like

$$\langle \psi|\mathcal{H}|\psi\rangle = m \quad \text{for fermions}, \qquad (6.33a)$$

$$\langle \phi|\mathcal{H}|\phi\rangle = M^2 \quad \text{for bosons}. \qquad (6.33b)$$

We have built a rather complicated and interesting theory involving interactions of scalars and fermions, and with a little modification, it serves as a pretty realistic model of how nucleons interact via pions. We will return to it briefly in chapter 8, when we discuss strong interactions. But for now, let us set this model aside and build our first fully realistic model: quantum electrodynamics.

Problems for Chapter 6

1. Simplify $\sum_s \bar{u}(p,s) M u(p,s)$ for the matrices $M = 1, \gamma_5, \gamma^\mu, \gamma_5\gamma^\mu$ and $\gamma^\mu\gamma^\nu$.

2. If $i\mathcal{M} = a\left[\bar{u}(p,s)(\not{p} + \not{p}')(1-\gamma_5)v(p',s')\right]$, where *a* is constant, simplify $\sum_{s,s'} |i\mathcal{M}|^2$ as much as possible. Assume the mass associated with *p* is *m*, so $p^2 = m^2$, and the mass associated with p' is 0.

3. For the pseudoscalar coupling with only *g* couplings, draw all six Feynman diagrams for $\psi\bar{\psi} \to \phi\phi\phi$. If you include the more general coupling of Fig. 6-7, what additional diagram(s) would you need to include?

4. Calculate the decay rate $\phi \to \psi\bar{\psi}$ if we have scalar instead of pseudoscalar couplings.

5. Write the full Feynman amplitude for $\psi(p)\phi(k) \to \psi(p')\phi(k')$ for pseudoscalar couplings. Show it can be simplified to an expression of the form $(\bar{u}'\not{k}u) f$, where *f* is some function of the momenta. Work out the differential cross-section $d\sigma/d\Omega$ in the cm frame. Let E_p and E_k be the energies of the initial particles, and $p = |\mathbf{p}| = |\mathbf{k}|$ be their common momenta.

6. Starting from eq. (6.26), work out the differential and total cross-section for $\psi(p_1)\bar{\psi}(p_2) \to \psi(p_3)\bar{\psi}(p_4)$ in the limit $m = M = 0$. The result should be very simple, and should be easy to integrate over angles (if it isn't, you made an error somewhere). For an added challenge, let $m \neq 0$ but keep $M = 0$.

7. Write the full Feynman amplitude for $\phi(k)\phi(k') \to \psi(p)\psi^*(p')$ for pseudoscalar couplings. Show it can be simplified to an expression of the form $(\bar{u}\not{k}v') f$, where *f* is some function of the momenta. Work out the differential cross-section $d\sigma/d\Omega$ in the cm frame.

8. Write down the full Feynman amplitude for $\psi\psi \to \psi\psi$ in general, keeping careful track of relative sign. Let p_1 and p_2 be the initial momenta and p_3 and p_4 the final momenta. Calculate the differential and total cross-section in the approximation $m = 0$ and $E \ll M$.

9. Repeat question 8 for $m = M = 0$. For an added challenge, let $m \neq 0$ but $M = 0$.

10. Suppose we have two fermions ψ_1 and ψ_2 with masses m_1 and m_2, each of which has pseudoscalar couplings to the ϕ of mass *M*, with strength g_1 and g_2, as sketched in Fig. 6-10. What is the total decay rate of the ϕ?

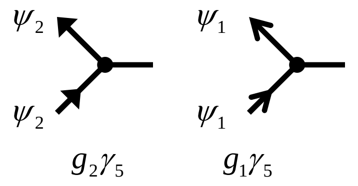


Figure 6-10: The couplings for problems 10 through 13.

11. Find the cross-section for $\psi_1\bar{\psi}_1 \to \psi_2\bar{\psi}_2$ in the theory of Fig. 6-10. Assume we are not near the ϕ resonance. Don't make any other approximations or assumptions.

12. Using the interactions in Fig. 6-10, find the differential and total cross section in the center of mass frame for $\psi_1\psi_2 \to \psi_1\psi_2$. Let the initial four momenta be p_1 and p_2 and the final four momenta be p_1' and p_2', and let $p = |\mathbf{p}_1| = |\mathbf{p}_2|$, and the initial energies E_1 and E_2.

13. Find the differential and total cross-sections for $\psi_1\bar{\psi}_2 \to \psi_1\bar{\psi}_2$ and $\bar{\psi}_1\bar{\psi}_2 \to \bar{\psi}_1\bar{\psi}_2$. Label the momenta the same way you did in problem 10. I strongly recommend you compare the result of $\sum_s |i\mathcal{M}|^2$ between this problem and problem 10 before completing the problem.

14. Consider a theory with two fermions ψ_1 and ψ_2 with masses m_1 and m_2, with a pseudoscalar ϕ with mass $M = 0$ that converts one to the other, so there are two couplings as sketched in Fig. 6-11. Calculate the decay rate $\psi_2 \to \psi_1\phi$.

15. In a theory with the couplings of Fig. 6-11, calculate the decay rate for $\phi \to \psi_1\psi_2^*$ and $\phi \to \psi_2\psi_1^*$, but this time assume $M > m_2 > m_1 = 0$.

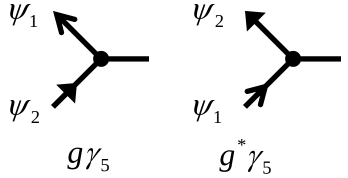


Figure 6-11: The couplings for problems 14 through 17.

16. Calculate the differential and total cross section for $\psi_1(p_1)\psi_2^*(p_2) \to \psi_1(p_1')\psi_2^*(p_2')$ in the theory with coupling described in Fig. 6-11. Let the initial energies be E_1 and E_2, and their common momenta p. Keep all three masses arbitrary.

17. Calculate the differential and total cross section for $\psi_1(p_1)\psi_1^*(p_1') \to \psi_2(p_2)\psi_2^*(p_2')$ in the theory with coupling described in Fig. 6-11. Keep all three masses arbitrary. Let E be the initial energies of each of the initial particles.

VII. Quantum Electrodynamics

Quantum electrodynamics, or QED for short, is the most successfully tested scientific theory to date. The theory has very few parameters, one for the mass of any particle you introduce, and one for the charge of that particle. Other than that, the theory is completely non-arbitrary. With its rigorous demand of gauge invariance, there is no flexibility concerning the coupling.

Before we get started, a word is in order about units. Most physicists do not use units where $\hbar = c = \varepsilon_0 = 1$, and we must nonetheless communicate with people who have not seen the light. One way to facilitate communication, as well as computation when we have to turn things into experimental numbers, is to use the *fine structure constant* α, which is dimensionless (in everyone's units), and in SI units it is given by

$$\alpha = \frac{e^2}{4\pi\varepsilon_0 \hbar c} \approx \frac{1}{137.036} .$$

We will rarely need this kind of precision, so we'll round off, and in our units

$$\alpha = \frac{e^2}{4\pi} \approx \frac{1}{137} . \tag{7.1}$$

Whenever possible, we will write answers in terms of α, not *e*.

A. Polarization Vectors and Electromagnetic Waves

The first thing we need to talk about is how we will describe a photon, a single particle of light energy. We write, in general, an electromagnetic wave in the form

$$A^\mu(x) = \varepsilon^\mu e^{-iq\cdot x} , \tag{7.2}$$

where ε^μ is the *polarization vector* of the electromagnetic wave. We then find the electromagnetic fields using eq. (2.26), so we have

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = i\left(q_\nu \varepsilon_\mu - \varepsilon_\nu q_\mu\right) e^{-iq\cdot x} . \tag{7.3}$$

We want this to be a *free* electromagnetic wave, so we should have it solve eq. (2.29) with no source, so

$$0 = \partial_\mu F^{\mu\nu} = i\partial_\mu \left[\left(q^\nu \varepsilon^\mu - \varepsilon^\nu q^\mu\right) e^{-iq\cdot x}\right] = \left[q^\nu \left(\varepsilon \cdot q\right) - \varepsilon^\nu q^2\right] e^{-iq\cdot x} ,$$

$$q^\nu\left(\varepsilon \cdot q\right) = \varepsilon^\nu q^2 . \tag{7.4}$$

Eq. (7.4) seems to suggest that we want solutions with ε^ν and q^ν proportional to each other. Let's look at this possibility briefly. Suppose $\varepsilon^\nu = cq^\nu$, with *c* some constant. Then we see immediately from eq. (7.3) that we would have $F_{\mu\nu} = 0$, but this isn't a solution, it isn't anything! Indeed, we could perform a gauge transformation eq. (2.31) using $\chi(x) = ice^{-iq\cdot x}$, which would lead to $A_\mu = 0$. Clearly, this is unacceptable.

The only other way to find solutions to eq. (7.4) is if we simultaneously demand that both $q^2 = 0$ and $\varepsilon \cdot q = 0$. So we have

$$q^2 = 0, \quad \text{or} \quad \left(q^0\right)^2 = \mathbf{q}^2, \tag{7.5a}$$

$$\varepsilon \cdot q = 0 \quad \text{or} \quad \varepsilon^0 q^0 = \boldsymbol{\varepsilon}\cdot\mathbf{q}. \tag{7.5b}$$

The former equation tells us the photon is massless.

Now, for any polarization vector ε^μ, we can make a gauge transformation which will change ε^μ to $\varepsilon^\mu + ck^\mu$. We now make a specific *gauge choice* to make things as simple as possible. In *Coulomb gauge* we choose *c* so that the time component of ε^μ vanishes, $\varepsilon^0 = 0$, and therefore only the space part exists, $\varepsilon^\mu = (0, \boldsymbol{\varepsilon})$. This will ensure that $A^0 = \Phi = 0$. Combined with eq. (7.5b), we therefore have

$$\varepsilon^0 = 0, \quad \boldsymbol{\varepsilon}\cdot\mathbf{q} = 0. \tag{7.6}$$

We will need to calculate the energy of the wave given by eq. (7.2). The first step is to find the electric and magnetic fields, which are given by eqs. (2.22) and (2.23):

$$\begin{aligned}\mathbf{E} &= -\partial_0 \mathbf{A} - \nabla A^0 = iq^0 \boldsymbol{\varepsilon} e^{-iq\cdot x}, \\ \mathbf{B} &= \nabla\times\mathbf{A} = i\left(\mathbf{q}\times\boldsymbol{\varepsilon}\right)e^{-iq\cdot x}.\end{aligned} \tag{7.7}$$

Now we need to find the energy of an electromagnetic wave. For *real* waves, the energy in SI units is

$$E = \int d^3\mathbf{x}\left(\frac{1}{2}\varepsilon_0 \mathbf{E}^2 + \frac{1}{2\mu_0}\mathbf{B}^2\right).$$

But we are dealing with complex waves, and we also are working in units where $\varepsilon_0 = \mu_0 = 1$. How you deal with the complex numbers is kind of complicated, but let me give you the answer without proof:

$$E = \int d^3\mathbf{x}\left(\mathbf{E}\cdot\mathbf{E}^* + \mathbf{B}\cdot\mathbf{B}^*\right). \tag{7.8}$$

Substituting eqs. (7.7) in, and integrating over a finite volume *V* and doing a lot of fancy work with the vectors, we find

$$\begin{aligned} E &= \int d^3\mathbf{x}\left[\left(q^0\boldsymbol{\varepsilon}\right)\cdot\left(q^0\boldsymbol{\varepsilon}^*\right) + \left(\mathbf{q}\times\boldsymbol{\varepsilon}\right)\cdot\left(\mathbf{q}\times\boldsymbol{\varepsilon}^*\right)\right] = V\left[\left(q^0\right)^2 \boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon}^* - \mathbf{q}\cdot\left(\boldsymbol{\varepsilon}\times\left(\mathbf{q}\times\boldsymbol{\varepsilon}^*\right)\right)\right] \\ &= V\left[\left(q^0\right)^2 \boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon}^* + \mathbf{q}\cdot\left(\mathbf{q}\left(\boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon}^*\right) - \boldsymbol{\varepsilon}^*\left(\mathbf{q}\cdot\boldsymbol{\varepsilon}\right)\right)\right] = V\left[\left(q^0\right)^2 + \mathbf{q}^2\right]\left(\boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon}^*\right), \\ E &= 2V\left(q^0\right)^2\left(\boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon}^*\right). \end{aligned} \tag{7.9}$$

In the course of deriving this formula, we used eqs. (7.5a) and (7.6), as well as some nifty vector identities:

$$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) \quad \text{and} \quad \mathbf{a}\times(\mathbf{b}\times\mathbf{c})=\mathbf{b}(\mathbf{a}\cdot\mathbf{c})-\mathbf{c}(\mathbf{a}\cdot\mathbf{b}).$$

We'd now like to use (7.7) to get the normalization right. A state vector $|\gamma,\mathbf{q},\sigma\rangle$, where γ represents the photon, **q** is the three-momentum, and σ is the polarization, would be properly normalized so that $\langle\gamma,\mathbf{q},\sigma|\gamma,\mathbf{q},\sigma\rangle=1$ and would have energy

$$\langle\gamma,\mathbf{q},\sigma|H_0|\gamma,\mathbf{q},\sigma\rangle=q^0 .$$

But we want relativistically normalized states, which have $\langle\gamma,q,\sigma|\gamma,q,\sigma\rangle=2q^0V$, so that $|\gamma,q,\sigma\rangle=\sqrt{2q^0V}\,|\gamma,\mathbf{q},\sigma\rangle$, and this implies

$$\langle\gamma,q,\sigma|H_0|\gamma,q,\sigma\rangle=2\left(q^0\right)^2V . \tag{7.10}$$

Comparison with (7.9) tells us we want

$$\boldsymbol{\varepsilon}\cdot\boldsymbol{\varepsilon}^*=1. \tag{7.11}$$

This will be our normalization condition.

The Coulomb gauge choice and transverse condition eqs. (7.6) leave two linearly independent directions that an electromagnetic wave can be polarized. We name these the polarizations, which we denote by $\sigma=1,2$, defined in such a way that

$$\varepsilon(q,\sigma)\cdot\varepsilon^*(q,\tau)=-\delta_{\sigma\tau} , \tag{7.12a}$$

$$\varepsilon(q,\sigma)\cdot q=0 , \tag{7.12b}$$

together with the choice $\varepsilon^0=0$. The minus sign in (7.12a) occurs because of the minus sign in the space part of the metric. The two equations (7.12) are in fact independent of gauge. Suppose, for example, we had **q** in the x^3-direction, then in Coulomb gauge we could pick

$$\varepsilon(q,1)=(0,1,0,0) \quad \text{and} \quad \varepsilon(q,2)=(0,0,1,0).$$

Although we will normally work with only real polarizations, it should be understood that it is sometimes convenient to choose complex ones. But we may sometimes get lazy and ignore the complex conjugates, meaning we are implicitly assuming we are using real polarizations.

B. Feynman Rules for Quantum Electrodynamics

We are now ready to see if we can't figure out the Feynman rules for QED. For the moment, we will include only fermions and photons, though we will allow there to be several different types of fermions. We will continue to indicate fermions by arrows going to the right and anti-fermions by arrows going to the left. We will denote photons by squiggly lines, reminiscent of electromagnetic waves, as sketching in Fig. 7-1. If we have electrons, for example, we might be interested in matrix elements of the form

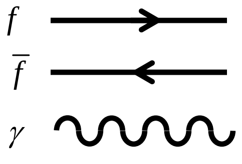


Figure 7-1: Symbols used in QED.

$$\left\langle e^-,p',s'\middle|\mathcal{H}\middle|e^-,p,s;\gamma,q,\sigma\right\rangle .$$

This matrix element will undoubtedly involve a factor of $\bar{u}(p',s')$, a $u(p,s)$ and the polarization vector $\varepsilon^\mu(q,\sigma)$. If you look at the derivation of eq. (3.64), you can see that there isn't really much choice about the remainder, and we see that we have

$$\left\langle e^-,p',s'\middle|\mathcal{H}\middle|e^-,p,s;\gamma,q,\sigma\right\rangle = -e\bar{u}(p',s')\gamma^\mu u(p,s)\varepsilon_\mu(q,\sigma) .$$

The factor of $-e$ comes from the electron's charge. For an arbitrary fermion *f* with charge *Qe*, this would become

$$\left\langle f,p',s'\middle|\mathcal{H}\middle|f,p,s;\gamma,q,\sigma\right\rangle = Qe\bar{u}(p',s')\gamma^\mu u(p,s)\varepsilon_\mu(q,\sigma) . \qquad (7.13)$$

The Feynman rule for the coupling won't be this complicated, though, because as before, we will associate the Dirac spinors $u(p,s)$ and $\bar{u}(p',s')$ with the fermion lines themselves, and we will also associated the $\varepsilon_\mu(q,\sigma)$ with the photon line. Multiplying by the usual factor of $-i$, the Feynman rule for our vertex will just be a factor of $-iQe\gamma^\mu$.

It is easy to see that, because photon interactions have an associated factor of $\varepsilon_\mu(q,\sigma)$ when on the right, as in eq. (7.13), we will get a factor of $\varepsilon_\mu(q,\sigma)$ for an initial photon and $\varepsilon_\mu^*(q,\sigma)$ for final photons. Harder to figure out is what will happen to the propagator. Not surprisingly (keeping in mind the photon is massless), we get a factor of i/q^2, but we should additionally get some factors coming from the polarizations, much as we did for the fermion propagator in the previous chapter. We therefore would expect our photon propagator to look something like

$$\frac{i}{q^2}\sum_\sigma \varepsilon_\mu^*(\mathbf{q},\sigma)\varepsilon_\nu(\mathbf{q},\sigma) . \qquad (7.14)$$

Unfortunately, the polarization sum in eq. (7.14) will depend on the choice of gauge. Indeed if we use Coulomb gauge, the expression won't even be Lorentz invariant. You can see this immediately in that if you pick $\mu=0$ or $\nu=0$, the polarization vector component automatically vanishes. This suggests that (7.14) will lead to horrible non-Lorentz invariant and non-gauge invariant Feynman rules. This will be a mess to deal with.

Fortunately, thanks to a proof by Feynman, there is a clever and complicated way to average over many gauges and develop a propagator that *is* Lorentz invariant. Doing so reduces the sum to the simple expression

$$\sum_\sigma \varepsilon_\mu^*(\mathbf{q},\sigma)\varepsilon_\nu(\mathbf{q},\sigma) = -g_{\mu\nu} . \qquad (7.15)$$

Because of the simplicity of eq. (7.15), this choice of *Feynman gauge* vastly simplifies our computations. It will lead to a propagator for the photon of $-ig_{\mu\nu}/q^2$. Eq. (7.15) will also be handy when we sum or average over polarizations, much as we summed or averaged over spins when discussing fermions in the previous chapter. Other rules for QED are pretty much identical with those in the previous chapter, and will not be rederived.

We are now ready to state our Feynman rules for QED. Rather than actually listing them, I will present them only diagrammatically, as illustrated in Fig. 7-2. In addition to the rules illustrated in Fig. 7-2, there is also still the rule about adding diagrams that differ by swapping an external boson (photon), or subtracting them if they differ by swapping an external fermion. It is also worth mentioning that as our theories get more realistic, it will sometimes be necessary to label internal lines with the type of particle, since there may be multiple potential charged fermions in a diagram. We have labeled the ends of the photon propagators and coupling with A_μ or A_ν, to remind us it is associated with the vector potential *A*, and more importantly, because there will be a Lorentz index associated with these interactions.

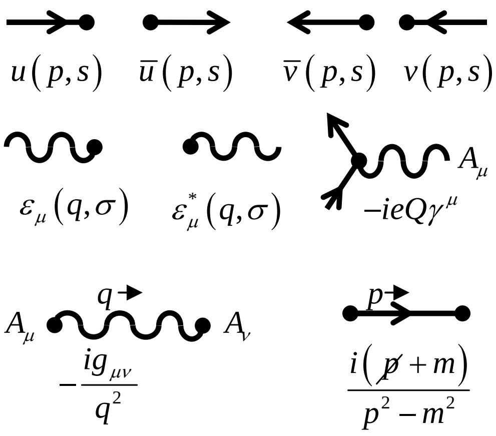


Figure 7-2: The Feynman rules for QED with fermions.

Of course, the real world contains more charged particles than just electrons, but fortunately, the electron is so light that you can perform experiments that are way too low to produce any other electromagnetically interacting particles. Nonetheless, more massive particles can appear in loop diagrams, and some experiments even at low energy are so sensitive they can detect the presence of these much more massive particles, even though they are present only as virtual particles. When you include the effects of these massive particles, the predictions of QED have in every case been consistently confirmed.

We will find that we often encounter expressions where a series of Dirac matrices are sandwiched between Dirac matrices with their common index contracted. The following identities, proven in homework problem 7.2, will prove useful:

$$\gamma^\mu\gamma_\mu = 4\,, \tag{7.16a}$$

$$\gamma^\mu\gamma^\alpha\gamma_\mu = -2\gamma^\alpha\,, \tag{7.16b}$$

$$\gamma^\mu\gamma^\alpha\gamma^\beta\gamma_\mu = 4g^{\alpha\beta}\,, \tag{7.16c}$$

$$\gamma^\mu\gamma^\alpha\gamma^\beta\gamma^\nu\gamma_\mu = -2\gamma^\nu\gamma^\beta\gamma^\alpha\,. \tag{7.16d}$$

C. Pair Production in QED

Let's start with a straightforward computation for the process $e^+e^- \to \bar{f}\,f$, where *f* is any fermion distinct from the electron. The electron has charge $-e$, and the fermion has charge Qe. There is one Feynman diagram, sketched in Fig. 7-3.

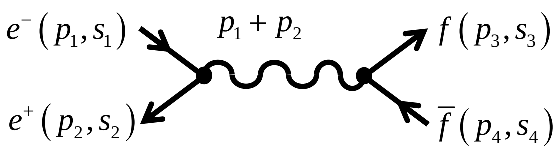


Figure 7-3: The Feynman diagram for pair creation in QED.

This diagram has two fermions lines that need to be traced. The one with the electron will yield a factor of $ie\left(\bar{v}_2\gamma^\mu u_1\right)$. The fermion *f* will contribute $-ieQ\left(\bar{u}_3\gamma^\nu v_4\right)$. There is a photon propagator connecting them, so we have

$$i\mathcal{M} = e^2Q\left(\bar{v}_2\gamma^\mu u_1\right)\left(\bar{u}_3\gamma^\nu v_4\right)\frac{-ig_{\mu\nu}}{\left(p_1+p_2\right)^2} = -\frac{ie^2Q}{s}\left(\bar{v}_2\gamma^\mu u_1\right)\left(\bar{u}_3\gamma_\mu v_4\right). \tag{7.17}$$

The complex conjugate of this expression is

$$\left(i\mathcal{M}\right)^* = \frac{ie^2Q}{s}\left(\bar{u}_1\gamma^\nu v_2\right)\left(\bar{v}_4\gamma_\nu u_1\right). \tag{7.18}$$

Although there was no real reason to change μ's to ν's between (7.17) and (7.18), we are about to multiply these expressions, in which case we have to be careful that we use different indices to keep track of the double sum. Now we multiply them together to give

$$\left|i\mathcal{M}\right|^2 = -\frac{e^4Q^2}{s^2}\left(\bar{v}_2\gamma^\mu u_1\right)\left(\bar{u}_1\gamma^\nu v_2\right)\left(\bar{u}_3\gamma_\mu v_4\right)\left(\bar{v}_4\gamma_\nu u_1\right).$$

As usual, we would like to sum over the final state spins and average over the initial state spins. This gives us a factor of $\frac{1}{4}$, and we use eqs. (3.50) to do the sums, turning sums of Dirac spinors into traces. We will immediately make an approximation, however. Because the electron is 200 times lighter than any other charged fermion, the electron mass is negligible, and we simply treat it as zero. We therefore have

$$\tfrac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 = \frac{e^4Q^2}{4s^2}\operatorname{Tr}\left(\not{p}_2\gamma^\mu\not{p}_1\gamma^\nu\right)\operatorname{Tr}\left[\left(\not{p}_3+m\right)\gamma_\mu\left(\not{p}_4-m\right)\gamma_\nu\right], \tag{7.19}$$

where *m* is the mass of the fermion. In homework problem 7.6 you will complete the computation for finite mass, but for these notes we will *also* assume the fermion *f* is very light compared to the energy, and therefore we have

$$\begin{aligned}
\tfrac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 &= \frac{e^4Q^2}{4s^2}\operatorname{Tr}\left(\not{p}_2\gamma^\mu\not{p}_1\gamma^\nu\right)\operatorname{Tr}\left(\not{p}_3\gamma_\mu\not{p}_4\gamma_\nu\right)\\
&= \frac{4e^4Q^2}{s^2}\left(p_2^\mu p_1^\nu + p_1^\mu p_2^\nu - p_1\cdot p_2 g^{\mu\nu}\right)\left(p_{3\mu}p_{4\nu} + p_{4\mu}p_{3\nu} - p_3\cdot p_4 g_{\mu\nu}\right)\\
&= \frac{4e^4Q^2}{s^2}\begin{bmatrix}2\left(p_2\cdot p_3\right)\left(p_1\cdot p_4\right)+2\left(p_1\cdot p_3\right)\left(p_2\cdot p_4\right)\\ +\left(p_1\cdot p_2\right)\left(p_3\cdot p_4\right)\left(-1-1-1+g^{\mu\nu}g_{\mu\nu}\right)\end{bmatrix},
\end{aligned}$$

$$\tfrac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 = \frac{8e^4Q^2}{s^2}\left[\left(p_2\cdot p_3\right)\left(p_1\cdot p_4\right)+\left(p_1\cdot p_3\right)\left(p_2\cdot p_4\right)\right]. \tag{7.20}$$

We used homework problem 2.3 to simplify our expression.

We now want to write out an explicit form for (7.20). We will assume the electron and positron each have energy *E* in the center of mass frame, and they are coming in along the x^3-axis. This will be the momentum as well. The final state particles also have this energy and momentum, since we are also treating them as massless, but their direction will be arbitrary but back to back. We therefore have

$$p_1^\mu = E\left(1,0,0,1\right), \quad p_3^\mu = E\left(1,\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta\right),$$
$$p_2^\mu = E\left(1,0,0,-1\right), \quad p_4^\mu = E\left(1,-\sin\theta\cos\phi,-\sin\theta\sin\phi,-\cos\theta\right).$$

So we have

$$p_2\cdot p_3 = p_1\cdot p_4 = E^2\left(1+\cos\theta\right), \quad p_1\cdot p_3 = p_2\cdot p_4 = E^2\left(1-\cos\theta\right), \quad s=4E^2 .$$

Substituting these into (7.20), we have

$$\tfrac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 = \frac{e^4Q^2}{2E^4}\left[E^4\left(1+\cos\theta\right)^2 + E^4\left(1-\cos\theta\right)^2\right] = e^4Q^2\left(1+\cos^2\theta\right).$$

Using our usual formulas eqs. (4.35) and (4.39), we find

$$\sigma = \frac{D}{4E^2\left|\Delta\mathbf{v}\right|} = \frac{1}{8E^2}\frac{E}{16\pi^2\left(2E\right)}\int d\Omega\left|i\mathcal{M}\right|^2 = \frac{e^4Q^2}{256\pi^2E^2}\int d\Omega\left(1+\cos^2\theta\right),$$
$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2Q^2}{16E^2}\left(1+\cos^2\theta\right), \quad \sigma = \frac{\pi\alpha^2Q^2}{3E^2} = \frac{4\pi\alpha^2Q^2}{3s}. \tag{7.21}$$

We have written our final answer in terms of the fine structure constant $\alpha = e^2/4\pi$.

Of course, eq. (7.21) is, strictly speaking, only valid if $E \gg m$. If $E < M$, the cross-section is zero, since it is impossible to create the *f* fermion. It can be shown that as *E* exceeds *m*, the cross section rises rapidly, quickly approaching the value (7.21), as you will demonstrate in problem 7.6.

This computation played an important role in the development of the standard model. The simple computation of the cross-section (7.21) assured that no matter what the particles did after they were created, one could measure the total production of a certain class of particles and determine the sum of Q^2. When quarks were first proposed with fractional charge, there was understandably a great deal of doubt about the claim, but when it was found that (7.21) fit the observed cross-section fairly well, with $\sum Q^2$ working out to a non-integer, many physicists began to take the idea of quarks seriously (see problem 7.5).

We explicitly stated early on that the fermion *f* must be something other than an electron. What if we want to do electron/positron scattering, $e^+e^- \to e^+e^-$? Then there is a second diagram, as illustrated in Fig. 7-4. You can obtain Fig. 7-4 from Fig. 7-3 by switching the two arrows coming out of the diagram ($e^-\left(p_3,s_3\right)$ and $e^+\left(p_2,s_2\right)$), and since they differ by switching a fermion line, there is a relative minus sign between them. The total Feynman amplitude is then

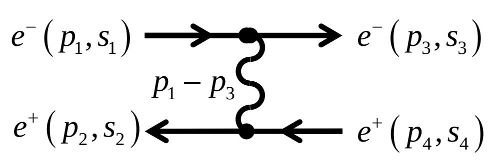


Figure 7-4: The second Feynman diagram contributing to $e^+e^- \to e^+e^-$

$$i\mathcal{M} = ie^2\left[\frac{\left(\bar{v}_2\gamma^\mu u_1\right)\left(\bar{u}_3\gamma_\mu v_4\right)}{\left(p_1+p_2\right)^2} - \frac{\left(\bar{u}_3\gamma^\mu u_1\right)\left(\bar{v}_2\gamma_\mu v_4\right)}{\left(p_1-p_3\right)^2}\right],$$

$$\left(i\mathcal{M}\right)^* = -ie^2\left[\frac{\left(\bar{u}_1\gamma^\nu v_2\right)\left(\bar{v}_4\gamma_\nu u_3\right)}{\left(p_1+p_2\right)^2} - \frac{\left(\bar{u}_1\gamma^\nu u_3\right)\left(\bar{v}_4\gamma_\nu v_2\right)}{\left(p_1-p_3\right)^2}\right],$$

$$\left|i\mathcal{M}\right|^2 = \frac{e^4}{4}\left\{\begin{array}{l} \dfrac{\left(\bar{v}_2\gamma^\mu u_1\right)\left(\bar{u}_1\gamma^\nu v_2\right)\left(\bar{u}_3\gamma_\mu v_4\right)\left(\bar{v}_4\gamma_\nu u_3\right)}{\left(p_1\cdot p_2\right)^2} + \dfrac{\left(\bar{u}_3\gamma^\mu u_1\right)\left(\bar{u}_1\gamma^\nu u_3\right)\left(\bar{v}_2\gamma_\mu v_4\right)\left(\bar{v}_4\gamma_\nu v_2\right)}{\left(p_1\cdot p_3\right)^2} \\ + \dfrac{\left(\bar{v}_2\gamma^\mu u_1\right)\left(\bar{u}_1\gamma^\nu u_3\right)\left(\bar{u}_3\gamma_\mu v_4\right)\left(\bar{v}_4\gamma_\nu u_3\right)}{\left(p_1\cdot p_2\right)\left(p_1\cdot p_3\right)} + \dfrac{\left(\bar{u}_3\gamma^\mu u_1\right)\left(\bar{u}_1\gamma^\nu v_2\right)\left(\bar{v}_2\gamma_\mu v_4\right)\left(\bar{v}_4\gamma_\nu u_3\right)}{\left(p_1\cdot p_2\right)\left(p_1\cdot p_3\right)} \end{array}\right\}. \quad (7.22)$$

We now sum and average over spins to yield the expression

$$\frac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 = \frac{e^4}{16}\left\{\begin{array}{l} \dfrac{\operatorname{Tr}\left(\not{p}_2\gamma^\mu \not{p}_1\gamma^\nu\right)\operatorname{Tr}\left(\not{p}_3\gamma_\mu \not{p}_4\gamma_\nu\right)}{\left(p_1\cdot p_2\right)^2} + \dfrac{\operatorname{Tr}\left(\not{p}_3\gamma^\mu \not{p}_1\gamma^\nu\right)\operatorname{Tr}\left(\not{p}_2\gamma_\mu \not{p}_4\gamma_\nu\right)}{\left(p_1\cdot p_3\right)^2} \\ + \dfrac{\operatorname{Tr}\left(\not{p}_2\gamma^\mu \not{p}_1\gamma^\nu \not{p}_3\gamma_\mu \not{p}_4\gamma_\nu\right)}{\left(p_1\cdot p_2\right)\left(p_1\cdot p_3\right)} + \dfrac{\operatorname{Tr}\left(\not{p}_3\gamma^\mu \not{p}_1\gamma^\nu \not{p}_2\gamma_\mu \not{p}_4\gamma_\nu\right)}{\left(p_1\cdot p_2\right)\left(p_1\cdot p_3\right)} \end{array}\right\}. \quad (7.23)$$

We have some nasty traces to deal with. The first two terms are very similar to what we dealt with before. We find

$$\operatorname{Tr}\left(\not{p}_2\gamma^\mu \not{p}_1\gamma^\nu\right)\operatorname{Tr}\left(\not{p}_3\gamma_\mu \not{p}_4\gamma_\nu\right) = 32\left(p_1\cdot p_4\right)\left(p_2\cdot p_3\right) + 32\left(p_1\cdot p_3\right)\left(p_2\cdot p_4\right), \quad (7.24a)$$

$$\operatorname{Tr}\left(\not{p}_3\gamma^\mu \not{p}_1\gamma^\nu\right)\operatorname{Tr}\left(\not{p}_2\gamma_\mu \not{p}_4\gamma_\nu\right) = 32\left(p_1\cdot p_4\right)\left(p_2\cdot p_3\right) + 32\left(p_1\cdot p_2\right)\left(p_3\cdot p_4\right). \quad (7.24b)$$

The remaining traces can be dealt with using eqs. (7.16c) and (7.16d):

$$\begin{aligned} \operatorname{Tr}\left(\not{p}_2\gamma^\mu \not{p}_1\gamma^\nu \not{p}_3\gamma_\mu \not{p}_4\gamma_\nu\right) &= -2\operatorname{Tr}\left(\not{p}_2\not{p}_3\gamma^\nu \not{p}_1\not{p}_4\gamma_\nu\right) = -8\left(p_1\cdot p_4\right)\operatorname{Tr}\left(\not{p}_2\not{p}_3\right), \\ &= -32\left(p_1\cdot p_4\right)\left(p_2\cdot p_3\right), \end{aligned} \quad (7.25a)$$

$$\begin{aligned} \operatorname{Tr}\left(\not{p}_3\gamma^\mu \not{p}_1\gamma^\nu \not{p}_2\gamma_\mu \not{p}_4\gamma_\nu\right) &= -2\operatorname{Tr}\left(\not{p}_3\not{p}_2\gamma^\nu \not{p}_1\not{p}_4\gamma_\nu\right) = -8\left(p_1\cdot p_4\right)\operatorname{Tr}\left(\not{p}_3\not{p}_2\right), \\ &= -32\left(p_1\cdot p_4\right)\left(p_2\cdot p_3\right). \end{aligned} \quad (7.25b)$$

Substituting eqs. (7.24) and (7.25) into (7.23), we have

$$\frac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 = 2e^4\left\{\begin{array}{c} \dfrac{\left(p_1\cdot p_4\right)\left(p_2\cdot p_3\right)+\left(p_1\cdot p_3\right)\left(p_2\cdot p_4\right)}{\left(p_1\cdot p_2\right)^2} \\ + \dfrac{\left(p_1\cdot p_4\right)\left(p_2\cdot p_3\right)+\left(p_1\cdot p_2\right)\left(p_3\cdot p_4\right)}{\left(p_1\cdot p_3\right)^2} - \dfrac{2\left(p_1\cdot p_4\right)\left(p_2\cdot p_3\right)}{\left(p_1\cdot p_2\right)\left(p_1\cdot p_3\right)} \end{array}\right\},$$

$$\frac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 = 2e^4\left\{\begin{array}{l} \dfrac{E^4\left(1+\cos\theta\right)^2 + E^4\left(1-\cos\theta\right)^2}{4E^4} \\ + \dfrac{E^4\left(1+\cos\theta\right)^2 + 4E^4}{E^4\left(1-\cos\theta\right)^2} - \dfrac{2E^4\left(1+\cos\theta\right)^2}{2E^4\left(1-\cos\theta\right)} \end{array}\right\} = e^4\left(\frac{3+\cos^2\theta}{1-\cos\theta}\right)^2.$$

In a straightforward manner, this allows us to find the differential cross-section,

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2 Q^2}{16E^2}\left(\frac{3+\cos^2\theta}{1-\cos\theta}\right)^2 .$$

If we attempt to integrate this expression, however, we will get infinity, because of the vanishing denominator as $\cos\theta \to 1$. This is clearly coming from the contribution from Fig. 7-4. Reintroducing the electron mass does nothing to eliminate the infinity; it is present because of the long range nature of the electromagnetic force. The particles will always deflect by at least a small angle, even if they miss by a wide margin. Experimentally, the cross-section is effectively finite, because sufficiently small angles of deflection are not detectable, and are indistinguishable from the particles missing each other. Effectively, there is a cutoff in the $\cos\theta$ integration.

D. Electron-Positron Annihilation: Summing over Polarizations

Another process that can occur is that an electron and positron annihilate to make a pair of photons. The two diagrams that are relevant are sketched in Fig. 7-5. It should be obvious that these diagrams differ only by the switching of the photon lines, and since photons are bosons, they get a relative plus sign.

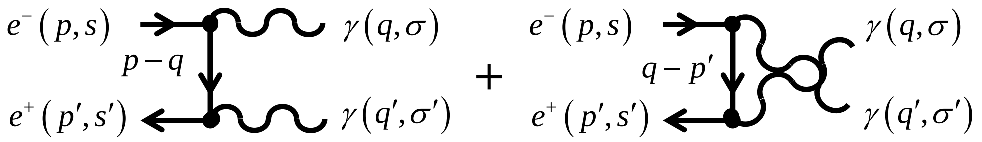


Figure 7-5: The Feynman diagram for pair e$^+$e$^-$ annihilation in QED.

The resulting Feynman amplitude will be

$$i\mathcal{M} = (ie)^2 \varepsilon_\mu^* \varepsilon_\nu'^* \left[\bar{v}'\gamma^\nu \frac{i(\not{p}-\not{q}+m)}{(p-q)^2-m^2}\gamma^\mu u + \bar{v}'\gamma^\mu \frac{i(\not{q}-\not{p}'+m)}{(q-p')^2-m^2}\gamma^\nu u \right]. \tag{7.26}$$

Let us ignore the electron mass here and throughout, which simplifies this to

$$i\mathcal{M} = ie^2 \varepsilon_\mu^* \varepsilon_\nu'^* \bar{v}' \left[\gamma^\nu \frac{\not{p}-\not{q}}{2p\cdot q}\gamma^\mu + \gamma^\mu \frac{\not{q}-\not{p}'}{2p'\cdot q}\gamma^\nu \right] u . \tag{7.27}$$

The complex conjugate of (7.27) is

$$(i\mathcal{M})^* = -ie^2 \varepsilon_\alpha \varepsilon_\beta' \bar{u} \left[\gamma^\alpha \frac{\not{p}-\not{q}}{2p\cdot q}\gamma^\beta + \gamma^\beta \frac{\not{q}-\not{p}'}{2p'\cdot q}\gamma^\alpha \right] v'. \tag{7.28}$$

Multiplying eqs. (7.27) and (7.28), we have

$$|i\mathcal{M}|^2 = e^4 \varepsilon_\mu^* \varepsilon_\alpha \varepsilon_\nu'^* \varepsilon_\beta' \bar{v}' \left[\gamma^\nu \frac{\not{p}-\not{q}}{2p\cdot q}\gamma^\mu + \gamma^\mu \frac{\not{q}-\not{p}'}{2p'\cdot q}\gamma^\nu \right] u\bar{u} \left[\gamma^\alpha \frac{\not{p}-\not{q}}{2p\cdot q}\gamma^\beta + \gamma^\beta \frac{\not{q}-\not{p}'}{2p'\cdot q}\gamma^\alpha \right] v' .$$

(7.29)

Let's assume we are doing an unpolarized cross-section, so that the spins of the initial particles are uncontrolled, and the polarizations of the final photons are unmeasured. In this case, you average over initial spins, but you also sum over final polarizations. We can use eq. (7.15) to then simplify the resulting expression. We end up having

$$\tfrac{1}{4}\sum_{\substack{\text{spins,}\\ \text{pols}}} \left|i\mathcal{M}\right|^2 = \frac{e^4}{16}\mathrm{Tr}\left(\not{p}'\left[\gamma^\nu \frac{\not{p}-\not{q}}{p\cdot q}\gamma^\mu + \gamma^\mu \frac{\not{q}-\not{p}'}{p'\cdot q}\gamma^\nu\right]\not{p}\left[\gamma_\mu \frac{\not{p}-\not{q}}{p\cdot q}\gamma_\nu + \gamma_\nu \frac{\not{q}-\not{p}'}{p'\cdot q}\gamma_\mu\right]\right). \tag{7.30}$$

We now need to expand this out and work out all the traces. We can save a lot of work by realizing that, for example, $\not{p}^2 = p^2 = m^2 = 0$. Then we find, for example,

$$\begin{aligned}\mathrm{Tr}\left[\not{p}'\gamma^\nu\left(\not{p}-\not{q}\right)\gamma^\mu\not{p}\gamma_\mu\left(\not{p}-\not{q}\right)\gamma_\nu\right] &= -2\mathrm{Tr}\left[\not{p}'\gamma^\nu\left(\not{p}-\not{q}\right)\not{p}\left(\not{p}-\not{q}\right)\gamma_\nu\right]\\ &= -2\mathrm{Tr}\left(\not{p}'\gamma^\nu\not{q}\not{p}\not{q}\gamma_\nu\right) = 4\mathrm{Tr}\left(\not{p}'\not{q}\not{p}\not{q}\right) = 32\left(p'\cdot q\right)\left(p\cdot q\right),\end{aligned} \tag{7.31a}$$

$$\begin{aligned}\mathrm{Tr}\left[\not{p}'\gamma^\nu\left(\not{p}-\not{q}\right)\gamma^\mu\not{p}\gamma_\nu\left(\not{q}-\not{p}'\right)\gamma_\mu\right] &= -2\mathrm{Tr}\left[\not{p}'\not{p}\gamma^\mu\left(\not{p}-\not{q}\right)\left(\not{q}-\not{p}'\right)\gamma_\mu\right]\\ &= -8\left(p-q\right)\cdot\left(q-p'\right)\mathrm{Tr}\left(\not{p}'\not{p}\right) = 32\left(p\cdot p'-q\cdot p-q\cdot p'\right)\left(p\cdot p'\right),\end{aligned} \tag{7.31b}$$

$$\begin{aligned}\mathrm{Tr}\left[\not{p}'\gamma^\mu\left(\not{q}-\not{p}'\right)\gamma^\nu\not{p}\gamma_\mu\left(\not{p}-\not{q}\right)\gamma_\nu\right] &= -2\mathrm{Tr}\left[\not{p}'\not{p}\gamma^\nu\left(\not{q}-\not{p}'\right)\left(\not{p}-\not{q}\right)\gamma_\nu\right]\\ &= -8\left(q-p'\right)\cdot\left(p-q\right)\mathrm{Tr}\left(\not{p}'\not{p}\right) = 32\left(p\cdot p'-q\cdot p-q\cdot p'\right)\left(p\cdot p'\right),\end{aligned} \tag{7.31c}$$

$$\begin{aligned}\mathrm{Tr}\left[\not{p}'\gamma^\mu\left(\not{q}-\not{p}'\right)\gamma^\nu\not{p}\gamma_\nu\left(\not{q}-\not{p}'\right)\gamma_\mu\right] &= -2\mathrm{Tr}\left[\not{p}'\gamma^\mu\left(\not{q}-\not{p}'\right)\not{p}\left(\not{q}-\not{p}'\right)\gamma_\mu\right]\\ &= 4\mathrm{Tr}\left[\not{p}'\left(\not{q}-\not{p}'\right)\not{p}\left(\not{q}-\not{p}'\right)\right] = 4\mathrm{Tr}\left(\not{p}'\not{q}\not{p}\not{q}\right) = 32\left(p'\cdot q\right)\left(p\cdot q\right).\end{aligned} \tag{7.31d}$$

Eqs. (7.31b) and (7.31c) can be further simplified by noting that

$$2\left(p\cdot p'-q\cdot p-q\cdot p'\right) = \left(p+p'-q\right)^2 = q'^2 = 0.$$

so they both vanish. Substituting eqs. (7.31a) and (7.31d) into (7.30) then yields

$$\tfrac{1}{4}\sum_{\substack{\text{spins,}\\ \text{pols}}} \left|i\mathcal{M}\right|^2 = 2e^4\left[\frac{\left(p'\cdot q\right)\left(p\cdot q\right)}{\left(p\cdot q\right)^2} + \frac{\left(p'\cdot q\right)\left(p\cdot q\right)}{\left(p\cdot q'\right)^2}\right]. \tag{7.32}$$

We are ready to start working out all the various dot products. Since we are treating all particles as massless, we can write our various momenta immediately:

$$\begin{aligned}p^\mu &= E\left(1,0,0,1\right), \quad q^\mu = E\left(1,\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta\right),\\ p'^\mu &= E\left(1,0,0,-1\right), \quad q'^\mu = E\left(1,-\sin\theta\cos\phi,-\sin\theta\sin\phi,-\cos\theta\right).\end{aligned}$$

This allows us to get our dot products, which are

$$p\cdot q = E^2\left(1-\cos\theta\right), \quad p\cdot q' = p'\cdot q = E^2\left(1+\cos\theta\right). \tag{7.33}$$

Substituting (7.33) into (7.32), we have

$$\tfrac{1}{4}\sum_{\substack{\text{spins,}\\ \text{pols}}} \left|i\mathcal{M}\right|^2 = 2e^4\left(\frac{1+\cos\theta}{1-\cos\theta}+\frac{1-\cos\theta}{1+\cos\theta}\right) = 4e^4\,\frac{1+\cos^2\theta}{1-\cos^2\theta}\,.$$

In the usual way, we then work towards the cross-section,

$$\sigma = \frac{D}{4E^2\left|\Delta\mathbf{v}\right|} = \frac{1}{8E^2}\frac{E}{16\pi^2\left(2E\right)}\int d\Omega\left|i\mathcal{M}\right|^2 = \frac{e^4}{64\pi^2E^2}\int\frac{1+\cos^2\theta}{1-\cos^2\theta}d\Omega,$$

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E^2}\frac{1+\cos^2\theta}{1-\cos^2\theta}\,. \tag{7.34}$$

There are two difficulties with completing this problem, however. The first is that we must recognize the presence of identical particles in the final state. This is easily handled by simply dividing by two. The other problem is that the integral of eq. (7.34) diverges at the boundaries $\cos\theta = \pm 1$. This is not a real divergence, but the result of treating the electron as massless. If you put in masses and redo the calculation, the differential cross-section works out to

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4Ep}\left[\frac{E^2+p^2\cos^2\theta}{E^2-p^2\cos^2\theta} - \frac{m^4\left(E^2+p^2\cos^2\theta\right)}{E^2\left(E^2-p^2\cos^2\theta\right)^2} + \frac{m^2\left(E^2+p^2\right)}{E^2\left(E^2-p^2\cos^2\theta\right)}\right].$$

With considerable pain, this can then be integrated over angles to yield

$$\sigma = \pi\alpha^2\left[\frac{1}{p^2}\left(1+\frac{m^2}{E^2}-\frac{m^4}{2E^4}\right)\tanh^{-1}\left(\frac{p}{E}\right) - \frac{1}{2pE}\left(1+\frac{m^2}{E^2}\right)\right].$$

This expression can be rewritten in terms of s using $\tfrac{1}{4}s = E^2 = p^2 + m^2$.

