

PHY 362L - Subatomic Physics

Michael V. De Palatis

Spring 2006

Contents

1	Introduction	2
1.1	What Do We Want Out of Physics?	2
1.2	Experimental Versus Theoretical Problems	2
1.3	An Introduction to the Standard Model	2
1.4	The Four Fundamental Interactions	3
1.4.1	Gravity	3
1.4.2	Electromagnetism	3
1.4.3	Strong	3
1.4.4	Weak	3
1.4.5	Mediators of the Fundamental Interactions	4
1.5	Dimensions and Units	4
2	Relativity	6
2.1	Relativistic Kinematics	6
2.2	Four-Vectors	6
2.2.1	Multiplying Four-Vectors	7
2.2.2	Going Between Frames	7
2.2.3	Two Colliding Particles in the Center of Momentum Frame	8
2.2.4	Other Invariant Variables	9
3	Conservation Laws	10
3.1	Conservation Laws and Quantum Mechanics	10
3.1.1	Translations	11
3.1.2	Rotations	12
3.2	Other Conservation Laws	12
3.3	Reflection Invariance	13
3.3.1	Example: Sinusoidal Wavefunctions	15
3.3.2	Example: Hydrogen Atom	15
3.3.3	Example: Pions	16
4	Scattering	17
4.1	History	17
4.2	Rutherford's Nuclear Model	17
4.3	Mean Free Path Between Scatters	19
4.4	Rutherford's Calculation	19
4.5	Mott Scattering	21
4.6	Rutherford Scattering Using Time-Dependent Perturbation Theory	22

5	The Nucleus	27
5.1	Nuclear Masses	27
5.2	Nucleon Interaction	27
5.3	The Liquid Drop Model	27
5.4	Nuclear Instability and Decay	30
5.4.1	Decay Rates	30
5.4.2	Units of Measurement	30
5.4.3	Multiple Decay Branches	31
5.5	Alpha Decay	32
5.6	Nuclear Reactions and Cross Sections	34
5.7	Fission, Chain Reactions, and Moderation	35
5.8	Nuclear Fusion	35
5.9	Beta Decay	37
5.9.1	The Neutrino	37
5.9.2	Fermi's Calculation	38
5.9.3	Parity Nonconservation	42
6	Angular Momentum and Raising and Lowering Operators	44
6.1	Commutators	44
6.2	Raising and Lowering Operators	44
6.3	Addition of Angular Momenta	45
7	Charge Conjugation and Time Reversal Symmetries	47
7.1	Charge Conjugation	47
7.1.1	Positronium	48
7.2	CP Conservation and Violation	48
7.3	Time Reversal Invariance	48
8	Rotations, $SU(2)$, and Isospin	50
8.1	Rotations	50
8.1.1	An Example	50
8.2	Isospin	51
8.2.1	Example: Deuteron	52
8.2.2	π^\pm -Nucleon Scattering	53
9	Relativistic Quantum Mechanics	55
9.1	Lagrangians and the Universe	55
9.2	Relativistic Quantum Mechanics	56
9.3	The Dirac Equation	57
9.3.1	Motivation	57
9.3.2	Dirac Gamma Matrices and the Covariant Form	58
9.3.3	Conserved Current	60
9.3.4	Other Conserved Currents and γ^5	60
9.3.5	The Dirac Equation Lagrangian	61
10	Quantum Electrodynamics	63
10.1	Maxwell's Equations	63
10.2	Electromagnetic Lagrangian	64
10.3	Symmetries	64
10.4	Quantization of the Electromagnetic Field	65
10.5	Local Gauge Invariance	66
10.6	Interactions	67