One interesting thing worth talking about is how we can tell if a Feynman amplitude is gauge invariant or not. In general, when we make a gauge choice, we are choosing the polarization vectors ε, but we can choose a different gauge, in which case ε will be modified to $\varepsilon^\mu \to \varepsilon^\mu + cq^\mu$. But the result of any computation must be independent of gauge choice. It follows that if we replace ε^μ by q^μ in any Feynman amplitude, the Feynman amplitude should vanish. We can do this to any external photon line.

Let's try it in this particular case. We will make the substitution $\varepsilon^\mu \to q^\mu$ in eq. (7.26), keeping the mass terms this time. We rewrite the propagator in the "proof" version, eq. (6.11), which yields

$$i\mathcal{M}\left(\varepsilon\to q\right) = \left(ie\right)^2\varepsilon'^*_\nu\left[\bar{v}'\gamma^\nu\frac{i}{\not{p}-\not{q}-m}\not{q}u + \bar{v}'\not{q}\frac{i}{\not{q}-\not{p}'-m}\gamma^\nu u\right]. \tag{7.35}$$

Now, we know that $\not{p}u = mu$ and $\bar{v}'\not{p}' = -\bar{v}'m$. It therefore follows that

$$\not{q}u = \left(\not{q}-\not{p}+m\right)u \quad \text{and} \quad \bar{v}'\not{q} = \bar{v}'\left(\not{q}-\not{p}'-m\right).$$

Substituting these into (7.35), we see that

$$\begin{aligned} i\mathcal{M}(\varepsilon\to q) &= (ie)^2\,\varepsilon'^{*}_{\nu}\left[\bar{v}'\gamma^{\nu}\frac{i}{\not p-\not q-m}\left(\not q-\not p+m\right)u+\bar{v}'\left(\not q-\not p'-m\right)\frac{i}{\not q-\not p'-m}\gamma^{\nu}u\right] \\ &= (ie)^2\,\varepsilon'^{*}_{\nu}i\left[-\bar{v}'\gamma^{\nu}u+\bar{v}'\gamma^{\nu}u\right]=0\,. \end{aligned}$$

Hence we can see that this amplitude is, in fact, gauge invariant. A similar proof can be applied if we replace $\varepsilon'^{\mu}\to q'^{\mu}$.

Probably the most important lesson we have learned from this section is that polarizations, like spins, must be summed over when a photon is in the final state, and averaged when in the initial state. We will encounter similar aspects concerning spins, polarizations, and colors when we study quarks and gluons in Chapter 9.

E. Loops in QED

For the most part, I will avoid discussing the effects of loops, but there are a few effects worth mentioning. As a rule of thumb, the contribution from a loop diagram in QED will be smaller than a non-loop diagram by about a factor of $\alpha/4\pi$, so usually the effects are fairly small, but in some cases they are quite measurable. You may recall from section 3F that the Dirac theory predicted that the magnetic dipole moment for the electron is given by eq. (3.62) with $g = 2$. The computation there was correct but incomplete, because there are additional loop diagrams that contribute to the electric dipole moment of the fermion in question. For example, the contribution from the loop diagram shown in Fig. 7-6 adds an additional term of α/π to g. This increases g to $g = 2.0023228$, a value accurate for the electron or muon to a few parts per million. By going to higher loops, and including the effects of intermediate particles, including strongly and weakly interacting particles, the match between theory and experiment is much better. The experimental value for the electron and muon are

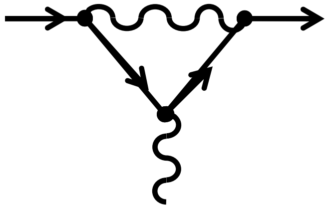


Figure 7-6: The one loop diagram modifying the interaction of a charged fermion with a background magnetic field. The incoming photon line at the bottom represents the magnetic field.

$$g_e = 2.0023193041882(15),\quad g_\mu = 2.0023318418(13),$$

with the number in parentheses representing the error in the last two digits. The theoretical value matches to better than one part per trillion for the electron, with essentially all the error attributable to our ignorance of the exact value of α. This represents the most accurate match between theory and experiment for any theory, and has gone far to convince people of the power and validity of QED. For the muon, the theoretical value is noticeably lower than the experimental, but close enough that it could be due to errors in the theoretical calculation or the experiment. Still, the numbers match to within a few parts per billion.

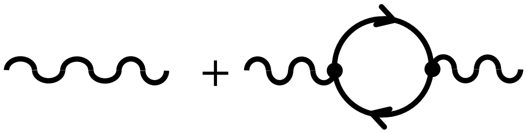


Figure 7-7: To any photon propagator (left) must be added loop diagrams, such as the one at right. These have the effect of increasing the strength of the electromagnetic interaction at high energy.

Another important effect is *running of coupling constants*. Consider any photon propagator, to which we must add diagrams with loops, as illustrated in Fig. 7-7. The net effect is that the electromagnetic coupling e effectively increases at high energy, or short distance. One way of explaining this is that there are virtual electron-positron pairs filling the universe. If you set down, say, a positive charge somewhere, any nearby electron-positron pairs will align themselves with the electron closer to the positive charge, partly screening or canceling that charge. The effect is that far away from the charge, the charge is partially screened, or reduced, whereas at close distances you more fully see the total "true" charge. Whether this qualitative description is accurate is doubtful, but it is agreed that at high energy the electromagnetic coupling acts stronger. In particular, for experiments done around the Z – mass, around 90 GeV, the fine structure constant is in the neighborhood of $\alpha \approx \frac{1}{128}$, as opposed to its low energy value closer to $\alpha \approx \frac{1}{137}$. If you continued up to still higher energies, the coupling would get stronger still, but long before that happens QED is supplanted by the electroweak theory, as described in chapter 10.

F. Other Interactions

In the previous two chapters, we started with restricted rules with only one simple vertex, and eventually generalized to the possibility of other types of interactions. Are there other interactions in QED? For a renormalizable theory of QED with fermions, there are not. As in the previous chapter, you can't have more than two fermions in a fundamental matrix element without breaking renormalizability. You also can't have interactions involving pure photons, like matrix elements of the form $\langle 0|\mathcal{H}|\gamma\gamma\gamma\rangle$ or $\langle 0|\mathcal{H}|\gamma\gamma\gamma\gamma\rangle$, because there is no gauge-invariant and Lorentz invariant way of writing such an expression. The same is true of a mass contribution to the photon, $\langle 0|\mathcal{H}|\gamma\gamma\rangle$. So in fact, the Feynman rules in Fig. 7-2 are actually complete, and require no further modification.

Although the standard model contains no charged scalar particles, such particles certainly might exist, in which case we would have to develop appropriate rules for QED with charged scalars. Let's denote our charged scalars as lines with solid arrows on them, to distinguish them from fermions, and give them a charge Q. The QED rules for charged scalars are given in Fig. 7-8. Note something new occurs – one of the rules involves the momenta of the external particles. The Feynman rules assume the momentum is flowing along the direction of the arrow. If this is not the case, we must change the sign of the corresponding momentum.

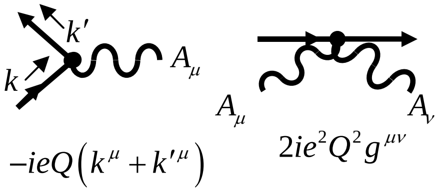


Figure 7-8: The interaction Feynman rules for a theory with a charged scalar.

Let's do a simple calculation with charged scalars. Suppose we try to pair create them from electron-positron collisions. There is one diagram, sketched in Fig. 7-9. The Feynman amplitude is

$$i\mathcal{M}=(ie)(-ieQ)\left(\bar{v}'\gamma^{\mu}u\right)\frac{-ig_{\mu\nu}\left(k^{\nu}-k'^{\nu}\right)}{\left(p+p'\right)^{2}}. \tag{7.36}$$

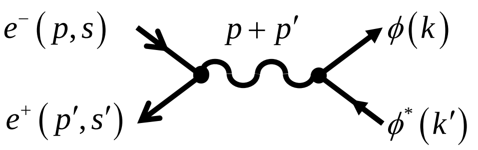


Figure 7-9: The Feynman diagram for scalar pair creation in QED.

We square and average over the incoming spins, which, ignoring the electron masses, yields

$$\frac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^{2}=\frac{e^{4}Q^{2}}{4s^{2}}\operatorname{Tr}\left[\not{p}'\left(\not{k}'-\not{k}\right)\not{p}\left(\not{k}'-\not{k}\right)\right]. \tag{7.37}$$

To simplify, use conservation of momentum to substitute $\not{k}'=\not{p}+\not{p}'-\not{k}$, and then recall that $\not{p}^{2}=p^{2}=0$ and $\not{p}'^{2}=p'^{2}=0$. So we have

$$\frac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^{2}=\frac{e^{4}Q^{2}}{4s^{2}}\operatorname{Tr}\left[\not{p}'\left(\not{p}+\not{p}'-2\not{k}\right)\not{p}\left(\not{p}+\not{p}'-2\not{k}\right)\right]=\frac{e^{4}Q^{2}}{s^{2}}\operatorname{Tr}\left(\not{p}'\not{k}\not{p}\not{k}\right),$$

$$\frac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^{2}=\frac{4e^{4}Q^{2}}{s^{2}}\left[2\left(p'\cdot k\right)\left(p\cdot k\right)-\left(p\cdot p'\right)k^{2}\right]. \tag{7.38}$$

Working in the center-of-mass frame, let E be the energy of the electron and positron, which will also be their momentum, then E will also be the energy of the charged scalars, and let k be their momentum and M their mass. Then we have

$$p^{\mu}=\left(E,0,0,E\right),\quad p'^{\mu}=\left(E,0,0,-E\right),\quad k^{\mu}=\left(E,k\sin\theta\cos\phi,k\sin\theta\sin\phi,k\cos\theta\right).$$

The dot products we need are therefore

$$p\cdot k=E\left(E-k\cos\theta\right),\quad p'\cdot k=E\left(E+k\cos\theta\right),\quad k^{2}=M^{2},\quad p\cdot p'=2E^{2}. \tag{7.39}$$

Substituting (7.39) into (7.38) we then have

$$\frac{1}{4}\sum_{\text{spins}}\left|i\mathcal{M}\right|^{2}=\frac{4e^{4}Q^{2}}{16E^{4}}\left[2E^{2}\left(E^{2}-k^{2}\cos^{2}\theta\right)-2E^{2}M^{2}\right]=\frac{e^{4}Q^{2}k^{2}}{2E^{2}}\left(1-\cos^{2}\theta\right). \tag{7.40}$$

We now find the cross-section in the usual way

$$\sigma=\frac{1}{4\left|E'\mathbf{p}-E\mathbf{p}'\right|}\frac{k}{16\pi^{2}\left(2E\right)}\int\tfrac{1}{4}\sum_{\text{spins}}\left|\mathcal{M}\right|^{2}d\Omega=\frac{e^{4}Q^{2}k^{3}}{512\pi^{2}E^{5}}\int\left(1-\cos^{2}\theta\right)d\Omega,$$

$$\frac{d\sigma}{d\Omega}=\frac{\alpha^{2}Q^{2}k^{3}}{32E^{5}}\left(1-\cos^{2}\theta\right)\quad\text{and}\quad\sigma=\frac{\pi\alpha^{2}Q^{2}k^{3}}{12E^{5}}=\frac{\pi\alpha^{2}Q^{2}}{3s}\left(1-\frac{4M^{2}}{s}\right)^{3/2}.$$

We now have a description of all electromagnetic interactions for any *renormalizable* theory of scalars or fermions. We have not yet dealt with charged vector bosons, such as the *W*-bosons, which will arise in chapter 10. But before we deal with *W*'s, it is probably a good idea to start talking about some of the other forces and interactions we encounter in nature.

<u>Problems for Chapter 7</u>

1. Consider two photon momenta with an angle θ between them. For definiteness, let $q_1^\mu = q_1(1,0,0,1)$ and $q_2^\mu = q_2(1,\sin\theta,0,\cos\theta)$. Find the two polarizations $\varepsilon(q_1,1)$ and $\boldsymbol{\varepsilon}(q_1,2)$ such that $q_1\cdot\varepsilon(q_1,\sigma)=0$ and $\varepsilon(q_1,\sigma)\cdot\varepsilon(q_1,\tau)=-\delta_{\sigma\tau}$. Repeat for q_2. Now find $\sum_{\sigma_1,\sigma_2}\left|\varepsilon(q_1,\sigma_1)\cdot\varepsilon(q_2,\sigma_2)\right|^2$.

2. Prove eqs. (7.2).

3. Draw all six Feynman diagrams for the scattering $e^+\gamma \to e^+\gamma\gamma$.

4. At tree level, photons do not scatter from each other. Draw all six one-loop diagrams whereby the process $\gamma\gamma\to\gamma\gamma$ can occur. Treat the electron as the only relevant charged particle for this purpose.

5. Calculate the total cross section ratio $\sigma\left(e^+e^-\to q\bar{q}\right)/\sigma\left(e^+e^-\to\mu^+\mu^-\right)$, summed over all quarks (multiply each contribution by 3 for colors), treating the quark as massless if $\sqrt{s}>2m$, and of course ignoring it if $\sqrt{s}<2m$. Do so in the range $2\text{ GeV}<\sqrt{s}<3\text{ GeV}$ (ignore the c quark), $5\text{ GeV}<\sqrt{s}<9\text{ GeV}$ (include the c quark) and $11\text{ GeV}<\sqrt{s}<50\text{ GeV}$ (include the b quark). Compare to the experimental values illustrated in Fig. 7-10 below. What do you think is going on at the other energies shown?

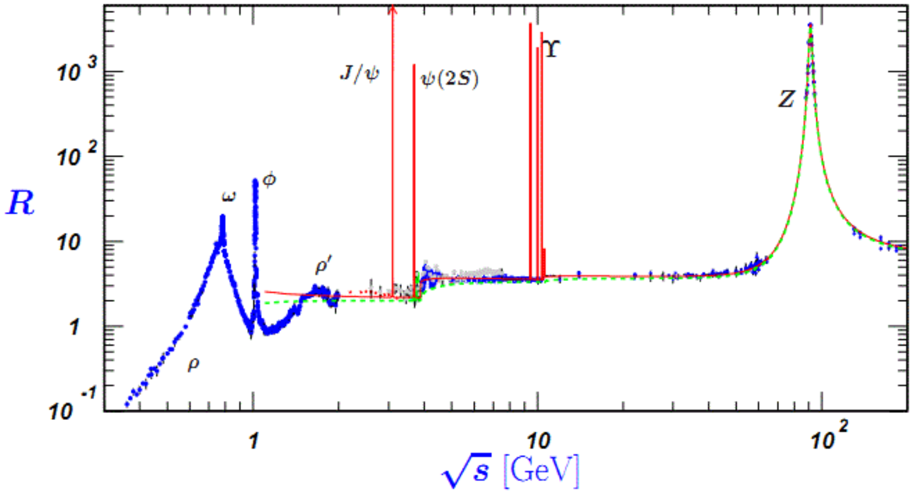


Figure 7-10: The ratio $R=\sigma\left(e^+e^-\to\text{hadrons}\right)/\sigma\left(e^+e^-\to\mu^+\mu^-\right)$. The cross-section to hadrons is believed to be caused by $\sigma\left(e^+e^-\to\text{quarks}\right)$.

6. Calculate the differential and total cross-section for $e^+e^- \to f\,\bar{f}$, treating the electron mass as zero, but not ignoring the fermion mass *m*. Sketch the result for $3\sigma E^2/\pi Q^2\alpha^2$ as a function of m/E. It should be 1 if *m* = 0.

7. Find the differential cross-section for the process $e^- f \to e^- f$, treating both the electron and *f* as massless. Note that it diverges at $\theta = 0$. Explain this qualitatively. Calculate the total cross-section, assuming there is a minimum angle $\theta_{\min}$ at which the scattering can be observed.

8. Find the differential cross-section for the process $e^-e^- \to e^-e^-$, treating the electron as massless. Note that it diverges both at $\theta = 0$ and $\theta = \pi$. Explain this qualitatively, and compare/contrast with problem 7.

9. Write the complete Feynman amplitude for the process $e^-(p)\gamma(q) \to e^-(p')\gamma(q')$. Then make the approximation that the electron is massless. Calculate the differential cross-section. The total cross-section will come out infinite, but this is an artifact of treating the electron as massless.

10. Consider again the Feynman amplitude for $e^-(p)\gamma(q) \to e^-(p')\gamma(q')$. Show that if you replace the polarization of one of the photons (either one – your choice) with its corresponding momentum, the Feynman amplitude becomes exactly zero. Do not treat the electron as massless.

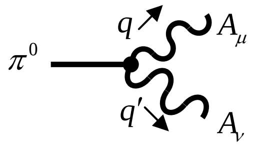
11. Consider the Feynman amplitude for $\phi^+(k)\gamma(q) \to \phi^+(k')\gamma(q')$, where ϕ^+ is a charged scalar. Show that if you replace the polarization of one of the photons (either one – your choice) with its corresponding momentum, the Feynman amplitude becomes exactly zero.

12. If the photon were massive, there would be *three* polarizations of the photon, and the sum over polarizations rule would be changed to

$$\sum_{\sigma=1}^{3} \varepsilon_\mu^*(\mathbf{q},\sigma)\,\varepsilon_\nu(\mathbf{q},\sigma) = -g_{\mu\nu} + q_\mu q_\nu / M^2 ,$$

where *M* is the photon mass. Calculate the decay rate $\gamma \to e^+e^-$, assuming no other Feynman rules change. Include the electron mass *m*.

13. The π^0 is not a fundamental particle, but made of quark/anti-quark pairs. It decays by the process $\pi^0 \to \gamma\gamma$. The coupling responsible is non-renormalizable, and is given in Fig. 7-11, where *g* is the coupling. Find the rate for this decay.



$ig\varepsilon^{\mu\alpha\nu\beta}q_\alpha q'_\beta$

Figure 7-11: The interaction responsible for π^0 decay.

VIII. The Strong Force

Quantum electrodynamics is a very successful theory, but it is clear there are other things going on. After all, the nucleus contains only positively charged protons and neutral neutrons, and yet it is tightly bound together. This suggests the presence of at least one additional force, called the *strong force*.

A. Forces and Particles

As we learned in the previous chapter, electromagnetic interactions always involve photons. True, in a process like $e^+e^- \to \mu^+\mu^-$ there are no photons in the initial nor the final state, but we understand that a virtual photon carries the energy from the initial to the final state. Charged particles feel electromagnetic forces, as well as particles that are neutral but are composed of charged particles, such as the neutron.

There are also *strong* interactions that bind together the nucleons (protons and neutrons) that comprise the nucleus. Not only is this force distinct from electromagnetism, it is much more powerful, since it can easily overcome the electric repulsion of the protons. This force is believed to also be responsible for many particle decays, such as, for example, the decay of the ρ meson to pairs of pions, say $\rho^+ \to \pi^+\pi^0$. This has a width of $\Gamma = 149 \text{ MeV}$, or about 19% of its mass, which means it has a lifetime of about $\tau = 4.4\times10^{-24}$ s. By contrast, the charged kaon, which often decays to pions, $K^+ \to \pi^+\pi^0$, has a mean lifetime of about $\tau = 1.24\times10^{-8}$ s, more than fifteen orders of magnitude higher. It seemed likely that such decays are governed by some other type of interaction, which we name *weak* interactions. The Standard Model is a complete theory that incorporates these three types of interactions – strong, weak, and electromagnetic.

Not all particles feel the strong force. The electron in an atom is well-described assuming the *only* influence of the nucleus is its electromagnetic interactions. We will therefore classify all particles into *hadrons*, particles that feel the strong forces, and all other particles. The fermions that do not feel strong forces are called *leptons*, for which there are six, and the bosons that don't feel strong forces don't have a categorical name.

When the theory of strong interactions was first being worked out, the list of strongly interacting particles was relatively short, and how much was known about them was also short. Now the particle data group makes a complete list of known particles (about 90 percent of which are hadrons) and much about their interactions every two years. An incomplete list can be found on the next page. A free copy of the particle physics booklet and complete table can be obtained for free from **http://pdg.lbl.gov/**. The particles' charges are implied by their names. Their total isospin I and I_3 values and their strangeness S (see section D) are also included.

A proton in isolation is, as far as we can tell, completely stable. Decay processes like $p^+ \to e^+\pi^0$ do not violate charge, momentum, or energy conservation. There must be some other quantity that is conserved that is protecting it. This quantity is called *baryon number*. A proton and a neutron each have baryon number 1.

Strongly interacting hadrons are subdivided based on their baryon number. Particles that carry baryon number +1 are called *baryons*. Their anti-particles carry baryon number -1 and are called *anti-baryons*. There are also a group of particles that carry baryon number 0 that are called *mesons*. Baryon number, like electric charge, is conserved, but unlike electric charge, there is no force associated with it. For every baryon listed in the tables, there is a corresponding anti-baryon. Every baryon listed has baryon number $B = +1$, and every meson has baryon number $B = 0$, while the anti-baryons have $B = -1$. In the standard model, baryon number is strictly conserved, but in many extensions of it baryon number is not conserved. Experimental searches for proton decay are ongoing.

One of the first things we note about this table is that the particles are grouped into sets of one to four particles with identical spins and similar masses, but different charges. The closeness of these masses is not a coincidence, and is a result of an approximate symmetry called *isospin*, the subject of the next section.

Mesons ($m < 900$)

All masses in MeV

Name	*Mass*	*Spin*	*I*	I_3	*S*
π^+	*139*	*0*	*1*	*+1*	*0*
π^0	*135*	*0*	*1*	*0*	*0*
π^-	*140*	*0*	*1*	*-1*	*0*
K^+	*494*	*0*	*½*	*+½*	*+1*
K^0	*498*	*0*	*½*	*–½*	*+1*
$\overline{K}^0$	*498*	*0*	*½*	*+½*	*–1*
K^-	*494*	*0*	*½*	*–½*	*–1*
η^0	*548*	*0*	*0*	*0*	*0*
ρ^+	*775*	*1*	*1*	*+1*	*0*
ρ^0	*775*	*1*	*1*	*0*	*0*
ρ^-	*775*	*1*	*1*	*–1*	*0*
ω^0	*783*	*1*	*0*	*0*	*0*
K^{*+}	*896*	*1*	*½*	*+½*	*+1*
K^{*0}	*892*	*1*	*½*	*–½*	*+1*
$\overline{K}^{*0}$	*892*	*1*	*½*	*+½*	*–1*
K^{*-}	*896*	*1*	*½*	*–½*	*–1*

Baryons ($m \lesssim 1500$)

All masses in MeV

Name	Mass	Spin	*I*	I_3	*S*
p^+	938	½	½	+½	0
n^0	940	½	½	–½	0
Λ^0	1116	½	0	0	–1
Σ^+	1189	½	1	+1	–1
Σ^0	1193	½	1	0	–1
Σ^-	1197	½	1	–1	–1
Δ^{++}	1231	3/2	3/2	+3/2	0
Δ^+	1232	3/2	3/2	+1/2	0
Δ^0	1234	3/2	3/2	-1/2	0
Δ^-	1235	3/2	3/2	-3/2	0
Ξ^0	1315	½	½	+½	–2
Ξ^-	1321	½	½	–½	–2
Σ^{*+}	1383	3/2	1	+1	–1
Σ^{*0}	1384	3/2	1	0	–1
Σ^{*-}	1387	3/2	1	–1	–1
Λ^0(1405)	1406	½	0	0	–1
N^+(1440)	1440?	½	½	+½	0
N^0(1440)	1440?	½	½	–½	0
Ξ^{*0}	1532	3/2	½	+½	–2
Ξ^{*-}	1535	3/2	½	–½	–2

B. Isospin

Even before the discovery of the multiple particles appearing in the particle data book, it was noted that there were certain similarities of various isotopes. The near-matching mass of the proton and neutron, for example, was repeated with the near-matching mass of the ^{3}H and ^{3}He isotopes composed of two protons and one neutron, or vice-versa. The pattern was repeated, not only in the ground state of various isotopes, but also for various excited states. This suggested that there was some sort of symmetry relating the proton and neutron.

Let's label protons and neutrons as if they were two different variations of the same particle, so we have

$$\left|p^+\right\rangle = \left|N_1\right\rangle, \quad \left|n^0\right\rangle = \left|N_2\right\rangle . \tag{8.1}$$

We will now let $2I_1$ be the operator that turns protons to neutrons and neutrons to protons, while leaving the momentum and spin unchanged, so

$$2I_1\left|N_1, p, s\right\rangle = \left|N_2, p, s\right\rangle, \quad 2I_1\left|N_2, p, s\right\rangle = \left|N_1, p, s\right\rangle . \tag{8.2}$$

The reason it is called $2I_1$ will be apparent later. Because the momenta or spin are not changed, we will tend to ignore these and mostly keep track of only the particle label. Let's write out (8.2) and its Hermitian conjugate in this abbreviated notation:

$$2I_1\left|N_1\right\rangle = \left|N_2\right\rangle, \quad 2I_1\left|N_2\right\rangle = \left|N_1\right\rangle, \qquad \left\langle N_1\right|2I_1 = \left\langle N_2\right|, \quad \left\langle N_2\right|2I_1 = \left\langle N_1\right| . \tag{8.3}$$

Now, the fact that the proton and neutron are interchangeable must mean that this operator commutes with the Hamiltonian, so $\left[\mathcal{H}, 2I_1\right] = 0$.

I will now use this information to prove that the proton and neutron have (approximately) the same mass. Recall that the mass of a fermion can be thought of as an interaction, $m_\psi = \left\langle\psi\right|\mathcal{H}\left|\psi\right\rangle$. We therefore have

$$m_p = \left\langle N_1\right|\mathcal{H}\left|N_1\right\rangle = \left\langle N_2\right|\left(2I_1\right)\mathcal{H}\left|N_1\right\rangle = \left\langle N_2\right|\mathcal{H}\left(2I_1\right)\left|N_1\right\rangle = \left\langle N_2\right|\mathcal{H}\left|N_2\right\rangle = m_n . \tag{8.4}$$

It's useful to think of I_1 as a matrix, which in the $\left(N_1, N_2\right)$ basis would be given by

$$I_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \tfrac{1}{2}\sigma_1 . \tag{8.5}$$

But I_1 isn't the only operator in this "space" that commutes with the Hamiltonian. So does charge, and baryon number. For the proton/neutron system, let's define $I_3 = Q - \tfrac{1}{2}B$, so

$$I_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \tfrac{1}{2}\sigma_3 . \tag{8.6}$$

It further follows that the commutator of I_3 and I_1 will also commute with the Hamiltonian. This commutator is given by

$$I_2 \equiv i\left[I_1, I_3\right] = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \tfrac{1}{2}\sigma_2 . \tag{8.7}$$

In general, an isospin operator acts on a nucleon state according to

$$I_a \left| N_i \right\rangle = \left| N_j \right\rangle \left(I_a \right)^j{}_i \, , \tag{8.8}$$

where $\left(I_a \right)^j{}_i$ is the row *j* and column *i* component of the matrix I_a.

The three matrices I_a have what should be familiar-looking commutation relations:

$$\left[I_a , I_b \right] = i\varepsilon_{abc} I_c \, . \tag{8.9}$$

These are identical to the commutation relations for spin operators, eq. (2.63). Because of the similarity to spin of these operators, and their initial use for understanding isotopes, the resulting symmetry is called *isotopic spin*, or *isospin* for short.

Let's spend a long paragraph discussing the mathematics of isospin. The three operators I_a are called the *generators* of the symmetry. Consider an arbitrary matrix of the form

$$U = \exp\left(-i\theta_a I_a \right) , \tag{8.10}$$

where θ_a are three arbitrary real numbers. Because the generators are all Hermitian matrices, it follows that

$$U^\dagger U = \exp\left(i\theta_a I_a \right) \exp\left(-i\theta_a I_a \right) = 1 \, . \tag{8.11}$$

This is the definition of a *unitary* matrix. It is not hard to show that the product of two unitary matrices is a unitary matrix, that the inverse of a unitary matrix is unitary, and that the identity is a unitary matrix. These properties, together with the associative law of multiplication for matrices, prove that the set of unitary matrices forms a mathematical object called a *group*. The set of unitary matrices of order 2 is called U(2). In addition, there is a mathematical theorem that relates the determinant of an exponential of a matrix to its trace, namely

$$\det\left[\exp\left(M \right) \right] = \exp\left[\mathrm{Tr}\left(M \right) \right] . \tag{8.12}$$

Since all three I_a's are traceless, it is then easy to show that

$$\det\left(U \right) = \det\left[\exp\left(-i\theta_a I_a \right) \right] = \exp\left[-i\theta_a \mathrm{Tr}\left(I_a \right) \right] = \exp\left(0 \right) = 1 \, . \tag{8.13}$$

Unitary matrices with determinant 1 are called *special unitary* matrices. The set of all such rank-two matrices is called SU(2). Isospin is an SU(2) approximate symmetry of strong interactions. We will almost never work with the matrices *U*, we will use only the I_a's.

In addition to the three isospin generators already defined, it is useful to also define the isospin raising and lowering operators in analogy with eq. (2.66), so that

$$I_\pm = I_1 \pm iI_2 \, , \quad I_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} , \quad I_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} . \tag{8.14}$$

Note that these operators are not Hermitian, they are Hermitian conjugates of each other. Their effects on the nucleon states are

$$I_+ \left| p^+ \right\rangle = 0 \, , \quad I_+ \left| n^0 \right\rangle = \left| p^+ \right\rangle , \quad I_- \left| p^+ \right\rangle = \left| n^0 \right\rangle , \quad I_- \left| n^0 \right\rangle = 0 \, . \tag{8.15}$$

C. Isospin and Isotopes

Let's try to understand how isospin can relate systems with more than one particle. Consider some combination of protons and neutrons, like, say $\left|pp\right\rangle = \left|N_1 N_1\right\rangle$. It is pretty obvious that it has baryon number 2 and charge 2, so $I_3 = B - \frac{1}{2}Q = +1$. This gives us a pretty good idea of how isospin operators work on multiple particles – it acts separately on each of them, so

$$I_a \left|N_i N_j\right\rangle = \left|N_k N_j\right\rangle \left(I_a\right)^k{}_i + \left|N_i N_k\right\rangle \left(I_a\right)^k{}_j . \tag{8.16}$$

This is probably easier to understand using an example. We have, for example,

$$I_- \left|p^+ p^+\right\rangle = \left|n^0 p^+\right\rangle + \left|p^+ n^0\right\rangle . \tag{8.17}$$

Let's define the following three states:

$$\left|1,+1\right\rangle = \left|p^+ p^+\right\rangle , \quad \left|1,0\right\rangle = \tfrac{1}{\sqrt{2}}\left(\left|n^0 p^+\right\rangle + \left|p^+ n^0\right\rangle\right) , \quad \left|1,-1\right\rangle = \left|n^0 n^0\right\rangle . \tag{8.18}$$

We also note that we can switch between the various states by using $I_\pm$. Specifically,

$$I_+\left|1,+1\right\rangle = 0 , \quad I_+\left|1,0\right\rangle = \sqrt{2}\left|1,+1\right\rangle , \quad I_+\left|1,-1\right\rangle = \sqrt{2}\left|1,0\right\rangle , \tag{8.19a}$$

$$I_-\left|1,+1\right\rangle = \sqrt{2}\left|1,0\right\rangle , \quad I_-\left|1,0\right\rangle = \sqrt{2}\left|1,-1\right\rangle , \quad I_-\left|1,-1\right\rangle = 0 . \tag{8.19b}$$

The Hermitian conjugates of eq. (8.19b) are

$$\left\langle 1,+1\right| I_+ = \sqrt{2}\left\langle 1,0\right| \quad \left\langle 1,0\right| I_+ = \sqrt{2}\left\langle 1,-1\right| , \quad \text{and} \quad \left\langle 1,-1\right| I_+ = 0 . \tag{8.20}$$

We can then use eqs. (8.19a) and (8.20) to relate the energies of our three states:

$$\begin{aligned}\left\langle 1,+1\right|\mathcal{H}\left|1,+1\right\rangle &= \tfrac{1}{\sqrt{2}}\left\langle 1,+1\right|\mathcal{H}I_+\left|1,0\right\rangle = \tfrac{1}{\sqrt{2}}\left\langle 1,+1\right|I_+\mathcal{H}\left|1,0\right\rangle = \left\langle 1,0\right|\mathcal{H}\left|1,0\right\rangle \\ &= \tfrac{1}{\sqrt{2}}\left\langle 1,0\right|\mathcal{H}I_+\left|1,-1\right\rangle = \tfrac{1}{\sqrt{2}}\left\langle 1,0\right|I_+\mathcal{H}\left|1,-1\right\rangle = \left\langle 1,-1\right|\mathcal{H}\left|1,-1\right\rangle .\end{aligned} \tag{8.21}$$

We now have a prediction. For every bound state of two protons, there will be a bound state of a proton plus neutron, plus a bound state of two neutrons, with the same energy. I don't think any of these states actually exist, but you get the idea.

There is one other way to combine a proton and a neutron, namely

$$\left|0,0\right\rangle = \tfrac{1}{\sqrt{2}}\left(\left|n^0 p^+\right\rangle - \left|p^+ n^0\right\rangle\right) . \tag{8.22}$$

If you let any of the isospin operators act on this state, you will get zero exactly. Hence there can also be bound proton-neutron states that are unrelated to any other isotopes. In fact, the deuteron, the nucleus of ^{2}H, is exactly such a state.

What is going on here is the addition of isospin, in exact analogy of the addition of spin in quantum mechanics. Just as when you add two spin $\frac{1}{2}$ objects you can get something with isospin 0 or 1. You will note that in eqs. (8.18) and (8.22) I labeled the states as $\left|I, I_3\right\rangle$. In this notation, the proton and neutron would be $\left|p^+\right\rangle = \left|\frac{1}{2},+\frac{1}{2}\right\rangle$ and $\left|n^0\right\rangle = \left|\frac{1}{2},-\frac{1}{2}\right\rangle$.

In section 2G we worked out general expressions for the spin generators for arbitrary spin, but it is helpful to have them all together in case you need them. Up to isospin $\frac{3}{2}$ they are:

$$0: \quad I_3 = I_\pm = \begin{pmatrix} 0 \end{pmatrix}. \tag{8.23a}$$

$$\tfrac{1}{2}: \quad I_3 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}, \quad I_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad I_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{8.23b}$$

$$1: \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad I_+ = \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix}, \quad I_- = \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix}, \tag{8.23c}$$

$$\tfrac{3}{2}: \quad I_3 = \begin{pmatrix} \frac{3}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{3}{2} \end{pmatrix}, \quad I_+ = \begin{pmatrix} 0 & \sqrt{3} & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & \sqrt{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad I_- = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \sqrt{3} & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix}. \tag{8.23d}$$

We will note for now that I_+ always increases the charge by one, while I_- decreases it by one. These various choices for the isospin generators are called the different *representations* of the isospin generators. Similarly, when we get to SU(3), we will be working with a variety of representations.

D. Isospin for Other Particles

We'd like to figure out what isospin does to the various particles and anti-particles in the table. First, consider an anti-nucleon. It has the opposite charge and baryon number of its corresponding particle, and therefore it will have the opposite I_3 value, so

$$I_3 \left| \bar{p}^- \right\rangle = -\tfrac{1}{2} \left| \bar{p}^- \right\rangle, \quad I_3 \left| \bar{n}^0 \right\rangle = \tfrac{1}{2} \left| \bar{n}^0 \right\rangle. \tag{8.24}$$

We see immediately that when dealing with anti-particles, we will have to introduce a minus sign. To keep track of this difference, let's denote anti-nucleons by the notation

$$\left| \bar{p}^- \right\rangle = \left| \bar{N}^1 \right\rangle, \quad \left| \bar{n}^0 \right\rangle = \left| \bar{N}^2 \right\rangle, \tag{8.25}$$

and the up index reminds us that it works a different way.

What about $I_\pm$? We know that I_+ increase charge by 1. It follows that, for example, $I_+ \left| \bar{n} \right\rangle = 0$, because there is no anti-nucleon with charge +1. In contrast, $I_+ \left| \bar{p} \right\rangle$ could produce an $\left| \bar{n} \right\rangle$, but we know there is already a minus sign, so we conclude $I_+ \left| \bar{p} \right\rangle = -\left| \bar{n} \right\rangle$. We can work out I_- similarly, and we find

$$I_+ \left| \bar{n}^0 \right\rangle = 0, \quad I_+ \left| \bar{p}^- \right\rangle = -\left| \bar{n}^0 \right\rangle, \qquad I_- \left| \bar{n}^0 \right\rangle = -\left| \bar{p}^- \right\rangle, \quad I_- \left| \bar{p}^- \right\rangle = 0. \tag{8.26}$$

Indeed, it isn't too hard to figure out from this that in general

$$I_a\left|\bar{N}^i\right\rangle = -\left(I_a\right)^i{}_j\left|\bar{N}^j\right\rangle. \tag{8.27}$$

Technically, the antinucleons can be thought of as another isospin $\frac{1}{2}$ set. We therefore rarely use up indices with isospin, but having expressions like eq. (8.27) will prove handy when we generalize to the group SU(3)$_F$.

Because isospin is an approximate symmetry of the strong interactions of the proton and neutron, it must also be a symmetry of any particles that interact with them. Hence all strongly interacting particles should come in isospin *multiplets*, just as nuclei do. The particles in the tables on page 121 have been gathered into isospin multiplets and labeled by their I and I_3 eigenvalues. You can deduce how they fit together by the arrangement in the tables. For example, the four Δ baryons form an isospin $\frac{3}{2}$ multiplet, with the individual particles related by

$$\left|\Delta^{++}\right\rangle = \left|\tfrac{3}{2},+\tfrac{3}{2}\right\rangle, \quad \left|\Delta^{+}\right\rangle = \left|\tfrac{3}{2},+\tfrac{1}{2}\right\rangle, \quad \left|\Delta^{0}\right\rangle = \left|\tfrac{3}{2},-\tfrac{1}{2}\right\rangle, \quad \left|\Delta^{-}\right\rangle = \left|\tfrac{3}{2},-\tfrac{3}{2}\right\rangle. \tag{8.28}$$

We can then use the matrices given in (8.23) to relate these particles, their masses, and their interactions. For example, to show that the various Δ's have the same mass, we would show

$$m\left(\Delta^{++}\right) = \left\langle\Delta^{++}\right|\mathcal{H}\left|\Delta^{++}\right\rangle = \frac{1}{\sqrt{3}}\left\langle\Delta^{++}\right|\mathcal{H}I_+\left|\Delta^{+}\right\rangle = \frac{1}{\sqrt{3}}\left\langle\Delta^{++}\right|I_+\mathcal{H}\left|\Delta^{+}\right\rangle. \tag{8.29}$$

Now, to find $\left\langle\Delta^{++}\right|I_+$, first note that $I_-\left|\Delta^{++}\right\rangle = \sqrt{3}\left|\Delta^{+}\right\rangle$, or taking the Hermitian conjugate, we have

$$\left\langle\Delta^{++}\right|I_+ = \sqrt{3}\left\langle\Delta^{+}\right|.$$

Substituting this into eq. (8.29), we find

$$m\left(\Delta^{++}\right) = \left\langle\Delta^{+}\right|\mathcal{H}\left|\Delta^{+}\right\rangle = m\left(\Delta^{+}\right).$$

We could then continue this process to show that all the Δ's have the same mass.

We can also use isospin to relate decay rates. For example, consider the decays $\Delta^{++}\to p^+\pi^+$, $\Delta^+\to n^0\pi^+$ and $\Delta^+\to p^+\pi^0$. We start by noting that

$$\left\langle p^+\pi^0\right|\mathcal{H}\left|\Delta^{+}\right\rangle = \frac{1}{\sqrt{3}}\left\langle p^+\pi^0\right|\mathcal{H}I_-\left|\Delta^{++}\right\rangle = \frac{1}{\sqrt{3}}\left\langle p^+\pi^0\right|I_-\mathcal{H}\left|\Delta^{++}\right\rangle, \tag{8.30a}$$

$$\left\langle n^0\pi^+\right|\mathcal{H}\left|\Delta^{+}\right\rangle = \frac{1}{\sqrt{3}}\left\langle n^0\pi^+\right|\mathcal{H}I_-\left|\Delta^{++}\right\rangle = \frac{1}{\sqrt{3}}\left\langle n^0\pi^+\right|I_-\mathcal{H}\left|\Delta^{++}\right\rangle. \tag{8.30b}$$

Now, we will need the state relations

$$I_+\left|p^+\pi^0\right\rangle = 0+\sqrt{2}\left|p^+\pi^+\right\rangle \quad\Rightarrow\quad \left\langle p^+\pi^0\right|I_+ = \sqrt{2}\left\langle p^+\pi^+\right|, \tag{8.31a}$$

$$I_+\left|n^0\pi^+\right\rangle = \left|p^+\pi^+\right\rangle + 0 \quad\Rightarrow\quad \left\langle n^0\pi^+\right|I_+ = \left\langle p^+\pi^+\right|. \tag{8.31b}$$

Substituting eqs. (8.31) into (8.30), we conclude

$$\left\langle p^+\pi^0\right|\mathcal{H}\left|\Delta^{+}\right\rangle = \sqrt{\tfrac{2}{3}}\left\langle p^+\pi^+\right|\mathcal{H}\left|\Delta^{++}\right\rangle \quad\text{and}\quad \left\langle n^0\pi^+\right|\mathcal{H}\left|\Delta^{+}\right\rangle = \sqrt{\tfrac{1}{3}}\left\langle p^+\pi^+\right|\mathcal{H}\left|\Delta^{++}\right\rangle. \tag{8.32}$$

Now, we still have no idea what any of these matrix elements are. But suppose we go ahead and calculate the corresponding decay rate. We would square this amplitude and then go on to calculate the decay rates in exactly the same way for all three cases. We therefore conclude that

$$\frac{\Gamma\left(\Delta^+ \to p^+\pi^0\right)}{\Gamma\left(\Delta^{++} \to p^+\pi^+\right)} = \frac{2}{3} \quad \text{and} \quad \frac{\Gamma\left(\Delta^+ \to n^0\pi^+\right)}{\Gamma\left(\Delta^{++} \to p^+\pi^+\right)} = \frac{1}{3}. \tag{8.33}$$

We can conclude that the total decay rate for these two particles is identical (which is hardly surprising), and if we know that the Δ decays pretty much always to nucleon plus pion, we can work out the branching ratios for the Δ^+.