10.7	Feynman Diagrams and Feynman Rules	69
10.7.1	Higher Order Diagrams	72
10.8	Magnetic Fields	73
11	The Weak Interaction	74
11.1	Pion Decay	74
11.2	The Weak Lagrangian	75
11.3	Cross Sections	77
11.3.1	Calculating Cross Sections — Revisited	77
11.3.2	Four-Point Interaction	78
11.4	Charged and Neutral Currents	79
11.4.1	Beta Decay Revisited	80
11.4.2	Neutral Current	80
11.5	The Electroweak Interaction	81
12	Quarks and Leptons	85
12.1	Quarks	85
12.1.1	Meson Decay	85
12.2	Meson Oscillation	86
12.3	CP Violation	88
12.4	Mixing in the Lepton Sector	89
12.5	The Strong Interaction	89
12.5.1	Finding Quarks	90
12.5.2	Quantum Chromodynamics	91
12.5.3	Strong Potential	91
12.6	A Summary of the Standard Model	91

Chapter 1

Introduction

1.1 What Do We Want Out of Physics?

What is the goal of studying the fundamental building blocks of matter? Is it to find truth? If so, then what really is truth? The main goal is to be able to describe the entire universe as it is today up to initial conditions. In that respect, subatomic physics as it stands today contains mostly models without theories (an analagous situation would be Kepler’s laws; they allow one to calculate the positions of the planets give particular parameters, but do not explain how they got where they are now).

1.2 Experimental Versus Theoretical Problems

There can be a significant distinction between experimental problems and theoretical ones. Examples of experimental problems include the precession of the perihelion of Mercury¹ and blackbody radiation.

Theoretical problems, on the other hand, are unique in that the only problems that arise are in a sense “imaginary.” That is, by imagining an experiment (which may not even be possible to perform), one can show that a theory or model *must* break down.

1.3 An Introduction to the Standard Model

We know today that all the matter we see is made up of two types of particles: Quarks and leptons. Of these, there are exactly three generations². They are organized as shown in Table 1.1.

Quarks	u	c	t
	d	s	b
Leptons	e^-	μ^-	τ^-
	ν_e	ν_μ	ν_τ

Table 1.1: The Standard Model. Each particle above also has an associated antiparticle.

Each of these particles is spin-1/2, making them all fermions. This is significant because this makes the Pauli exclusion principle come into effect for these most³ fundamental building blocks of matter.

The quarks have either charge $+2/3e$ or $-1/3e$: The “up type quarks” (i.e., the up, charmed, and top quarks) have the former, while the “down type quarks” (i.e., the down, strange, and bottom quarks) have

¹This was in fact the main reason why Einstein’s general relativity eventually came to be adopted as the “true” theory of gravity over Newton’s.

²Why three? Nobody knows.

³Probably not really, but close enough for our purposes.

the latter. The leptons either have charge $-e$ (in the case of the electron, muon, and tau lepton) or 0 (in the case of the neutrinos).

A number of quarks can be put together to form a composite particle. Such particles are known as *hadrons*. A hadron consisting of a quark-antiquark pair (schematically, $q\bar{q}$) is called a *meson*. A hadron consisting of three quarks (qqq) is called a *baryon*. For example, the proton consists of two up quarks and a down quark (uud) and the neutron consists of an up quark and two down quarks (udd). Given the charges of the quarks, we see that the proton does in fact have a net charge of $+e$ while the neutron has no net charge.

1.4 The Four Fundamental Interactions

The following is a brief outline of the four fundamental interactions⁴ in nature.

1.4.1 Gravity

Gravity is the force that holds the universe together. Yet despite this fact, it is exceptionally weak on the small scale. Between two protons, the gravitational force is

$$F = \frac{GM_p^2}{4\pi\hbar c} \sim 10^{-40}.$$

The gravitational force also has an infinite range and drops off like r^{-2} .

1.4.2 Electromagnetism

The electromagnetic force holds atoms and molecules together. We represent its strength with the finestructure constant:

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \approx \frac{1}{137}.$$

Like gravity, electromagnetism has an infinite range and drops off like r^{-2} . Due to being so much stronger than gravity, electromagnetism tends to be the most important interaction in our everyday lives. The electromagnetic force couples to two charges, either $+$ or $-$.

1.4.3 Strong

The strong interaction holds quarks together. Unlike the above two interactions, the strong force gets stronger as r increases, though only up to about 10^{-15} m. That is,

$$\alpha_S = \begin{cases} \sim 1 & \text{if } r \sim 10^{-15} \text{ m} \\ < 1 & \text{if } r \ll 10^{-15} \text{ m} \end{cases}$$

The strong interaction couples to “color,” with the three colors being (r)ed, (g)reen, and (b)lue.

1.4.4 Weak

Unlike the other three interactions, the weak force holds *nothing* together. As anticlimactic as that may seem, the weak force is responsible primarily for nuclear decay. Its strength is measured by

$$\alpha_W \sim \left(\frac{mc}{\hbar}\right)^2 \frac{G_F}{\hbar c} \sim 10^{-15}$$

and has a range on the order of 10^{-5} . The weak interaction couples to the “weak hypercharge.”

⁴Oftentimes, we call say the “four fundamental forces,” but the usage of the word “interactions” is really more appropriate.

1.4.5 Mediators of the Fundamental Interactions

Of the four fundamental interactions, all four are felt by the quarks whereas the leptons feel all but the strong (so that is to say, leptons have no color). Each force is (at least in theory) mediated by a boson (that is, the mediating particles have integral spin).

- The electromagnetic force is mediated by the *photon*, which is represented by the Greek letter γ . The photon has zero mass and spin 1.
- The strong force is mediated by the *gluon*, represented by the letter g . Gluons have zero mass and zero spin.
- The weak force is mediated by the W^+ , W^- , and Z^0 bosons. Each has a mass on the order of 90 GeV/c and spin 1.
- The gravitational force is theorized to be mediated by gravitons. If these particles exist, they are thought to have zero mass and spin 2.

1.5 Dimensions and Units

In nuclear and particle physics, we often choose to use units other than SI due to their convenience. We start with energy, which we measure in electron volts (eV). Typically, in nuclear physics we deal in the MeV (10^6 eV) scale while in particle physics we reach GeV (10^9) and even TeV (10^{12}). In both cases, compare this to atomic physics, which is in the several eV range (e.g., the ground state energy of hydrogen at -13.6 eV).

The typical length scale is the *femtometer*, or as it is more commonly called, the *fermi*. $1 \text{ fm} = 10^{-15} \text{ m}$. By Einstein's famous formula,

$$E = mc^2, \tag{1.1}$$

so we measure mass in eV/c^2 , typically referring to MeV/c^2 or GeV/c^2 .

As for momentum, the “master formula” relating energy, mass, and momentum is

$$E^2 = p^2c^2 + m^2c^4, \tag{1.2}$$

and so we see that momentum is to be in units of eV/c (usually MeV/c or GeV/c).

It is often easier to do our calculations in so-called “natural units.” That is, we set

$$\begin{aligned} c &= 1 \\ \hbar &= 1. \end{aligned}$$

So, for example, when reporting a mass, we would simply drop the “ c^2 .” When we want to get “real” answers, we can get back to “real” units using the easy to remember combination

$$\hbar c = 197 \text{ MeV fm} = 0.197 \text{ GeV fm}. \tag{1.3}$$

In terms of the common units,

$$\begin{aligned} c &= 3 \times 10^{23} \text{ fm/s} \\ \hbar &= 6.58 \times 10^{-22} \text{ MeV s}. \end{aligned}$$

As an example of using natural units, consider the uncertainty principle. In the case of “real” units, we write for the energy-time uncertainty principle that

$$\Delta E \Delta t \geq \frac{\hbar}{2},$$

so now we drop the \hbar to write

$$\Delta E \Delta t \geq \frac{1}{2}.$$

Thus we see that in natural units, energy is measured in s^{-1} . Similarly, from the Heisenberg uncertainty principle, we see that momentum is measured in units of $(\text{length})^{-1}$.

Chapter 2

Relativity

In many topics in physics, we need not even consider relativistic effects. However, we recall the rule of thumb that when the kinetic energy of a particle is comparable to its mass, we have no choice but to take into account relativity. In subatomic physics, this is nearly always the case.