It would be nice to have a complete theory of strong interactions in which the matrix elements are parameterized and we are able to do specific calculations. However, this is difficult. In fact, the particles that we have been describing (baryons and mesons) are composite particles, and hence the interactions can be complicated functions of the momenta (the theory is not renormalizable). In particular, if you collide particles at very high energy, the constituent quarks get all shaken up, and generally the particles you are working with fall apart, creating myriad new particles. It is virtually impossible to describe such interactions in terms of the interactions of specific baryons and mesons.

However, at low energy, baryons and nucleons can often remain intact, and it is possible to make a detailed theory of such particles. In particular, it is often a good approximation to work in a theory consisting of just pions and nucleons. Such a theory has just two masses, the mass of the pion and the nucleon. In addition, there are four interactions involving a pion and a nucleon in the initial state and a nucleon in the final state, namely

$$\left\langle p^+\middle|\mathcal{H}\middle|p^+,\pi^0\right\rangle,\quad \left\langle n^0\middle|\mathcal{H}\middle|n^0,\pi^0\right\rangle,\quad \left\langle n^0\middle|\mathcal{H}\middle|p^+,\pi^-\right\rangle,\quad \left\langle p^+\middle|\mathcal{H}\middle|n^0,\pi^+\right\rangle.$$

These matrix elements are related by isospin. For example,

$$\begin{aligned}\left\langle p^+\middle|\mathcal{H}\middle|p^+,\pi^0\right\rangle &= \sqrt{\tfrac{1}{2}}\left\langle p^+\middle|\mathcal{H}I_+\middle|p^+,\pi^-\right\rangle = \sqrt{\tfrac{1}{2}}\left\langle p^+\middle|I_+\mathcal{H}\middle|p^+,\pi^-\right\rangle = \sqrt{\tfrac{1}{2}}\left\langle n^0\middle|\mathcal{H}\middle|p^+,\pi^-\right\rangle,\\ \left\langle p^+\middle|\mathcal{H}\middle|p^+,\pi^0\right\rangle &= \left\langle p^+\middle|\mathcal{H}I_+\middle|n^0,\pi^0\right\rangle = \left\langle p^+\middle|I_+\mathcal{H}\middle|n^0,\pi^0\right\rangle = \left\langle n^0\middle|\mathcal{H}\middle|n^0,\pi^0\right\rangle.\end{aligned}$$

We can in fact relate all four matrix elements, and we find

$$\left\langle p^+\middle|\mathcal{H}\middle|p^+,\pi^0\right\rangle = \left\langle n^0\middle|\mathcal{H}\middle|n^0,\pi^0\right\rangle = \frac{1}{\sqrt{2}}\left\langle n^0\middle|\mathcal{H}\middle|p^+,\pi^-\right\rangle = \frac{1}{\sqrt{2}}\left\langle p^+\middle|\mathcal{H}\middle|n^0,\pi^+\right\rangle. \tag{8.34}$$

We can now get more specific about the interactions. Experiment shows that the pion is a pseudoscalar particle and has pseudoscalar couplings. In particular, the couplings to a good approximation are just of the form $ig\gamma_5$. Because there are four non-zero matrix elements, there are four Feynman rules, but they all have the same couplings. We have denoted a nucleon by a

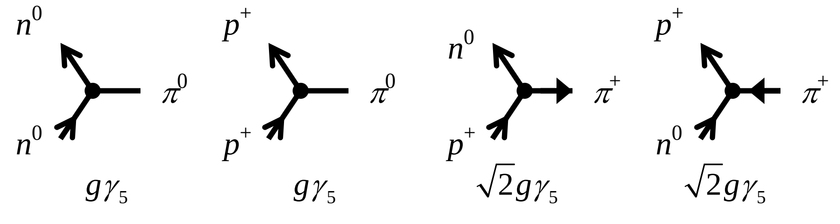


Figure 8-1: The four interactions of the pion-nucleon theory.

line with an open arrow, a π^+ with a closed arrow, and a π^0 by no arrow. The Feynman vertices are listed in Fig. 8-1. In addition, there are the rules for boson and fermion propagators, and the rule for subtracting diagrams with swapped external fermion lines. Note that though the theory has five different particles, there are only three parameters: the nucleon mass, the pion mass, and the single coupling. This illustrates the power of internal symmetries (in this case isospin) to simplify a problem.

Before moving onwards, it is worthwhile to note that it is always possible to work with *only* the spin $\frac{1}{2}$ representation of the isospin group, if we so desire. We showed in section C that we could, for example, get isospin 1 by combining two isospin $\frac{1}{2}$ nucleons. We could similarly describe any isospin 1 set of particles by putting two indices, say $\left|{}_{ij}\right\rangle$, or isospin $\frac{3}{2}$ but putting three indices down, $\left|{}_{ijk}\right\rangle$. In this notation, for example, the three pions would be described as

$$\left|\pi^+\right\rangle = \left|\pi_{11}\right\rangle, \quad \left|\pi^0\right\rangle = \tfrac{1}{\sqrt{2}}\left(\left|\pi_{12}\right\rangle + \left|\pi_{21}\right\rangle\right), \quad \left|\pi^-\right\rangle = \left|\pi_{22}\right\rangle. \tag{8.35}$$

We could then easily compute the effects of the various operators; for example

$$I_-\left|\pi^+\right\rangle = I_-\left|\pi_{11}\right\rangle = \left|\pi_{21}\right\rangle + \left|\pi_{12}\right\rangle = \sqrt{2}\left|\pi^0\right\rangle. \tag{8.36}$$

A state with isospin zero would just be described with no indices, like $\left|\eta\right\rangle$. The advantage of this approach is that we need never work with matrices larger than 2×2, and we work only with the easily understood matrices given by eq. (8.23b). This advantage will loom large when we get to SU(3) symmetry, since otherwise we would be forced to work with 8×8 or 10×10 matrices, a nightmarish proposition.

It should be recognized that isospin is only an *approximate* symmetry of nature. The general assumption is that the Hamiltonian contains multiple terms, but the one responsible for the strong interactions respects, or approximately respects, isospin symmetry. It is certainly the case that electromagnetism does *not* respect isospin symmetry, since the proton and neutron do not have the same electric charge. Weak interactions also do not care about isospin. Hence the π^0 decays electromagnetically, while the $\pi^\pm$ decay via weak interactions.

It should be noted that there are other quantities that are conserved by strong interactions. Among these are baryon number and electric charge. One particularly useful combination is *strangeness*, which for these particles is given by

$$S \equiv 2\left(Q - I_3\right) - B. \tag{8.37}$$

Note that strangeness is constant across each multiplet. The strangeness values are included in the tables on page 121. Strangeness vanishes for the most common/stable particles (nucleons, pions, etc.). Strangeness is conserved by both strong forces and electromagnetism, but not weak interactions.

Isospin goes a long way to interrelating the many known particles, but not far enough. We need a bigger symmetry that can relate even more particles, and to this we turn to the group SU(3).

E. SU(3)

Before we discuss SU(3), let’s review SU(2). The group SU(2) consists of all 2×2 unitary matrices with determinant one. They can be formed from the three generators given by eqs. (8.5), (8.6), and (8.7). These matrices are chosen so that

$$\mathrm{Tr}\left(T_a T_b\right) = \tfrac{1}{2}\delta_{ab} \;, \tag{8.38}$$

where I have decided to call these generators T_a for this section. The generators T_a act on the two basis states $\left|{}_i\right\rangle$ for $i=1,2$, and they also can act on states with up indices $\left|{}^i\right\rangle$. We can figure out how it works on any combination of up and down indices with the help of eqs. (8.16) and (8.27), namely

$$\begin{aligned} T_a \left| {}^{j_1 j_2 \cdots j_M}_{i_1 i_2 \cdots i_N} \right\rangle = {} & \left| {}^{j_1 j_2 \cdots j_M}_{k i_2 \cdots i_N} \right\rangle \left(T_a\right)^k{}_{i_1} + \left| {}^{j_1 j_2 \cdots j_M}_{i_1 k \cdots i_N} \right\rangle \left(T_a\right)^k{}_{i_2} + \cdots + \left| {}^{j_1 j_2 \cdots j_M}_{i_1 i_2 \cdots k} \right\rangle \left(T_a\right)^k{}_{i_N} \\ & - \left(T_a\right)^{j_1}{}_k \left| {}^{k j_2 \cdots j_M}_{i_1 i_2 \cdots i_N} \right\rangle - \left(T_a\right)^{j_2}{}_k \left| {}^{j_1 k \cdots j_M}_{i_1 i_2 \cdots i_N} \right\rangle - \cdots - \left(T_a\right)^{j_M}{}_k \left| {}^{j_1 j_2 \cdots k}_{i_1 i_2 \cdots i_N} \right\rangle . \end{aligned} \tag{8.39}$$

Don’t let this formidable formula scare you, it just means that there is one term for each index up or down, and T_a acts on just one index at a time. It should be clearer when we actually need it. In SU(2), there is little motivation to use up indices, but it will be valuable in SU(3).

We now want to generalize the concept of SU(2) and develop the group SU(3), the set of special unitary 3×3 matrices. The generators for SU(3) will be 3×3 traceless Hermitian matrices that satisfy (8.38). There are eight of them, and they are given explicitly by

$$\begin{gathered} T_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad T_4 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad T_6 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\ T_2 = \frac{1}{2}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad T_5 = \frac{1}{2}\begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad T_7 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \end{gathered} \tag{8.40a}$$

$$T_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad T_8 = \frac{1}{2\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{8.40b}$$

These matrices (or actually, twice them) are called the *Gell-Mann* matrices, after Murray Gell-Mann.

It is not difficult to show that the commutator of any two of these matrices yields some linear combination of the same matrices, so that

$$\left[T_a, T_b\right] = i f_{abc} T_c \;, \tag{8.41}$$

where the f_{abc}’s are a set of numbers called the *structure constants*. These can be worked out without a great deal of difficulty by simply computing commutators. For example,

$$\left[T_4,T_5\right]=\frac{1}{2}\begin{pmatrix} i & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -i \end{pmatrix}=i\left(\tfrac{1}{2}T_3+\tfrac{\sqrt{3}}{2}T_8\right).$$

From this we can deduce that $f_{453}=\frac{1}{2}$, $f_{458}=\frac{1}{2}\sqrt{3}$, and the rest of the f_{45c}'s vanish. The structure constants play the same role in SU(3) that the ε_{abc}'s did in SU(2), except the indices take on eight values instead of three. Like ε_{abc}, they are completely anti-symmetric.

In isospin, we can diagonalize I_3 and list particles within an isospin representation by their I_3 value. Similarly, in SU(3), it is possible to diagonalize *both* T_3 and T_8, since they commute with each other. In isospin, we then paired the remaining two isospin generators into $I_1 \pm iI_2$. In SU(3), similar combinations will prove easier to work with. The logical combinations turn out to be $T_1 \pm iT_2$, $T_4 \pm iT_5$ and $T_6 \pm iT_7$. I will give them names like $T_{i\to j}$, where *i* and *j* are distinct integers chosen from 1, 2, 3. The six resulting operators are

$$T_{1\to 2}=\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}=T_{2\to 1}^{\dagger}, \quad T_{1\to 3}=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}=T_{3\to 1}^{\dagger}, \quad T_{2\to 3}=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}=T_{3\to 2}^{\dagger}. \qquad (8.42)$$

These expressions will be easier to understand when we use them. These operators replace the eight appearing in eq. (8.40a), which we won't actually be using. Note that these matrices are not Hermitian; in fact, $T_{i\to j}^{\dagger}=T_{j\to i}$.

Just as with the group SU(2), where there are many possible spins or representations, for SU(3) there are many possible representations as well. Like with SU(2), these could be represented by larger and larger matrices akin to eqs. (8.23), but it is easier to use *only* these three by three matrices and deal with larger representations by working with multiple indices that can be either down or up. The eight operators T_a act on an arbitrary state with up and down indices according to eq. (8.39). What this means for T_3 and T_8 is that the T_3 value and T_8 value is just the sum of the values for the down indices minus the value for the up indices. For example, suppose we had a state like $\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle$. We would have

$$T_3\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle=\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle 0-\left(-\tfrac{1}{2}\right)\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle=\tfrac{1}{2}\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle, \quad T_8\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle=\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle\left(-\tfrac{1}{\sqrt{3}}\right)-\tfrac{1}{2\sqrt{3}}\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle=-\tfrac{\sqrt{3}}{2}\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle.$$

For $T_{i\to j}$, you get one term for each index. For a down index, it would change *i* to *j* if the index is *i*, otherwise it yields zero. For an up index, it would change *j* to *i* if the index is *j* and also introduce a minus sign, otherwise it yields zero. For example, let's figure out what all six of the $T_{i\to j}$'s do to $\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle$:

$$T_{1\to 2}\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle=0-\left|\begin{smallmatrix}1\\3\end{smallmatrix}\right\rangle, \quad T_{1\to 3}\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle=0-0, \quad T_{2\to 3}\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle=0-0,$$

$$T_{2\to 1}\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle=0-0, \quad T_{3\to 1}\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle=\left|\begin{smallmatrix}2\\1\end{smallmatrix}\right\rangle-0, \quad T_{3\to 2}\left|\begin{smallmatrix}2\\3\end{smallmatrix}\right\rangle=\left|\begin{smallmatrix}2\\2\end{smallmatrix}\right\rangle-\left|\begin{smallmatrix}3\\3\end{smallmatrix}\right\rangle.$$

We now have all we need to know about the group SU(3). It is time to learn what it tells us about particle physics.

F. SU(3) Flavor Symmetry

It now seems sensible to ask whether SU(3) is a good approximate symmetry of strong interactions. At first it seems unlikely, since we would expect there to be larger groups of particles with nearly identical masses. Looking over our particle masses, there are relatively few opportunities to find nearly identical masses that are not explained by isospin. However, there are two reasons we should not despair. First of all, if you study isospin in more detail, you discover that the approximation that the masses are equal is a much worse approximation than that the interactions are equal. If this follows through into SU(3), then we would expect rather poor equality of mass, and somewhat better equality of interactions. The second reason we shouldn't give up is that even though the masses of, say, the lightest mesons are rather different, the lightest ones all have certain things in common. For example, they are all spin zero and have pseudoscalar couplings. There are other more technical similarities. It therefore seems sensible to try putting them together into representations of SU(3). In this context, it is called SU(3) flavor symmetry, where *flavor* refers to the different variety of baryons and mesons (later, quarks), and is abbreviated $SU(3)_F$.

The reason this was so difficult for anyone dumber than Gell-Mann (i.e. everyone) to recognize is that the light particles (it turns out) do not fit into the simple three-dimensional representation of $SU(3)_F$. Specifically, the eight lightest mesons don't correspond to states of the form $\left|{}_i\right\rangle$, with one index down, but to $\left|{}^j_i\right\rangle$, with one index up and one down. Let's call all these states $\left|M_i^j\right\rangle$. Specifically,

$$\begin{array}{ccc} & \left|K^0\right\rangle=\left|M_2^3\right\rangle, \qquad \left|K^+\right\rangle=\left|M_1^3\right\rangle, & \\ -\left|\pi^-\right\rangle=\left|M_2^1\right\rangle, & \begin{array}{c} -\left|\pi^0\right\rangle=\frac{1}{\sqrt{2}}\left(\left|M_1^1\right\rangle-\left|M_2^2\right\rangle\right), \\ \left|\eta\right\rangle=\frac{1}{\sqrt{6}}\left(\left|M_1^1\right\rangle+\left|M_2^2\right\rangle-2\left|M_3^3\right\rangle\right), \end{array} & \left|\pi^+\right\rangle=\left|M_1^2\right\rangle, \\ & -\left|K^-\right\rangle=\left|M_3^1\right\rangle, \qquad \left|\bar{K}^0\right\rangle=\left|M_3^2\right\rangle. & \end{array} \tag{8.43}$$

I have tried to arrange eqs. (8.43) so they make an approximate hexagon shape. The scattered minus signs are a matter of convention, and appear here to make our notation consistent with section D. It can be checked that the eight generators of SU(3) simply take these states into each other. The eight particles or *octet* can be categorized by their T_3 and T_8 values, as sketched in Fig. 8-2. When you make such a graph, it makes a perfectly regular hexagon with two particles at the origin.

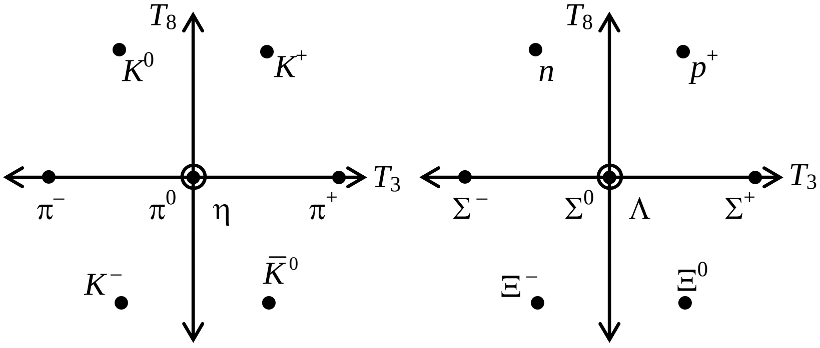


Figure 8-2: The arrangement of the eight lightest spin-0 mesons (left) and spin-1/2 baryons (right) into octets, or 8's, according to Gell-Mann's eightfold way. The circle with a dot signifies there are two particles represented at that point.

This way of organizing particles was called the *eight-fold way* by Gell-Mann, jokingly comparing it to Buddhism. In a nearly identical manner, the eight lightest spin $\frac{1}{2}$ baryons can similarly be organized into a set of eight particles, specifically

$$\begin{gathered} \left|n\right\rangle = \left|B_2^3\right\rangle, \qquad \left|p^+\right\rangle = \left|B_1^3\right\rangle, \\ -\left|\Sigma^-\right\rangle = \left|B_2^1\right\rangle, \quad \begin{matrix} -\left|\Sigma^0\right\rangle = \frac{1}{\sqrt{2}}\left(\left|B_1^1\right\rangle - \left|B_2^2\right\rangle\right), \\ \left|\Lambda\right\rangle = \frac{1}{\sqrt{6}}\left(\left|B_1^1\right\rangle + \left|B_2^2\right\rangle - 2\left|B_3^3\right\rangle\right), \end{matrix} \quad \left|\Sigma^+\right\rangle = \left|B_1^2\right\rangle, \\ -\left|\Xi^-\right\rangle = \left|B_3^1\right\rangle, \qquad \left|\Xi^0\right\rangle = \left|B_3^2\right\rangle. \end{gathered} \tag{8.44}$$

Of course, these eight particles also form a nice hexagon as before as included in Fig. 8-2. The eight lightest spin-1 mesons also form another pattern of eight.

We note that isospin symmetry is just a part of SU(3)$_F$ symmetry. Specifically, T_1, T_2 and T_3 correspond to I_1, I_2, and I_3. Mathematically, we say that isospin is a *subgroup* of SU(3)$_F$.

Do we need other representations than the eightfold way? Yes, it turns out we do. The presence of the Δ's tells you that you need something with larger isospins in it. It turns out that the Δ's and other particles fit into a representation with three lowered indices, of the form $\left|{}_{ijk}\right\rangle$. I will call these $\left|B_{ijk}^*\right\rangle$, and it turns out you can include all of the spin-$\frac{3}{2}$ baryons into such a representation. Specifically, we find

$$\begin{gathered} \left|\Delta^-\right\rangle = \left|B_{222}^*\right\rangle, \; \left|\Delta^0\right\rangle = \tfrac{1}{\sqrt{3}}\left(\left|B_{122}^*\right\rangle + \left|B_{212}^*\right\rangle + \left|B_{221}^*\right\rangle\right), \; \left|\Delta^+\right\rangle = \tfrac{1}{\sqrt{3}}\left(\left|B_{112}^*\right\rangle + \left|B_{121}^*\right\rangle + \left|B_{211}^*\right\rangle\right), \; \left|\Delta^{++}\right\rangle = \left|B_{111}^*\right\rangle, \\ \left|\Sigma^{*-}\right\rangle = \tfrac{1}{\sqrt{3}}\left(\left|B_{223}^*\right\rangle + \left|B_{232}^*\right\rangle + \left|B_{322}^*\right\rangle\right), \qquad \left|\Sigma^{*+}\right\rangle = \tfrac{1}{\sqrt{3}}\left(\left|B_{113}^*\right\rangle + \left|B_{131}^*\right\rangle + \left|B_{311}^*\right\rangle\right), \\ \left|\Sigma^{*0}\right\rangle = \tfrac{1}{\sqrt{6}}\left(\left|B_{123}^*\right\rangle + \left|B_{132}^*\right\rangle + \left|B_{213}^*\right\rangle + \left|B_{231}^*\right\rangle + \left|B_{312}^*\right\rangle + \left|B_{321}^*\right\rangle\right), \\ \left|\Xi^{*-}\right\rangle = \tfrac{1}{\sqrt{3}}\left(\left|B_{233}^*\right\rangle + \left|B_{323}^*\right\rangle + \left|B_{332}^*\right\rangle\right), \quad \left|\Xi^{*0}\right\rangle = \tfrac{1}{\sqrt{3}}\left(\left|B_{133}^*\right\rangle + \left|B_{313}^*\right\rangle + \left|B_{331}^*\right\rangle\right). \end{gathered} \tag{8.45}$$

If you act on most of these particles with the T_a's you will find that they produce only the nine states given here. The only exception is that $T_{2\to3}\left|\Xi^{*-}\right\rangle$ and $T_{1\to3}\left|\Xi^{*0}\right\rangle$ both yield $\sqrt{3}\left|B_{333}^*\right\rangle$. Gell-Mann noticed this, and predicted the existence of a new particle, the Ω^-, with charge -1 and strangeness -3, given by

$$\left|\Omega^-\right\rangle = \left|B_{333}^*\right\rangle. \tag{8.46}$$

Not only was he able to predict these properties, he was able to make a pretty good estimate of the mass, and when it was discovered, physicists felt they were really on to something. Suddenly everyone needed to learn group theory, so they could understand the obscure mathematics that Gell-Mann was using to make these predictions.

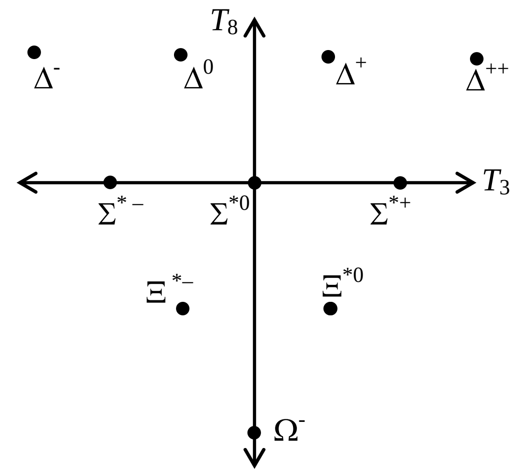


Figure 8-3: The arrangement of the ten lightest spin-3/2 baryons into decuplets, or 10's, according to Gell-Mann's tenfold way.

If we make a plot of T_3 vs. T_8 for the resulting particles, we find the ten spin-$\frac{3}{2}$ baryons (or *decuplet*) make a nice equilateral triangle, with the Ω^- filling out the last corner, as sketched in Fig. 8-3. This method of incorporating baryons was called the *ten-fold way*. So we agree that SU(3)$_F$ is a good way to categorize particles. But what other predictions can it make?

What would happen if SU(3)$_F$ were an exact symmetry of nature? First of all, all of the eight lightest mesons would be degenerate in mass. For example, using the explicit form of $\left|\pi^+\right\rangle$ and $\left|\bar{K}^0\right\rangle$, we see that $T_{1\to3}\left|\pi^+\right\rangle = \left|\bar{K}^0\right\rangle$ and $T_{3\to1}\left|\bar{K}^0\right\rangle = \left|\pi^+\right\rangle$, the latter implying $\left\langle\bar{K}^0\right|T_{1\to3} = \left\langle\pi^+\right|$. If $T_{1\to3}$ commutes with the Hamiltonian, then

$$\left\langle\bar{K}^0\right|\mathcal{H}\left|\bar{K}^0\right\rangle = \left\langle\bar{K}^0\right|\mathcal{H}T_{1\to3}\left|\pi^+\right\rangle = \left\langle\bar{K}^0\right|T_{1\to3}\mathcal{H}\left|\pi^+\right\rangle = \left\langle\pi^+\right|\mathcal{H}\left|\pi^+\right\rangle.$$

Recalling that for bosons, matrix elements can be interpreted as masses squared, this would imply that $m_K^2 = m_\pi^2$. This equation is false by about a factor of twenty, which caused some detractors of SU(3) to claim that it was valid only in the limit 1 = 20. In fact, there is a natural mass scale of about 1 GeV for strong interactions, so the best that can be said for this relation is that both sides are small compared to this scale.

It turns out that one can do much better by including some corrections. This you will explore in homework problem 8.10 and 8.11. But SU(3)$_F$ is going to work better for interactions than it is for masses, so let's stop worrying about masses and start working on interactions.

Consider, for example, the coupling of the Δ particles to nucleons and pions. We found in section D that

$$\left\langle p^+\pi^+\right|\mathcal{H}\left|\Delta^{++}\right\rangle = \sqrt{\tfrac{3}{2}}\left\langle p^+\pi^0\right|\mathcal{H}\left|\Delta^+\right\rangle = \sqrt{3}\left\langle n^0\pi^+\right|\mathcal{H}\left|\Delta^+\right\rangle.$$

These relations can be obtained as usual using the commutation of $\mathcal{H}$ with $T_{2\to1} = I_+$, as we did in section D. But SU(3)$_F$ gives us even more information. For example, it is easy to work out that

$$T_{3\to1}\left|p^+,\bar{K}^0\right\rangle = T_{3\to1}\left|B_1^3,M_3^2\right\rangle = 0 - 0 + \left|B_1^3,M_1^2\right\rangle - 0 = \left|p^+,\pi^+\right\rangle,$$

or taking the Hermitian conjugate of this, we have

$$\left\langle p^+,\bar{K}^0\right|T_{1\to3} = \left\langle p^+,\pi^+\right|.$$

We therefore can conclude that

$$\begin{aligned}\left\langle p^+\pi^+\right|\mathcal{H}\left|\Delta^{++}\right\rangle &= \left\langle p^+,\bar{K}^0\right|T_{1\to3}\mathcal{H}\left|\Delta^{++}\right\rangle = \left\langle p^+,\bar{K}^0\right|\mathcal{H}T_{1\to3}\left|\Delta^{++}\right\rangle = \left\langle p^+,\bar{K}^0\right|\mathcal{H}T_{1\to3}\left|B_{111}^*\right\rangle \\ &= \left\langle p^+,\bar{K}^0\right|\mathcal{H}\left(\left|B_{311}^*\right\rangle + \left|B_{131}^*\right\rangle + \left|B_{113}^*\right\rangle\right) = \sqrt{3}\left\langle p^+,\bar{K}^0\right|\mathcal{H}\left|\Sigma^{*+}\right\rangle.\end{aligned}$$

Indeed, with some effort, *all* of the matrix elements of the form $\left\langle BM\right|\mathcal{H}\left|B^*\right\rangle$ can be related to each other. It is important to realize, however, that this does not imply, say, $\Gamma\left(\Sigma^{*+}\to p^+\bar{K}^0\right) = \frac{1}{3}\Gamma\left(\Delta^{++}\to p^+\pi^+\right)$. Indeed, the Σ^* is too light to decay to a kaon plus nucleon. SU(3) symmetry relates the Feynman amplitudes, not the decay rates. You must work out the decay rates from the Feynman amplitudes to get relations between rates.

Problems for Chapter 8

1. Define *hypercharge* (distinct from weak hypercharge) as $Y = 2\left(Q - I_3\right)$. It will be constant across isospin multiplets. Work *Y* out for the eight lightest mesons and all listed baryons on page 121. Using a combination of hypercharge, baryon number, and energy conservation, prove that the Λ, Σ, and Ξ baryons all cannot decay via strong interactions, while the Δ, Σ^* and Ξ^* all can decay strongly.

2. Since strong interactions preserve both *Q* and I_3, they must also preserve *Y*. Make a plot of I_3 vs. *Y* for the eight lightest spin-0 mesons, drawing a dot for each meson and labeling it by the particle name, and circling it if there are two at a particular point. Repeat for the eight lightest spin-1/2 baryons. Then do it for the nine lightest spin-3/2 baryons. Look up the Ω^- baryon from the particle data group and add it to the diagram. Comment on any similarities, differences, symmetries, etc. of the diagrams you end up with.

3. Use isospin symmetry to show that the masses of the three Σ baryons should all be equal. Do the same for the three π mesons. Keep in mind that a matrix element for a fermion means something different than that for a boson, so your proofs should be (slightly) different.

4. Which matrix elements of the form $\left\langle \Lambda\pi \middle| \mathcal{H} \middle| \Sigma^* \right\rangle$ could be non-zero, based on charge conservation? Relate the three non-zero matrix elements using isospin symmetry, and make a prediction for the relative rates for the decays $\Gamma\left(\Sigma^* \to \Lambda\pi\right)$.

5. Write the equation $\left\langle K^{*+} \middle| \left[\mathcal{H}, I_-\right] \middle| K^+, \pi^+ \right\rangle$ out explicitly, and use it to predict the relative decay rates $\Gamma\left(K^{*+} \to K^+\pi^0\right)$ and $\Gamma\left(K^{*+} \to K^0\pi^+\right)$. Then write out $\left\langle K^{*0} \middle| \left[\mathcal{H}, I_+\right] \middle| K^0, \pi^- \right\rangle$, and use it to predict the relative decay rates of the K^{*0}.

6. Calculate the differential and total cross-section for $\pi^+ n^0 \to \pi^+ n^0$, working in the cm frame, using the Feynman rules of Fig. 8-1.

7. Find $\left[T_6, T_7\right]$, and use this to deduce f_{67c} for all *c*.

8. Prove that $\operatorname{Tr}\left(\left[T_a, T_b\right] T_c\right) = \frac{1}{2} i f_{abc}$. Using this or otherwise, show that f_{abc} is completely anti-symmetric, that is, you get a minus if you interchange any pair of indices. For this purpose, it is sufficient to prove $f_{abc} = -f_{bac} = -f_{acb}$.

9. Work out $T_{1\to3}\left|B^*\right\rangle$ for all nine states in eq. (8.45). Then copy Fig. 8-3, and draw an arrow showing what happens to each of these states. For example, since $T_{1\to3}\left|\Xi^{*0}\right\rangle = \sqrt{3}\left|\Omega^-\right\rangle$, you would draw an arrow from $\left|\Xi^{*0}\right\rangle$ to $\left|\Omega^-\right\rangle$.

10. In this problem we will derive the Gell-Mann Okubo mass relations for SU(3)$_F$. Because T_8 is the only generator that commutes with the isospin subgroup (generated by T_1, T_2 and T_3) it makes sense that the mass splittings have something to do with T_8. Assume the baryons in the decuplet have mass given by

$$\left\langle B^*_{ijk} \middle| \mathcal{H} \middle| B^*_{\ell mn} \right\rangle = X \delta^i_\ell \delta^j_m \delta^k_n + Y \left(T_8\right)^i_\ell \delta^j_m \delta^k_n \ ,$$

where X and Y are constants.

(a) Write the masses of the Δ, Σ^*, Ξ^*, and Ω^- in terms of X and Y.

(b) Find a linear relationship between the masses of the Δ, Σ^*, and Ξ^*. Check that it works pretty well.

(c) Predict the Ω^- mass. This was a prediction made by Gell-Mann and Okubo, and convinced everyone they needed to learn to use this symmetry.

11. Assume the masses of the baryons in the octet are given by

$$\left\langle B^j_i \middle| \mathcal{H} \middle| B^\ell_k \right\rangle = X \delta^i_k \delta^\ell_j + Y \left(T_8\right)^i_k \delta^\ell_j + Z \delta^i_k \left(T_8\right)^\ell_j \ ,$$

where X, Y and Z are constants.

(a) Find formulas for m_N, m_Λ, and m_Σ and m_Ξ in terms of X, Y, and Z.

(b) Eliminate X, Y, and Z to show that $2m_N + 2m_\Xi = 3m_\Lambda + m_\Sigma$. Demonstrate that this works pretty well.

(c) Since the mesons are also in an octet, we might think this implies $4m_K = 3m_\eta + m_\pi$ (where we used the fact that the kaons have the same masses as their anti-particles). Use this to "predict" the kaon mass, and show that this doesn't work very well. What went wrong? (hint – what do Hamiltonian matrix elements represent for bosons?) Fix the formula and show that it now works better.

12. Find a generator $T_{i \to j}$ of SU(3)$_F$ such that $\left|\Sigma^{*+}\right\rangle \propto T_{i\to j}\left|\Delta^{++}\right\rangle$ and determine the proportionality constant. Use SU(3)$_F$ symmetry to relate two of the relevant matrix elements for the $\left|\Sigma^{*+}\right\rangle$ decays to the single matrix element responsible for $\left|\Delta^{++}\right\rangle$ decay. Use another generator to relate the remaining matrix element to the rest. Assuming the final state particles all have the same mass for the three different decays, predict the relative decay rates for the three decays. For which of the three decays will the relationship you just derived be a good approximation? Now look up the masses of the various possibilities and see if you can predict which of the decays actually will be the fastest. Look up the actual branching ratio for Σ^{*+}, and compare with your results.

IX. Quarks and Quantum Chromodynamics

Gell-Mann's SU(3)$_F$ symmetry was an important organizing principle for strong interactions, and it served much the same purpose as the periodic table did for chemistry. It noticed patterns relating various particles, but it didn't really explain them. Why was SU(3) the symmetry? It turned out that only a few representations of SU(3) were ever used. For mesons, the only representations used were the 8 representation, and the 1 representation (which has the property $T_a \left| \ \right\rangle = 0$ for all a). For baryons, the representations used were the 8, the 10 and the 1. For anti-baryons, the indices that were down become up and vice versa, and it turns out the 8, the $\overline{10}$ and the 1 were the only ones used (the $\overline{10}$ looks like an upside-down 10). Why are only the 8 and 1 used for mesons? Why are only the 8, 10, and 1 used for baryons? More importantly, how can we explain strong interactions? Fortunately, rapid progress was being made, and the first major advance was the introduction of quarks.

A. Quarks

In playing with SU(3) symmetry, it makes a lot of sense to start simple, and the simplest non-trivial representation of SU(3) involves a three-dimensional representation corresponding to just one index down, $\left|{}_i\right\rangle$. Shortly after the discovery of SU(3)$_F$ symmetry, Murray Gell-Mann and George Zweig independently came up with a simple way of understanding why only these particular representations occur. They assumed that all baryons were made of three particles called *quarks*, named by Gell-Mann after a quote from James Joyce's *Finnegan's Wake*.[1] Mesons would be made of one quark and one anti-quark, and the anti-baryons from three anti-quarks. The three quarks could be logically called $\left|q_1\right\rangle, \left|q_2\right\rangle, \left|q_3\right\rangle$, but since $\left|q_1\right\rangle$ and $\left|q_2\right\rangle$ correspond to isospin $+\frac{1}{2}$ and $-\frac{1}{2}$ respectively, they were named the *up* and *down* quarks. Because $\left|q_3\right\rangle$ turned out to correspond to the property called strangeness, it is called the *strange* quark. For every quark there is also an anti-quark. The strangeness is just the number of anti-strange quarks minus the number of strange quarks.[2] A plot of the T_3 and T_8 values for the quarks and anti-quarks is given in Fig. 9-1. They are all assumed to be spin $\frac{1}{2}$. The quarks are generally abbreviated *u*, *d*, and *s*, with their anti-quarks denoted $\bar{u}$, $\bar{d}$, and $\bar{s}$. Since a quark is one-third of a baryon, quarks must have

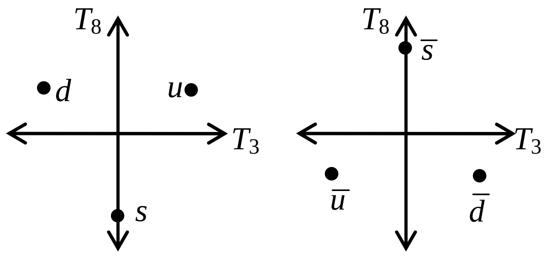


Figure 9-1: The arrangement of the up, down, and strange quarks (left) and their anti-quarks (right).

---

[1] Zweig proposed the name *aces*. Don't you think quark is a cooler name?

[2] Obviously, strangeness should be defined with the opposite sign, but the convention is established, and it's too late to change it.

baryon number $\frac{1}{3}$, which looks strange enough. Even stranger, the charges can be worked out from the charges of the baryons and mesons they comprise. They work out to:

$$Q\left|u\right\rangle = +\tfrac{2}{3}\left|u\right\rangle, \quad Q\left|d\right\rangle = -\tfrac{1}{3}\left|d\right\rangle, \quad Q\left|s\right\rangle = -\tfrac{1}{3}\left|s\right\rangle. \qquad (9.1)$$

Fractional charges certainly looked odd, and this may be why Gell-Mann initially suggested that quarks should not necessarily be thought of as actual particles, but rather as a mathematical tool for understanding SU(3)$_F$ symmetry. However, we now firmly believe that quarks are real.

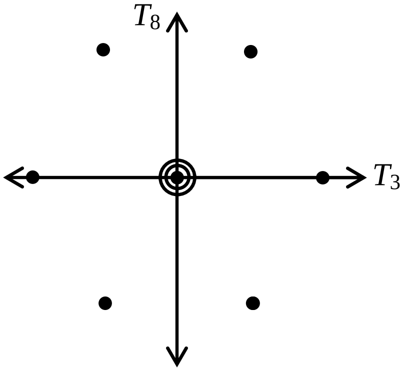


Figure 9-2: The prediction of the possible combinations of T_3 and T_8 values when you make a meson from quark plus anti-quark.

If you take one quark and one anti-quark to make a meson, there are nine possible combinations. The resulting T_3 and T_8 values are drawn in Fig. 9-2. It doesn't take much imagination to see that the mesons then correspond to an 8 with a 1 on top of it. Similarly, if you take three quarks, there are 27 possible combinations, as sketched in Fig. 9-3. It takes only a little more imagination to see that the baryons correspond to a 10, two 8's, and a 1. In group theory, we would write these relations as[1]

$$\begin{aligned} 3\otimes\overline{3} &= 8\oplus 1, \\ 3\otimes 3\otimes 3 &= 10\oplus 8\oplus 8\oplus 1. \end{aligned} \qquad (9.2)$$

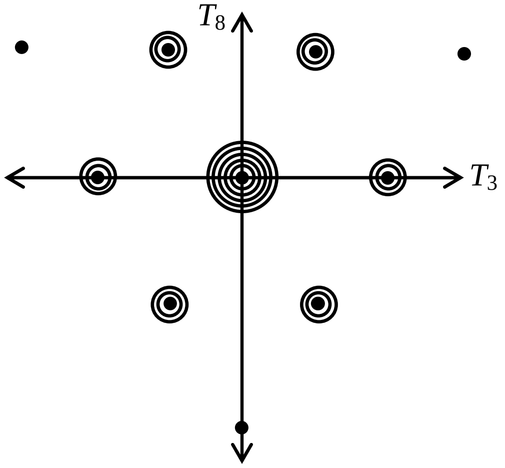


Figure 9-3: The prediction of the possible combinations of T_3 and T_8 values when you make a a baryon from three quarks.

We won't go over enough group theory for this to necessarily make sense to you, but trust me, in the end, the proof is no more sophisticated than using the diagrams in Figs. 9-2 and 9-3. The conclusion is that the quark model successfully predicts why only these representations of SU(3)$_F$ were needed.

Given all this information, how do we explain the partial success of flavor SU(3)$_F$ and the even better success of isospin? Basically, let us assume that strong interactions are indifferent to the type of quark we have: up, down, and strange quarks all have exactly the same strong interactions. But let's imagine that the strange quark is substantially heavier than the other two. Then objects comprised of these various particles will often behave similarly, the main difference being the increased mass when a strange quark is present. The interchange of up with down quarks (isospin) will be an excellent approximation. In contrast, the exchange of an up or down quark for a strange quark (SU(3)$_F$) will be a much poorer approximation, though most of that error will concern the masses, not the interactions. The different types of quarks are called different *flavors*.

---

[1] If it makes you feel any better, when I was taking particle physics my first year in graduate school, the professor wrote equations like eq. (9.2) without explaining what they meant. A fellow students asked me if I understood $3\otimes 3\otimes 3 = 10\oplus 8\oplus 8\oplus 1$. "No," I replied, "but I understand $3\times 3\times 3 = 10+8+8+1$."

Originally, only three quarks were proposed. Due to a problem involving weak interactions (see chapter 10), Glashow, Illiopoulos and Maiani in 1970 predicted the existence of the charm quark, denoted *c*, another charge $+\frac{2}{3}$ quark, discovered in 1974. Later, two more quarks, the bottom (*b*) with charge $-\frac{1}{3}$ and top (*t*) with charge $+\frac{2}{3}$ were predicted and later discovered.[1] For complicated reasons, we believe there are no undiscovered quarks within the Standard Model.

Given the existence of all these additional quarks, is it time to start discussing the group SU(4)$_F$ or maybe even SU(6)$_F$? Though SU(3)$_F$ works reasonably well, basically because all three quarks are lighter than the scale at which strong interactions become strong (about 1 GeV), the charm quark is comparable to this scale, while the bottom and top quarks are much heavier than this scale. This makes objects containing *b* and *t* quarks very unlike their light counterparts. The group SU(4)$_F$ has some modest value to organize particles into representations, though our drawings would get much more complicated. But SU(4)$_F$ does nothing to help us predict masses, and very little to predict interactions. So it isn't very helpful, and we stop with SU(3)$_F$.

We haven't really discussed exactly how one can put together quarks to make various particles. For mesons, you can pretty much read off the quark content from eq. (8.43), where an up index represents a quark and a down index an anti-quark. For example a π^0 is

$$-\left|\pi^0\right\rangle = \tfrac{1}{\sqrt{2}}\left(\left|u\bar{u}\right\rangle - \left|d\bar{d}\right\rangle\right). \tag{9.3}$$

It may look odd that it is a superposition of different kinds of particles, but quantum mechanically there is nothing wrong with this. When you combine a quark and anti-quark, they each have spin $\frac{1}{2}$. If there is no orbital angular momentum, this would predict that mesons should have spin 0 or spin 1, and indeed, all the lightest ones do. Some of the heavier ones do not, but presumably, this just means the quark anti-quark pair are orbiting each other, giving them extra kinetic energy and increasing the meson mass.

Baryons should be formed from three spin $\frac{1}{2}$ particles, and therefore can have total spin $\frac{1}{2}$ or $\frac{3}{2}$, which is in fact what the lightest baryons are. It is pretty easy to read off the quark content for the spin $\frac{3}{2}$ B^*'s from eq. (8.45); for example,

$$\left|\Delta^{++}\right\rangle = \left|uuu\right\rangle. \tag{9.4}$$

It is a little more complicated for the *B*'s, because the spins have to partially cancel, so we won't worry about them.

But this brings us to a problem. Looking at eq. (9.4), or indeed, any of the B^*'s from eq. (8.45), it is clear that if we look at just the flavor of the three quarks, it is completely symmetric. Furthermore, these are light spin $\frac{3}{2}$ baryons. The way you make a light spin $\frac{3}{2}$ baryon is to combine the three spins in a completely symmetric way. This tells us that all three quarks are in the same quantum state, which violates the Pauli Exclusion Principle. In fact, you can get a pretty good idea of which baryons will be lightest based on the assumption that quarks *only* combine in completely symmetric combinations. This created such a problem that many abandoned the quark model. Others, desperate to save it, invented a concept called color.

---

[1] Initially, these quarks were often named "truth" and "beauty." Remember, it was the 70's in California.

B. Color

How can we reconcile the fact that we have apparently three identical quarks in the same quantum state for the Δ^{++}, or indeed, in a completely symmetric state for any of the B^* baryons? The answer is that the quarks must not in fact be identical, there must be some hidden property which distinguishes them, which we will call “color.” Let’s assume there are in fact *three* different up quarks, which we will cleverly name u_1, u_2 and u_3. or sometimes, we’ll call them red, green, and blue up quarks. The state function can be made anti-symmetric by anti-symmetrizing the color variable, specifically

$$\left|\Delta^{++}\right\rangle = \tfrac{1}{\sqrt{6}}\left(\left|u_1u_2u_3\right\rangle + \left|u_2u_3u_1\right\rangle + \left|u_3u_1u_2\right\rangle - \left|u_1u_3u_2\right\rangle - \left|u_2u_1u_3\right\rangle - \left|u_3u_2u_1\right\rangle\right) = \tfrac{1}{\sqrt{6}}\varepsilon^{ijk}\left|u_iu_ju_k\right\rangle .$$

Of course, such an introduction of a new quantum number seems incredibly poorly motivated. But it is now believed that all quarks come in three colors. Hence there are really eighteen quarks, not six. All three up quarks have identical mass, charge, and so on.

The obvious question to ask is why the three up quarks (or three colors of any quark) should have identical properties, such as mass and charge. Recall that when we were first looking at strong interactions, we noticed that the proton and neutron had nearly identical masses and rather similar properties under strong interactions. We assumed there was an approximate symmetry of nature which related them, at least as far as strong interactions are concerned. We will make the same type of assumption here. Specifically, we will assume that there is an SU(3) color symmetry, called SU(3)$_C$ symmetry, that relates the three up quarks, so that we can define the eight generators T_a which act on the quarks according to

$$T_a\left|q_i\right\rangle = \left|q_j\right\rangle\left(T_a\right)^j{}_i\,, \tag{9.5}$$

where the T_a’s are given by eqs. (8.40). Anti-quarks will come in three anti-colors (called anti-red, anti-blue, and anti-green), and are labeled $\left|\bar{q}^i\right\rangle$. The color operators will transform them according to

$$T_a\left|\bar{q}^i\right\rangle = -\left(T_a\right)^i{}_j\left|\bar{q}^j\right\rangle . \tag{9.6}$$

With eqs. (9.5) and (9.6), you can work out how the T_a’s affect any combination of quarks and anti-quarks.

It cannot be emphasized too strongly that this SU(3)$_C$ is distinct from SU(3)$_F$ discussed in the last chapter. SU(3)$_F$ relates different types of quarks (up, down, and strange) whereas SU(3)$_C$ relates the different color quarks of one type. SU(3)$_C$ symmetry is also an exact symmetry. The three types of up quark have exactly the same mass and charge.

If the T_a's commute with the Hamiltonian, can we find various states that have the same mass? In particular, suppose we let any of the T_a's, say $T_{1\to 2}$, act on $\left|\Delta^{++}\right\rangle$. What do we get? It is easy to see that because of the anti-symmetric nature of the color properties of the $\left|\Delta^{++}\right\rangle$ you get zero. In other words, the $\left|\Delta^{++}\right\rangle$ is in the trivial representation of SU(3)$_C$, so that all of the T_a's vanish when they act on it. Hence the combination of three quarks is *colorless*, even though each of the quarks carry color, much as the combination of red, green, and blue light yields white light.

You may recall that I said in the last section that you can understand which of the baryons are light by assuming the quarks come in a completely symmetric wave function. Since they must actually be in an anti-symmetric wave function, it follows that the color part of the wave function must be completely anti-symmetric. In other words, the color part of any baryon is always the same, explicitly

$$\left|B\right\rangle = \tfrac{1}{\sqrt{6}}\varepsilon^{ijk}\left|{}_{ijk}\right\rangle . \tag{9.7}$$

In other words, baryons come only in the one-dimensional representation of SU(3). Therefore, color never relates the masses or interactions of any baryons. Of course, the same is true of the anti-baryons.

What about mesons? A meson consists of a quark and an anti-quark, so under color it has one index up and one down. The two particles are obviously distinct, and hence symmetry will not help us at all. It can be shown that if you combine a quark and an anti-quark, the resulting combination could result in either eight or one identical mass particles. In fact, the eight is not used. The combination of a quark and an anti-quark can be colorless. For example, a π^+ meson consists of an up and anti-down quarks, and the color combination is

$$\left|\pi^+\right\rangle = \tfrac{1}{\sqrt{3}}\left(\left|u_1\bar{d}^1\right\rangle + \left|u_2\bar{d}^2\right\rangle + \left|u_3\bar{d}^3\right\rangle\right). \tag{9.8}$$

Then if you act on $\left|\pi^+\right\rangle$ with any of the generators T_a, it can be shown (since T_a is traceless) that you get zero. This rule generalizes to any meson, so the color part of the meson state is

$$\left|M\right\rangle = \tfrac{1}{\sqrt{3}}\left|{}^{i}_{i}\right\rangle . \tag{9.9}$$

In other words, in both the case of mesons and baryons, the only combinations that occur are colorless. Indeed, if you treat this rule as a guiding principle (only colorless combinations occur), a natural consequence is that quarks combine only in the patterns three quarks, three anti-quarks, or quarks with anti-quarks. In other words, SU(3)$_C$ both solves the problem of getting the symmetry properties right and of allowing a simple rephrasing of the rules for combining quarks.

It may seem at the moment that we haven't accomplished very much with the introduction of SU(3)$_C$. We solved the problem of anti-symmetrizing fermion wave functions, but at the cost of tripling the number of quarks. We then "explained" the rules for combining quarks by demanding that they combine in colorless combinations. But the replacement of one ad-hoc assumption by another can hardly be considered progress.

C. Quantum Chromodynamics

What we would like to have is an explanation in terms of forces and so on why only colorless combinations occur. This will now be our goal. Let me illustrate what is going on by making an analogy with electromagnetism. Specifically, suppose that we had not yet discovered electromagnetic forces. Nonetheless, we found that when interactions of any type occur, a quantity called "charge" is conserved, and we go around labeling particles by what charge they have (-1 for the electron, etc.). The conservation of charge means that if we multiply all quantum states $\left|\Psi\right\rangle$ by a phase proportional to their charge, specifically

$$\left|\Psi\right\rangle \to e^{-iQ\theta}\left|\Psi\right\rangle ,$$

matrix elements of the Hamiltonian density $\mathcal{H}$ will remain unchanged.

Let me go further. Suppose that we start looking at the universe around us, and discover that it is made almost exclusively of neutral atoms, which have $Q\left|\Psi\right\rangle = 0$. At the moment this seems like just an arbitrary rule, but it is really a clue that charge is not just a conserved quantity, it actually corresponds to a force in nature (electromagnetism). Thus, by the observation that most objects are neutral, I conclude that there is some sort of electromagnetic force that energetically favors neutral combinations of particles (opposites attract).

I would now like to carry this over to the color force. We have eight generators T_a, each of which, like Q, commutes with $\mathcal{H}$. Looking around us, we find only combinations of particles for which $T_a\left|\Psi\right\rangle = 0$. Is it possible that these generators correspond to forces as well? It turns out that they do, and the force is called the color force, or sometimes the strong force. I will now attempt to develop a theory of the color force in strict analogy to quantum electrodynamics. This theory is called *quantum chromodynamics* (QCD).

The first step in developing QCD is to go all the way back to the Dirac equation, which was

$$\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi = 0. \tag{9.10}$$

To get the electromagnetic interactions, we replaced the derivative by a covariant derivative

$$\partial_{\mu} \to D_{\mu} = \partial_{\mu} + ieQA_{\mu}. \tag{9.11}$$

We now just want to copy this prescription. We will replace Q by T_a, because these are the generators of color, and we will rename the coupling e into g_s (s for strong), so that we don't confuse it with the electromagnetic coupling. Finally, instead of one electromagnetic field A_{μ}, there will be eight *gluon* fields A_{μ}^{a}, one for each generator T_a. The covariant derivative for QCD will be

$$\partial_{\mu} \to D_{\mu} = \partial_{\mu} + ig_sT_aA_{\mu}^{a}. \tag{9.12}$$

Thus the Dirac equation for a single quark interacting with a background gluon field would be

$$\left(i\not{D} - m\right)\Psi = 0,$$
$$\left(i\not{\partial} - g_sT_aA_{\mu}^{a}\gamma^{\mu} - m\right)\Psi = 0. \tag{9.13}$$

This equation takes a while to understand. For example, Ψ has twelve components: three (for each color) times four (the usual four components). The Dirac matrices act only on the four usual pieces of Ψ, while T_a acts only on color. If you are familiar with tensor notation, we should think of γ^{μ} as $\gamma^{\mu}\otimes 1$ and T_a as $1\otimes T_a$.

The covariant derivative results in a theory with gauge invariance, just as it did in QED. Recall that in the Dirac equation with electromagnetism, we could simultaneously change the phase of $\Psi\left(x\right)$ provided we also perform a gauge transformation on $A^{\mu}\left(x\right)$. Specifically, we needed to make the simultaneous transformation given by eqs. (3.56):

$$\Psi \to \Psi' = e^{-iQe\chi}\Psi \quad \text{and} \quad A_\mu \to A'_\mu = A_\mu + \partial_\mu \chi \, . \tag{9.14}$$

Now, to make the notation as clear as possible, define

$$U \equiv e^{-iQe\chi} \, . \tag{9.15}$$

Then we can rewrite eq. (9.14) in the form

$$\Psi \to \Psi' = U\Psi \quad \text{and} \quad QA_\mu \to QA'_\mu = QA_\mu + \frac{i}{e} U^* \partial_\mu U \, . \tag{9.16}$$

The number *U* is a number of magnitude 1, so $U^* U = 1$. We can think of *U* as a unitary matrix of size 1×1. For this reason, electromagnetism is described as a U(1)$_Q$ gauge theory, with the *Q* reminding us that the electron field couplings are proportional to the electric charge *Q*.

In a similar manner, quantum chromodynamics is an SU(3)$_C$ gauge theory. Let $U(x)$ be an arbitrary function of space-time whose values are SU(3)$_C$ matrices. It can be shown that the QCD Dirac equation is invariant under the transformation

$$\Psi \to \Psi' = U\Psi \quad \text{and} \quad T_a A^a_\mu \to T_a A'^a_\mu = T_a A^a_\mu + \frac{i}{g_s} U^\dagger \partial_\mu U \, . \tag{9.17}$$

Once you have the idea, you can build gauge theories for any group you want (technically, any Lie Algebra). But SU(3) is the one that describes strong interactions, so I will leave it at that.

We need to discuss the color field strengths, the equivalent of the electric and magnetic fields for the gluon fields. Recall that the electromagnetic field could be deduced from eq. (3.54), repeated here:

$$\left[D_\mu, D_\nu \right] \Psi = ieQF_{\mu\nu} \Psi \, .$$

By analogy, let's define the chromoelectric and chromomagnetic fields $G^a_{\mu\nu}$ by the relation

$$\left[D_\mu, D_\nu \right] \Psi = ig_s T_c G^c_{\mu\nu} \Psi \, . \tag{9.18}$$

We now write out the covariant derivatives using eq. (9.12). We have

$$\begin{aligned} D_\mu D_\nu \Psi &= \partial_\mu \partial_\nu \Psi + ig_s T_a A^a_\mu \partial_\nu \Psi + ig_s T_b \partial_\mu \left(A^b_\nu \Psi \right) - g_s^2 T_a T_b A^a_\mu A^b_\nu \Psi \\ &= \partial_\mu \partial_\nu \Psi + ig_s T_a A^a_\mu \partial_\nu \Psi + ig_s T_b A^b_\nu \partial_\mu \Psi + ig_s T_b \left(\partial_\mu A^b_\nu \right) \Psi - g_s^2 T_a T_b A^a_\mu A^b_\nu \Psi \; , \\ D_\nu D_\mu \Psi &= \partial_\mu \partial_\nu \Psi + ig_s T_a A^a_\mu \partial_\nu \Psi + ig_s T_b A^b_\nu \partial_\mu \Psi + ig_s T_a \left(\partial_\nu A^a_\mu \right) \Psi - g_s^2 T_b T_a A^a_\mu A^b_\nu \Psi \; , \\ \left[D_\mu, D_\nu \right] \Psi &= ig_s \left\{ T_b \left(\partial_\mu A^b_\nu \right) - T_a \left(\partial_\nu A^a_\mu \right) \right\} \Psi - g_s^2 \left[T_a, T_b \right] A^a_\mu A^b_\nu \Psi \; . \end{aligned} \tag{9.19}$$

We now use the commutation relations eq. (8.41) and change some of the dummy indices *a* and *b* to *c* to write this as

$$\left[D_\mu, D_\nu \right] \Psi = ig_s T_c \left(\partial_\mu A^c_\nu - \partial_\nu A^c_\mu \right) \Psi - ig_s^2 f_{abc} T_c A^a_\mu A^b_\nu \Psi \; . \tag{9.20}$$

Comparison with eq. (9.18) then tells us

$$G^c_{\mu\nu} = \partial_\mu A^c_\nu - \partial_\nu A^c_\mu - g_s f_{abc} A^a_\mu A^b_\nu \; . \tag{9.21}$$

Something remarkable happened here. Unlike the electromagnetic fields, the color fields are non-linear. This makes them inherently more difficult. The presence of nonlinear effects means there will be new interactions where gluon fields interact with themselves. It would be as if electric fields themselves had charge, and created more electric fields. It will make QCD more difficult to actually do computations in. But now that we have all the important interactions, it's time to start discussing Feynman diagrams for QCD.

D. Feynman rules for QCD

As in all previous theories, fermions (quarks) will be denoted by lines with arrows on them. But we don't just have one type of quark, we have six different types, and they come in three colors. In principle, every quark line should be labeled by what flavor and what color *i* the quark is, where *i* takes on three values. In fact, the colors of quark lines are rarely labeled, but they are understood. We also need a symbol for the gluon. To show it is different than the photon, and to signify that it is a stronger force, we'll use a spring-like symbol. Though there are not multiple "flavors" of quarks, they do come in eight colors, so a label *a* should in principle (but rarely in practice) specify which gluon we are dealing with. These two symbols are sketched in Fig. 9-4.

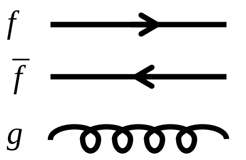


Figure 9-4: Symbols used in QCD.

Most of the Feynman rules for QCD are easy to understand. Since the quark is a fermion, it has the usual fermion propagator, and also the usual rules for external lines. The gluon is a spin-1 boson, massless like the photon (a mass term would violate gauge invariance), and therefore has the exact same propagator as the photon. The quark-gluon coupling is very similar to the fermion-photon coupling, except that $e \to g_s$ and Q is replaced by T_a. The component of this matrix will depend on the colors of the gluon and the incoming and outgoing quarks. It should be recognized that the gluon-quark coupling never changes the flavor of quarks, but it does change their color. There are rules for dealing with external fermion lines and external gluon lines, just like in QED.

The surprising new interactions are the gluon self-couplings, which come in a three- and four-gluon vertex.

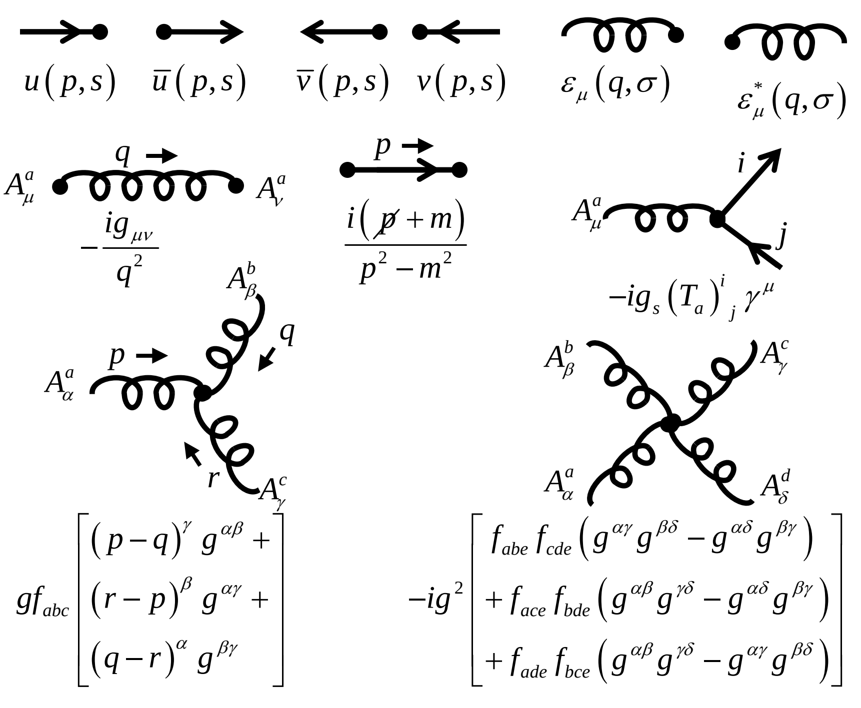


Figure 9-5: The Feynman rules for QCD.

The expressions are very complex, so we won't be working with them much. You're welcome. The full QCD Feynman diagram rules are listed in Fig. 9-5.

Now that we have Feynman rules, we are prepared to do calculations. But wait! We never experimentally have free quarks and gluons in the initial or final state, so it seems that our rules are useless. How can I justify doing any QCD calculation if the particles I'm considering don't ever occur alone?

Perhaps an analogy will help. Suppose I collide two automobiles. The way to think of the collision depends on what energy you are colliding at. At 10 m/sec, it is two automobiles. At 10^3 m/sec, it is essentially a collection of largely disconnected automobile parts. The welds and so on that hold the pieces together at such an energy are essentially irrelevant. By the time you reach 10^4 m/sec, it is a collection of atoms; chemical bonds are all broken. At 10^7 m/sec, it is electrons and nuclei. At 10^8 m/sec it is protons, neutrons, and electrons. And at 2.9×10^8 m/sec, it is electrons and quarks.

Having roughly justified why and when it makes sense to perform a QCD perturbative calculation, let me do one. Suppose I have a proton colliding with an anti-proton. Each of these particles has various things inside, including up and anti-up quarks. Let's imagine we know the probability of finding an up quark inside one and an anti-up quark inside the other. What is the cross-section for, say, producing a charm quark and a charm anti-quark?

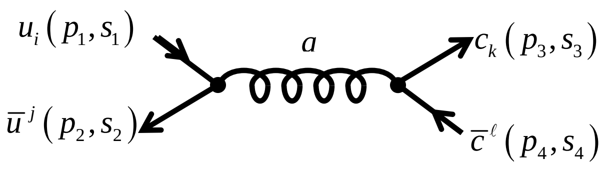


Figure 9-6: The Feynman diagram for the process $u\bar{u}\to c\bar{c}$.

The relevant Feynman diagram is sketched in Fig. 9-6. There is only one diagram, though it corresponds to eight different possible gluons that contribute to the propagator. All the quarks and the gluon have been labeled by their corresponding color. The Feynman amplitude is given by

$$i\mathcal{M}=\left(-ig_s\right)^2\left(T_a\right)^j{}_i\left(T_a\right)^k{}_\ell\left(\bar{v}_2\gamma^\nu u_1\right)\left(\bar{u}_3\gamma^\mu v_4\right)\frac{-ig_{\mu\nu}}{\left(p_1+p_2\right)^2}=\frac{ig_s^2}{s}\left(T_a\right)^j{}_i\left(T_a\right)^k{}_\ell\left(\bar{v}_2\gamma^\mu u_1\right)\left(\bar{u}_3\gamma_\mu v_4\right). \tag{9.22}$$

Note that we used the same subscript *a* on each of the T_a's, because it's the same gluon connecting the two vertices. Note that *a*, since it is repeated, is summed over, $a=1,\ldots,8$. Except for the change $e^2\to g_s^2$ and the two factors of T_a, this expression is identical to what we encountered in the section on QED, eq. (7.17). Thus if we are working well above the charm mass, the cross-section can be written without further work as

$$\sigma_{ijk\ell}=\frac{4\pi\alpha_s^2}{3s}\left(T_a\right)^j{}_i\left(T_a\right)^k{}_\ell\left[\left(T_b\right)^j{}_i\right]^*\left[\left(T_b\right)^k{}_\ell\right]^*,\qquad\text{(no sum on } i,j,k,\ell) \tag{9.23}$$

where, in analogy with electromagnetism, we define

$$\alpha_s\equiv\frac{g_s^2}{4\pi}. \tag{9.24}$$

Taking advantage of the fact that the T_a's are Hermitian, we can rewrite eq. (9.23) in the slightly more palatable form

$$\sigma_{ijk\ell} = \frac{4\pi\alpha_s^2}{3s} \left(T_a\right)^j{}_i \left(T_b\right)^i{}_j \left(T_a\right)^k{}_\ell \left(T_b\right)^\ell{}_k . \qquad \text{(no sum on } i, j, k, \ell\text{)} \tag{9.25}$$

This would be the cross-section if the quark/anti-quark pair were coming in with a known color. In fact, the quark and anti-quark are each contained inside a proton and an anti-proton. We have already said that the color of the quarks inside a proton are completely anti-symmetric, and therefore they each have an equal probability of having any of three colors. We should therefore average over initial colors. Similarly, if we only ask how many total charm quarks we produce, not how many of each color, we should sum over all possible values. Thus we replace eq. (9.25) with

$$\sigma = \tfrac{1}{9} \sum_{i,j,k,\ell} \sigma_{ijk\ell} = \frac{4\pi\alpha_s^2}{27s} \left(T_a\right)^j{}_i \left(T_b\right)^i{}_j \left(T_a\right)^k{}_\ell \left(T_b\right)^\ell{}_k = \frac{4\pi\alpha_s^2}{27s} \mathrm{Tr}\left(T_a T_b\right) \mathrm{Tr}\left(T_a T_b\right) . \tag{9.26}$$

The traces can be worked out from eq. (8.38) and then the sums over the eight gluon colors yield

$$\sigma = \frac{4\pi\alpha_s^2}{27s} \frac{1}{2}\delta_{ab} \frac{1}{2}\delta_{ab} = \frac{8\pi\alpha_s^2}{27s} . \tag{9.27}$$

In a similar manner, we can often take over QED calculations when we need to do QCD calculations. There remains the problem of how you figure out the probability of finding a quark of a given type inside a proton, and how you know what its momentum is, but this is beyond the scope of this class.

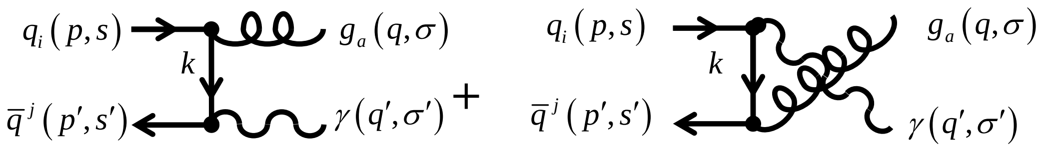


Figure 9-7: The Feynman diagram for the process $q\bar q \to g\gamma$.

One comment is worth making comparing and contrasting QED and QCD. In QED, a particle's identity never changes, not even its color. Effectively, this means that the QED coupling for quarks should involve a factor something like $-ieQ\delta^i{}_j$, because the color can't change. For example, suppose we were calculating the cross-section for quark-antiquark annihilation to photon plus gluon. There would be two diagrams, as sketched in Fig. 9-7. The resulting Feynman amplitude would be

$$i\mathcal{M} = \left(-ieQ\right)\left(-ieg_s\right)\varepsilon_\mu^*\varepsilon_\nu'^* \left[\delta^j{}_k \left(T_a\right)^k{}_i \bar v' \gamma^\nu \frac{i\left(\not p - \not q - m\right)}{\left(p-q\right)^2 - m^2} \gamma^\mu u + \left(T_a\right)^j{}_k \delta^k{}_i \bar v' \gamma^\mu \frac{i\left(\not q - \not p' - m\right)}{\left(q-p'\right)^2 - m^2} \gamma^\nu u \right],$$

$$i\mathcal{M} = -ieQg_s \left(T_a\right)^j{}_i \varepsilon_\mu^*\varepsilon_\nu'^* \left[\bar v' \gamma^\nu \frac{i\left(\not p - \not q - m\right)}{\left(p-q\right)^2 - m^2} \gamma^\mu u + \bar v' \gamma^\mu \frac{i\left(\not q - \not p' - m\right)}{\left(q-p'\right)^2 - m^2} \gamma^\nu u \right] . \tag{9.28}$$

We see this is nearly identical with the electron-positron annihilation we calculated in section 7D. By comparing this amplitude to eq. (7.27) we can jump immediately to the differential cross-section eq. (7.34), with a couple of minor adjustments, and we find

$$\frac{d\sigma_{ija}}{d\Omega} = \frac{\alpha\alpha_s Q^2}{4E^2} \left| \left(T_a\right)^j{}_i \right|^2 \frac{1+\cos^2\theta}{1-\cos^2\theta} . \tag{9.29}$$

As usual, we actually want to sum over the outgoing colors *a*, and average over the incoming colors *i* and *j* to give us

$$\frac{d\sigma}{d\Omega} = \frac{1}{9}\sum_{ija}\frac{d\sigma_{ija}}{d\Omega} = \frac{\alpha\alpha_s Q^2}{36E^2}\left(T_a\right)^j{}_i\left(T_a\right)^i{}_j\frac{1+\cos^2\theta}{1-\cos^2\theta} = \frac{\alpha\alpha_s Q^2}{36E^2}\sum_a \mathrm{Tr}\left(T_a^2\right)\frac{1+\cos^2\theta}{1-\cos^2\theta}$$

$$= \frac{\alpha\alpha_s Q^2}{36E^2}\frac{8}{2}\frac{1+\cos^2\theta}{1-\cos^2\theta} = \frac{\alpha\alpha_s Q^2}{9E^2}\frac{1+\cos^2\theta}{1-\cos^2\theta} = \frac{4\alpha\alpha_s Q^2}{9s}\frac{1+\cos^2\theta}{1-\cos^2\theta} .$$

The total cross-section is apparently infinite, but this infinity is caused by a combination of invalid assumptions, such as treating the quarks as massless (which they aren't quite) and treating them as free particles before and after the collision (which they definitely aren't). I won't be worrying about these details here.

E. Loops in QCD

It is a bit premature to jump into loops in QCD – but we should probably do it, since it gives us some insight. You may have noticed by now that I have not told you what the value of the strong coupling constant, or perhaps more relevant, α_s is. The reason is that, like α, its value runs, and changes with scale. At the *Z*-mass, a common reference point, its value is about

$$\alpha_s\left(M_Z^2\right) = 0.1184 \pm 0.0007 . \tag{9.30}$$

Right off the bat, this is telling you the strong coupling is some fifteen times stronger at this scale. Still, perturbation theory should still have some validity, since loop effects will be smaller by about a factor of $\alpha_s/4\pi$, or about the 1 percent level. Actually, because there are so many possible color combinations, the effects are enhanced a bit, and it is really the few percent level.

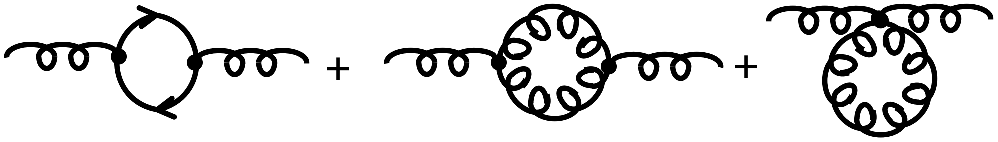


Figure 9-8: Three of the loop diagrams contributing to the running of the QCD coupling constant.

But remember that loops also cause the coupling constant to effectively be a function of energy. It turns out in this case there are three loops that are relevant, one for any colored quark going in a loop, and two others for loops of gluons, as sketched in Fig. 9-8.[1] What is perhaps a little surprising is that the gluon loops have the opposite effects of fermion loops, and they dominate the result, so instead of the coupling getting stronger as you go up in energy, instead it

---

[1] Actually, it's a little worse than this, because there are extra rules for loops in QCD. You don't want to know, and I'd have to look it up, because it's pretty messy.

gets stronger as you go down in energy, and weaker as you go up. The fact that the force gets weaker as you increase in energy is called *asymptotic freedom*, since at sufficiently high energies the particles become more and more like free particles (and perturbation theory becomes increasingly reliable). Nonetheless, for reasonably accurate results, computations need to include loops. Computations of QCD processes for actual comparison with experiment are typically carried out to the two or three loop level. Even then, they are only used at high energies.

What makes things truly messy is low energies, where QCD exhibits *infrared slavery*. As you go down in energy, the strong coupling gets stronger, and by the time you are at 1 GeV it is around $\alpha_s = 0.5$. At this point perturbation theory becomes almost completely worthless.

There *are* techniques for doing QCD calculations that do not rely on perturbation theory. Lattice QCD performs computations by dividing space and time into a discrete set of lattice points. Unfortunately, the amount of computation grows as an exponential function of the number of points, so even modest computations require vast computer resources. Typically, computations require as input the strength of the strong coupling constant together with all the quark masses, though only the up, down, and strange are needed for most low-energy computations. Hence some subset of observed quantities (say, the proton, neutron, kaon, and pion masses) are used to determine the masses and coupling, and then all other properties and masses are computed. Such results can often be obtained to within a few percent or better. But it *is* very computationally intense, and doesn't necessarily provide insight into what is going on. Instead of relying on this, I will attempt to provide some qualitative insight into QCD.

F. Quark Confinement

I have worked hard to make an analogy between QED and QCD, but clearly the analogy is not perfect. In QED, it is true that neutral atoms are favored over ions, but it is not true that ions and electrons can never occur. In contrast, free quarks, or any colored combination of quarks, have never been seen. What is different about QCD? Most obvious is the much stronger coupling constant, but this only makes it more difficult. What is probably even more important is the fact that the gluon fields *themselves* have color charge. Let's try to understand how this might lead to confinement.

Consider first positronium, an electrically bound state of an electron and a positron, as sketched in Fig. 9-9. When they are close together, there is a relatively strong force between them, because the electric field lines that bind them are close together, meaning they are strong and have a lot of energy associated with them. But as they separate, the electric field spreads out, until eventually the electric field between them is so weak the force effectively

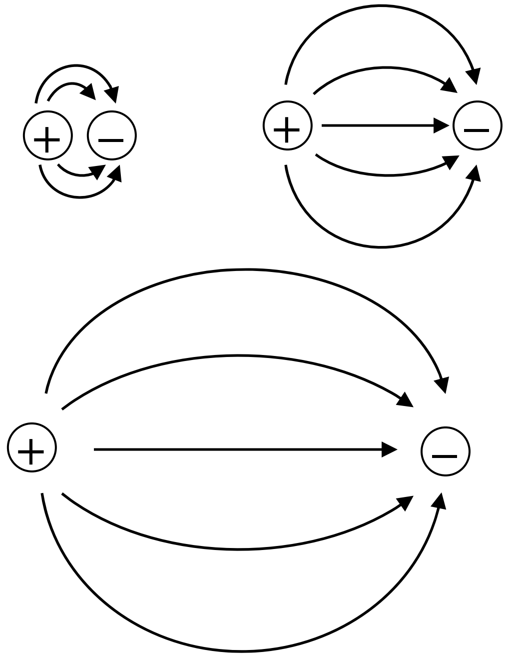


Figure 9-9: As you separate an electron and positron, the electric fields spread out, until eventually the force is very small, and they are essentially isolated particles.

vanishes. Hence the force drops off with distance, and eventually the electron and positron are no longer confined.

In contrast, consider a meson, a bound state of a quark and anti-quark, as sketched in Fig. 9-10. When they are close together, it isn't that different from positronium, though the force is stronger. But as you separate them, the gluon field lines between them attract each other, and therefore instead of spreading out, they end up bound together in a *flux tube*. Because the density of the fields remains constant, the force required for further separation never diminishes, and therefore the energy is proportional to the separation. Eventually, the energy becomes so great that it becomes energetically favorable to produce a quark-antiquark pair, which allows the flux tube to snap back and create two mesons, as shown in Fig. 9-10. Hence it is impossible to produce an isolated quark or anti-quark.

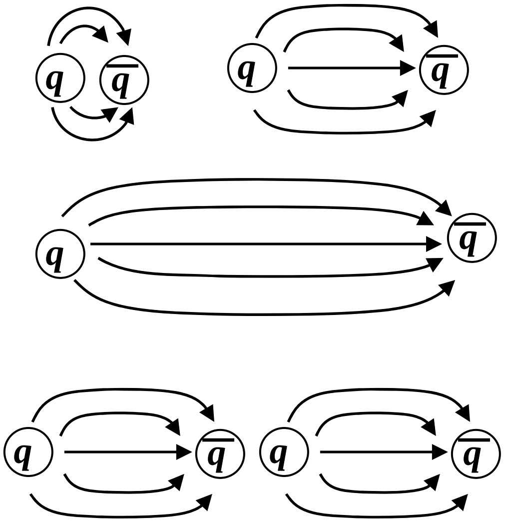


Figure 9-10: A meson (upper left) starts to separate (upper right), but the quark-antiquark pair is connected by a flux tube (middle) that eventually produces a new quark/antiquark pair that makes two mesons (bottom).

Given the theory of QCD, what can we say about baryons and mesons? In broad terms, it is possible to understand why quarks combine only into threes or singly with anti-quarks in terms of opposites attract. Indeed, for some of the heavy quarks, such as the charm or bottom quarks, you can make rather accurate predictions about the bound states of quarks and anti-quarks by making an analogy with positronium. The reason this works for heavy quarks is that large mass implies high energy, so that strong interactions are not so terribly strong, and the formulas of electromagnetism can work pretty well.

What can we say about bound states of light quarks? For example, a proton consists of two up quarks and a down quark. Is its mass close to the sum of these masses? No, it is much bigger. The reason is rather complex, but roughly speaking, the effective mass of a quark has two contributions, one coming from its true mass, and the other coming from the gluon field lines it is attached to. The sum of these is called the *constituent quark mass,* and for the three light quarks (up, down, strange), the main contribution comes from the gluon fields. All the quarks have the same color couplings, so the color field part of the mass is about the same for them. Because the up, down, and strange quark are all very light compared to this gluon field part of their mass, their constituent quark masses are all about the same. If this mass were exactly the same, strong interactions would not be able to distinguish between the three quarks, and $SU(3)_F$ would be an exact symmetry of strong interactions. Because the up and down quarks are much lighter than the strange, isospin symmetry is a much better approximation than $SU(3)_F$ symmetry.

Now, let's talk a little about hadron colliders like the LHC. First of all, pretending a proton is a combination of two up and one down quark is pretty naïve. A large part of the energy, perhaps even a majority, is in the gluon fields. Furthermore, these gluons are strongly coupled to quarks, and therefore they are constantly making quark-antiquark pairs. Hence a proton is a horribly complex combination of quarks, anti-quarks, and gluons. The two up and one down quark are called the *valence* quarks, and about the best we can say is that there are two more up quarks than anti-up quarks in a proton, and one more down quark than anti-down quarks. At the extreme energies of the Large Hadron Collider, in fact, the non-valence quarks and gluons are about as likely to be involved in an interaction as one of the valence quarks. Hence it doesn't matter much whether you collide protons with anti-protons, or with protons, since any light quark or anti-quark or gluon you want will be in there somewhere. Protons are easier to produce in large quantities, so the LHC opted to do proton-proton collisions.

Now let's say you collide two protons, and one strongly interacting particle from one collides with another with the other to produce something new and interesting – maybe a bottom anti-bottom quark pair. The initial protons are now missing a colored particle of some sort, which means there is a flux tube connecting it with the interaction point. This flux tube will shatter to produce one or maybe several quark pairs, at which point the remnants tend to be moving along the same direction as the initial proton, which means most of this crud goes down the beam pipe and is lost. The two bottom quarks will head out in some other directions, also with flux tubes connecting them to the interaction point. This will similarly shatter, and a significant number of hadrons will come out in two directions, one each in the direction of each bottom quark. Many of these particles will themselves be unstable strong particles, which will decay producing even more particles. In the end, there will be two *jets* of particles, moving more or less in the direction of the original *b* quarks. So it tends to be jets we look for when we produce strongly interacting particles. Because strong interactions are strong, we tend to get a *lot* of jets, and most of them are caused by strong interactions that we understand. This makes hadron collider data difficult to analyze. Still, proton colliders reach higher energies than can be attained other ways, so it is the method of choice for exploring the highest energy collisions.

Problems for Chapter 9

1. Based on the charge and strangeness of the octet of lightest baryons, deduce what quarks are inside them (don't worry about symmetrizing, just write the answer as $|p\rangle = |uud\rangle$). It is acceptable to repeat.

2. Although the 1, 8, and 10 dimensional representations of SU(3)$_F$ are all that are used for *light* quarks, there are some objects containing heavy quarks that are different. The Λ_c particle is a heavy spin $\frac{1}{2}$ baryon with quark content $|\Lambda_c\rangle = \frac{1}{\sqrt{2}}\left(|cud\rangle - |cdu\rangle\right)$. The charmed quark is unaffected by SU(3)$_F$, so $T_a|c\rangle = 0$ for all *a*.
   (a) Suppose all the operators $T_{i\to j}$ are allowed to act repeatedly on $|\Lambda_c\rangle$. List all of the states that you can get. Make up names for them.
   (b) Identify the T_3 and T_8 values for all the states you found in part (a) (including $|\Lambda_c\rangle$). Make a plot of T_3 vs. T_8 for these states.

3. Draw all tree level Feynman diagrams for gluon-gluon scattering, $gg \to gg$. Don't write the amplitude, unless you enjoy headaches.

4. Draw all thirteen tree-level Feynman diagrams for the process $q\bar{q} \to ggg$, where q is any quark, and g is a gluon. You don't have to do anything with the diagrams.

5. Draw all three Feynman diagrams for the scattering of a gluon off an up quark, $u(p)g(q) \to u(p')g(q')$. Write the Feynman invariant amplitude for this process. You may treat all particles as massless.

6. An up quark is scattering off of a down quark, $u(p_1)d(p_2) \to u(p_3)d(p_4)$, due to strong interactions. Treat the quarks as massless, and assume the initial quarks are of random spin and color. Find the differential cross section, as well as the total cross-section assuming there is a minimum angle $\theta_{\min}$ for which the scattering is distinguishable from not scattering.

7. Find the cross section for $u\bar{u} \to b\bar{b}$ for b quarks, including only QCD contributions, but don't ignore the b quark mass. Eq. (9.22) is a good starting point.

8. In eq. (9.27), we found the cross section for $u\bar{u} \to c\bar{c}$, using only the tree-level QCD contribution. Add in the QED contribution to the amplitude. Convince yourself (and me) that the cross terms $(i\mathcal{M}_{QCD})(i\mathcal{M}_{QED})^*$ vanishes. Find the total cross-section.

9. Calculate the differential cross-section for $u(p)\gamma(q) \to u(p')g(q')$ in the center of mass frame. Average over all initial spins and colors, and sum over all final spins and colors. Treat all particles as massless.

10. Write down the Feynman amplitude for the process $u\bar{u} \to gg$. Make sure you include all three Feynman diagrams. You don't have to do anything else with it.

X. Weak Interactions

We now move on to the one remaining type of interactions – weak interactions. Weak interactions were first discovered in nuclear β-decay, and so our story starts there. It was realized very early on that there must be some sort of invisible particle also produced, which we call the neutrino, so let's start with a little bit of discussion of neutrinos.

A. Neutrinos

Beta decay is the process whereby a nucleus increases its charge *Z* by one while leaving its mass number *A* unchanged, and in the process emits an electron. We now understand that this implies that a neutron is changing to a proton by the emission of an electron, $n \to p^+ e^-$. This conserves electric charge, but there are numerous problems. First, there is one fermion on the left, and two on the right, so it is impossible to conserve angular momentum. It was also noted that the electron did not have a definite energy, but rather came out with a range of energies. Conservation of momentum was not a problem, since the proton (imbedded in a nucleus) could absorb this momentum. Wolfgang Pauli suggested in 1930 that it might be due to some sort of additional particle, a neutral fermion. Initially the particle was named the neutron, but the particle we now call the neutron was discovered in 1932, so it was renamed the *neutrino*, and we now write this decay as

$$n \to p^+ e^- \bar{\nu}_e \,. \tag{10.1}$$

In this case, the particle produced is actually an anti-neutrino, or more precisely, an electron anti-neutrino.[1] By studying the spectrum of electrons coming out, it is possible to put a limit on the mass of the neutrino. Experiments along these lines continue today, and indicated the electron neutrino is lighter than about 1 eV.

From the start it was recognized that neutrinos would interact very weakly with matter. Neutrinos of ordinary energies (a few MeV) can easily pass through millions of kilometers of normal matter and still have very little chance of interacting. At first it was thought this would prevent direct detection of the neutrino, but in 1942 the process given by eq. (10.1) was essentially reversed, and neutrinos were detected by the process

$$\bar{\nu}_e p^+ \to n e^+ \,. \tag{10.2}$$

Later experiments demonstrated that the neutrino and anti-neutrino were distinct particles – but more on this in chapter 12.

The next important advance in weak interactions came with the discovery of the muon, a sort of heavier electron. It decays almost all the time into an electron, $\mu^- \to e^-$, but the decay clearly conserved neither energy nor momentum. Indeed, there is a range of energies for the

---

[1] Throughout this chapter I will assume the neutrino is massless, and will use the Standard Model characterization of the neutrinos as electron neutrino, muon neutrino, and tau neutrino. Now that we know neutrinos have mass, the situation is much more confusing.

resulting electron, which implies not a single invisible particle, but two invisible particles, and we now would describe this process as

$$\mu^- \to e^- \nu_\mu \bar{\nu}_e \, . \tag{10.3}$$

Shortly thereafter the pions were discovered. The charged pions decay almost always to a muon,

$$\pi^- \to \mu^- \bar{\nu}_\mu \quad \text{and} \quad \pi^+ \to \mu^+ \nu_\mu \, . \tag{10.4}$$

Pions are easily produced by smashing protons into fixed targets. As the protons collide with the nuclei, they produce a host of strongly interacting particles, most of which decay quickly into pions. If the pions are allowed to decay (which takes about 26 ns on average), and then you put some sort of shielding material that stops ordinary particles (like a lot of rock), you will have a powerful beam of almost pure muon neutrinos and anti-neutrinos. These muon neutrinos can then be discovered via the process

$$\nu_\mu n \to p^+ \mu^- \quad \text{and} \quad \bar{\nu}_\mu p^+ \to n \mu^+ \, . \tag{10.5}$$

The fact that this process created muons, not electrons, indicated that the two neutrinos, ν_μ and ν_e, were distinct particles.

We are starting to see a pattern here. Each of the two charged leptons (electron and muon) seem to have associated with them a neutrino. In a weak interaction, a charged lepton gets changed into a neutrino, or vice versa, or some sort of equivalent process. This changes the charge by one unit. Meanwhile, some other particle changes its charge by one unit the other way; for example, a proton changes into a neutron or vice versa. Physicists began talking about weak interactions in terms of *weak charged current* interactions.

Now, when talking about protons, neutrons, or pions, the situation is a little confusing. At this point quarks hadn't even been discovered, so it is not surprising they had some difficulty getting that to work out, and it will take us a while to figure out what is going on. Let us focus exclusively on the lepton sector; specifically the muon decay, $\mu^- \to e^- \nu_\mu \bar{\nu}_e$.

B. Fermi Theory and Parity Violation

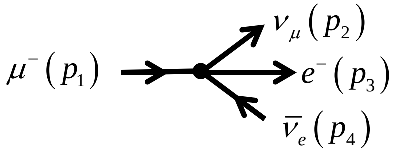


Figure 10-1: The Feynman diagram for muon decay in the Fermi theory.

The first attempt to explain weak decays was proposed by Enrico Fermi. Though the language was different, and he was focusing on weak interactions, for muon decay his proposal would amount to describing the decay in terms of an interaction sketched in Fig. 10-1, a fundamental interaction between four fermions. It clearly made sense to put the electron with the electron neutrino, and the muon with the muon neutrino, and his theory effectively assumed the interaction gave rise to a Feynman amplitude

$$i\mathcal{M} = -i\sqrt{2}G_F \left(\bar{u}_2 \gamma^\mu u_1\right)\left(\bar{u}_3 \gamma_\mu v_4\right), \tag{10.6}$$

where G_F is a new fundamental constant called *Fermi's constant* whose value is

$$G_F = 1.16637 \times 10^{-5} \text{ GeV}^{-2} \, . \tag{10.7}$$

Exactly how that factor of $\sqrt{2}$ crept into eq. (10.6) is unclear to me. The theory gave qualitatively correct behaviors; for example, it predicted that weak interactions would increase as you increased the energy, but there were a lot of details of the spectrum that didn't work out. Most physicists (correctly) assumed the error was that Fermi had assumed the weak coupling in eq. (10.6) was of the vector-vector type, which we might call $i\mathcal{M}_{VV}$; that is, both the matrices sandwiched between the Dirac spinors in eq. (10.6) were vectors. Other interactions were tried (see problem 3.3 to understand how they are categorized), such as

$$i\mathcal{M}_{AA} \sim \left(\bar{u}_2\gamma^\mu\gamma_5 u_1\right)\left(\bar{u}_3\gamma_\mu\gamma_5 v_4\right), \quad i\mathcal{M}_{SS} \sim \left(\bar{u}_2 u_1\right)\left(\bar{u}_3 v_4\right), \quad i\mathcal{M}_{PP} \sim \left(\bar{u}_2\gamma_5 u_1\right)\left(\bar{u}_3\gamma_5 v_4\right).$$

There is also a tensor-tensor term $i\mathcal{M}_{TT}$ which I won't bother you with. These five possibilities were the only expression that respected Lorentz invariance, including parity. Other possibilities, like $i\mathcal{M}_{VA}$, respected proper Lorentz transformations but not parity, and hence were ignored. Various conclusions were drawn about the interactions, then discarded when they didn't fit one spectrum or another. No one could figure out the right combination.

At the same time, new particles were being discovered. One pair of mesons was particularly troubling, at the time called the τ^+ and θ^+ (and their anti-particles).[1] The problem was that these particles had exactly the same mass, lifetime, production mechanisms and so on. The τ^+ decayed to three pions and the θ^+ to two. Since it was already known that the pions were pseudoscalars, this proved that the τ^+ was a pseudoscalar and the θ^+ a scalar, and hence they were different particles. But why were they so similar?

Eventually it occurred to Lee and Yang that they *could* be the same particle – all that had to be assumed was that weak interactions didn't respect parity. Experiments were done on nuclear beta decay – and suddenly it was realized that parity was violated, in fact, maximally violated. In the absence of parity violation, neutrinos should be produced equally likely to have left or right helicity. But they are apparently always produced with left helicity. We now call the single particle the K^+, and we have discarded both the names τ^+ and θ^+.

With this discovery, it was quickly realized that other possible interactions existed. This led to the correct form of the interaction, namely

$$i\mathcal{M} = -i\tfrac{1}{\sqrt{2}}G_F\left[\bar{u}_2\gamma^\mu\left(1-\gamma_5\right)u_1\right]\left[\bar{u}_3\gamma_\mu\left(1-\gamma_5\right)v_4\right]. \tag{10.8}$$

Because each piece is the difference between the vector and axial-vector couplings, we now say that weak interactions are $V - A$ interactions.

Let's calculate the total rate of muon decay from eq. (10.8). The complex conjugate of eq. (10.8) is

$$\left(i\mathcal{M}\right)^* = i\tfrac{1}{\sqrt{2}}G_F\left[\bar{u}_1\left(1+\gamma_5\right)\gamma^\nu u_2\right]\left[\bar{v}_4\left(1+\gamma_5\right)\gamma_\nu u_3\right]. \tag{10.9}$$

To get eq. (10.9) right, we have to remember that not only do we reverse the order of factors when we bar them, we have to remember that $\bar{\gamma}_5 = -\gamma_5$. Also note that the Lorentz index μ got changed to ν, so as to avoid confusion when we take the product, which is

$$\left|i\mathcal{M}\right|^2 = \tfrac{1}{2}G_F^2\left[\bar{u}_2\gamma^\mu\left(1-\gamma_5\right)u_1\bar{u}_1\left(1+\gamma_5\right)\gamma^\nu u_2\right]\left[\bar{u}_3\gamma_\mu\left(1-\gamma_5\right)v_4\bar{v}_4\left(1+\gamma_5\right)\gamma_\nu u_3\right]. \tag{10.11}$$

---

[1] This has nothing to do with the τ-lepton, which was discovered later.

We now sum over all final spins and average over initial spins. The neutrinos are effectively massless, and the electron is 200 times lighter than the muon, so we will treat it as massless as well.

$$\tfrac{1}{2}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 = \tfrac{1}{4}G_F^2\,\mathrm{Tr}\left[\left(\not{p}_2+m\right)\gamma^\mu\left(1-\gamma_5\right)\not{p}_1\left(1+\gamma_5\right)\gamma^\nu\right]\mathrm{Tr}\left[\not{p}_3\gamma_\mu\left(1-\gamma_5\right)\not{p}_4\left(1+\gamma_5\right)\gamma_\nu\right]. \tag{10.12}$$

Now, we can use the fact that an odd number of Dirac matrices (γ_5 doesn't count) always has a vanishing trace to get rid of the *m* term. Furthermore, we can use the fact that γ_5 anti-commutes with all the Dirac matrices and that $\gamma_5^2 = 1$ to show that

$$\left(1-\gamma_5\right)\not{p}\left(1+\gamma_5\right) = \left(1-\gamma_5\right)^2\not{p} = 2\left(1-\gamma_5\right)\not{p}\,. \tag{10.13}$$

Substituting this into eq. (10.12), and working the γ_5's to the left, we find

$$\begin{aligned}\tfrac{1}{2}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 &= G_F^2\,\mathrm{Tr}\left[\left(1-\gamma_5\right)\not{p}_2\gamma^\mu\not{p}_1\gamma^\nu\right]\mathrm{Tr}\left[\left(1-\gamma_5\right)\not{p}_3\gamma_\mu\not{p}_4\gamma_\nu\right]\\ &= 16G_F^2\left[p_2^\mu p_1^\nu + p_1^\mu p_2^\nu - g^{\mu\nu}\left(p_1\cdot p_2\right) - i\varepsilon^{\alpha\mu\beta\nu}p_{2\alpha}p_{1\beta}\right]\times\\ &\quad\times\left[p_{3\mu}p_{4\nu} + p_{4\mu}p_{3\nu} - g_{\mu\nu}\left(p_3\cdot p_4\right) - i\varepsilon^{\gamma}{}_{\mu}{}^{\delta}{}_{\nu}\,p_{3\gamma}p_{4\delta}\right].\end{aligned}$$

At this point you pretty much have a nightmare, but it is not hard to see that because the ε terms are anti-symmetric on the indices μ and ν, and the other terms are symmetric, the cross terms go away, and this simplifies to

$$\tfrac{1}{2}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 = 16G_F^2\left[2\left(p_2\cdot p_3\right)\left(p_1\cdot p_4\right) + 2\left(p_2\cdot p_4\right)\left(p_1\cdot p_3\right) - \varepsilon^{\alpha\mu\beta\nu}p_{2\alpha}p_{1\beta}\varepsilon^{\gamma}{}_{\mu}{}^{\delta}{}_{\nu}\,p_{3\gamma}p_{4\delta}\right]. \tag{10.14}$$

The last term can be simplified with the fact that the Levi-Civita symbol is completely anti-symmetric, together with eq. (2.17a), to show that

$$\begin{aligned}\varepsilon^{\alpha\mu\beta\nu}p_{2\alpha}p_{1\beta}\varepsilon^{\gamma}{}_{\mu}{}^{\delta}{}_{\nu}\,p_{3\gamma}p_{4\delta} &= \varepsilon^{\alpha\beta\mu\nu}\varepsilon^{\gamma\delta}{}_{\mu\nu}\,p_{2\alpha}p_{1\beta}p_{3\gamma}p_{4\delta} = 2\left(-g^{\alpha\gamma}g^{\beta\delta} + g^{\alpha\delta}g^{\beta\gamma}\right)p_{2\alpha}p_{1\beta}p_{3\gamma}p_{4\delta}\\ &= -2\left(p_1\cdot p_4\right)\left(p_2\cdot p_3\right) + 2\left(p_1\cdot p_3\right)\left(p_2\cdot p_4\right).\end{aligned}$$

Substituting this into eq. (10.14), we have

$$\tfrac{1}{2}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 = 64G_F^2\left(p_2\cdot p_3\right)\left(p_1\cdot p_4\right).$$

From this can be computed the decay rate, which you did in problem 4.11, to find

$$\Gamma = \tfrac{1}{192\pi}G_F^2 m_\mu^5\,. \tag{10.15}$$

Note the factor of m_μ^5 in this result. If we had a much heavier particle, the decay would go a lot faster. This suggests that weak interactions are in fact weak only at low energies. In fact, by the time you reach 100 GeV in energy, they are as strong or stronger than electromagnetic interactions.

C. The *W*-boson and its leptonic couplings

Even from the start, it was understood that it was likely that there was some sort of particle mediating the weak interactions. Note that the Fermi coupling has dimensions of mass to the minus two, and hence the weak interaction is non-renormalizable. In fact, if you calculate cross-sections using the Fermi theory, the cross-sections rise with energy. If you think of cross-sections as probability, the probability for scattering rises and eventually passes one. This should be impossible.

This all suggests that the Fermi theory is simply an effective theory. All we have to do is imagine that there is some sort of heavy particle, which we'll call *W* for weak, that acts as an intermediary between the fermions. Fig. 10-2 sketches the appropriate Feynman diagram in this case, with the wavy line representing the *W*- particle. We have labeled the line to avoid confusion with the photon. By the form of the coupling in eq. (10.8), it must be a spin 1 or vector particle. It is easy to see from the diagram that it must carry a charge of -1, but it has an anti-particle as well, so there are two particles $W^\pm$.[1] If we assume the *W* is a fairly heavy particle, then the propagator will probably have a factor of M_W^2 in the denominator, suppressing the amplitude. Hence the smallness of weak interactions is just a consequence of the large mass of the *W*.

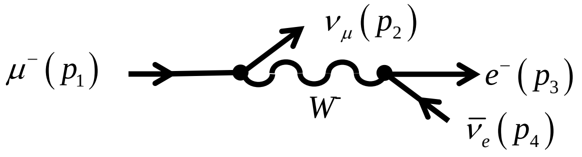


Figure 10-2: The Feynman diagram for muon decay in via the *W*-boson.

There are some pretty odd things about the coupling of the *W*. It must be proportional to $1-\gamma_5$, which in the chiral basis with which we are working is given by

$$\tfrac{1}{2}(1-\gamma_5)=\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}. \tag{10.16}$$

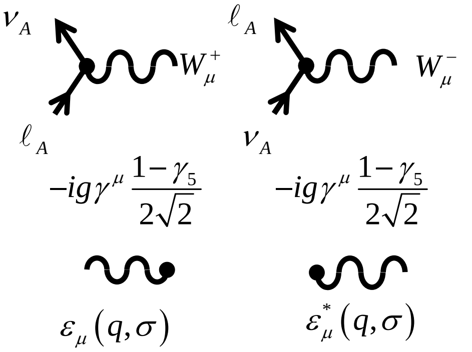


Figure 10-3: The fundamental *W* couplings to leptons, and the factors for external lines of *W*'s.

This means it connects only with Ψ_L and not Ψ_R. Another thing that is odd about the *W* is that it connects particles of different types, like an electron and its associated neutrino. This didn't happen in electromagnetism, where the photon only connects particles of the same type. In QCD, gluons connect quarks with different *colors*, but there is an exact symmetry relating these quarks, and they have the same mass and charge, etc.

The fundamental interaction of the *W* with leptons is as sketched in Fig. 10-3, with the corresponding rule. You can use any of the charged leptons for ℓ_A, (*e*, μ or τ) and the corresponding neutrino. The coupling will be called *g*. The factor of $1/2\sqrt{2}$ will make more sense later. Because the *W* is not equivalent to its anti-particle, there are actually two diagrams. I have

---

[1] For *all* particles, it is a matter of convention which is the "particle" and which is the "anti-particle," There is a standard convention that the quarks, the negatively charged leptons and their corresponding neutrinos are the particles. For the *W*-boson, there is no standard convention about which is the particle.

labeled the *W* lines as if they were going into the diagram, though they could just as easily be coming out.

In addition to these couplings, there will once again be polarization factors for any external *W* bosons, ε_μ for any initial state boson, and ε_μ^* for any final state boson. Polarization works a little differently for massive vector fields, however. Consider, for example, a *W* at rest, with momentum $q^\mu = \left(M_W, 0, 0, 0\right)$. The polarizations must satisfy $p \cdot \varepsilon = 0$ and $\varepsilon^2 = -1$, but in this case, that corresponds to three possible polarizations, which could be chosen to be unit vectors in the x^1, x^2, and x^3-directions, for example. We will sometimes need to sum over polarizations. For a particle at rest, this is pretty easy to do:

$$\sum_{\sigma=1}^{3} \varepsilon_\mu\left(\mathbf{q}=0,\sigma\right)\varepsilon_\nu^*\left(\mathbf{q}=0,\sigma\right) = \begin{cases} 1 & \text{if } \mu = \nu \neq 0\,, \\ 0 & \text{otherwise.} \end{cases} \tag{10.17}$$

We need to generalize (10.17) to arbitrary momentum. Noting that eq. (10.17) is nearly the matrix $-g_{\mu\nu}$, we can figure out that we could rewrite eq. (10.17) in the form

$$\sum_{\sigma} \varepsilon_\mu\left(q,\sigma\right)\varepsilon_\nu^*\left(q,\sigma\right) = -g_{\mu\nu} + \frac{q_\mu q_\nu}{M_W^2}\,. \tag{10.18}$$

Noting that the right side is a nice Lorentz covariant expression, we conclude that this must be the correct expression for all momenta, not just $\mathbf{q} = 0$. We would use this rule when summing over final polarizations, or averaging over initial ones. But remember, there are *three* polarizations, so if you average over initial polarizations, you include a factor of $\frac{1}{3}$.

We also need to work out the propagator for the *W*. The denominator will be $q^2 - M_W^2$, and in the numerator will be a factor of eq. (10.18). The resulting propagator will then be

$$\frac{i\left(-g_{\mu\nu} + q_\mu q_\nu \big/ M_W^2\right)}{q^2 - M_W^2}\,. \tag{10.19}$$

We will hardly use this expression, so don't let its frightening appearance bother you.

Now, let's calculate the Feynman amplitude for Fig. 10-2 using our new rules. The momentum is $q = p_1 - p_2$, but making the assumption that it is small compared to M_W, the propagator eq. (10.19) simplifies to $ig_{\mu\nu} \big/ M_W^2$, and we find the Feynman amplitude:

$$i\mathcal{M} = \left(\frac{-ig}{2\sqrt{2}}\right)^2 \frac{i}{M_W^2}\left[\bar{u}_2\gamma^\mu\left(1-\gamma_5\right)u_1\right]\left[\bar{u}_3\gamma_\mu\left(1-\gamma_5\right)v_4\right]. \tag{10.20}$$

Comparison of eqs. (10.8) and (10.20) then tell you

$$G_F = \frac{g^2}{4\sqrt{2}M_W^2}\,. \tag{10.21}$$

Knowledge of G_F tells us the ratio of the coupling to the *W* mass, but neither of these quantities separately.

D. Weak Interactions as Gauge Couplings

In QCD, gluons change quarks of one color into another. We were able to incorporate this into the Dirac equation (9.14) by expanding the wave function Ψ to incorporate more than one color. Can we accomplish the same thing with the weak interactions? This will be difficult, because the *W*-couplings only affect the left-handed part of the field. The first step is to split the free Dirac equation (3.16) into two pieces using (3.10), so that we have

$$i\not{\partial}\Psi_R = m\Psi_L\,, \quad i\not{\partial}\Psi_L = m\Psi_R\,. \tag{10.22}$$

There will be two pairs of equations, one for the electron and one for the neutrino. Let's replace Ψ by ℓ for the charged lepton, and by ν for the neutrino. We'll assume the neutrino is massless. Then eq. (10.22) becomes four equations:

$$i\not{\partial}\ell_L = m\ell_R\,, \quad i\not{\partial}\ell_R = m\ell_L\,, \quad i\not{\partial}\nu_L = 0, \quad i\not{\partial}\nu_R = 0\,. \tag{10.23}$$

The right-handed neutrino actually has no weak couplings, and is completely disconnected from everything else. In fact, in the Standard Model, it doesn't exist at all. So let's just drop it, which reduces us to three equations. Now, we are trying to connect the left-handed lepton and left-handed neutrino. We need to put them together into a single wave function, something like

$$\Psi_L = \begin{pmatrix} \nu_L \\ \ell_L \end{pmatrix}. \tag{10.24}$$

Then eqs. (10.23) would become, in terms of these,

$$i\not{\partial}\Psi_L - \begin{pmatrix} 0 \\ m \end{pmatrix}\ell_R = 0\,, \tag{10.25a}$$

$$i\not{\partial}\ell_R - \begin{pmatrix} 0 & m \end{pmatrix}\Psi_L = 0\,. \tag{10.25b}$$

We are now ready to include the couplings of the $W_\mu^\pm$. The $W_\mu^\pm$-field converts $\ell_L \leftrightarrow \nu_L$. This implies it should only appear in eq. (10.25a), which gets modified to

$$\left[i\not{\partial} - \frac{g}{2\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\not{W}^+ - \frac{g}{2\sqrt{2}}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\not{W}^- \right]\Psi_L - \begin{pmatrix} 0 \\ m \end{pmatrix}\ell_R = 0\,. \tag{10.26}$$

Now, to make this look more like a gauge coupling, let's take the $W_\mu^\pm$ fields and write them in terms of W_μ^1 and W_μ^2, in terms of which

$$W_\mu^\pm = \tfrac{1}{\sqrt{2}}\left(W_\mu^1 \mp iW_\mu^2 \right). \tag{10.27}$$

Then eq. (10.26) becomes

$$\left[i\not{\partial} - gT_1\not{W}^1 - gT_2\not{W}^2 \right]\Psi_L - \begin{pmatrix} 0 \\ m \end{pmatrix}\ell_R = 0\,, \tag{10.28}$$

where

$$T_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad T_2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{10.29}$$

These are two of the three generators of the symmetry SU(2), comparable to the matrices I_1 and I_2 appearing in eqs. (8.5) and (8.6). This suggests that we might be able to make a gauge theory out of weak interactions if we used the gauge group SU(2), which in this context is called *weak* SU(2), which we write as SU(2)$_L$ to remind us that it only affects the left-handed part of the leptons (and later, quarks). This only makes sense if we have a third field W_μ^3 which couples via a third generator T_3 given by

$$T_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{10.30}$$

Then eq. (10.28) must become

$$\left[i\partial\!\!\!/ - gT_a W\!\!\!/^a\right]\Psi_L - \begin{pmatrix} 0 \\ m \end{pmatrix}\ell_R = 0, \tag{10.31}$$

Eq. (10.31) has an implied sum on *a* from 1 to 3. We now have a derivative that is gauge invariant under the group SU(2)$_L$. Define the covariant derivative

$$D_\mu = \partial_\mu + igT_a W_\mu^a. \tag{10.32}$$

Then we can rewrite eqs. (10.31) and (10.25b) as

$$iD\!\!\!\!/\,\Psi_L - \begin{pmatrix} 0 \\ m \end{pmatrix}\ell_R = 0, \tag{10.33a}$$

$$iD\!\!\!\!/\,\ell_R - \begin{pmatrix} 0 & m \end{pmatrix}\Psi_L = 0. \tag{10.33b}$$

At first glimpse, it looks like we mistakenly wrote the covariant derivative in eq. (10.33b), but we just have to specify that the field ℓ_R isn't affected by SU(2)$_L$; that is, for ℓ_R, we use the trivial matrices

$$T_1 = T_2 = T_3 = (0). \tag{10.34}$$

We would say that Ψ_L is in the *doublet* or spin $\frac{1}{2}$ representation of weak SU(2)$_L$, while ℓ_R is in the *singlet* or spin 0 representation of weak SU(2) $_L$.

The derivative terms in eqs. (10.33) are certainly gauge invariant under SU(2)$_L$, but the mass terms are not. Let us set aside the problem of the lepton masses for now. There is another problem we *are* willing to tackle right now. The two fields appearing in Ψ_L have different charges, $Q = 0$ for the neutrino, and $Q = -1$ for the charged lepton. If we were to attempt to modify the covariant derivative, say change it to $D_\mu = \partial_\mu + igT_a W_\mu^a + ieQA_\mu$, we have the problem that *Q* would be a matrix, and would not commute with T_1 and T_2. In fact, for these two fields, it is easy to see that $Q = T_3 - \frac{1}{2}$, and that gives us an idea. Maybe the electric charge is a combination of T_3 and some other type of charge, with its own gauge coupling.

E. Electroweak Theory

Let's assume that in addition to SU(2)$_L$ gauge couplings, there is an additional U(1) coupling. We'll call the field B_μ, its coupling strength g' and its associated charge *weak hypercharge*, denoted *Y*, and the gauge group will be called U(1)$_Y$. We will arrange *Y* so that we always have

$$Q = T_3 + Y\ . \tag{10.35}$$

We therefore have $Y = -\frac{1}{2}$ for Ψ_L, and $Y = -1$ for ℓ_R. Our covariant derivative is

$$D_\mu = \partial_\mu + igT_a W_\mu^a + ig'YB_\mu\ . \tag{10.36}$$

We'd like to see our photon interactions coming out of eq. (10.36), which should look like $ieQA_\mu$. To try to make this work out, define two new fields A_μ and Z_μ which are an arbitrary mixture of W_μ^3 and B_μ. Specifically, let's write

$$\begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix}, \tag{10.37}$$

where θ_W is an angle called the *weak mixing angle*.[1] Then a part of the covariant derivative can be rewritten as

$$\begin{aligned} gT_3W_\mu^3 + g'YB_\mu &= gT_3W_\mu^3 + g'\left(Q - T_3\right)B_\mu \\ &= gT_3\left(\sin\theta_W A_\mu + \cos\theta_W Z_\mu\right) + g'\left(Q - T_3\right)\left(\cos\theta_W A_\mu - \sin\theta_W Z_\mu\right) \\ &= \left[\left(g\sin\theta_W - g'\cos\theta_W\right)T_3 + g'\cos\theta_W Q\right]A_\mu + \left[\left(g\cos\theta_W + g'\sin\theta_W\right)T_3 - g'\sin\theta_W Q\right]Z_\mu\ . \end{aligned} \tag{10.38}$$

The electromagnetic coupling will turn out right if we make sure that

$$\tan\theta_W = g'/g\ , \tag{10.39a}$$

$$e = g'\cos\theta_W = g\sin\theta_W\ . \tag{10.39b}$$

Then eq. (10.38) can be rewritten as

$$gT_3W_\mu^3 + g'YB_\mu = eQA_\mu + \frac{e}{\cos\theta_W \sin\theta_W}\left[T_3 - Q\sin^2\theta_W\right]Z_\mu\ . \tag{10.40}$$

Substituting eq. (10.40) back into (10.36), we have

$$D_\mu = \partial_\mu + igT_1W_\mu^1 + igT_2W_\mu^2 + ieQA_\mu + \frac{ie}{\sin\theta_W \cos\theta_W}\left(T_3 - Q\sin^2\theta_W\right)Z_\mu\ . \tag{10.41}$$

The W_μ^1 and W_μ^2 fields will lead to the $W_\mu^\pm$ couplings we already know about. We notice that we have succeeded in *also* obtaining the photon coupling in the same theory. Because this model contains both weak and electromagnetic interactions, it is called the *electroweak theory*.

---

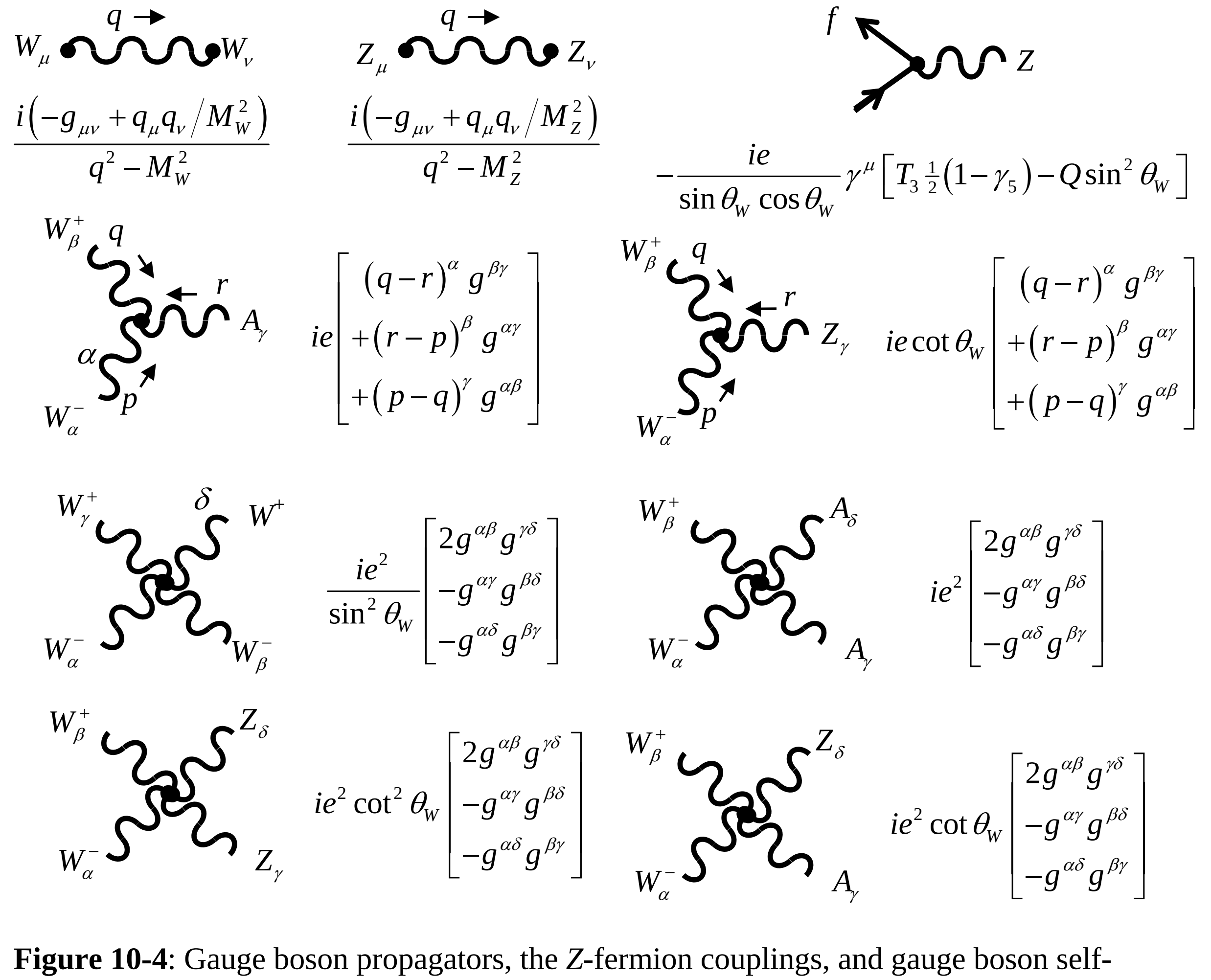
[1] It is sometimes called the Weinberg angle, but it was first introduced by Glashow, not Weinberg.

Because it contains the gauge groups SU(2)$_L$ and U(1)$_Y$, the gauge group is called SU(2)$_L$ × U(1)$_Y$.

The best feature of a theory is if it predicts the existence of new phenomena. We note that another field was necessary – the Z_μ – and it should produce its own interactions. Its couplings can be deduced from eq. (10.41), and have been included in Fig. 10-4. We have put in the explicit $\frac{1}{2}(1-\gamma_5)$ that keeps track of the fact that the T_3 eigenvalue applies only to the left-handed part, for which we have $T_3=+\frac{1}{2}$ for the neutrinos and .$T_3=-\frac{1}{2}$ for the charged leptons. The Q values, however, are the same for the left- and right-handed pieces. Since the W- and Z-bosons both have mass, they have similar propagators. It turns out there is a simple relation between the Z and W mass, which we will derive in the next chapter (eq. 11.18), which is

$$M_W = M_Z\cos\theta_W \, . \tag{10.42}$$

Because the gauge bosons come partly from an SU(2)$_L$ theory, in which the generators T_a do not commute, there will also be three- and four-boson interactions, just as we had for gluons in section 9D. I have included them in Fig. 10-4 as well. For interactions, the W boson lines are labeled as if the particle is going into the diagram.



$$\frac{i\left(-g_{\mu\nu}+q_\mu q_\nu/M_W^2\right)}{q^2-M_W^2}$$

$$\frac{i\left(-g_{\mu\nu}+q_\mu q_\nu/M_Z^2\right)}{q^2-M_Z^2}$$

$$-\frac{ie}{\sin\theta_W\cos\theta_W}\gamma^\mu\left[T_3\tfrac{1}{2}(1-\gamma_5)-Q\sin^2\theta_W\right]$$

$$ie\begin{bmatrix}(q-r)^\alpha g^{\beta\gamma}\\+(r-p)^\beta g^{\alpha\gamma}\\+(p-q)^\gamma g^{\alpha\beta}\end{bmatrix}$$

$$ie\cot\theta_W\begin{bmatrix}(q-r)^\alpha g^{\beta\gamma}\\+(r-p)^\beta g^{\alpha\gamma}\\+(p-q)^\gamma g^{\alpha\beta}\end{bmatrix}$$

$$\frac{ie^2}{\sin^2\theta_W}\begin{bmatrix}2g^{\alpha\beta}g^{\gamma\delta}\\-g^{\alpha\gamma}g^{\beta\delta}\\-g^{\alpha\delta}g^{\beta\gamma}\end{bmatrix}$$

$$ie^2\begin{bmatrix}2g^{\alpha\beta}g^{\gamma\delta}\\-g^{\alpha\gamma}g^{\beta\delta}\\-g^{\alpha\delta}g^{\beta\gamma}\end{bmatrix}$$

$$ie^2\cot^2\theta_W\begin{bmatrix}2g^{\alpha\beta}g^{\gamma\delta}\\-g^{\alpha\gamma}g^{\beta\delta}\\-g^{\alpha\delta}g^{\beta\gamma}\end{bmatrix}$$

$$ie^2\cot\theta_W\begin{bmatrix}2g^{\alpha\beta}g^{\gamma\delta}\\-g^{\alpha\gamma}g^{\beta\delta}\\-g^{\alpha\delta}g^{\beta\gamma}\end{bmatrix}$$

Figure 10-4: Gauge boson propagators, the Z-fermion couplings, and gauge boson self-couplings for electroweak theory.

F. Weak Neutral Currents

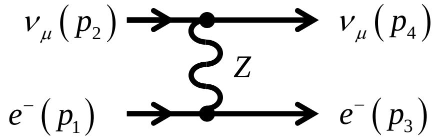


Figure 10-5: A neutral current process for scattering neutrinos.

Fig. 10-5 gives an example of an interaction that must be present in the electroweak theory. Such interactions are called *neutral currents*: neutral because the particle charges are unchanged by the interaction with the *Z*. Let's calculate this cross-section. To make things simple, let's assume we are working at energies well above the electron mass, but also well below the *Z*-mass. The Feynman amplitude for this scattering is

$$i\mathcal{M} = \frac{(ie)(-ie)}{(4\sin\theta_W\cos\theta_W)^2}\left[\bar{u}_4\gamma^\mu(1-\gamma_5)u_2\right]\left[\bar{u}_3\gamma^\nu(-1+\gamma_5+4\sin^2\theta_W)u_1\right]\frac{i\left(-g_{\mu\nu}+q_\mu q_\nu/M_Z^2\right)}{q^2-M_z^2},$$

$$i\mathcal{M} = \frac{ie^2\left[\bar{u}_4\gamma^\mu(1-\gamma_5)u_3\right]\left[\bar{u}_2\gamma^\nu(1-\gamma_5-4\sin^2\theta_W)u_1\right]}{16\sin^2\theta_W\cos^2\theta_W\left(q^2-M_Z^2\right)}\left(-g_{\mu\nu}+\frac{q_\mu q_\nu}{M_Z^2}\right), \tag{10.43}$$

where $q = p_2 - p_4$ is the momentum of the virtual *Z*. The first step is to demonstrate that the $q_\mu q_\nu/M_Z^2$ in the propagator does not, in fact, contribute. Keeping in mind that the neutrinos are massless (and, if it comes to it, we are also treating the electrons as massless), we have

$$\bar{u}_4\not{q}(1-\gamma_5)u_2 = \bar{u}_4(\not{p}_2-\not{p}_4)(1-\gamma_5)u_2 = \bar{u}_4\not{p}_2(1-\gamma_5)u_2 = \bar{u}_4(1+\gamma_5)\not{p}_2u_2 = 0\,.$$

We could also have eliminated this term since we are well below the *Z*-mass, so that $q_\mu q_\nu/M_Z^2$ is small. Indeed, we will also approximate $q^2 - M_Z^2 = -M_Z^2$. Eq. (10.43) then becomes

$$i\mathcal{M} = \frac{ie^2}{16\sin^2\theta_W\cos^2\theta_W M_Z^2}\left[\bar{u}_4\gamma^\mu(1-\gamma_5)u_2\right]\left[\bar{u}_3\gamma_\mu(a-\gamma_5)u_1\right], \tag{10.44a}$$

$$a \equiv 1-4\sin^2\theta_W\,. \tag{10.44b}$$

The coefficient out front can be simplified using eqs. (10.39b), (10.42), and (10.21), so that this simplifies to

$$i\mathcal{M} = \tfrac{1}{2\sqrt{2}}G_F\left[\bar{u}_4\gamma^\mu(1-\gamma_5)u_2\right]\left[\bar{u}_3\gamma_\mu(a-\gamma_5)u_1\right]. \tag{10.45}$$

We would then multiply by the complex conjugate, and then sum or average over all spins. This would then yield

$$\tfrac{1}{2}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 = \tfrac{1}{16}\mathrm{G}_F^2\mathrm{Tr}\left[\not{p}_4\gamma^\mu(1-\gamma_5)\not{p}_2(1+\gamma_5)\gamma^\nu\right]\mathrm{Tr}\left[\not{p}_3\gamma_\mu(a-\gamma_5)\not{p}_1(a+\gamma_5)\gamma_\nu\right]. \tag{10.46}$$

There is a subtle point we need to discuss in eq. (10.46). We normally sum over outgoing spins and average over initial spins. Logically, this would imply that we should want a factor of $\frac{1}{4}$ out front, not $\frac{1}{2}$. Did I make a mistake? Remember that in the Standard Model, neutrinos only come in a left-handed variety. Hence the incoming muon neutrinos must have only one chirality. This suggests that the factor of $\frac{1}{2}$ is correct, since the electron spin is still being averaged over, but it seems I erred in summing over the neutrino spins, so the sum on spins should be done only over the electrons, not the neutrinos. However, the matrix element for the right-handed neutrinos automatically vanishes, because of the factors of $(1-\gamma_5)$, so that by summing on the neutrino spins, I am only adding zeros. So eq. (10.46) is correct.

We now bring all the factors with γ_5's over to the left by anti-commuting them with the other Dirac matrices, which yields

$$\begin{aligned}
\frac{1}{2}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 &= \tfrac{1}{16}G_F^2\operatorname{Tr}\left[(1-\gamma_5)^2\,\not{p}_4\gamma^\mu\not{p}_2\gamma^\nu\right]\operatorname{Tr}\left[(a-\gamma_5)^2\,\not{p}_3\gamma_\mu\not{p}_1\gamma_\nu\right] \\
&= \tfrac{1}{8}G_F^2\operatorname{Tr}\left[(1-\gamma_5)\,\not{p}_4\gamma^\mu\not{p}_2\gamma^\nu\right]\operatorname{Tr}\left[\left(1+a^2-2a\gamma_5\right)\not{p}_3\gamma_\mu\not{p}_1\gamma_\nu\right] \\
&= 2G_F^2\left(p_4^\mu p_2^\nu + p_2^\mu p_4^\nu - g^{\mu\nu}p_2\cdot p_4 + i\varepsilon^{\alpha\mu\beta\nu}p_{4\alpha}p_{2\beta}\right)\times \\
&\quad\left[\left(1+a^2\right)\left(p_{3\mu}p_{1\nu}+p_{1\mu}p_{3\nu}-g_{\mu\nu}p_1\cdot p_3\right)+2ia\varepsilon^{\sigma}{}_{\mu}{}^{\tau}{}_{\nu}p_{3\sigma}p_{1\tau}\right] \\
&= 4G_F^2\left(1+a^2\right)\left[(p_4\cdot p_3)(p_1\cdot p_2)+(p_4\cdot p_1)(p_2\cdot p_3)\right] \\
&\quad+8G_F^2 a\left[(p_4\cdot p_3)(p_1\cdot p_2)-(p_4\cdot p_1)(p_2\cdot p_3)\right],
\end{aligned}$$

$$\frac{1}{2}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 = 4G_F^2\left[(1+a)^2(p_4\cdot p_3)(p_1\cdot p_2)+(1-a)^2(p_4\cdot p_1)(p_2\cdot p_3)\right]. \tag{10.47}$$

In the center of mass frame, let the energies be E, and let the scattering angle be θ. Then we have

$$p_4\cdot p_3 = p_1\cdot p_2 = 2E^2 \quad \text{and} \quad p_4\cdot p_1 = p_2\cdot p_3 = E^2(1+\cos\theta).$$

Substituting each of these into eq. (10.47), and also using eq. (10.44b), we have

$$\frac{1}{2}\sum_{\text{spins}}\left|i\mathcal{M}\right|^2 = 64G_F^2\left[\left(1-2\sin^2\theta_W\right)^2E^4+\sin^4\theta_W E^4(1+\cos\theta)^2\right]. \tag{10.48}$$

We then proceed to find the cross-section using our usual cross-section and two-body final state formulas, eqs. (4.35) and (4.39):

$$\begin{aligned}
\sigma &= \frac{1}{4(pE+pE)}\frac{p}{16\pi^2(2E)}64G_F^2\int\left[\left(1-2\sin^2\theta_W\right)^2E^4+\sin^4\theta_W E^4(1+\cos\theta)^2\right]d\Omega \\
&= \frac{G_F^2E^2}{2\pi}\int_{-1}^{1}\left[\left(1-2\sin^2\theta_W\right)^2+\sin^4\theta_W(1+\cos\theta)^2\right]d\cos\theta, \\
\sigma &= \tfrac{1}{\pi}G_F^2E^2\left[1-4\sin^2\theta_W+\tfrac{16}{3}\sin^4\theta_W\right] = \tfrac{1}{\pi}G_F^2 s\left[\tfrac{1}{4}-\sin^2\theta_W+\tfrac{4}{3}\sin^4\theta_W\right].
\end{aligned} \tag{10.49}$$

Measurement of the cross section (or, more likely, the differential cross-section) then allows you to measure the weak mixing angle, which turns out to have a value of about

$$\sin^2\theta_W \approx 0.2312. \tag{10.50}$$

Because the gauge couplings g and g' run, just like the electromagnetic coupling e (see section 7E), this angle is weakly a function of energy. The value in eq. (10.50) is correct at the Z mass.

G. Weak Interactions and the Hadrons

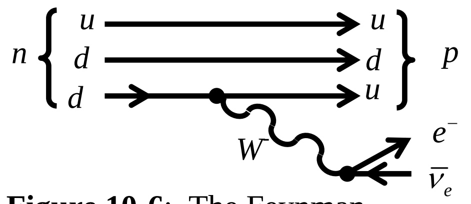


Figure 10-6: The Feynman diagram for neutron decay.

Not surprisingly, reaching an understanding of how weak interactions work for strongly interacting particles was harder than it was for leptons. Initially, it wasn't known that the proton and neutron were not elementary, and the first attempts to describe weak interactions were built in terms of these particles. But we know of the existence of quarks, and therefore we can take advantage of this right from the start. For example, nuclear beta decay, $n \to p^+ e^- \bar{\nu}_e$, presumably involves the conversion of a down quark to an up quark, as sketched in Fig. 10-6. We just have to work out what sort of coupling is appropriate for the down/up link. It isn't obvious how to experimentally determine this, because the direction of spin of the down quark isn't related in an obvious way to the spin of the neutron and proton, but the most obvious thing to do was to simply copy what worked in the lepton sector. Let's ignore the mass of the quarks for now, since we will be dealing with those in chapter 11, when we consider the Higgs boson.

Let's assume that once again only the left-handed quarks get involved, and our first step is to put u_L and d_L into an SU(2) doublet, much as we did for the neutrino and charged leptons, but we'll leave the right-handed fields u_R and d_R alone. We'll call the combination q_L,

$$q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}. \tag{10.51}$$

Our covariant Dirac equations would then look like (ignoring masses)

$$i\not{D} q_L = 0, \quad i\not{D} u_R = 0, \quad i\not{D} d_R = 0. \tag{10.52}$$

Unlike the neutrino, we have kept the d_R field, since we know down quarks actually *do* have mass, and are coupled to other particles via the strong and electromagnetic interactions anyway. The covariant derivative is given by eqs. (10.36) or (10.41), repeated here

$$D_\mu = \partial_\mu + igT_a W_\mu^a + ig'YB_\mu = \partial_\mu + igT_1 W_\mu^1 + igT_2 W_\mu^2 + ieQA_\mu + \frac{ie}{\sin\theta_W \cos\theta_W}\left(T_3 - Q\sin^2\theta_W\right) Z_\mu .$$

As in the lepton sector, we use the matrices given by eqs. (10.29) and (10.30) when T_a acts on q_L, but use the trivial matrices $T_a = (0)$ when acting on u_R or d_R. As for the weak hypercharge Y, we can use the formula $Q = T_3 + Y$ to deduce Y for each type of particle:

$$Yq_L = +\tfrac{1}{6} q_L , \quad Yu_R = +\tfrac{2}{3} u_R , \quad Yd_R = -\tfrac{1}{3} d_R . \tag{10.53}$$

Based on these assumptions, you actually get a pretty good theory. The $W^\pm$ will have quark-sector couplings just like the leptons, corresponding to diagrams like Fig. 10-3, but with the up and down quarks replacing the neutrino and lepton respectively. We can now change up quarks to down, and vice versa, via weak interactions. A lot of interactions could be predicted pretty accurately using this simple formalism. For example, the decay $\pi^- \to \mu^- \bar{\nu}_\mu$ could be explained by the process sketched in Fig. 10-7.

But we are definitely missing something. After all, there are also weak interactions that convert strange quarks (at the time, these were the only three quarks known). Consider, for example, the decay $K^- \to \mu^- \bar{\nu}_\mu$, This can be explained almost identically to $\pi^- \to \mu^- \bar{\nu}_\mu$, if we only assume a similar coupling.

The kaon decay is slightly faster than pion decay. But the kaon is quite a bit heavier, so if you assume the coupling is comparable, it should be *much* faster. This implies that the up/strange coupling must be noticeably smaller than the up/down coupling, about a factor of 4 or 5 smaller. But it's not immediately obvious what the right approach is. Do we just write one more equation for the strange quark, putting it in another doublet, with the up quark again?

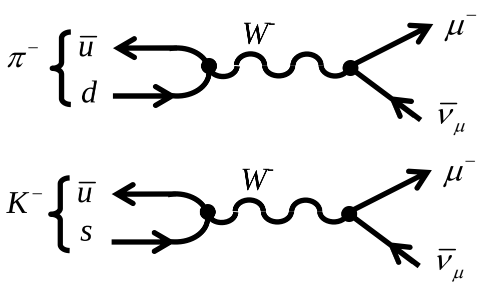


Figure 10-7: Some Feynman diagrams for charged pion (top) and kaon (bottom) decays.

The solution was worked out by Nicola Cabibbo. He surmised that the combination eq. (10.51) was probably wrong. He decided that up should be paired, not with down, but with some combination of down and strange. He replaced eq. (10.51) with

$$q_L = \begin{pmatrix} u_L \\ d_L \cos\theta_C + s_L \sin\theta_C \end{pmatrix}, \tag{10.54}$$

where θ_C is the *Cabibbo angle*. It's value is $\theta_C \approx 13°$. This implies that the $u \leftrightarrow d$ vertex would be suppressed by about $\cos\theta_C = 0.97$ while the $u \leftrightarrow s$ would be suppressed by about $\sin\theta_C = 0.225$, and this fit well with experiment.

For the *Z*-couplings, however, it didn't work as well. It turned out that the *Z*-coupling should have turned *d*-quarks into *s*-quarks, or vice versa. This rate should be suppressed, because of the mixing, by a factor of $\sin\theta_C \cos\theta_C = 0.22$. This predicts, for example, the decay $K^- \to \pi^- e^+ e^-$, with a rate comparable to the rate $K^- \to \pi^0 e^- \bar{\nu}_e$, as sketched in Fig. 10-8. But the former rate is suppressed by a factor of more than 10^5 compared to the latter. Such *flavor changing neutral currents* were rarely, if ever, seen.

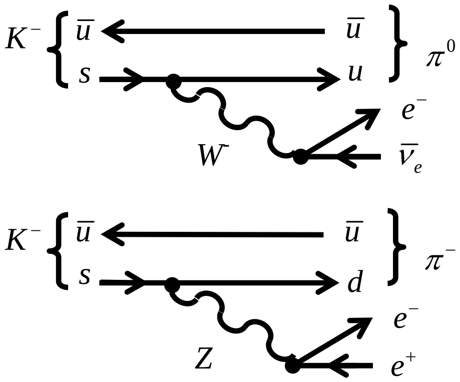


Figure 10-8: The upper decay of the kaon, $K^- \to \pi^0 e^- \bar{\nu}_e$, is common. The bottom decay, $K^- \to \pi^- e^+ e^-$, is about a 10^5 times more rare.

Fortunately, Glashow, Iliopoulos, and Maiani came up with an explanation, now called the *GIM mechanism*. Their idea was that there was a *second* doublet involving the orthogonal combinations of the down and strange quark to a new quark, dubbed *charm*, which was like a heavier up quark. In this case, eq. (10.46) would become the two equations

$$q_{L1} = \begin{pmatrix} u_L \\ d_L \cos\theta_C + s_L \sin\theta_C \end{pmatrix}, \quad q_{L2} = \begin{pmatrix} c_L \\ s_L \cos\theta_C - d_L \sin\theta_C \end{pmatrix}. \tag{10.55}$$

In addition, there would be a c_R particle, not in a doublet. Eqs. (10.51) would become

$$i\not{D}q_{L1} = i\not{D}q_{L2} = 0\,, \quad i\not{D}u_R = i\not{D}c_R = 0\,, \quad i\not{D}d_R = i\not{D}s_R = 0\,. \tag{10.56}$$

In eqs. (10.56), we have grouped together pairs of similar equations, because, for example, q_{L1} and q_{L2} behave exactly the same way, as do the pair u_R and c_R, and the pair d_R and s_R. There is a repeating pattern to everything, and we have two *generations*.

How does eq. (10.55) solve the problem of flavor changing neutral currents? It turns out that the contribution of the *Z*-coupling from $i\not{D}q_{L1}$ to $s \leftrightarrow d$ transitions is exactly canceled by that from $i\not{D}q_{L2}$. To make this slightly more obvious, consider if instead of eqs. (10.55), we had defined

$$q'_{L1} = \begin{pmatrix} u_L \cos\theta_C - c_L \sin\theta_C \\ d_L \end{pmatrix}, \quad q'_{L2} = \begin{pmatrix} c_L \cos\theta_C + u_L \sin\theta_C \\ s_L \end{pmatrix}. \tag{10.57}$$

Then we could rewrite the first two equations of eq. (10.56) in terms of these, *i.e.*,

$$i\not{D}q_{L1} = i\not{D}q_{L2} = 0 \quad \Leftrightarrow \quad \not{D}q'_{L1} = i\not{D}q'_{L2} = 0. \tag{10.58}$$

It is only a matter of convention that we normally write our doublet combinations in the form of eq. (10.55); we could just as easily claim that it is the up-type quarks that mix. Indeed, this tells us that the mixing only affects how the up- and down-type quarks interconvert via *W*-emission and absorption; the neutral currents mediated by the *Z* don't care about the mixing.

For reasons I will elucidate in chapter 11, Kobayashi and Maskawa suggested that it might be necessary to introduce a third generation, two new quarks (and corresponding leptons), the bottom and top, and this complicates the situation further. Let us denote the generations by a generation index *A* that runs from 1 to 3. We'll call the *u, c,* and *t* quarks up-type quarks, and the three others down-type quarks, and we'll denote them

$$\left(u_1, u_2, u_3\right) = \left(u, c, t\right), \qquad \left(d_1, d_2, d_3\right) = \left(d, s, b\right). \tag{10.59}$$

Then we define a set of three left-handed doublets

$$q_{LA} = \begin{pmatrix} u_{LA} \\ V_{AB} d_{LB} \end{pmatrix}, \tag{10.60}$$

where V_{AB} is a set of nine complex numbers such that *V* forms a matrix, called the *Cabibbo-Kobayashi-Maskawa matrix,* or *CKM matrix* for short. The CKM matrix is unitary, so $V^\dagger V = 1$. In fact, it is a completely arbitrary unitary matrix, but as we will discuss in chapter 11, some of the parameters describing it are irrelevant, so it turns out that it can be chosen in such a way to be described in terms of four parameters, normally written as three angles called θ_{12}, θ_{13}, and θ_{23}, and one phase called δ, in terms of which the CKM matrix takes the form:

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{10.61}$$

where we have abbreviated $s_{AB} \equiv \sin\theta_{AB}$ and $c_{AB} \equiv \cos\theta_{AB}$. The magnitudes of the elements of the CKM matrix take on the values

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97427(15) & 0.22534(65) & 0.00351(15) \\ 0.22520(65) & 0.97344(16) & 0.0412(8) \\ 0.00867(30) & 0.0404(8) & 0.999146(33) \end{pmatrix}, \qquad (10.62)$$

where the errors in the final digit(s) are given in parentheses at the end. We note that this matrix is close to the identity matrix; the only significant non-zero off-diagonal matrix elements being V_{us} and V_{cd}. This tells us that θ_{13} and θ_{23} are small angles; in the limit these are zero, the third generation would decouple from the other two and θ_{12} would correspond to the Cabibbo angle θ_C.

The V matrix only affects the W-couplings of the quarks, so that the Z-couplings are still given by the expression in Fig. 10-4. There are two closely-related rules for the W-couplings for quarks, sketched in Fig. 10-9. Technically, the quarks are colored as well, so both the W- and Z-coupling includes an additional δ_i^j for the corresponding color indices. We have used our knowledge that $g = e/\sin\theta_W$ to eliminate the coupling g, which we should also do back in Fig. 10-3. Indeed, we have summarized all the rules for weak interactions on the back cover, and have done so there.

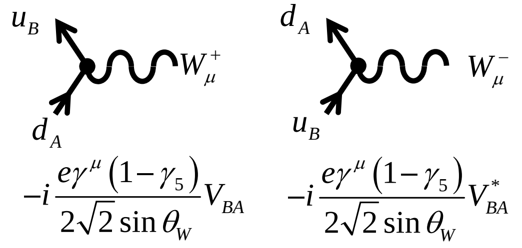


Figure 10-9: The fundamental W couplings to quarks.

We have continued to ignore the masses of the gauge bosons, the charged leptons and the quarks. We will deal with these problems in chapter 11.

H. Decay of the W

Let's see if we can figure out the decay rate of the W-boson, for which there are multiple decay channels. We will focus only on those containing two final state fermions, as sketched in Fig. 10-10. There are basically two slightly different diagrams, one corresponding to leptonic decays, the other to quark decays. We'll do the lepton first. For example, for the decay $\Gamma\left(W^- \to e^-\bar{\nu}_e\right)$, the Feynman amplitude is

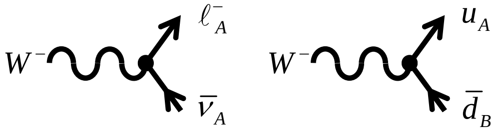


Figure 10-10: The leptonic (left) and quark (right) decays of the W-boson.

$$i\mathcal{M} = -\frac{ie}{2\sqrt{2}\sin\theta_W}\left[\bar{u}\gamma^\mu(1-\gamma_5)v'\right]\varepsilon_\mu ,$$

$$\left|i\mathcal{M}\right|^2 = \frac{e^2}{8\sin^2\theta_W}\left[\bar{u}\gamma^\mu(1-\gamma_5)v'\bar{v}'(1+\gamma_5)\gamma^\nu u\right]\varepsilon_\mu\varepsilon_\nu^* . \qquad (10.63)$$

We will average over polarizations. Since the neutrino is effectively massless and the electron has a small mass, we will ignore the mass terms. Letting the charged lepton have momentum p

and the neutrino p' and the W $q = p + p'$, and using eq. (10.18) for the sum of polarizations, we have

$$\frac{1}{3}\sum_{\substack{\text{spins,}\\\text{pols}}} \left|i\mathcal{M}\right|^2 = \frac{e^2}{24\sin^2\theta_W}\operatorname{Tr}\left[\not{p}\gamma^\mu\left(1-\gamma_5\right)\not{p}'\left(1+\gamma_5\right)\gamma^\nu\right]\left(-g_{\mu\nu}+\frac{q_\mu q_\nu}{M_W^2}\right). \qquad (10.64)$$

We bring the $1\pm\gamma_5$ all the way to the left in the usual way, and we have

$$\begin{aligned}\frac{1}{3}\sum\left|i\mathcal{M}\right|^2 &= \frac{e^2}{24\sin^2\theta_W}\operatorname{Tr}\left[\left(1-\gamma_5\right)^2\not{p}\gamma^\mu\not{p}'\gamma^\nu\right]\left(-g_{\mu\nu}+\frac{q_\mu q_\nu}{M_W^2}\right)\\ &= \frac{e^2}{12\sin^2\theta_W}\left\{-\operatorname{Tr}\left[\left(1-\gamma_5\right)\not{p}\gamma^\mu\not{p}'\gamma_\mu\right]+\frac{1}{M_W^2}\operatorname{Tr}\left[\left(1-\gamma_5\right)\not{p}\left(\not{p}+\not{p}'\right)\not{p}'\left(\not{p}+\not{p}'\right)\right]\right\}\\ &= \frac{e^2}{12\sin^2\theta_W}\left\{2\operatorname{Tr}\left[\left(1-\gamma_5\right)\not{p}\not{p}'\right]+\frac{1}{M_W^2}\operatorname{Tr}\left[\left(1-\gamma_5\right)\left(p^2\not{p}'+\not{p}p'^2\right)\left(\not{p}+\not{p}'\right)\right]\right\},\end{aligned}$$

$$\frac{1}{3}\sum\left|i\mathcal{M}\right|^2 = \frac{2e^2\left(p\cdot p'\right)}{3\sin^2\theta_w} = \frac{e^2M_W^2}{3\sin^2\theta_W}, \qquad (10.65)$$

where in the last step we used $M_W^2 = q^2 = \left(p+p'\right)^2 = 2p\cdot p'$. It is now a straightforward step to proceed to the decay rate using eqs. (4.38) and (4.35):

$$\Gamma\left(W^-\to e^-\bar{\nu}_e\right) = \frac{1}{2M_W}\frac{p}{16\pi^2M_W}\int d\Omega\tfrac{1}{3}\sum\left|i\mathcal{M}\right|^2 = \frac{1}{2M_W}\frac{\frac{1}{2}M_W}{16\pi^2M_W}\frac{4\pi e^2M_W^2}{3\sin^2\theta_W},$$

$$\Gamma\left(W^-\to e^-\bar{\nu}_e\right) = \frac{\alpha M_W}{12\sin^2\theta_W}. \qquad (10.66)$$

At the last step we wrote our answer in terms of the fine structure constant α.

Of course, eq. (10.66) represents only one of three possible leptonic final states. But the muon and tau have identical couplings, and even the tau has $m_\tau^2/m_W^2 \approx 5\times10^{-4}$, so unless we are expecting exceedingly high precision, they will have essentially the same decay rate

$$\Gamma\left(W^-\to\mu^-\bar{\nu}_\mu\right) = \Gamma\left(W^-\to\tau^-\bar{\nu}_\tau\right) = \frac{\alpha M_W}{12\sin^2\theta_W}. \qquad (10.67)$$

Our attention now turns to the quarks. Let's start with $\Gamma\left(W^-\to d\bar{u}\right)$. The mass of the up and down quarks are again small, and the coupling is the same save for a factor of V_{ud}. However, the quarks come in three colors. The *W*-coupling does not change color, so assuming the *u* and *d* quarks are the same color, we would have

$$\Gamma\left(W^-\to d\bar{u}\right) = \frac{\alpha M_W}{12\sin^2\theta_W}V_{ud}V_{ud}^* \qquad \text{(one color).}$$

To get the rate to all three colors, we simply multiply by three. We therefore have

$$\Gamma\left(W^- \to d\bar{u}\right) = \frac{\alpha M_W}{4\sin^2\theta_W} V_{ud} V_{ud}^* . \tag{10.68}$$

So far so good. We now need to consider all possible quark decays The top quark is too heavy, so we exclude it, but the possible *W* quark decays are

$$\begin{aligned} &W^- \to d\bar{u}, \quad W^- \to s\bar{u}, \quad W^- \to b\bar{u}, \\ &W^- \to d\bar{c}, \quad W^- \to s\bar{c}, \quad W^- \to b\bar{c}. \end{aligned} \tag{10.69}$$

Can we treat all these decays as massless? The up, down, strange, and charm quark are all light enough that any effects from their mass are probably well below one percent. The bottom quark has $m_b^2 / m_W^2 \approx 0.004$, which is getting to around a percent, but we see from eq. (10.62) that V_{ub} and V_{cb} are already small, so the small mass correction to an unlikely decay branch implies we are pretty safe ignoring this mass. Now, experimentally, any of these strong decays will almost immediately become a spray of hadrons, so the six decays in eq. (10.69) will be pretty much experimentally indistinguishable. We therefore sum them, and find

$$\Gamma\left(W^- \to \text{hadrons}\right) = \frac{\alpha M_W}{4\sin^2\theta_W}\begin{pmatrix} V_{ud}V_{ud}^* + V_{us}V_{us}^* + V_{ub}V_{ub}^* \\ +V_{cd}V_{cd}^* + V_{cs}V_{cs}^* + V_{cb}V_{cb}^* \end{pmatrix}. \tag{10.70}$$

At this point you might think it is time to turn to the experimentally determined values of the CKM matrix, eq. (10.62), but there is a better way. Keep in mind that *V* is a unitary matrix. This implies $VV^\dagger = 1$. Looking at the *uu* component of this matrix product, we see that

$$1 = \left(VV^\dagger\right)_{uu} = V_{uA}V_{Au}^\dagger = V_{uA}V_{uA}^* = V_{ud}V_{ud}^* + V_{us}V_{us}^* + V_{ub}V_{ub}^* . \tag{10.71}$$

Similarly,

$$1 = V_{cd}V_{cd}^* + V_{cs}V_{cs}^* + V_{cb}V_{cb}^* . \tag{10.72}$$

Substituting eqs. (10.71) and (10.72) into eq. (10.70), we have

$$\Gamma\left(W^- \to \text{hadrons}\right) = \frac{\alpha M_W}{2\sin^2\theta_W}. \tag{10.73}$$

There are other processes, such as $\Gamma\left(W^- \to d\bar{u}g\right)$, but this will tend to be suppressed by about a factor of $\alpha_S / 4\pi$, and at the weak scale, eq. (9.27) tells us this puts these corrections at the percent level. Loops, especially of gluons, will similarly distort these values.

Let's put everything together and get the total decay rate and the branching ratios. The total decay rate, to leading order, is the sum of eqs. (10.66), (10.67), and (10.73). Using the weak scale value of $\alpha \approx \frac{1}{128}$, we then have

$$\Gamma\left(W^-\right) = \frac{3\alpha M_W}{4\sin^2\theta_W} = 2.04 \text{ GeV} .$$

The experimental value is $2.08 \pm .04$ GeV, so we did pretty well. The branching ratio is the ratio of the individual decay rates to the total, which then works out to

$$\mathrm{BR}\left(W^- \to e^-\bar{\nu}_e\right) = \mathrm{BR}\left(W^- \to \mu^-\bar{\nu}_\mu\right) = \mathrm{BR}\left(W^- \to \tau^-\bar{\nu}_\tau\right) = \tfrac{1}{9}, \quad \mathrm{BR}\left(W^- \to \text{hadrons}\right) = \tfrac{2}{3}.$$

The experimental values for the first three are about 10.8%, 10.6%, and 11.2%, pretty close to our predicted value of 11.1%. I don't know the source of the discrepancy, though part of it may be experimental error. The hadronic branching ratio is 67.6%, probably increased from our expectations because we didn't include cases like $\Gamma\left(W^- \to d\bar{u}g\right)$. Apparently they somehow managed to pull out those cases where there is a charm quark, which roughly comes out to 33%, about what we would predict.

Computations like these should convince you that a gauge theory of weak interactions works pretty well, at least at tree level. Unfortunately, to match reality, we had to introduce masses for the *W* and *Z*, as well as for the charged leptons and the six quarks. At the moment, we do not know where these terms came from, nor do we understand why, for example, there is a complicated CKM matrix that appears for the quarks, but nothing comparable for the leptons. And that's because there is still one ingredient missing in the Standard Model – the Higgs boson.

Problems for Chapter 10

1. Find the following decay rates for the τ^- lepton: $\Gamma\left(\tau^- \to \nu_\tau e^-\bar{\nu}_e\right)$, $\Gamma\left(\tau^- \to \nu_\tau \mu^-\bar{\nu}_\mu\right)$, $\Gamma\left(\tau^- \to \nu_\tau d\bar{u}\right)$, $\Gamma\left(\tau^- \to \nu_\tau s\bar{u}\right)$ (don't forget colors). Assume all final particles are massless. Calculate the total decay rate, and the branching ratio for each of the first two decays, and compare with the experimental values. The decay rates are easily computed with minor modifications of eq. (10.15). Don't be bothered if you don't get exactly the right answer.

2. For a free bottom quark, calculate the partial decay rates $\Gamma\left(b \to c\ell^-\bar{\nu}_\ell\right)$ and $\Gamma\left(b \to c\bar{u}_A d_B\right)$ where ℓ^- is any one charged lepton, and $\bar{u}_A$ is the *u* or *c* anti-quark, and d_B is the down or strange quark. Treat the bottom quark as if it had mass 5 GeV, and treat all other particles as massless. Find the branching ratios and the total decay rate and lifetime. Compare to the experimental leptonic branching ratio and total lifetime for the B^+ meson (which actually contains a $\bar{b}$ quark, so you'll be searching for an anti-lepton).

3. Find the decay rate for top decay, $t \to W^+ b$, neglecting the bottom mass, but including the *W*-mass, in GeV. Note that even though this is considered a "weak" decay, it has a very large rate (a GeV rate is faster than typical strong interactions).

4. Draw all tree-level Feynman diagrams for the process $e^+e^- \to W^+W^-$. You do not need to compute them.

5. Draw all tree-level Feynman diagrams for the process $e^-\gamma \to W^-\nu_e$, and write the corresponding amplitude.

6. Draw all tree-level Feynman diagrams for the process $e^+e^- \to Z^0Z^0$, and write the corresponding amplitude.

7. A muon neutrino with energy $E = 10.0$ MeV is attempting to scatter off of a stationary electron. What is s? What is the cross-section? How far would the neutrino have to travel on average before scattering if it were traveling through water?

8. Find the amplitude for $e^-\nu_\mu \to \mu^-\nu_e$. Assume the energies are well below the W-mass, so you can ignore the momentum in the denominator of the W-propagator. Treat all four leptons as massless. Write your answer exclusively in terms of G_F and $\sin^2\theta_W$. Find the differential and total cross section for this scattering.

9. Calculate the differential and total cross-section in the center of mass frame for $e^-\bar{\nu}_e \to \mu^-\bar{\nu}_\mu$. Ignore all lepton masses. Do not assume anything about the energy of the initial particles compared to the W-boson, other than that you are not too close to resonance.

10. A π^+ meson is a combination of an up quark and down anti-quark, and therefore can become a W^+ boson. We can't predict the corresponding coupling because the pion is a strongly interacting particle, but based on Lorentz invariance we can conclude that the corresponding Feynman rule would be something of the form $f_\pi k^\mu$, where k is the momentum of the pion, as sketched in Fig. 10-11. Calculate the decay rate $\Gamma\left(\pi^+ \to \mu^+\nu_\mu\right)$ and $\Gamma\left(\pi^+ \to e^+\nu_e\right)$. Treat the neutrinos as massless, but do *not* ignore the other masses, even the electron. Find the ratio of these two decays, and compare with the experimental branching ratios.

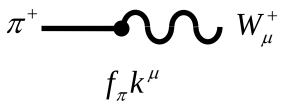


Figure 10-11: The Feynman diagram responsible for changing a pion to a virtual W-boson.

11. The K^+ meson decays *often* exactly the same way as the π^+ meson. The only difference is that the Feynman rule for the coupling is of the form $f_K k^\mu$. Using the formulas from problem 10, find formulas for $\Gamma\left(K^+ \to \mu^+\nu_\mu\right)$ and $\Gamma\left(K^+ \to e^+\nu_e\right)$. By comparing the rates to the experimental values, determine the ratio f_K / f_π. You may find the solution to problem 4-9 helpful.

12. Determine the decay rates $\Gamma\left(Z \to f\,\bar{f}\right)$ for each possible final state fermion. Treat the fermions as massless, except for top, which is too heavy. Find the total decay rate for the Z, and the branching ratio to all neutrinos, any one lepton, and hadrons, and compare to the experimental values.

13. Write the amplitude for $e^+e^- \to \nu_\mu\bar{\nu}_\mu$ in the center of mass frame. Treat the electron as massless, and show that the $q_\mu q_\nu / M_Z^2$ term in the Z-propagator does not contribute. Find the differential and total cross section as a function of s. Do not assume the energy is small compared to the Z-mass, but you can assume you aren't too close to the Z-resonance. At what center of mass energy $\sqrt{s}$ is the cross-section as large as a typical QED cross-section $4\pi\alpha^2/3s$? At infinite energy, how does this weak cross section compare to $4\pi\alpha^2/3s$?

XI. The Standard Model

We mentioned in the previous section that our gauge symmetry works, except for the mass terms. But the charged leptons have mass, as do all the quarks, and we can't just ignore this mass. We also noted that the *W* and *Z* bosons have mass as well. But massive gauge bosons are not gauge invariant. Hence our whole program for trying to write weak interactions as a gauge theory seems to be failing. What shall we do? Let's tackle the lepton masses first.

A. The Higgs Field

First look at eqs. (10.33), reproduced here:

$$i\not{D}\Psi_L - \begin{pmatrix} 0 \\ m \end{pmatrix} \ell_R = 0, \quad i\not{D}\ell_R - \begin{pmatrix} 0 & m \end{pmatrix} \Psi_L = 0.$$

The mass matrices are Hermitian conjugates, as they must be to lead to a Hermitian Hamiltonian. Now, we reach back to section 3H, where we pointed out that a scalar field that is non-zero can, when interacting with a fermion, change that fermion's mass, or in this case, give it mass. Let's imagine there is a new field Φ, called the *Higgs field* that is a scalar and an SU(2)$_L$ doublet, which we will write as

$$\Phi = \begin{pmatrix} \Phi_+ \\ \Phi_0 \end{pmatrix}. \tag{11.1}$$

We will replace the mass with the scalar field Φ, throwing in a coupling that we will name $k\sqrt{2}$, so we have[1]

$$i\not{D}\Psi_L - k\sqrt{2}\Phi\ell_R = 0, \quad i\not{D}\ell_R - k\sqrt{2}\Phi^\dagger\Psi_L = 0. \tag{11.2}$$

If we can somehow cleverly arrange for Φ to have a non-zero value, then this will look just like a mass term. In other words, we want the Higgs field to take the value

$$\Phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{11.3}$$

What do we know about this Higgs field? We can deduce a lot from eq. (11.2). For example, we know that if it is to remain gauge invariant under the U(1)$_Y$ symmetry, the equation should remain unchanged if we multiply each field by $e^{-iY\theta}$, where *Y* is the weak hypercharge, and θ is arbitrary. We know that Ψ_L has $Y = -\frac{1}{2}$, and ℓ_R has $Y = -1$, so to make it work out Φ will have $Y = +\frac{1}{2}$. Similarly, we know from SU(2)$_L$ symmetry that when we perform an SU(2)$_L$ transformation, Ψ_L gets multiplied by $\exp(-i\theta_a T_a)$, while ℓ_R is unchanged, so it must be that Φ also transforms this way. Basically, Φ is an SU(2)$_L$ doublet with hypercharge

---

[1] The factor of $\sqrt{2}$ in this equation and later is related to the difference between real fields and complex ones. Just trust me.

$Y = +\frac{1}{2}$. This allows us, for example, to determine that the Φ's two components have charge $Q = T_3 + Y$, or +1 for Φ_+ and 0 for Φ_0.

Okay, but we haven't solved the problem of how we manage eq. (11.3). To figure this out, we need an expression for the potential energy density $V(\Phi)$ coming from the field Φ. Suppose the potential takes the form

$$V(\Phi) = \tfrac{1}{2}\lambda\left(\Phi^\dagger\Phi - \tfrac{1}{2}v^2\right)^2 . \qquad (11.4)$$

The potential is only a function of $|\Phi| \equiv \sqrt{\Phi^\dagger\Phi}$, and is sketched in Fig. 11-1. Then $V(\Phi)$ can easily be shown to have a minimum when $|\Phi| = v/\sqrt{2}$. Hence eq. (11.3) is, in fact, the minimum of the potential.

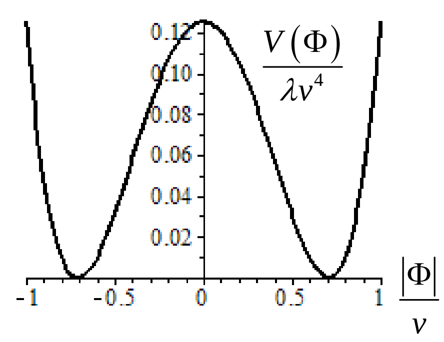


Figure 11-1: The potential $V(\Phi)$ given by eq. (11.4).

Now, eq. (11.4) looks pretty arbitrary, but in fact it is only slightly so. Just as we can't have more than four scalars in a matrix element like $\langle 0|\mathcal{H}|\Phi\Phi\Phi\Phi\rangle$ in a renormalizable theory, the potential $V(\Phi)$ has to be a polynomial of no higher than fourth order in Φ. The only gauge-invariant combinations, to this order, are a constant term, $\Phi^\dagger\Phi$ and $(\Phi^\dagger\Phi)^2$. Since a constant term is irrelevant, there are effectively only two parameters describing a completely general potential $V(\Phi)$. The only arbitrariness involved in eq. (11.4) was to make the $\Phi^\dagger\Phi$ coefficient negative, assuring that $\Phi = 0$ is a local maximum, not a minimum. Since the minimum is at $\Phi \neq 0$, the value of this field is not invariant under the SU(2)$_L$ × U(1)$_Y$ symmetry. We say that this symmetry is *spontaneously broken*.

Now, how can we be sure Φ will end up with a non-zero value in exactly the right spot? We can't. In fact, because of the SU(2)$_L$ × U(1)$_Y$ symmetry, it is guaranteed that you can put *v* into Φ_0 or Φ_+ or in fact a linear combination of the two (even a complex linear combination). But the same symmetry comes to our rescue. Whatever direction nature chooses to put *v* in, we can perform an SU(2)$_L$ × U(1)$_Y$ transformation and move it where we want it. So we'll simply arrange it in the form of eq. (11.3) and go on. Once we substitute eq. (11.3) into (11.2), it will become eq. (10.33), with $m = kv$, and we have solved the problem of lepton masses.

Okay, so it works for one lepton. Will it work for three? All we need is to generalize eq. (11.2) for three generations, so we have

$$i\not{D}(\Psi_L)_A - k_A\sqrt{2}\Phi(\ell_R)_A = 0, \quad i\not{D}(\ell_R)_A - k_A\sqrt{2}\Phi^\dagger(\Psi_L)_A = 0. \qquad (11.5)$$

This works perfectly. We need three different real coupling constants k_A to make the lepton masses come out right. But I want to try to generalize eq. (11.5) even more.

There is no reason you can't do more complicated things in the lepton sector than eq. (11.5). In particular, you could make the Higgs-couplings be off diagonal, or even be complex numbers. The most general form of eq. (11.5) would be

$$i\not{D}(\Psi_L)_A - k_{AB}\sqrt{2}\Phi(\ell_R)_B = 0, \quad i\not{D}(\ell_R)_B - k^*_{AB}\sqrt{2}\Phi^\dagger(\Psi_L)_A = 0. \qquad (11.6)$$

The couplings k_{AB} are now a completely arbitrary and unconstrained complex matrix. In this theory, there are not three real parameters describing the coupling of the Higgs, but eighteen, representing the nine complex numbers in a 3×3 matrix. What a nightmare!

It turns out that eq. (11.6) is not in fact more general than eq. (11.5). An arbitrary square complex matrix of any size can be written as the combination of a real, diagonal, and positive matrix and two unitary matrices. Typically, we would write this as

$$k = U_1^\dagger k_d U_2 \ .$$

where k_d is real and diagonal. Written in terms of components, this would mean

$$k_{AB} = U_{1AC}^\dagger k_C U_{2CB} \ , \tag{11.7}$$

where there is an implied sum on *C*. The k_C numbers in the middle are three positive real numbers. Substitute this into eqs. (11.6), which become

$$i\not{D}\left(\Psi_L\right)_A - \sqrt{2}U_{1AC}^\dagger k_C U_{2CB}\Phi\left(\ell_R\right)_B = 0 \ , \tag{11.8a}$$

$$i\not{D}\left(\ell_R\right)_B - \left(U_{1AC}^\dagger k_C U_{2CB}\right)^* \sqrt{2}\Phi^\dagger\left(\Psi_L\right)_A = 0 \ . \tag{11.8b}$$

Now, multiply eq. (11.8a) by U_{1DA} and eq. (11.8b) by U_{2DB}. We now take advantage of the fact that U_1 and U_2 are unitary matrices. We have

$$\begin{aligned} i\not{D}\left[U_{1DA}\left(\Psi_L\right)_A\right] - \sqrt{2}U_{1DA}U_{1AC}^\dagger k_C U_{2CB}\Phi\left(\ell_R\right)_B = 0, \\ i\not{D}\left[U_{1DA}\left(\Psi_L\right)_A\right] - \sqrt{2}k_D\Phi\left[U_{2CB}\left(\ell_R\right)_B\right] = 0 \ , \end{aligned} \tag{11.9a}$$

$$\begin{aligned} i\not{D}\left[U_{2DB}\left(\ell_R\right)_B\right] - U_{2DB}U_{2BC}^\dagger k_C U_{1CA}\sqrt{2}\Phi^\dagger\left(\Psi_L\right)_A = 0 \ , \\ i\not{D}\left[U_{2DB}\left(\ell_R\right)_B\right] - k_D\sqrt{2}\Phi^\dagger\left[U_{1DA}\left(\Psi_L\right)_A\right] = 0. \end{aligned} \tag{11.9b}$$

If we now define two new wave functions

$$\left(\Psi_L'\right)_D \equiv U_{1DA}\left(\Psi_L\right)_A \ , \quad \left(\ell_R'\right)_D \equiv U_{2DA}\left(\ell_R\right)_A \ ,$$

then eqs. (11.9) will become

$$i\not{D}\left(\Psi_L'\right)_D - \sqrt{2}k_D\Phi\left(\ell_R'\right)_D = 0 \ , \quad i\not{D}\left(\ell_R'\right)_D - k_D\sqrt{2}\Phi^\dagger\left(\Psi_L'\right)_D = 0 \ . \tag{11.10}$$

Thus eqs. (11.6) can always be rewritten in the form of eqs. (11.10). Comparison with eqs. (11.5) now shows us that eqs. (11.5) are in fact completely general. There may, in some abstract sense, be eighteen parameters of the Higgs doublet coupling, but by redefining our left-handed lepton doublets Ψ_L and charged right handed leptons ℓ_R, one can show that only three of them matter. Furthermore, if we knew what *v* is (which we will soon), we could immediately relate those couplings to the three charged lepton masses.

B. Quark Masses from Higgs Couplings

We would like to recreate the successes of the previous section in the quark sector; that is, we'd like to show how the Higgs field can give rise to masses for the quarks. Our task is a bit more difficult this time, because both up and down quarks require mass. We therefore need to include a u_R field as well. Because the up quark requires mass, we will need a field $\tilde{\Phi}$ which has a non-zero background value in its upper components, not its lower component. Let us write down the equivalent of the completely general Higgs doublet couplings analogous to eq. (11.6), so we have

$$i\not{D}\left(q_L\right)_A - f_{AB}\sqrt{2}\tilde{\Phi}\left(u_R\right)_B - h_{AB}\sqrt{2}\Phi\left(d_R\right)_B = 0, \tag{11.11a}$$

$$i\not{D}\left(u_R\right)_B - f_{AB}^*\sqrt{2}\tilde{\Phi}^\dagger\left(q_L\right)_A = 0, \tag{11.11b}$$

$$i\not{D}\left(d_R\right)_B - h_{AB}^*\sqrt{2}\Phi^\dagger\left(q_L\right)_A = 0. \tag{11.11c}$$

First, let's figure out what we can about our two scalar fields Φ and $\tilde{\Phi}$. The only way for eqs. (11.11) to make sense is if both Φ and $\tilde{\Phi}$ are SU(2)$_L$ doublets again. The weak hypercharges of q_L, u_R and d_R are given by $Y = +\frac{1}{6}$, $Y = +\frac{2}{3}$ and $Y = -\frac{1}{3}$ allow us to deduce that Φ has $Y = +\frac{1}{2}$ and $\tilde{\Phi}$ has $Y = -\frac{1}{2}$. Hence Φ has the same SU(2)$_L$ × U(1)$_Y$ structure as it did before, and is the same Higgs doublet we used in section C. But what about $\tilde{\Phi}$? Do we need to introduce a second Higgs-like doublet?

The answer is no. Recall that the anti-particle of an SU(2)$_L$ doublet is still a doublet, but any U(1)$_Y$ charge will get reversed. In fact, the anti-particles corresponding to the field Φ work perfectly to be $\tilde{\Phi}$, and we don't need anything new. Define $\tilde{\Phi}$ in terms of the complex conjugate of Φ, so that

$$\tilde{\Phi} = \begin{pmatrix} \Phi_0^* \\ -\Phi_+^* \end{pmatrix}. \tag{11.12}$$

Then it can be shown that

$$\Phi \to \exp\left(-iT_a\theta_a\right)\Phi \quad \Leftrightarrow \quad \tilde{\Phi} \to \exp\left(-iT_a\theta_a\right)\Phi,$$

which means that it is still transforming as a doublet. Then eqs. (11.11) will be gauge invariant under SU(2)$_L$ × U(1)$_Y$ symmetry. We don't need to introduce any additional fields.

Naively, the matrices *f* and *h* are arbitrary 3×3 complex matrices, and therefore we would need 54 real parameters to describe them. But we already understand from section A that this is an illusion, since we can redefine our fields q_L, u_R and d_R to make many of these parameters disappear. As before, we write *f* and *h* as the product of real and diagonal matrices with two unitary matrices, so we have

$$f = U_3^\dagger f_d U_4, \quad \text{and} \quad h = U_5^\dagger h_d U_6. \tag{11.13}$$

We'd now like to do the same trick we did before. But we have a problem this time, because to diagonalize the up quark masses we want to define $\left(q_L'\right)_D \equiv U_{3DA}\left(q_L\right)_A$, while to diagonalize the down quark masses we want to define $\left(q_L'\right)_D \equiv U_{5DA}\left(q_L\right)_A$. We can't do both, so we simply have to recognize that there will remain one undiagonalized mass. Let's define

$$\left(q_L'\right)_D \equiv U_{3DA}\left(q_L\right)_A \quad \left(u_R'\right)_D \equiv U_{4DA}\left(u_R\right)_A\,, \quad \left(d_R'\right)_D \equiv U_{6DA}\left(d_R\right)_A \tag{11.14}$$

Then eqs. (11.11) will work out to

$$i\not{D}\left(q_L'\right)_A - f_A\sqrt{2}\tilde{\Phi}\left(u_R'\right)_A - \left(U_{3AB}U_{5BC}^\dagger\right)h_C\sqrt{2}\Phi\left(d_R'\right)_C = 0\,, \tag{11.15a}$$

$$i\not{D}\left(u_R'\right)_A - f_A\sqrt{2}\tilde{\Phi}^\dagger\left(q_L'\right)_A = 0\,, \tag{11.15b}$$

$$i\not{D}\left(d_R'\right)_A - h_A\left(U_{5AB}U_{3BC}^\dagger\right)\sqrt{2}\Phi^\dagger\left(q_L'\right)_C = 0. \tag{11.15c}$$

The parameters f_A and h_C are now just three real parameters. If we define the matrix

$$V = U_3U_5^\dagger\,, \tag{11.16}$$

then eqs. (11.15) become

$$i\not{D}\left(q_L'\right)_A - f_A\sqrt{2}\tilde{\Phi}\left(u_R'\right)_A - V_{AC}h_C\sqrt{2}\tilde{\Phi}\left(d_R'\right)_C = 0\,, \tag{11.17a}$$

$$i\not{D}\left(u_R'\right)_A - f_A\sqrt{2}\tilde{\Phi}^\dagger\left(q_L'\right)_A = 0\,, \tag{11.17b}$$

$$i\not{D}\left(d_R'\right)_A - h_AV_{AC}^\dagger\sqrt{2}\Phi^\dagger\left(q_L'\right)_C = 0. \tag{11.17c}$$

I hope it is clear from eqs. (11.17a) that the purpose of *V* is to figure out which particular combination of the q_L''s is paired with the d_R''s to make a mass term for the down-type quarks. In fact, the only place where *V* will become relevant is when you are connecting *d*'s with the *u*'s by interacting with the *W*-bosons, because this is the only time we care about which combination of down quarks is paired with each up quark. In fact, *V* is nothing other than the CKM matrix encountered in section 10G.

At the moment, it looks like *V* is a completely arbitrary unitary 3×3 matrix, which one can show still has no less than nine real parameters to describe it. However, there is still a bit of freedom we can choose to use to redefine our fields without messing up the form of eqs. (11.17a). If we define yet another set of fields by multiplying by some arbitrary phases, so that

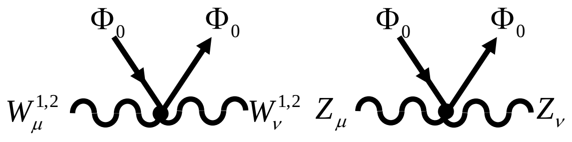
$$\left(q_L''\right)_A = e^{i\alpha_A}\left(q_L'\right)_A\,, \quad \left(u_R''\right)_A = e^{i\alpha_A}\left(u_R'\right)_A\,, \quad \left(d_R''\right)_A = e^{i\beta_A}\left(d_R'\right)_A\,,$$

then we find that eqs. (11.17) are unaltered, except that a lot of phases change in *V*. This allows one to simplify the CKM matrix so it has only a single phase, and one can choose to put it in the form of eq. (10.61).

The bottom line is that if you introduce the Higgs doublet, give it a potential of the form eq. (11.4) (which is almost completely general), it will automatically acquire a non-zero value throughout space, and if you then introduce the most general possible scalar couplings into the Dirac equation for the leptons and quarks, it will automatically yield mass terms of the form of eq. (11.5) for the leptons and of the form of eq. (11.17) for the quarks, including a generic CKM matrix of the form of eq. (10.53). As far as we can tell, all possible interactions that could be there are there. We have succeeded in figuring out where the masses of all the fermions come from, as well as understanding all their weak interactions.

C. Gauge Boson Masses

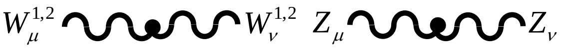
Having successfully given the fermions masses, what about the *W* and *Z* masses? Consider for the moment the gauge couplings of the Φ itself, specifically, the Φ_0 portion of the field. We haven't talked much about gauge couplings of scalars, but if you look back at section 7F, specifically Fig. 7-7, you see that in electromagnetism, charged scalars have an interaction with the photon. The field Φ_0 has $Q = 0$, but since it is a part of an SU(2) doublet, it will have couplings to W^1, W^2, and *Z*. The couplings we are interested in at the moment are those that connect Φ_0 with two gauge bosons. These are sketched in Fig. 11-2, with what should be the corresponding rule.



$$\frac{2ie^2 g^{\mu\nu}}{4\sin^2\theta_W} \qquad \frac{2ie^2 g^{\mu\nu}}{4\sin^2\theta_W\cos^2\theta_W}$$

Figure 11-2: The Feynman diagrams with two gauge bosons and two Φ_0 's.

And now, we recall from section 5D that if we are perturbing around the wrong minimum, we can replace any external scalar line with the value of the field about which we should be perturbing and then get rid of the external lines. This means eliminating the Φ_0 lines from Fig. 11-2, and multiplying by two factors of $v/\sqrt{2}$. This leads to two new rules, sketched in Fig. 11-3.



$$\frac{ie^2 v^2 g^{\mu\nu}}{4\sin^2\theta_W} \qquad \frac{ie^2 v^2 g^{\mu\nu}}{4\sin^2\theta_W\cos^2\theta_W}$$

Figure 11-3: The gauge boson mass terms the result from replacing the external Φ_0 lines in Fig. 11-2 with $v/\sqrt{2}$.

And what do these new interactions mean? We have found for both scalars and fermions that a term with just two lines on it corresponds to a mass term, and that's exactly what happens here. The $W^{1,2}$ masses become the $W^\pm$ masses, and we find

$$M_W^2 = \frac{e^2 v^2}{4\sin^2\theta_W} \quad \text{and} \quad M_Z^2 = \frac{e^2 v^2}{4\sin^2\theta_W\cos^2\theta_W}, \tag{11.18}$$

from which eq. (10.42), $M_W = M_Z\cos\theta_W$, is trivial. From eq. (10.39b) we see that $e/\sin\theta_W = g$, and then using eq. (10.21) we have

$$G_F = \frac{g^2}{4\sqrt{2}}\frac{4}{g^2 v^2} = \frac{1}{v^2\sqrt{2}}. \tag{11.19}$$

We can then substitute the explicit value $G_F = 1.16637\times 10^{-5}\ \text{GeV}^2$ and find

$$v = 246.2\ \text{GeV}. \tag{11.20}$$

Since G_F determines the muon decay rate, the value of *v* was one of the first aspects of this theory that was worked out. We already found the weak mixing angle θ_W, given in eq. (10.50). We then can find all the Higgs boson couplings to fermions; they are always m_f/v. But before we think we are done with the Standard Model, there is one more particle we should talk about.

D. The Higgs Boson

Up to now we have been treating the Higgs field as a classical field, much as in chapter 3 we treated the electromagnetic field as classical. Indeed, we have given it even less consideration, since we assumed the Higgs field is always at its minimum point given by eq. (11.3). There is no reason to assume this is always true. Let us rewrite eq. (11.3) including the possibility that Φ is not necessarily at its minimum, so that

$$\Phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \tag{11.21}$$

where H is a real field. Then the potential eq. (11.4) can be rewritten as

$$V(H) = \tfrac{1}{8}\lambda\left(H^2 + 2Hv\right)^2 = \tfrac{1}{2}\lambda v^2 H^2 + \tfrac{1}{2}\lambda v H^3 + \tfrac{1}{8}\lambda H^4 . \tag{11.22}$$

The first term in eq. (11.22) is a mass term for the resulting particle, which we will call the *Higgs boson*. The remaining terms describe self-interactions of the Higgs. Its mass is given by

$$m_H^2 = \lambda v^2 . \tag{11.23}$$

Hence the way to measure the coupling λ is to simply measure the Higgs mass. Until 2012, the Higgs mass was unknown, but on July 4, 2012, the Large Hadron Collider announced the discovery of a particle with a mass of about 126 GeV consistent with the Higgs boson. Though we don't know for sure that it *is* the Higgs boson, most physicists believe it is.

But wait! In eq. (11.21), why did we only allow Φ_0, the second component of Φ, to vary? Come to think of it, why did we not allow H to be complex? Logically, there should be three other degrees of freedom, the $H^\pm$ with electric charge ± 1 and another particle H' corresponding to the imaginary part of Φ_0. Did we forget about these?

The answer is no, but it gets a little complicated. Recall when we found the minimum of Φ, we argued that all possible "directions" were equally possible, but we could make a gauge transformation in such a way to assure that the minimum was in the real Φ_0 direction. Because gauge transformations are local, that is, they can be performed by different angles/amounts at every point in spacetime, we can do this point by point, and hence demand that Φ takes the form of eq. (11.21). So there will not be particles H' and $H^\pm$. The only actual particle coming from the Higgs doublet is the Higgs particle H.

So, what became of the degrees of freedom associated with the three fields we got rid of by performing our gauge transformation? Did they disappear into thin air? No, they got *eaten*.[1] Recall that the $W^\pm$ and Z fields are massive, and when we discussed massive vector bosons in section 10C, we pointed out that they have *three* polarizations, not two. Basically, the three particles H' and $H^\pm$, which ordinarily would be three new particles, are eaten by the $W^\pm$ and Z, allowing them to acquire a third polarization, which represents the lost degrees of freedom of these particles. And when you eat something, you gain mass.

At this point, it remains only to work out the Feynman rules for diagrams including the Higgs boson. The Higgs boson is usually denoted by a dashed line. Because the Higgs boson is a neutral scalar boson, the propagator was worked out back in section 5B.

---

[1] Yes, the term *eaten* is the actual term serious physicists use to explain this process.

Because the Higgs field is responsible for all the fermion masses, it is not surprising that all fermions with mass couple to the Higgs particle. The coupling is just a factor of m_f/v. The massive gauge bosons also couple to the Higgs. Two of the relevant diagrams come from Fig. 11-2. Two more can be deduced by replacing a single external Higgs line with its background value *v*. Finally, the Higgs field couples to itself, as can be seen from the potential eq. (11.22). A complete list of the relevant interactions for the Higgs can be found in Fig. 11-4.

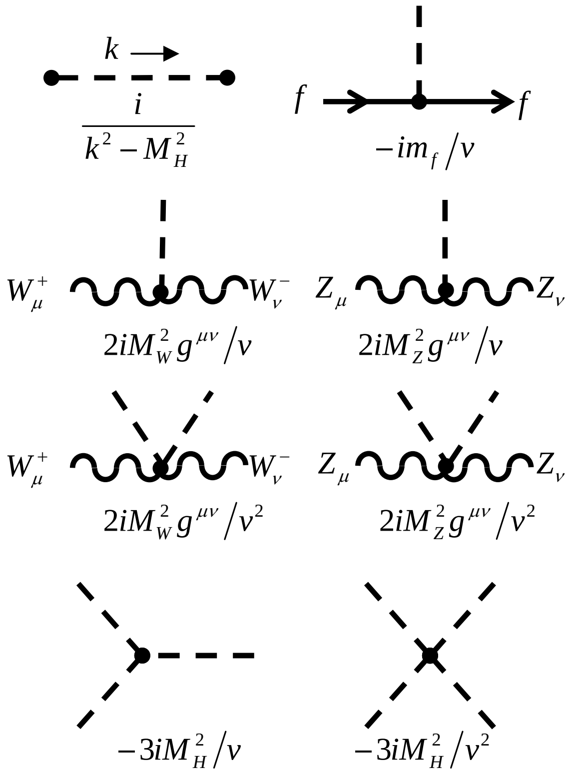


Figure 11-4: The Feynman diagrams involving the Higgs particle.

One of the things you notice is that the Higgs boson couplings generally are proportional to the masses of the particles it is coupling to. This makes it very hard to produce, because the objects we use in colliders (electrons and protons) are made of some of the lightest elementary particles (electron, up quark, down quark, gluons). This means you have to use a great number of collisions to produce the particle. One of the best processes for producing it is in electron-positron colliders, via the process $e^+e^- \to ZH$. Treating the electron as essentially massless, there is only one tree-level contribution, as sketched in Fig. 11-5. Unfortunately, the Large Electron-Positron collider (LEP) only worked up to energies of $\sqrt{s} = 209\ \text{GeV}$, which allowed detection up to about 115 GeV, about 10 GeV too low for LEP to make the discovery.

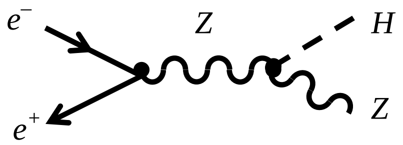


Figure 11-5: How to make a Higgs at a lepton collider.

The Large Hadron Collider collides protons. A variety of processes contribute to the production of Higgs bosons. Probably the most important is gluon fusion, sketched in Fig. 11-6. The Higgs, once produced, is most likely to decay to the particle with the largest mass light enough to appear as a decay product. This is most likely to be bottom quarks for a mass $m_H \approx 126\ \text{GeV}$. But bottom quarks are being produced in copious quantities from strong interactions, such as $gg \to b\bar{b}$, so this decay is tough to spot against the background. The most important decay that contributed to the 2012 discovery of the Higgs is the decay $H \to \gamma\gamma$, which also proceeds via a loop.

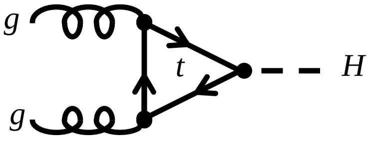


Figure 11-6: Gluons fuse via a virtual top quark loop to make the Higgs.

E. $\mathcal{CP}$ Violation and the Neutral Kaons

We already commented that weak interactions violate parity $\mathcal{P}$; for example, the W interaction only couples to the left-handed neutrinos and right-handed anti-neutrinos. It follows that charge conjugation $\mathcal{C}$ is also violated, since the anti-particle of a left-handed neutrino is a left-handed anti-neutrino, which if it exists at all does not have weak interactions. But we haven't discussed whether perhaps their combination, $\mathcal{CP}$, might be a symmetry of nature. By the $\mathcal{CPT}$ theorem, $\mathcal{CPT}$ is *always* a symmetry of nature, so if $\mathcal{CP}$ is a symmetry of nature, so will be time reversal $\mathcal{T}$. These are the only remaining symmetries that might still be valid. But how experimentally can we tell?

It turns out a particularly fruitful way to study this question is the neutral kaons, $\left|K^0\right\rangle = \left|d\bar{s}\right\rangle$ and $\left|\bar{K}^0\right\rangle = \left|s\bar{d}\right\rangle$. If we have one of these states at rest, the expectation value of the Hamiltonian would be the mass, so

$$\left\langle K^0\middle|H\middle|K^0\right\rangle = M \; . \tag{11.24}$$

By the anti-particle property, we can shift each of these to the other side and deduce

$$\left\langle \bar{K}^0\middle|H\middle|\bar{K}^0\right\rangle = M \; . \tag{11.25}$$

The largest contributors to eqs. (11.24) and (11.25) will be strong interactions, but the principle applies in general, even if we take into account weak interactions or any hypothetical other interactions.

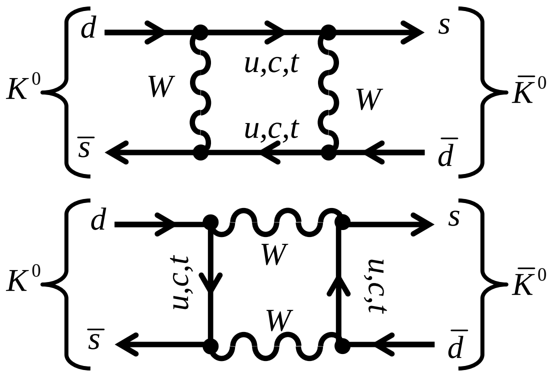


Figure 11-7: The two diagrams responsible for K^0 mixing with $\bar{K}^0$

Now, what about off-diagonal terms, like $\left\langle \bar{K}^0\middle|H\middle|K^0\right\rangle$. Will this be non-zero? Note that the strangeness of the two sides differs by two units. Only W-couplings can change strangeness, and only by one unit, so if this is non-zero at all, it would require at least two W-couplings. There *are* diagrams that cause this, such as the two box diagrams in Fig. 11-7, but because they are doubly weak, they will be small. So we write the resulting matrix elements as

$$\left\langle \bar{K}^0\middle|H\middle|K^0\right\rangle = \Delta \, , \quad \left\langle K^0\middle|H\middle|\bar{K}^0\right\rangle = \Delta^* \; . \tag{11.26}$$

They must be complex conjugates of each other because H is Hermitian.

It is not hard to see how $\mathcal{C}$, which turns anti-particles into particles, will affect each of these particles:

$$\mathcal{C}\left|K^0\right\rangle = \mathcal{C}\left|d\bar{s}\right\rangle = \left|\bar{d}s\right\rangle = -\left|s\bar{d}\right\rangle = -\left|\bar{K}^0\right\rangle , \quad \mathcal{C}\left|\bar{K}^0\right\rangle = \mathcal{C}\left|s\bar{d}\right\rangle = \left|\bar{s}d\right\rangle = -\left|d\bar{s}\right\rangle = -\left|K^0\right\rangle .$$

Keeping in mind that these particles are pseudoscalars, we then have

$$\mathcal{CP}\left|K^0\right\rangle = \left|\bar{K}^0\right\rangle, \quad \mathcal{CP}\left|\bar{K}^0\right\rangle = \left|K^0\right\rangle. \tag{11.27}$$

Keeping in mind that $\mathcal{C}$ and $\mathcal{P}$ are both their own Hermitian conjugate (basically, because doing either of them twice takes you back to where you started), we can also conclude that

$$\left\langle K^0\right|\mathcal{CP} = \left\langle \bar{K}^0\right|, \quad \left\langle \bar{K}^0\right|\mathcal{CP} = \left\langle K^0\right|. \tag{11.28}$$

Now, suppose $\mathcal{CP}$ is as symmetry of nature, and commutes with the Hamiltonian. Then

$$\Delta^* = \left\langle K^0\middle|H\middle|\bar{K}^0\right\rangle = \left\langle K^0\middle|H\mathcal{CP}\middle|K^0\right\rangle = \left\langle K^0\middle|\mathcal{CP}H\middle|K^0\right\rangle = \left\langle \bar{K}^0\middle|H\middle|K^0\right\rangle = \Delta\,, \tag{11.29}$$

so that Δ is real. In contrast, a complex value of Δ would imply that there is $\mathcal{CP}$ violation going on. This is a hallmark of $\mathcal{CP}$ violation – the presence of complex terms in the Hamiltonian.

Now, let's look at the Hamiltonian for the $\left\{\left|K^0\right\rangle,\left|\bar{K}^0\right\rangle\right\}$ system. As a matrix, this would look like

$$H = \begin{pmatrix} M & \Delta \\ \Delta & M \end{pmatrix}. \tag{11.30}$$

This matrix has two eigenstates,

$$\left|K_1\right\rangle = \tfrac{1}{\sqrt{2}}\left(\left|K^0\right\rangle + \left|\bar{K}^0\right\rangle\right), \quad \left|K_2\right\rangle = \tfrac{1}{\sqrt{2}}\left(\left|K^0\right\rangle - \left|\bar{K}^0\right\rangle\right). \tag{11.31}$$

These states have masses $M \pm \Delta$. So in fact, the states $\left|K^0\right\rangle$ and $\left|\bar{K}^0\right\rangle$ would *not* actually be the particles we would observe, instead we would observe the particles $\left|K_1\right\rangle$ and $\left|K_2\right\rangle$. These states are the eigenstates of $\mathcal{CP}$ with eigenvalues

$$\mathcal{CP}\left|K_1\right\rangle = +\left|K_1\right\rangle, \quad \mathcal{CP}\left|K_2\right\rangle = -\left|K_2\right\rangle. \tag{11.32}$$

Not only would we expect there to be two very slightly different mass particles, they would have rather different decays. Both must decay weakly (which means they are slow), but they often decay to pions. It isn't hard to see that states with two pions, $\left|\pi^0\pi^0\right\rangle$ or $\left|\pi^+\pi^-\right\rangle$, have eigenvalue +1 under $\mathcal{CP}$ symmetry, so only the state $\left|K_1\right\rangle$ can decay to them. The state $\left|K_2\right\rangle$ would have to do something more complicated, like decaying to three pions, but there is barely enough energy to make this happen, so it would be much slower. Hence we predict that the neutral kaons appearing in the table on page 121 are in fact a sham. There should be two neutral kaons, one of which lasts a short time and usually decays to two pions, and the other of which lasts a much longer time and never decays to two pions. How does this compare to reality?

In fact, there are two neutral kaons, which are called K_L and K_S (K-long and K-short). The K_L is heavier by about 3.5×10^{-6} eV. The K_S decays quickly, almost always to two pions. The K_L lasts about 570 times longer. It decays in a variety of ways, three pions among them, but about 0.2% of the time, it decays to two pions. It seems that K_S and K_L are almost exactly K_1 and

K_2, but not quite. The fact that K_L occasionally decays the wrong way tells us there must be $\mathcal{CP}$ violation in the Standard Model.

Where is that violation? Recall that the sign of $\mathcal{CP}$ violation is complex numbers in the Hamiltonian. If you look through all our Feynman diagrams, keeping in mind that Feynman rules come from multiplying Hamiltonian matrix elements by $-i$, you may eventually find it. In Fig. 10-9, the quark couplings to the W contain a factor from the CKM matrix, which for three generations takes the form of eq. (10.61) and contains complex numbers. It is not hard to show that the same would not occur if we only had two generations; in this case, the corresponding matrix (which we would call the Cabibbo matrix) would have only one real parameter, the Cabibbo angle. Hence such $\mathcal{CP}$ violation would not occur in the Standard model with two generations. In fact, $\mathcal{CP}$ violation in the kaon system was discovered before the bottom and top quarks (as well as the tau and its neutrino) and on this basis Kobayashi and Maskawa predicted the existence of the third generation. The smallness of this $\mathcal{CP}$ violation is evident theoretically because it requires the involvement of the third generation, but these mixings are very small. It is manifested experimentally in that the K_L only rarely decays to two pions.

The phase in the CKM matrix is the only source of $\mathcal{CP}$ violation in the Standard Model, but it is always possible that some other contribution is actually causing this effect. Research into this topic is ongoing. A very similar phenomenon occurs with the $\left|B^0\right\rangle = \left|b\bar{d}\right\rangle$ and $\left|\bar{B}^0\right\rangle = \left|d\bar{b}\right\rangle$ mesons, also under intense scrutiny.

Violation of $\mathcal{CP}$ symmetry may have played a crucial role in the existence of, well, you, for example. It is believed that in the very early universe, there was a very hot mixture of quarks and anti-quarks, with approximately or even exactly equal numbers. If they *started* in equal numbers, they would continue to this day in equal numbers, which means virtually all matter/anti-matter would have annihilated, *unless* there is some sort of interaction that treats particles and anti-particles differently. Such an interaction must violate $\mathcal{CP}$ symmetry. It is possible but not likely that the $\mathcal{CP}$ violation in the CKM matrix is the ultimate source of this discrepancy. But any interaction that changes the matter/anti-matter mix would also have to violate baryon number, and there is no baryon number violating interactions in the Standard Model. So either our idea of equal amounts of matter/anti-matter in the early universe is wrong, or there are new sources of $\mathcal{CP}$ violation beyond the CKM matrix.

F. The Standard Model Summarized

As a particle physicist, I would like to look at the standard model and ask: what is the simplest possible description of the particles and interactions that exist? To what extent are the interactions we see inevitable, and to what extent are they arbitrary? Now is a good time to give that summary.

The first step is to describe the gauge group, which is

$$\mathrm{SU}\left(3\right)_C \times \mathrm{SU}\left(2\right)_L \times \mathrm{U}\left(1\right)_Y . \tag{11.33}$$

The subscripts aren't really necessary, they are just a way of keeping track of the fact that SU(3) is associated with color, that SU(2) only affects the left-handed fermions, and U(1) is associated with weak hypercharge *Y*.

The next thing to list would be a complete list of the particles, and how they transform under these groups. There is just one scalar field, the Higgs, which does not have color, which we indicate by saying it is in the 1 (or trivial) representation of the color group, but it is an SU(2)$_L$ doublet and has $Y=+\frac{1}{2}$. We list this as:

$$\text{scalars:}\quad \left(1,2,+\tfrac{1}{2}\right). \tag{11.34}$$

Now we do the fermions. There is the left-handed quark doublet q_L and the right-handed singlets u_R and d_R, all of which are color triplets. There is the left handed lepton doublets Ψ_L and the right-handed charged lepton ℓ_L, all of which are color singlets. This pattern is repeated three times. So we could summarize this as

$$\begin{aligned}\text{left-fermions:}\quad & 3\left[\left(3,2,\tfrac{1}{6}\right)\oplus\left(1,2,-\tfrac{1}{2}\right)\right]\\ \text{right-fermions:}\quad & 3\left[\left(3,1,+\tfrac{2}{3}\right)\oplus\left(3,1,-\tfrac{1}{3}\right)\oplus\left(1,1,-1\right)\right].\end{aligned} \tag{11.35}$$

Given the gauge group eq. (11.33), the scalar content (11.34), and the fermion content (11.35), we then include every possible gauge-invariant, Lorentz-invariant, and renormalizable interaction possible. It turns out that there are only a finite number of possible interactions, and it then remains to specify the values that determine these interactions. There will be one coupling for each of the gauge groups: g_S, g, and g', though we usually exchange these for α_S, α, and $\sin^2\theta_W$ instead. Then we have the potential for the Higgs, eq. (11.4), with the two parameters v and λ, but I would trade λ for m_H. Next come the Higgs couplings to the nine massive quarks and charged leptons, which I would trade for those masses. And finally, we have the three generation mixing angles and one phase that appears in the CKM matrix, eq. (10.61). This totals eighteen real independent parameters:

$$\alpha_S,\alpha,\sin^2\theta_W,v,M_H,m_e,m_\mu,m_\tau,m_u,m_c,m_t,m_d,m_s,m_b,\theta_{12},\theta_{13},\theta_{23},\delta. \tag{11.36}$$

And thus we are done, and have completely specified the Standard Model.

It is entirely a matter of convention which objects we label "particles" and which "anti-particles." Recall that the anti-particle has the opposite helicity of the particle, so the anti-particle of a right-handed up quark would be left-handed. We could thereby move any particles in the right-fermion list of eq. (11.35) into the left-fermion list, or vice versa. An anti-particle would have the opposite hypercharge. The anti-particles in a doublet or "2" representation would still be a doublet, but the anti-particles of a "3" representation would be the "$\overline{3}$" representation. It has become conventional, for no particular reason, to list only left-fermions. Then we would summarize eqs. (11.33), (11.34) and (11.35) as

$$\begin{aligned}\text{gauge group:}\quad & \mathrm{SU}(3)\times\mathrm{SU}(2)\times\mathrm{U}(1).\\ \text{scalars:}\quad & \left(1,2,+\tfrac{1}{2}\right),\\ \text{left-fermions:}\quad & 3\left[\left(3,2,\tfrac{1}{6}\right)\oplus\left(\overline{3},1,-\tfrac{2}{3}\right)\oplus\left(\overline{3},1,+\tfrac{1}{3}\right)\oplus\left(1,2,-\tfrac{1}{2}\right)\oplus\left(1,1,+1\right)\right].\end{aligned} \tag{11.37}$$

In principle, one could communicate with a hyper-intelligent alien in an alternate universe, and having already explained that the universe has four dimensions of space-time, could then communicate eq. (11.37) to such an alien, and he would immediately respond by asking what the value of the eighteen relevant parameters are. This is, in a sense, the minimal possible description of the universe around us.

G. Conservation Laws in the Standard Model

Before we go on, let me discuss a couple of conservation laws that occur in the Standard Model. In addition to conservation of four-momentum and angular momentum, which are guaranteed by the way we made our interactions Lorentz invariant and translation independent, and electric charge and color, which are automatically assured by the unbroken SU(3)$_C$ and U(1)$_Q$ symmetries, there are several other quantities that are also automatically conserved. The first is baryon number. We don't think of baryons as elementary particles any more, but remember that quarks have $B = \frac{1}{3}$. There is simply no way to write down a gauge-invariant and renormalizable interaction that changes the number of quarks minus anti-quarks. It isn't terribly hard to write down *non*-gauge invariant interactions, such as the one sketched in Fig. 11-8. Such an interaction would cause proton decay through processes like $p^+ \to e^+\pi_0$, a process that is known to have a mean lifetime larger than about 10^{34} years. Like the four-fermion interactions in the Fermi theory of weak interactions, such a term would have a coefficient with units of M^{-2}. Apparently renormalizability is saving us from an otherwise disastrous instability of matter. A symmetry or conservation law that is not imposed, but comes about because of the demands of gauge invariance and renormalizability is called an *accidental symmetry*.

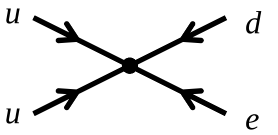


Figure 11-8: A non-renormalizable four fermion interaction that would lead to proton decay.

Baryon number is not the only accidental symmetry of the Standard Model. If you look at the interactions of one of the leptons, say the electron, you see that the only thing that can increase or decrease the number of electrons (minus anti-electrons) is the *W*-coupling, but this always creates a ν_e neutrino. If we define *electron family number* L_e as the total number of electrons plus electron neutrinos (minus anti-particles), then L_e is also conserved. Two more lepton family numbers, L_μ and L_τ, are also conserved, as are any combination of these. Hence there are four independent accidental symmetries of the Standard Model. As we will discover very soon, however, most of these are *not* actual conservation laws. Baryon number *B* is probably pretty well conserved, and it is possible that total lepton number $L = L_e + L_\mu + L_\tau$ is conserved, but the separate lepton family numbers are not conserved. But this leads us beyond the limits of the Standard Model.

And now – it's time for me to come clean on one big lie. I have been doing everything in the context of Feynman diagrams, a strictly perturbative approach. Perturbation theory includes effects that can be written as polynomials in the coupling constants. But there are known processes in the Standard Model that are non-perturbative. One of them is a configuration of the SU(2)$_L$ fields called a *sphaleron*. It is a little difficult to understand (at least for me), but a sphaleron is a gauge and Higgs field configuration which reshuffles the energies of all the fermions that are affected by SU(2)$_L$, causing them all to shift slightly in energy. Think of all the states like a hotel with an infinite number of rooms, with all the negative room numbers (negative energy states) filled and all the positive room numbers (positive energy) empty, as artistically rendered in Fig. 11-9. Now we ask every occupant to shift one room to the right, increasing their room number. This will create an occupied positive energy state without a corresponding negative energy state – we have just created a particle! It does not violate conservation of energy, since the field configuration that works this magic requires energy. It is similarly possible to have an anti-sphaleron, which is like shifting everyone one room to the left, and creates one anti-particle.

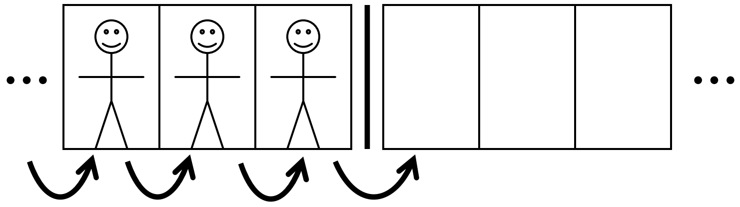


Figure 11-9: In an infinite hotel, suppose you start with all the negative room numbers occupied, and all the positive room numbers unoccupied. If everyone shifts one room to the right, one of the positive rooms can become occupied without creating a negative room vacancy.

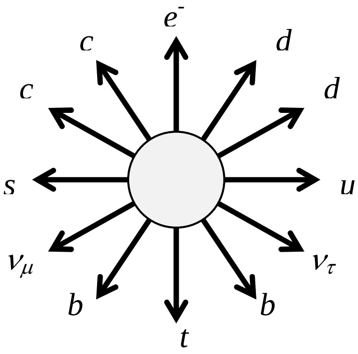


Figure 11-10: One of several possible particle combinations a sphaleron can create.

Now, for an SU(2)$_L$ sphaleron, it can be shown that you actually create one particle for each SU(2) doublet, so you get nine quarks (three families times three colors) and three leptons from a sphaleron. Each "particle" coming out can alternatively be thought of as an anti-particle going in. Fig. 11-10 shows one of many possible combinations of particles that can be produced. At low energies, the probability of such an interaction occurring is very small, suppressed by a factor of roughly

$$\Gamma \sim \exp\left(-16\pi^2/g^2\right) = \exp\left(-4\pi \sin^2 \theta_W/\alpha\right) \sim 10^{-173} . \tag{11.38}$$

This is way too slow to have any practical significance. At high temperatures, such as existed in the early universe, the exponential suppression does not exist, and it is believed that this process can occur.

If you look at the particles produced or annihilated by a sphaleron, you can see that it creates nine quarks, or three baryons, and one lepton from each of the three families. Hence baryon number and each of the three lepton family numbers are not conserved. But the following three combinations are conserved:

$$B - L , \quad L_e - L_\mu , \quad L_\mu - L_\tau . \tag{11.39}$$

Of course, the combination $L_e - L_\tau$ is also conserved, but this is just the sum of the last two. When we include neutrino masses, as we will in chapter twelve, only the combination $B - L$ has any chance of surviving.

H. Naturalness and Flexibility of the Standard Model

The Standard Model has been remarkably successful, but how sure are we that it is right? Is there anything inevitable about it? How much flexibility can we give it, and still have it match experiment?

Consider first the assignment of fermions to the gauge group, as given in eq. (11.37). We know all the particles listed exist, because we see them all. Could we add just one more? Suppose we were so foolish as to add just one more copy of, say, $\left(1,2,-\frac{1}{2}\right)$, a new left-handed lepton doublet, but without the corresponding right-handed charged lepton singlet. Recall that the Higgs field combines the left-handed doublet with the right-handed singlet to give it a mass. Since there are only three right-handed singlets, one of the left-handed doublets will be left over. By analogy, if you have a town with four men and three women, one of the men is going to end up a bachelor. So clearly this won't work. Similarly, you can't just add another left-handed quark doublet or right-handed quark singlet without conflicting with experiment.

So, maybe we could add $\left(1,2,-\frac{1}{2}\right)\oplus\left(1,1,+1\right)$, which would lead to a new charged lepton and a new neutrino. In this four-lepton model of the standard model, the neutrinos would still all be massless, and therefore there would be a new channel for *Z*-decay, $Z\to\nu_4\bar{\nu}_4$. This would substantially increase the total decay rate, or width, of the *Z*, and would have been detected in the Stanford Linear Collider (SLC) or the Large Electron Positron Collider (LEP). So we know that's not right either.

How about the quarks? Suppose we add $\left(3,2,+\frac{1}{6}\right)\oplus\left(\overline{3},1,-\frac{2}{3}\right)\oplus\left(\overline{3},1,+\frac{1}{3}\right)$. This would yield another up type quark and down type quark, and if they are heavy enough, we might not yet have detected them. It turns out this doesn't work, and it has to do with renormalizability. There is a loop diagram involving gauge couplings of fermions as sketched in Fig. 11-11, which is called the triangle diagram. It turns out that this diagram sometimes yields infinity, and the infinity cannot be eliminated by the usual techniques of renormalization. The infinity is multiplied by a factor that depends on which type of gauge boson is coming in. For the case of the B_μ, associated with hypercharge, the infinity is multiplied by

$$\sum_f Y_f^3\,, \qquad (11.40)$$

where the sum is taken over all left-handed fermions (if you use right-handed fermions, like from eq. (11.35), you have to subtract them). Fortunately, by an amazing coincidence, this factor works out to exactly zero (see problem 11.5), but only if you have an equal number of generations of quarks and leptons. This is our only understanding of why there are equal numbers of quarks and leptons. Eq. (11.40) is simply one of many so-called *anomalies* that must vanish. There is a similar expression for any combination of three gauge bosons in the standard model, though many of them automatically vanish. We say that the Standard Model is anomaly free.

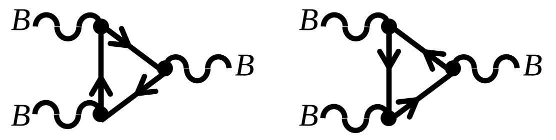


Figure 11-11: These triangle diagrams yield nonsensical infinite results unless eq. (11.40) vanishes.

Hence there really is less flexibility than there might seem in the assignments in eq. (11.37). Adding *just* a generation of quarks, but not a corresponding generation of leptons, causes eq. (11.40) to not vanish, and the standard model would have anomalies. Hence the number of generations has to match. We normally consider the four particles e^-,ν_e,u,d the first generation, with μ^-,ν_μ,c,s the second generation, but these associations are completely arbitrary. Though we understand that there have to be an equal number of generations, we do not understand why there are three generations, rather than four or two.

How about in the scalar sector? Could we add another Higgs field that looks like $\left(1,2,+\frac{1}{2}\right)$? The answer is yes. Anomalies do not arise from scalar particles, so such a new particle could certainly exist. In supersymmetry, for example (see section 12C), the fields Φ and $\tilde{\Phi}$ of section 11B are actually separate fields. Since each of them is a complex doublet, there are eight real degrees of freedom, but three of them get "eaten," leaving behind two scalar fields (like the Higgs and another Higgs), a neutral pseudoscalar, and a charged pair $H^\pm$. None of these other particles have been discovered, but eight days before I typed this, neither was the Higgs boson. In fact, it remains to be seen if the recently discovered particle around 126 GeV is truly the simplest Standard Model Higgs boson, or one of these several Higgs-like particles, or something else entirely. However, strictly speaking, any extension of the Higgs sector means we are going beyond the Standard Model

One thing we decidedly *don't* understand is what determines all the parameters in the Standard Model. Many of them seem to be of order one, such as the three gauge couplings, the dimensionless parameter λ in the Higgs potential, and the coupling of the Higgs to the top quark. Many seem to be small, such as the other eight Higgs couplings (manifested by the fact that $m_f \ll v$), and the three angles mixing the generations of quarks. The smallness of masses will be even more acute when we discuss neutrino masses in the next chapter. We think the smallness of these parameters must be telling us something, but we don't know what it is. It is just one of the puzzles that may help us to someday go beyond the Standard Model. Let us turn to some ideas of how to go beyond the Standard Model.

Problems for Chapter 11

1. Find a formula for the rate of Higgs decay to leptons $H \to \ell^+\ell^-$ or any quark $H \to \bar{q}q$, for arbitrary Higgs mass M_H and arbitrary fermion mass *m*. Don't forget colors, when appropriate. Why did I leave out the neutrinos? For the actual Higgs mass (126 GeV), determine which rate(s) dominate, and estimate the total decay rate in MeV. We currently can resolve the particles produced with an accuracy of about 1 GeV. Are we close to directly measuring the width Γ for the Higgs? You may find problem 6.4 useful.

2. Find a formula for the rate of Higgs decay $H \to W^+W^-$ and $H \to Z^0Z^0$ as a function of Higgs mass, assuming it is heavy enough for these to occur. Do not neglect any of the masses.

3. In Fig. 11-5, I drew only one Feynman diagram for the process $e^+e^- \to Z^0H$. Draw the other diagrams. Explain why they aren't relevant. Than write the Feynman amplitude for the relevant diagram. Simplify it insofar as possible. You don't have to do more.

4. The Higgs boson was actually discovered by the decay $H \to \gamma\gamma$. This process is impossible by tree diagrams. Draw some loop diagrams, including at least one with no fermion loop, that would contribute to this process. Speculate which diagram(s) you think might be most important.

5. It is sometimes said that a non-zero δ is what causes CP violation, because it makes the CKM matrix eq. (10.61) complex. It isn't quite this simple.
   (a) Argue that the CKM matrix is real also if $\delta = \pi$ or if $s_{13} = 0$.
   (b) The Standard Model is unchanged if the CKM matrix is multiplied on the left or right by just diagonal phases, for example, we can change it to $V \to U^* V U$, where
   $$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{i\delta} \end{pmatrix}, \quad U^* = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-i\delta} \end{pmatrix}.$$
   show that if $s_{23} = 0$, this will make the CKM matrix V change to a real matrix.
   (c) Find a different diagonal matrix U such that if $s_{12} = 0$, then $V \to U^* V U$ will make the CKM matrix V come out to a real matrix.

6. There are actually multiple quantities that have to vanish to avoid anomalies. Among them are:
   $$\sum Y^3, \quad \sum Y, \quad \sum Y T_3^2,$$
   where the sums are taken over all left-handed fermions, Y is the hypercharge, and T_3 is the SU(2)$_L$ operator.
   (a) Check that each of these vanishes for the standard model, eq. (11.37)
   (b) If you add just one more set of quarks $\left(2, 3, +\tfrac{1}{6}\right) \oplus \left(1, \overline{3}, -\tfrac{2}{3}\right) \oplus \left(1, \overline{3}, +\tfrac{1}{3}\right)$, which of these is ruined?

7. Are we sure that the neutrino is *exactly* neutral? Are we sure that the proton *exactly* cancels the electron's charge? Let's define the Higgs field to have $Y = +\tfrac{1}{2}$ (this just sets the overall scale), but then change the various charges of the fermions to

   left-fermions: $3\left[\left(3, 2, \tfrac{1}{6} + \delta\right) \oplus \left(\overline{3}, 1, -\tfrac{2}{3} - \delta\right) \oplus \left(\overline{3}, 1, +\tfrac{1}{3} - \delta\right) \oplus \left(1, 2, -\tfrac{1}{2} + \varepsilon\right) \oplus \left(1, 1, +1 - \varepsilon\right)\right].$

   This form is in fact required to make the Higgs couplings to work out.
   (a) Using $Q = T_3 + Y$, find the charge of the electron, the neutrino, the up quark, the down quark, the proton, and the hydrogen atom.
   (b) Using some useful combination of the three anomaly conditions in problem 5, show that, in fact, $\delta = \varepsilon = 0$.

XII. Beyond the Standard Model

The Standard Model has been highly successful, but it is not perfect. As we noted in chapter 1, it cannot explain gravity. This is a hard problem, and we're not sure how to solve it. But there is an easier problem: neutrino masses. Given the particle assignments in eq. (11.37), it is inevitable that neutrinos are massless, and indeed experiments that measure the end points in beta decay (finding the maximum energy of the electron) indicate that at least the electron neutrino is lighter than 2 eV or so. Direct measurements of the other neutrinos are much more difficult, but there is no hint of mass. Nonetheless, we now firmly believe neutrinos have mass, and this is because of neutrino oscillations.

A. Neutrino Oscillations

When we explained the origin of quark mass in sec. 11B, we found that we had to introduce a CKM matrix relating the mass states of the three down-like quarks to the three up-like quarks. This did not occur in the lepton sector, but only because neutrinos were massless. Since we are now contemplating neutrino masses, it seems likely that the mass eigenstates of the neutrino sector will not quite match those of the charged leptons.

To understand the consequences, consider first a system with only two generations. Just as up was paired not with down, but with the linear combination $d\sin\theta_C + s\sin\theta_C$ in the quarks, there will be similar pairing in the leptons. This means that $\left|\nu_e\right\rangle$, the quantum particle that is paired with the electron in weak interactions, is probably not an eigenstate of the Hamiltonian, but a superposition of two such states, so that

$$\begin{pmatrix}\left|\nu_e\right\rangle \\ \left|\nu_\mu\right\rangle\end{pmatrix} = \begin{pmatrix}\cos\theta & \sin\theta \\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\left|\nu_1\right\rangle \\ \left|\nu_2\right\rangle\end{pmatrix}. \tag{12.1}$$

The two states $\left|\nu_1\right\rangle$ and $\left|\nu_2\right\rangle$ will be *mass eigenstates*, and will have masses m_1 and m_2. Now, suppose we manage to make a neutrino, say, by beta decay, or some similar process, and it has momentum p. At $t = 0$, we can describe it as

$$\left|\Psi\left(t=0\right)\right\rangle = \left|\nu_1\right\rangle\cos\theta + \left|\nu_2\right\rangle\sin\theta. \tag{12.2}$$

The energy for each of these two states will be given by

$$E_i = \sqrt{p^2 + m_i^2}. \tag{12.3}$$

Because neutrinos are typically produced with energies in the range of MeV (from beta decay) to many GeV (from particle colliders), they are ultrarelativistic, and therefore their momentum is enormous compared to their mass. We can then Taylor expand eq. (12.3) and conclude

$$E_i = p + \frac{m_i^2}{2p}. \tag{12.4}$$

We know from ordinary quantum mechanics that a state with energy *E* will behave like $\left|\Psi(t)\right\rangle = e^{-iEt}\left|\Psi(0)\right\rangle$, so we can apply this to each term in eq. (12.2), and we conclude that

$$\left|\Psi(t)\right\rangle = \exp\left(-ipt - \frac{im_1^2 t}{2p}\right)\left|\nu_1\right\rangle\cos\theta + \exp\left(-ipt - \frac{im_2^2 t}{2p}\right)\left|\nu_2\right\rangle\sin\theta\,. \tag{12.5}$$

Now, suppose we go downstream and try to detect this neutrino. If the neutrino interacts via a charged current interaction, it might produce an electron or a muon. It can only produce an electron if it is an electron neutrino, so effectively, we are now measuring whether it is an electron neutrino or not. The probability of it being an electron neutrino is given by

$$\begin{aligned}
P(\nu_e \to \nu_e) &= \left|\left\langle \nu_e \middle| \Psi(t)\right\rangle\right|^2 \\
&= \left|\left(\cos\theta\left\langle\nu_1\right| + \sin\theta\left\langle\nu_2\right|\right)\left[\exp\left(-ipt - \frac{im_1^2 t}{2p}\right)\left|\nu_1\right\rangle\cos\theta + \exp\left(-ipt - \frac{im_2^2 t}{2p}\right)\left|\nu_2\right\rangle\sin\theta\right]\right|^2 \\
&= \left|e^{-ipt}\left[\cos^2\theta e^{-im_1^2 t/2p} + \sin^2\theta e^{-im_2^2 t/2p}\right]\right|^2 \\
&= \left[\cos^2\theta e^{im_1^2 t/2p} + \sin^2\theta e^{im_2^2 t/2p}\right]\left[\cos^2\theta e^{-im_1^2 t/2p} + \sin^2\theta e^{-im_2^2 t/2p}\right] \\
&= \cos^4\theta + \sin^4\theta + \sin^2\theta\cos^2\theta\left\{\exp\left[i\left(m_1^2 - m_2^2\right)t/2p\right] + \exp\left[i\left(m_2^2 - m_1^2\right)t/2p\right]\right\} \\
&= \cos^4\theta + \sin^4\theta + 2\sin^2\theta\cos^2\theta\cos\left(t\Delta m^2/2p\right),
\end{aligned}$$

$$P(\nu_e \to \nu_e) = 1 - \tfrac{1}{2}\sin^2(2\theta)\left[1 - \cos\left(\frac{t\Delta m^2}{2p}\right)\right]. \tag{12.6}$$

For relativistic neutrinos, we can approximate $t = L$, the distance traveled, and $p = E$, so this formula more commonly appears in terms of *L* and *E*. A similar formula can be found for $\nu_e \to \nu_\mu$, and we have

$$P(\nu_e \to \nu_e) = 1 - \tfrac{1}{2}\sin^2(2\theta)\left[1 - \cos\left(\frac{\Delta m^2}{2E}L\right)\right], \tag{12.7a}$$

$$P(\nu_e \to \nu_\mu) = \tfrac{1}{2}\sin^2(2\theta)\left[1 - \cos\left(\frac{\Delta m^2}{2E}L\right)\right]. \tag{12.7b}$$

Thus what starts as an electron neutrino can become a muon neutrino, but will then oscillate back to being an electron neutrino.

Early on it was demonstrated that electron neutrinos are distinct from muon neutrinos, but this was simply because the baseline *L* for such experiments were too short. If the mass-squared difference is small, the cosine will not oscillate rapidly enough, and eqs. (12.7a) and (12.7b) will simply yield 1 and 0 respectively.

The first experiments that indicated neutrinos were undergoing oscillations came from studying neutrinos that came from the Sun. Neutrinos are produced by a variety of processes, most notably

$$p^+ + p^+ = \left(p^+ n^0\right) + e^+ + \nu_e \,,$$

where $\left(p^+ n^0\right)$ is the deuteron, or nucleus of ^{2}H. The deuteron remains in the Sun to eventually get fused to helium, the positron annihilates with an electron to produce energy, and the neutrino travels at the speed of light out of the Sun with negligible chance of being stopped. A sufficiently large terrestrial detector can find enough neutrinos that we can determine if the Sun is, indeed, producing copious quantities of neutrinos. Early experiments could only see electron neutrinos, and the rate of neutrinos from the Sun was found to be approximately one-third of the rate expected. This was called the *Solar Neutrino Problem*.

There were many explanations, among them that the Sun was not actually undergoing fusion, but eventually the Sudbury Neutrino Observatory was built, which was able to see other neutrinos as well (via neutral current interactions), and it demonstrated that the Sun is making neutrinos at the rate we expect, but that most of them are not electron neutrinos by the time they reach the Earth. Particle physicists concluded that neutrinos were actually oscillating, and therefore have mass. The analysis is complex, because it was only realized in the 1980's that weak interactions with the electrons in the Sun shift the neutrino energies slightly, and most importantly, shift it differently for electron neutrinos than it would for other neutrinos (because of charged-current interactions) and hence simple formulas like eq. (12.5) are not in fact valid. But this phenomenon is still called neutrino oscillations.

Another clue that something interesting was going on in the neutrino sector came from cosmic ray experiments. Cosmic rays are produced when protons or simple nuclei slam into the upper atmosphere of the Earth from outer space. When these projectiles hit the nuclei of air molecules, a spray of hadrons results, which mostly ends up as pions. Neutral pions decay to photons, but charged pions decay via $\pi^- \to \mu^- \bar{\nu}_\mu$, and the muons then decay via $\mu^- \to \nu_\mu \bar{\nu}_e e^-$. Hence the total number of muon neutrinos or anti-neutrinos should be twice the number of electron neutrinos or anti-neutrinos. If you build a detector far enough underground, you can block out all the cosmic rays that would confuse the issue and see only neutrinos. It was found that for neutrinos going down, which had traveled through a few kilometers of air and rock before getting to the detector, this worked well, but for neutrinos coming up, which had been produced on the other side of the Earth and traveled thousands of kilometers, there was a shortage of muon neutrinos. It seemed likely that the muon neutrinos were oscillating into some other kind of neutrino. The mass squared difference required for atmospheric neutrino oscillations was, however, about an order of magnitude larger than that required for solar neutrinos.

Since then, a variety of experiments have made clear that neutrinos of one type become neutrinos of another type. We call the three mass eigenstates ν_1, ν_2 and ν_3, and specify which one is which by assuming that in weak doublets, the electron mostly goes with ν_1, the muon with ν_2, and the tau with ν_3. Then a combination of solar and terrestrial experiments have given us limits on the masses, namely,

$$m_2^2 - m_1^2 \approx 7.58\times10^{-5}\ \text{eV}^2 \,, \tag{12.8a}$$

$$\left|m_3^2 - m_1^2\right| \approx 2.35\times10^{-3}\ \text{eV}^2 \,. \tag{12.8b}$$

Eq. (12.8a) comes primarily from the study of solar neutrinos, which are affected by their interaction with the matter (electrons) in the Sun, and therefore we can determine the sign of the mass squared difference. Eq. (12.8b) comes from a variety of terrestrial experiments, which because of the large mass squared difference, is relatively unaffected by the presence of matter, and hence we are basically using equations like eq. (12.7). The analysis is much more complicated because of the involvement of all three neutrinos. Most theorists assume the neutrinos are arranged with mass $m_1 < m_2 < m_3$ (normal hierarchy), but it is also possible they are arranged as $m_3 < m_1 < m_2$ (inverted hierarchy). Assuming the normal hierarchy, eqs. (12.8) tell us that $m_3 > 0.048\text{ eV}$ and $m_2 > 0.0087\text{ eV}$, but do not actually tell us the mass. Our best guess is that these two neutrinos are of the order of these numbers, and m_1 is somewhat smaller than either of them, but we have no way of knowing.

Interactions of massive neutrinos via the *W*-bosons would involve a unitary matrix called *U*, akin to the CKM matrix *V* of quark interactions. It can be parameterized in a manner effectively identical to eq. (10.61), namely

$$U = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \qquad (12.9)$$

where, as before, $s_{ij} \equiv \sin\theta_{ij}$ and $c_{ij} \equiv \cos\theta_{ij}$, but all the angles and the phase δ are distinct parameters from those appearing in the quark sector. Experimentally, the three angles are roughly measured to be

$$\sin^2\theta_{12} \approx 0.31, \quad \sin^2\theta_{13} \approx 0.02, \quad \sin^2\theta_{23} \approx 0.42. \qquad (12.10)$$

Unlike the quark sector, the neutrino sector has a lot of mixing between the generations, so that *U* is not close to the identity matrix. The phase δ at present is completely unmeasured.

You might worry that all the rules of Feynman diagrams need to be changed because of these mixings. Technically, you are right, especially if you are going to measure which of the three neutrinos ν_1, ν_2, and ν_3 are being createed. But you can't realistically measure the energies accurately enough to determine which neutrino state you end up with, and your chance of capturing the neutrino afterwards is effectively nil. So you should sum over all the final neutrino states, and treat the neutrinos as massless, and the Feynman rules of the previous chapter can be used with impunity.

B. Dirac and Majorana Masses

So, assuming we have convinced ourselves that we need neutrino masses, we need to explain them. The logical thing to do would be to simply add a right-handed neutrino field, $(\nu_R)_A$, one for each family. Eqs. (11.6) would gain a new term, and there would be a third equation. We would find

$$i\not{D}(\Psi_L)_A - k_{AB}\sqrt{2}\Phi(\ell_R)_B - k'_{AB}\sqrt{2}\tilde{\Phi}(\nu_R)_B = 0, \qquad (12.11a)$$

$$i\not{D}(\ell_R)_B - k^*_{AB}\sqrt{2}\Phi^\dagger(\Psi_L)_A = 0, \qquad (12.11b)$$

$$i\not{D}(\nu_R)_B - k'^*_{AB}\sqrt{2}\tilde{\Phi}^\dagger(\Psi_L)_A = 0. \qquad (12.11c)$$

This works fine. These simple modifications of the standard model would be called *Dirac masses*. If eqs. (12.11) are all that is going on, then we need an additional seven parameters to explain things – the three neutrino masses and the three angles and one phase in eq. (12.9).

It is worth noting that the neutrino oscillations are definitely breaking the conservation of $L_e - L_\mu$ and $L_\mu - L_\tau$. If we ignore sphalerons, then the only conserved quantities beyond those imposed by gauge conditions are baryon number B and lepton number L. If we include sphalerons, then $B-L$ is the only conserved quantity.

Note that if neutrino masses are of the order of those given in eq. (12.8), and we use Dirac masses, then the couplings k' appearing in eqs. (12.11a) and (12.11c) must be very small, of order 10^{-13} or so. We have no explanation of why this should be so small. But then again, the electron has a coupling of order 10^{-6}, and we have no explanation for that either.

So, it seems we can account for neutrino masses simply by adding three new right-handed fields. Obviously, they won't have color, since they aren't strongly interacting, and they will be SU(2)$_L$ singlets, so they automatically have $T_3 = 0$, and we can deduce from $Q = T_3 + Y$ that they have $Y = 0$. Then the standard model plus neutrino masses would be described by

$$\begin{aligned}\text{gauge group:}\quad & \mathrm{SU}(3)\times\mathrm{SU}(2)\times\mathrm{U}(1).\\ \text{scalars:}\quad & \left(1,2,+\tfrac{1}{2}\right),\\ \text{left-fermions:}\quad & 3\begin{bmatrix}\left(3,2,\tfrac{1}{6}\right)\oplus\left(\overline{3},1,-\tfrac{2}{3}\right)\oplus\left(\overline{3},1,+\tfrac{1}{3}\right)\\ \oplus\left(1,2,-\tfrac{1}{2}\right)\oplus\left(1,1,+1\right)\oplus\left(1,1,0\right)\end{bmatrix}.\end{aligned} \tag{12.12}$$

Now, by analogy, we would just allow all possible interactions consistent with gauge invariance and renormalizability. And we would end up with eqs. (12.11). Except – it turns out we missed something.

Recall that mass terms are terms that connect left-handed and right-handed particles. Recall that the anti-particle of a right-handed particle is a left-handed particle. We therefore could consider adding a term to the Hamiltonian with a non-zero matrix element of the form

$$\left\langle \left(\overline{\nu_R}\right)_A \middle| \mathcal{H} \middle| \left(\nu_R\right)_B \right\rangle = M_{AB}\ , \tag{12.13}$$

where M is almost any matrix. We are being a bit sloppy here, as we are ignoring spin, but it is clearly possible to make eq. (12.13) Lorentz invariant, and it is invariant under the full SU(3)$_C$ × SU(2)$_L$ ×U(1)$_Y$ gauge symmetry. It isn't hard to show that M_{AB} must be symmetric, $M_{AB} = M_{BA}$, but it can be complex. If we include it, eq. (12.11c) would be modified to

$$i\not{D}\left(\nu_R\right)_B - k'^{*}_{AB}\sqrt{2}\tilde{\Phi}^\dagger\left(\Psi_L\right)_A - M_{AB}\left(\mathcal{C}\nu_R\right)_A = 0\,, \tag{12.14}$$

where $\mathcal{C}\nu_R$ is defined, in our usual chiral representation of the Dirac matrices, by eq. (3.27).

We aren't *required* to include mass terms coming from (12.13), but we *can*, and based on our experience with the Standard Model, maybe we *should*. Let's try to figure out the consequences if we do.

We note that the interaction in eq. (12.13) changes lepton number by two units, since it turns a neutrino into an anti-neutrino. Indeed, this means that there is no distinction between neutrinos and anti-neutrinos. Such a scenario involves what are called *Majorana* neutrinos. If L

is no longer conserved, then the only remaining accidental symmetry of the standard model, ignoring sphaleron effects, is baryon number B. If you include sphaleron effects, then there are no accidental symmetries.

The mass in eq. (12.13), as already explained, respects $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge symmetry, and therefore it need not be mediated by the Higgs field, absent from the last term of eq. (12.14). Thus the right-handed neutrino mass has nothing to do with the scale v = 246 GeV. We would expect the Dirac mass terms to be of order or less than 246 GeV, but the right-handed neutrinos could be heavier, even arbitrarily heavier. To understand the implications, let's imagine that we have just one generation. Then there are two neutrino particles, ν_L and ν_R, with the Dirac mass changing one into the other and the Majorana term connecting ν_R with itself. The resulting mass matrix would look something like

$$\begin{pmatrix} 0 & m \\ m & M \end{pmatrix}, \tag{12.15}$$

where m is the Dirac mass and M the right-handed Majorana mass. The actual masses of the particles that would result would be, approximately, the eigenvalues of eq. (12.15), which are

$$\tfrac{1}{2}\left(M \pm \sqrt{M^2 + 4m^2}\right). \tag{12.16}$$

Since we are going to assume $M \gg m$, we can Taylor expand eq. (12.16), and keep only the leading term. One expression comes out negative, but this can be fixed, so we'll just ignore the minus sign and conclude that the two masses are

$$m_H \approx M\,, \quad m_L \approx \frac{m^2}{M}\,. \tag{12.17}$$

The heavy one m_H is almost exclusively ν_R and its anti-particle, and hence will have very little weak interaction, and no other interactions. It would be hard to make in any collider, even if we had sufficient energy. The light one m_L would be almost exclusively ν_L and its anti-particle. It will have almost normal weak interactions. This mechanism is called the *see-saw mechanism*. Note that the heavier the heavy one is, the lighter the light one is. This might naturally explain why neutrinos are so light, because ν_R is heavy. Suppose, for example, that the heaviest neutrino we know of, with a mass of order 0.05 eV, is due to a Dirac mass m comparable to the top quark mass, $m \approx 170\ \mathrm{GeV}$, then the heavy counterpart would have a mass of order 10^{15} GeV, a number we will encounter again in section D. If this is the case, then the smallness of the neutrino masses is giving us clues to the physics at high energies, much higher than we can hope to achieve in colliders in the foreseeable future.

If neutrinos are their own anti-particles, than it has some interesting implications. There are, for example, certain nuclei that are energetically forbidden from undergoing beta decay, but they are energetically *allowed* to undergo *double beta decay*, for example,

$$^{76}\mathrm{Ge} \to {}^{76}\,\mathrm{Se} + e^- + e^- + \bar{\nu}_e + \bar{\nu}_e\,. \tag{12.18}$$

Because this is a double-weak interaction, it is very slow, and takes an average of 1.8×10^{21} years. But if neutrinos are their own anti-particles, then it is possible that one of the anti-neutrinos becomes a neutrino and is absorbed, and we have the process of *neutrinoless double beta decay,* so that

$$^{76}\mathrm{Ge} \to ^{76}\mathrm{Se} + e^- + e^- . \qquad (12.19)$$

This is experimentally distinguished from eq. (12.18) by the fact that the total energy of the two electrons will sum to the mass difference between $^{76}\mathrm{Ge}$ and $^{76}\mathrm{Se}$. Such neutrinoless double beta decay has never been observed, which puts a limit on the Majorana mass of the electron neutrino around 1 eV. But things get complicated, because the electron neutrino is not a mass eigenstate, and besides, the neutrinos are probably lighter than this anyway.

We now have two theories that explain neutrino masses – the Dirac one, with seven parameters, and the Dirac plus Majorana explanation, which, if I'm doing the counting correctly, has twelve more parameters than the Dirac one. We don't know which one to choose, so we don't know which should be the "Standard Model" of neutrino interactions. But these are not our only choices. We explained earlier that a combination of experimental and theoretical constraints demands that there be equal numbers of families of quarks and leptons. The same is not true of ν_R. You can add as many or as few copies as you want. You have to add at least two, to get two mass differences in the neutrinos, but you could add two (which would mean one neutrino is exactly massless) or add four (which means there is a "sterile" neutrino with no weak interactions at all). There are other ways to do it as well. You could introduce no right-handed neutrino fields ν_R at all, but instead introduce a new scalar field that is a triplet (spin 1) under SU(2)$_L$, in the representation $(1,3,+1)$, which would lead to Majorana masses for the left-handed neutrinos without any right-handed neutrinos at all. The possibilities are endless. Without clear guidance on the right way to include neutrino masses, we simply say they are not part of the Standard Model, and we have to wait for more information before we can make reasonable conjectures about which way to extend it.

I hope I have given you a taste for some of the possibilities that neutrino masses allow. But there are other puzzles, and some proposals have attempted to address these puzzles. Let's look at some of the alternatives.

C. Supersymmetry

We have previously discussed how coupling constants "run", that is, if you measure them at high energies, they effectively get stronger or weaker. Thus far we have only discussed this for the gauge couplings, but it is true of all the couplings of the standard model, including the Higgs self-coupling λ, and all the interactions of the Higgs with the fermions, Basically, it is found for all of these dimensionless parameters that they run logarithmically, so that, very crudely $\alpha\left(E^2\right) = \alpha + \alpha^2 \ln\left(E^2\right)$. Because of this slow evolution, you can go up many orders of magnitude in energy without the coupling changing too much. Our best guess is that at some very high energy, there is a deep theory we have not yet discovered that predicts the value of the fundamental couplings, but as you go down in energy, the couplings evolve to the values we observe. No problem.

The problem comes when we study the running of the Higgs mass term. This dimensionful parameter evolves because of loops of all kinds of particles, including diagrams like the one appearing in Fig. 12-1. The resulting dependence of the Higgs mass on energy looks like

$$M_H^2\left(E^2\right) = M_H^2 + \lambda E^2 \ , \tag{12.20}$$

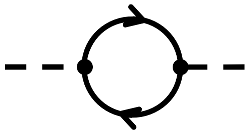


Figure 12-1: A one loop contribution to the Higgs mass. The contribution is dominated by the top quark, because it has the largest coupling.

where λ is a complicated combination of all the Higgs couplings. This implies that that as you go to very high energies, the effective Higgs mass is different, perhaps very different, from what we find at low energies.

This is unnerving because we believe that the Higgs mass probably has some value that comes about at very high energies, maybe 10^{16} GeV or so, but then we have to use eq. (12.20) (plus lots of corrections) to figure out what the low energy Higgs mass M_H^2 is. Now, there's nothing wrong, in principle, with just claiming that $M_H^2(E)$ is some arbitrary large number that just accidentally cancels λE^2 almost exactly to leave a Higgs mass of just 126 GeV, but that seems just a little too coincidental. How we get the mass to cancel so well is called the *hierarchy problem*.

There are lots of idea of how to get around this problem, but one of the most popular is *supersymmetry*, often abbreviated SUSY. SUSY is a symmetry that relates fermions and bosons. In a theory with exact supersymmetry, there is exactly one fermion for every boson, and the *superpartner* will have the same mass, charge, color, etc. Obviously, this doesn't represent reality, but let's run with it a bit. If supersymmetry were exact, then it turns out that for every diagram with a fermion running around it like Fig. 12-1, there is another with a boson running around it, and the two exactly cancel. Hence there will be no quadratic running of the Higgs mass as shown in eq. (12.20).

Okay, so what does this have to do with the real world? It is possible that, just like the gauge group $SU(2)_L \times U(1)_Y$ in the Standard Model, supersymmetry is broken. If this occurs, then there can be a splitting, and all the superpartners of the Standard Model particles could have different masses. This would spoil the *exact* cancellation that eliminates the term λE^2, but it would replace it with something like $\lambda \Delta M^2$, where ΔM^2 is the mass splitting between the Standard Model particles and their superpartners. This could naturally account for the smallness of the Higgs mass, but only if ΔM^2 is not a lot bigger than the weak scale, say, not bigger than about 1 TeV. The superpartners are usually named by prefixing the letters "s-" to the particle for fermions, while bosons usually have the suffix "-ino" added, more or less. The easiest particles to create and look for in supersymmetry tend to be the squarks, such as the stop squark, and the gluinos, because they have strong interactions. At present, the LHC should have already found squarks or gluinos if they are lighter than about a TeV. We have to seriously consider the possibility that supersymmetry is wrong, or at least it occurs at a high enough energy that it does not solve the hierarchy problem. Or maybe we'll get lucky, and find it in the next few months or few years at the LHC. Only time will tell.

Supersymmetry is incorporated into many attempts to include gravity, in which case it is called supersymmetric super gravity (SUSY SUGR, pronounced "Suzy Sugar"). Many of the infinities which plague ordinary gravitational quantization seem to magically cancel in SUSY SUGR, but no one has proven that they all do. Lots of ideas, including string theory (section E) incorporate SUSY SUGR.

D. Grand Unified Theories

Another popular idea for going beyond the Standard Model is Grand Unified Theories, or GUT's. Let us note first that in the Standard Model, we managed to combine apparently disparate particles, like the left-handed neutrino/charged lepton and the left-handed up quark/down quark, into $SU(2)_L$ doublets. The fact that they look so different was caused by the spontaneous symmetry breaking induced by the Higgs field. Perhaps there is some grander symmetry which would gather even more particles together into more complicated representations of groups.

Consider another fact. In the Standard Model, there are three fundamental gauge couplings, which we could describe by three fine structure constants α_1, α_2, and $\alpha_S = \alpha_3$. At the electroweak scale, these have different strengths, with $\alpha_1 < \alpha_2 < \alpha_3$. However, if we go up in energy, they all change. It turns out that in the Standard Model, α_1 increases, α_2 decreases very slightly, and α_3 decreases faster. Perhaps at some magic energy, they will all be the same strength. If so, they could all be some part of a bigger, grander gauge group. At the time these ideas were first proposed, the coupling constants were not well measured, and it was estimated that the couplings would unify at about 10^{15} GeV, but as you can see from Fig. 12-2, this doesn't quite work. If, on the other hand, supersymmetry is correct, then there should be a host of undiscovered superpartners just around the corner, which changes the running of the coupling constants. As far as we can tell, the coupling constants *do* converge on a common value at about 10^{16} GeV when you include the superpartners, also in Fig. 12-2. This suggests there might be lots of new physics at an energy of 10^{15} to 10^{16} GeV.

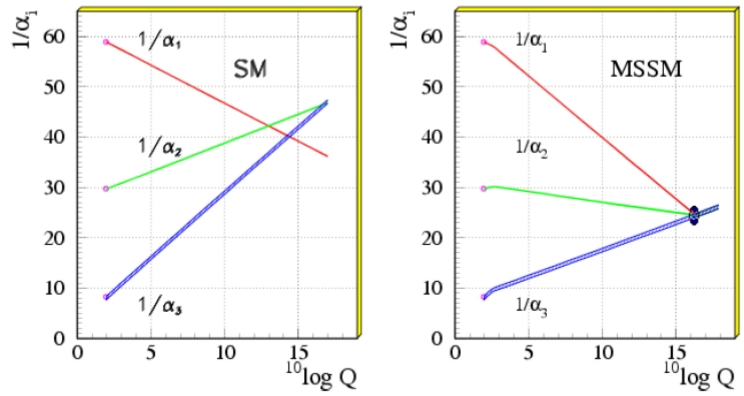


Figure 12-2: With just Standard Model particles, the three couplings do not quite unify around 10^{15} GeV (left), but when you add in superpartners, they unify nicely near 10^{16} GeV. Note that the reciprocals of the three couplings are plotted, in which the coupling constants are nearly linear functions of log*E*. Image stolen from D. I. Kazakov, **http://arxiv.org/abs/hep-ph/0012288**

We are not going to have colliders operating at this energy (called the *GUT scale*) any time soon. But that doesn't mean there are no observable consequences. As already pointed out in section B, right-handed neutrinos around this mass might be just right, via the see-saw mechanism, to give rise to the tiny neutrino masses we observe. In many GUTs, certainly the most popular ones, there are lots of heavy gauge bosons, similar to the *W*'s but with masses at the GUT scale. These intermediate particles can mediate interesting phenomena like proton decay.

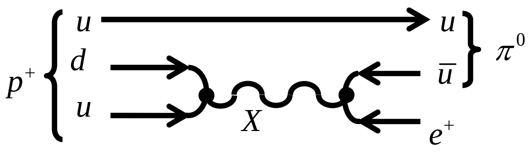


Figure 12-3: A Feynman diagram for proton decay in some GUT theories.

Fig. 12-3 shows a means by which proton decay might be mediated by a heavy *X* boson that would occur in many GUTs. We might vaguely approximate this process, somewhat similar to weak decays, as

$$\Gamma \sim \frac{m_p^5}{100 M_X^4}. \tag{12.21}$$

When GUTs were first developed, it was thought that the GUT scale was about 10^{15} GeV, which leads to a typical lifetime of order 10^{31} years, but with a GUT scale of 10^{16} GeV, we get a lifetime more like 10^{35} years. The actual limit for $p^+ \to e^+ \pi^0$ is about 10^{34} years. The actual rate could be faster or slower than our simple predictions, depending on the details of the model, but we have at least a chance of seeing proton decay in the foreseeable future, depending on the exact GUT scale.

E. Other Ideas

There are a lot of other ideas about how to describe physics beyond the Standard Model. One idea that has been extensively explored is the idea of *hidden dimensions*. In the simplest version of this, in addition to the ordinary three dimensions of space and one of time, there is an additional space-like direction, but this direction is periodic, and very short, so that if you move in this direction you quickly come back to your starting point, much as in the video game Asteroids or Pac-Man you can go off one side and come back on the other. Trust me, in general relativity, this makes sense. The scale for this direction must be very small, such that the wavelength of particles at any ordinary energy would be larger than the extra dimension. But as we increase energy, it is possible that we will be able to "see" this extra dimension, or its effects. We haven't seen any such effects at the LHC yet, which suggests the extra dimension is no larger than about 10^{-19} m or so.

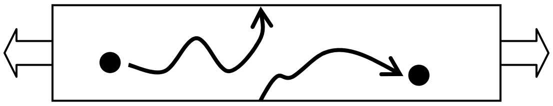


Figure 12-4: A very simple description of a universe with a small extra dimension. The universe is infinite or large in three space direction (represented by the horizontal direction) but very small in any additional directions (the vertical direction). An object moving in the extra direction would almost immediately loop around the extra dimension.

Other proposals include the *Randall-Sundrum model*, where there is an extra dimension which can be larger, perhaps even macroscopic, but we are trapped on a very thin four-dimensional slice called a "brane" (from membrane), much as the mythical shapes in Edwin Abbott's book *Flatland*. Spacetime is severely warped in the region off the membrane, and it's a good guess ordinary objects could not survive there. Such a model, as farfetched as it sounds, is one of the more promising approaches to explaining the hierarchy problem.

Another alternative that has generated a lot of interest is *string theory*. String theory first developed as a model of strong interactions, when quarks and gluons were not yet recognized as the correct theory of strong interactions. String theory posited that rather than thinking of elementary particles as point-like objects, perhaps they were little loops of stuff that were oscillating and gyrating. Different particles would correspond to different modes of vibration of the strings. It was recognized early on that string theory actually provided a self-consistent model that included quantum gravity, but that it only worked if supersymmetry is true and the universe has ten dimensions (four are the usual spacetime and the remainder are probably small, like what appears in Fig. 12-4). String theory was almost unique, in that there were believed to be only five different possible string theories, each of which had no tunable parameters, and hence one of these theories should, in principle, match the observed universe in all its particulars. Such a demonstration would indeed be remarkable, but thus far, string theory has only been demonstrated to predict gravity. There is no evidence in favor or in contradiction of the claim that it can also predict the rest of physics. In the mid 1990's, it was realized that there might be deep connections between the five string theories. It is now believed that all five string theories are different approximations to an even deeper theory called *M-theory*. M-theory exists only in a universe with a total of eleven space-time dimensions. Like string theory, at the moment M-theory has no known connection with experiment.

It is unfortunate that, so far, none of these theories has good testable predictions. A consistent theory that includes gravity plus all the couplings of the standard model should be a complete theory, in principle explaining all current observations and successfully predicting all future observations. Such a theory is called a *theory of everything*, or TOE for short. M-theory is probably the most popular proposal for a theory of everything (as of 2012), but other alternatives are being explored.

One of the problems confronting us is the extreme difficulty and cost of building higher energy accelerators. The LHC, currently operating at $\sqrt{s} = 8.0\text{ TeV}$ is designed to operate at $\sqrt{s} = 14.0\text{ TeV}$, probably starting around 2015, and as the luminosity is increased, it will adequately explore this high energy regime. If weak scale supersymmetry is correct, we should see it evidenced at the LHC. More precise measurements of the Higgs boson could be explored by producing them through the process $e^+e^- \to ZH$ if we built an e^+e^- collider operating at, say, $\sqrt{s} > 250\text{ GeV}$. But getting any information about particles heavier than about 10 TeV or so would require a much higher energy collider. It is unlikely we can fund such a collider unless there is a substantial advance in technology.

Given the difficulties of exploring high energies, many physicists favor a low energy approach. The properties of neutrinos are poorly understood, and we have only begun to understand the masses and mixings in the neutrino sector. Neutrino detectors tend to have very weak signals, and to suppress confusion from spurious sources, such as cosmic rays, they are normally housed deep underground. The neutrinos are typically produced by firing a bean of protons into stationary targets, which produce pions whose decay includes neutrinos.

F. Cosmology

There is, however, another approach to studying particles. It is known that cosmic rays, probably produced by supernova explosions, have energies as high as about $3\times10^{11}\text{ GeV}$. Since such particles are so rare, we cannot hope to collide pairs of such particles, let alone detect such collisions, but they will frequently strike stationary particles. A proton with this energy striking a stationary proton would have $\sqrt{s} \approx 750\text{ TeV}$, or two orders of magnitude above what we can reach in current colliders. Such a collision occurs in the upper atmosphere, and at best we can detect only a small fraction of the shower of particles produced by these monster cosmic rays. It is difficult to deduce what fundamental interactions might be occurring in these collisions.

Another promising approach comes from *cosmology*, the study of the universe as a whole, its past, present, and evolution. It is known, for example, that the majority of the gravitational mass of galaxies is non-luminous, or *dark matter*. The nature of this dark matter is unknown, but unlike ordinary matter (called *baryonic matter* in this context), dark matter does not accumulate in the disk and bulge of galaxies, but rather is spread out in a vast halo that surrounds the visible galaxy. There are strong arguments (one of which will appear shortly) that this matter is not baryonic. The most likely candidates are exotic particles, particles not contained in the standard model. One popular candidate is the *lightest supersymmetric partner*, or LSP, the lightest of the superpartners predicted by supersymmetry. The LSP is assumed to be some combination of the superpartners of the photon, the Z, and the Higgs bosons; it will be a neutral spin ½ fermion. Such an LSP would probably be its own anti-particle, and it is possible that we could detect it by watching for self-annihilations. Another popular candidate would be one of the right-handed neutrinos discussed in section 12B, though making this work would require some rather unnatural fine tuning of the parameters.

Right-handed neutrinos are speculative; left-handed neutrinos are not. Their interactions are known, and it is believed that in the early universe they would have been in thermal equilibrium with everything else. On this basis we can make a prediction of their abundance in the current universe. Their combined abundance should be comparable to the abundance of the photons that were also produced in the early universe, and now comprise a 2.7 K background radiation that pervades the universe. If the neutrino masses are great enough, gravity should draw them into galaxies, adding their mass to the galaxy and shaping it as well. They might even have more mass than all the ordinary baryonic matter, though this is unlikely. Limits from galaxy clustering (combined with a great deal of model assumptions) indicate that the sum of the neutrino masses is probably less than about 0.3 eV. Combining this with eqs. (12.8) suggests they are each less than about 0.1 eV.

One puzzle that is particularly confusing in cosmology is *dark energy*. There is considerable evidence that the Universe is expanding, and this expansion is accelerating. According to Einstein's General Theory of Relativity, this can happen if the universe is pervaded by some material with sufficiently negative pressure. Though there are many candidates for such dark energy, the simplest, which fits well with observations, is a *cosmological constant*. In the context of particle physics, the simplest cause for a cosmological constant would just be to add a constant term to the Hamiltonian density $\mathcal{H}$. Ordinarily, we treat a constant term in the Hamiltonian density as irrelevant in particle physics, but in General Relativity, such a term would contribute to the total energy, and hence would contribute to gravity. Experimentally, the value is of order $3\times10^{-11}\text{ eV}^4 = 3\times10^{-47}\text{ GeV}^4$, a ridiculously small value. By comparison, the constant term in the Higgs potential eq. (11.4) is $\frac{1}{8}M_H^2 v^2 \sim 10^5\text{ GeV}^4$. We really have no clue why this value is what it is.

Probably the most interesting interplay of particle physics and cosmology comes not from describing the current universe, but by trying to understand the very early universe. The universe is believed to have begun in a giant explosion called the *Big Bang*. According to the Big Bang theory, the universe was initially in a hot, nearly uniform state. Using Einstein's General Theory of Relativity, together with just a little particle physics, we can determine a formula relating the age of the universe t to the temperature T. During the early radiation-dominated era, this relationship is approximately

$$t \approx \left(\frac{\text{MeV}}{k_B T} \right)^2 \text{ s}, \tag{12.22}$$

where k_B is Boltzmann's constant.[1] This formula is roughly valid for $k_B T > 1 \text{ eV}$. Typically, at a temperature *T*, particle/anti-particle pairs of every particle with a mass $m < 3k_B T$ will be abundant. Hence particles which are now rare were abundant in the early universe. The presence of these particles affects all interactions. In principle, if we can detect remnants of these very early processes, we should be able to deduce some things about particle physics.

When temperatures were around 100 keV, it is believed that protons and neutrons that were produced by the Big Bang were able to fuse into simple nuclei, principally ^{4}He, while most of the heavier nuclei were created much later, in stars. The exact fraction of isotopes such as ^{4}He, ^{3}He, ^{2}H, ^{6}Li, and ^{7}Li, compared to the dominant ^{1}H are all sensitive to details of the early universe, such as the abundance of baryons and the number of neutrinos that exist. By studying the composition of those starts that have little or no heavier elements, we can estimate the fraction of these elements produced in *Big Bang nucleosynthesis*. Such estimates indicate that there are only three light weakly-interacting neutrinos (this observation slightly predates the same conclusion from measurements of the *Z*-width) and that the abundance of ordinary matter is far too little to comprise the dark matter.

Another possible clue into the nature of particle physics comes from the fact, previously mentioned in section 11E, that the universe is made of matter, not anti-matter. We have strong evidence from cosmic rays that at least out to the distance of nearby galaxies, the universe contains baryons, not anti-baryons. It is very difficult to imagine a mechanism that would separate these, so it is reasonable to guess that all galaxies in the visible universe contain a surplus of baryons. For every proton, we expect there is (approximately) one electron, so that the universe can be electrically neutral. We don't actually know about the neutrino sector; it could be that there is a large imbalance either direction between neutrinos and anti-neutrinos.

In models with Majorana neutrinos, it is not terribly difficult to have significant $\mathcal{C}P$ violation in the neutrino sector, principally having to do with the right-handed neutrinos. This could create an imbalance at an early stage between the numbers of neutrinos and antineutrinos. Sphalerons, which violate *B* and *L* could then convert this lepton asymmetry into a baryon asymmetry. As the universe cooled, sphalerons would disappear, freezing in that baryon asymmetry. This is probably the most popular scenario for creating the baryon asymmetry in the current universe.

Grand Unified Theories also violate baryon number. The theories are complex enough that they can easily contain sufficient $\mathcal{C}P$ breaking to account for the baryon asymmetry. In this scenario, up around a temperature of 10^{15} or 10^{16} GeV, when the Grand Unified fields were broken into the standard $\text{SU}(3)_C \times \text{SU}(2)_L \times \text{U}(1)_Y$ gauge group, the heavy gauge bosons would have created this asymmetry as they decayed, and this asymmetry was preserved until today.

Some of the other puzzles of cosmology have inspired a great deal of work in particle physics. Two of these are the *flatness* and *smoothness* problems. The flatness problem basically points out that the total density of all the "stuff" in the universe (ordinary matter, radiation, neutrinos, dark matter, dark energy) is, according to General Relativity, exactly right to create

---

[1] All sensible particle physicists use units where $k_B = 1$, but I saw no point in introducing this two pages before the end of the main text.

the amount of gravity needed to counteract the effects of the expansion of the universe. If this doesn't make sense to you, well, it's probably because you haven't had a course in General Relativity. The smoothness problem has to do with the fact that the universe, on the largest scales, looks essentially the same in all directions and all distances. This is seemingly impossible, since in a finite age universe, stuff on one side of the universe has never had a chance to interact with stuff on the other side of the universe.

A popular solution to these apparent problems is *inflation*, a theory where some field (generically called the *inflaton*), undergoes a transition, such that the energy density of the field is initially very high, but then it makes a pretty sudden transition to its current low value. General Relativity predicts that when the energy density is high, the universe will undergo exponential expansion, such that the distance between objects grows as $d \propto e^{Ht}$, where H is a constant related to the energy density of the inflaton. This allows parts of the Universe that were initially very close together (letting them smooth out) to move to very distant points, allowing the universe to be smooth on the scales we can see, and probably much farther. Inflation typically takes place at the GUT scale (10^{15} to 10^{16} GeV) or even higher.

Inflation has some interesting consequences. If inflation is valid, and tiny objects can grow to very large objects, then quantum effects (which usually dominate at very small scales) can grow and have effects on enormous scales. It is believed that this may be the source of structure in the present day universe. Quantum fluctuations on very small scales caused some portions of the early universe to be slightly denser or less dense than other nearby regions. Much later, these high density fluctuations will, through gravitational effects, gradually draw in nearby matter, ultimately forming high density regions that will become globular clusters, galaxies, groups, galaxy clusters and superclusters. Even before this structure forms, they will distort the cosmic microwave background radiation, and by studying tiny temperature and polarization effects of this radiation, we should be able to see the effects of inflation. Several spacecraft (most noticeably the Wilkinson Microwave Anisotropy Probe, or WMAP) have already indicated inflation fits the data pretty well. The European Planck spacecraft is scheduled to release its data by the end of 2012, and may well indicate that inflation is the correct solution.

Even the existence of the universe itself might someday yield to a proper understanding of particle physics. In particle physics, particles can appear out of nothingness. In any sensible quantum theory of gravity, it seems likely that space and time can appear out of nothingness. This is little more than speculation, but this event, sometimes called the Ultimate Free Lunch, may explain the existence of existence.

It is ironic that the study of the universe, and some of the largest structures in it, may be leading to an understanding of how physics behaves on scales far smaller than our direct experiments can hope to achieve. Who knows? A final connection, and an answer to the ultimate questions, like where the universe comes from, may come within our lifetime.

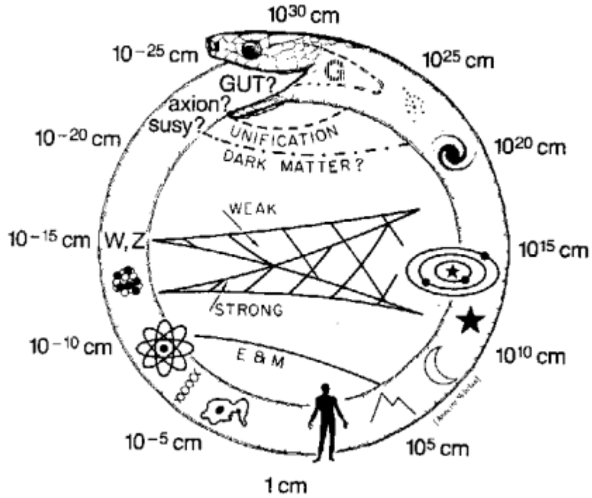


Figure 12-5: The cosmic Oroborus.

Problems for Chapter 12

1. Rewrite eq. (12.7b) in terms of L in km, Δm^2 in eV^2, and E in GeV. Then find a formula relating these three quantities at the shortest distance where L where oscillations have their first peak (when $P\left(\nu_e \to \nu_\mu\right)$ is at a maxium).

2. The initial Solar Neutrino Problem involved neutrinos from a rare reaction with an average energy of about 7.0 MeV. Suppose Earth just happens to be at the right distance from the Sun to be at the first peak, as calculated in problem 1. What would be the correct Δm^2 to account for the solar neutrino problem? This was called the *just-so* solution.

3. For the actual mass difference $m_2^2 - m_1^2$, find the distance for the first peak for a 7.0 MeV neutrino. Compare to the radius of the Sun.

4. For neutrinos with momentum 7.0 MeV, what is the energy difference between neutrinos ν_2 and ν_1? In the presence of a large number of electrons, the energy of an electron neutrino is shifted by an amount $\sqrt{2}G_F n_e$, where n_e is the number density of electrons. At what electron density n_e is this difference equal to the energy difference you just found? Give your answer in mol/cm^3.

5. Using the appropriate mass difference $m_3^2 - m_1^2$ for atmospheric neutrinos, for what energy E of neutrino would the first peak occur after neutrinos have traveled exactly one Earth diameter?

6. Three neutrinos with mass differences given by eqs. (12.8) have common energy $E = 10.0\text{ MeV}$. They travel across a distance of 10^{10} light years. How much difference in time is there between their arrival time?

7. Using eqs. (12.9) and (12.10), make a 3×3 table showing what the probability is for each of the three mass eigenstates $\nu_{1,2,3}$ to correspond to any of the three flavor eigenstates $\nu_{e,\mu,\tau}$. Check that the rows and columns add to approximately one. Since we don't know the phase, simply choose $\delta = 0$ to keep it simple.

8. In Section 12C, I got negative mass for some of the neutrinos and claimed "this can be fixed." Suppose Ψ satisfies the Free Dirac equation with *negative* mass, so $\left(i\not{\partial} + m\right)\Psi = 0$. Define $\Psi' = \gamma_5 \Psi$. Show that Ψ' satisfies the Free Dirac equation with positive mass.

9. The mean lifetime for ^{76}Ge to decay via neutrinoless double beta decay is known to be longer than about $2.7\times 10^{25}\text{ y}$. Assuming that natural germanium is 7.8% ^{76}Ge, how many kg of naturally occurring Ge would be required to get three decays per year?

10. Explain, insofar as you can easily do so, eq. (12.21) (hint: where did the powers of M_X come from? The rest is dimensional analysis and guessing). If the X-mass is of the order of the Planck scale from problem 1.9, how long does the proton live?

11. Using the crude time-temperature relation, eq. (12.22), estimate the following times. Convert to conventional units when times are much longer than a second:
    - (a) The Planck scale ($k_B T = M_P$ from problem 1.9).
    - (b) The GUT scale $k_B T = M_{GUT}$
    - (c) Electroweak symmetry was broken $k_B T = \frac{1}{3} M_W$
    - (d) Quarks become confined ($k_B T \approx 300\text{ MeV}$)
    - (e) Positrons annihilate with electrons $k_B T = \frac{1}{3} m_e$
    - (f) Primordial Nucleosynthesis
    - (g) Electrons combined with protons to make Hydrogen $k_B T = 0.3\text{ eV}$

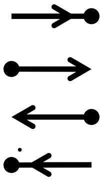
12. Using the crude time-temperature relation, eq. (12.22), estimate the following temperatures $k_B T$, when the Universe's age was the same as the mean lifetime of
    - (a) a muon
    - (b) a neutron
    - (c) a ^{7}Be nucleus (created during nucleosynthesis)
    - (d) a tritium ^{3}H nucleus
    - (e) an American human

Index

Feynman Rules for the Standard Model

External lines

Fermions (any)



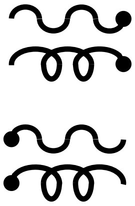
$u(p,s)$

$\bar{u}(p,s)$

$\bar{v}(p,s)$

$v(p,s)$

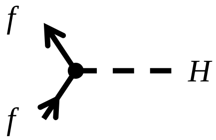
Spin -1 (any)



$\varepsilon_\mu(q,\sigma)$

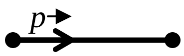
$\varepsilon^*_\mu(q,\sigma)$

Fermion-Higgs coupling

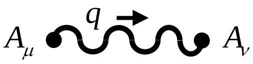


$-i\frac{m_f}{v}$

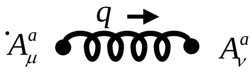
Propagators



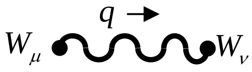
$$\frac{i(\not{p}+m)}{p^2-m^2}=\frac{i}{\not{p}-m}$$



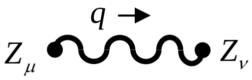
$$-\frac{ig_{\mu\nu}}{q^2}$$



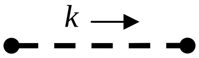
$$-\frac{ig_{\mu\nu}}{q^2}$$



$$\frac{i\left(-g_{\mu\nu}+q_\mu q_\nu/M_W^2\right)}{q^2-M_W^2}$$

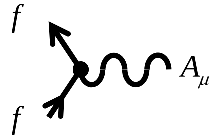


$$\frac{i\left(-g_{\mu\nu}+q_\mu q_\nu/M_Z^2\right)}{q^2-M_Z^2}$$

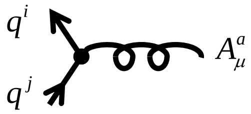


$$\frac{i}{k^2-M_H^2}$$

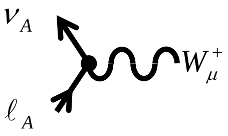
Fermion-gauge couplings



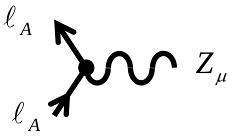
$-ieQ\gamma^\mu$



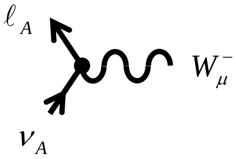
$-ig_s\left(T_a\right)^i{}_j\,\gamma^\mu$



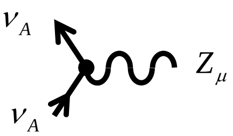
$$-i\frac{e\gamma^\mu(1-\gamma_5)}{2\sqrt{2}\sin\theta_W}$$



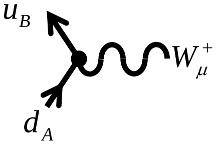
$$\frac{ie\gamma^\mu\left(1-4\sin^2\theta_W-\gamma_5\right)}{4\sin\theta_W\cos\theta_W}$$



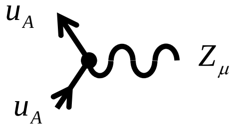
$$-i\frac{e\gamma^\mu(1-\gamma_5)}{2\sqrt{2}\sin\theta_W}$$



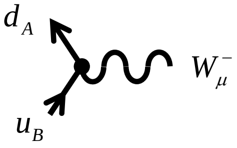
$$-\frac{ie\gamma^\mu(1-\gamma_5)}{4\sin\theta_W\cos\theta_W}$$



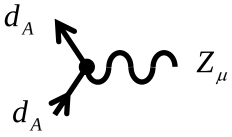
$$-i\frac{e\gamma^\mu(1-\gamma_5)}{2\sqrt{2}\sin\theta_W}V_{BA}$$



$$-\frac{ie\gamma^\mu\left(1-\frac{8}{3}\sin^2\theta_W-\gamma_5\right)}{4\sin\theta_W\cos\theta_W}$$

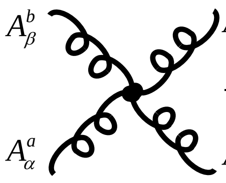


$$-i\frac{e\gamma^\mu(1-\gamma_5)}{2\sqrt{2}\sin\theta_W}V^*_{BA}$$

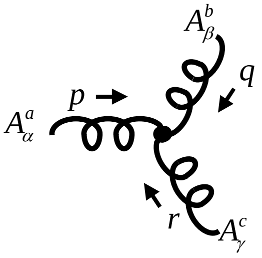


$$\frac{ie\gamma^\mu\left(1-\frac{4}{3}\sin^2\theta_W-\gamma_5\right)}{4\sin\theta_W\cos\theta_W}$$

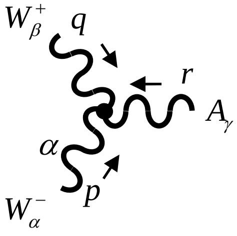
Gauge Boson Self-Couplings



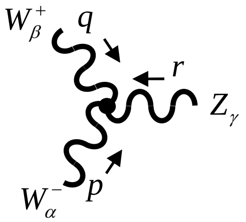
$$-ig^2\begin{bmatrix} f_{abe}f_{cde}\left(g^{\alpha\gamma}g^{\beta\delta}-g^{\alpha\delta}g^{\beta\gamma}\right)\\ +f_{ace}f_{bde}\left(g^{\alpha\beta}g^{\gamma\delta}-g^{\alpha\delta}g^{\beta\gamma}\right)\\ +f_{ade}f_{bce}\left(g^{\alpha\beta}g^{\gamma\delta}-g^{\alpha\gamma}g^{\beta\delta}\right)\end{bmatrix}$$



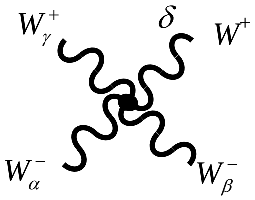
$$gf_{abc}\begin{bmatrix}(p-q)^\gamma g^{\alpha\beta}+\\ (r-p)^\beta g^{\alpha\gamma}+\\ (q-r)^\alpha g^{\beta\gamma}\end{bmatrix}$$



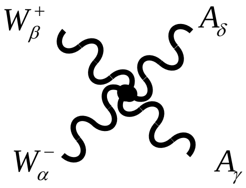
$$ie\begin{bmatrix}(q-r)^\alpha g^{\beta\gamma}\\ +(r-p)^\beta g^{\alpha\gamma}\\ +(p-q)^\gamma g^{\alpha\beta}\end{bmatrix}$$



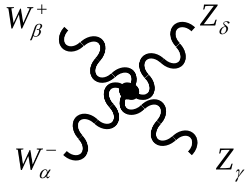
$$ie\cot\theta_W\begin{bmatrix}(q-r)^\alpha g^{\beta\gamma}\\ +(r-p)^\beta g^{\alpha\gamma}\\ +(p-q)^\gamma g^{\alpha\beta}\end{bmatrix}$$



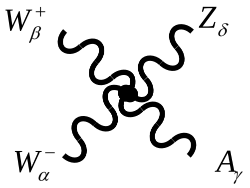
$$\frac{ie^2}{\sin^2\theta_W}\begin{bmatrix}2g^{\alpha\beta}g^{\gamma\delta}\\ -g^{\alpha\gamma}g^{\beta\delta}\\ -g^{\alpha\delta}g^{\beta\gamma}\end{bmatrix}$$



$$ie^2\begin{bmatrix}2g^{\alpha\beta}g^{\gamma\delta}\\ -g^{\alpha\gamma}g^{\beta\delta}\\ -g^{\alpha\delta}g^{\beta\gamma}\end{bmatrix}$$

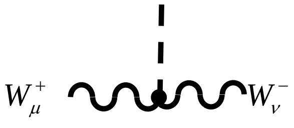


$$ie^2\cot^2\theta_W\begin{bmatrix}2g^{\alpha\beta}g^{\gamma\delta}\\ -g^{\alpha\gamma}g^{\beta\delta}\\ -g^{\alpha\delta}g^{\beta\gamma}\end{bmatrix}$$

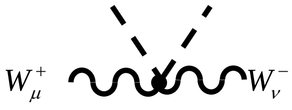


$$ie^2\cot\theta_W\begin{bmatrix}2g^{\alpha\beta}g^{\gamma\delta}\\ -g^{\alpha\gamma}g^{\beta\delta}\\ -g^{\alpha\delta}g^{\beta\gamma}\end{bmatrix}$$

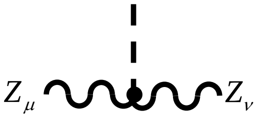
Higgs Gauge and Higgs Self-Couplings



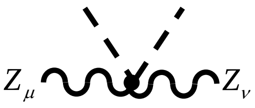
$$\frac{2iM_W^2g^{\mu\nu}}{v}$$



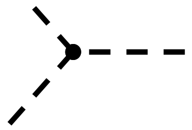
$$\frac{2iM_W^2g^{\mu\nu}}{v^2}$$



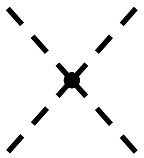
$$\frac{2iM_Z^2g^{\mu\nu}}{v}$$



$$\frac{2iM_Z^2g^{\mu\nu}}{v^2}$$



$$-\frac{3iM_H^2}{v}$$



$$-\frac{3iM_H^2}{v^2}$$