2.1 Relativistic Kinematics

The energy of a particle is related to its mass by Einstein’s famous formula

$$E = m\gamma, \tag{2.1}$$

(in natural units), where γ is the Lorentz factor

$$\gamma = \frac{1}{\sqrt{1 - \beta^2}} \tag{2.2}$$

with $\beta = v$ when in natural units.

Energy, mass, and momentum are all related by

$$E^2 = p^2 + m^2, \tag{2.3}$$

which implies that for massless particles, $E = p$ (e.g., the photon). Additionally, *only* massless particles can move with speed c , and they must *always* move with speed c .

2.2 Four-Vectors

Four-vectors are a useful type of vector that holds all the spacetime information of a particle. As such, it is invariant under coordinate transforms.

One particularly useful four-vector is the four-momentum vector, $\tilde{p} = (E, \vec{p})$ ¹. Note that this contains information both in the one time coordinate (via E) as well as all three space coordinates (via \vec{p}). Similarly, the four-position vector is (t, \vec{r}) . Another useful four-vector is that containing the scalar and vector potentials from electromagnetism: (ϕ, \vec{A}) . In general, any scalar quantity and some related three dimensional vector can be used in the form of a four-vector.

¹Note that any quantity with a tilde over it is a four-vector, whereas a quantity with an arrow over it is a “normal” vector.

2.2.1 Multiplying Four-Vectors

Given two four-vectors \tilde{A} and \tilde{B} , they can be multiplied as follows²:

$$\tilde{A} \cdot \tilde{B} = A^0 B^0 - \vec{A} \cdot \vec{B}.$$

The reason for the minus sign is not entirely straightforward, so what follows is a quick mathematical argument for it.

Another way we can write the above multiplication as³

$$\tilde{A} \cdot \tilde{B} = A_\mu B^\mu = A^\mu B_\mu = g_{\mu\nu} A^\mu B^\mu,$$

where the *metric*⁴ $g_{\mu\nu}$ is (here) for flat spacetime (*Minkowski space*):

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (2.4)$$

Using this, we see that

$$\begin{aligned} \tilde{p} \cdot \tilde{p} &= g_{\mu\nu} p^\mu p^\nu \\ &= p^0 p^0 - \vec{p} \cdot \vec{p} \\ &= E^2 - \vec{p} \cdot \vec{p} \\ &= E^2 - p^2 \\ &= m^2. \end{aligned}$$

This equation is *invariant*, and hence we call m the *invariant mass*.

2.2.2 Going Between Frames

Say we measure p and E in the lab, but we want to know their values in the center of momentum⁵ frame. Then we “boost” (along the z axis) into the new coordinate frame by using the Lorentz transformation

$$L = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \quad (2.5)$$

So, for example, in the “boosted” frame,

$$\begin{pmatrix} E' \\ p'_x \\ p'_y \\ p'_z \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} E \\ p_x \\ p_y \\ p_z \end{pmatrix}.$$

This gives

²The superscript 0 signifies the time component of each four-vector, while the arrowed quantity refers to the space component of the four-vector.

³The notation follows Einstein’s convention with implicit summing over $\mu = 0, 1, 2, 3$.

⁴A metric tells one how to measure distance. It can be found using the Einstein field equations, which are well beyond the scope of this course.

⁵Sometimes we say “center of mass,” though in subatomic physics, we almost always really mean “center of momentum.”

$$\begin{aligned}
E' &= \gamma E - \beta\gamma p_z \\
p'_x &= p_x \\
p'_y &= p_y \\
p'_z &= -\beta\gamma E + \gamma p_z
\end{aligned}$$

Note then that here, $E' < E$ (assume an ultrarelativistic particle, then $E \approx p_z$ and $\beta = 1$, so in this case, E' would approach zero).

To go the opposite way, the $-\beta\gamma$ terms in Equation 2.5 drop the negative sign.

2.2.3 Two Colliding Particles in the Center of Momentum Frame

For two colliding particles in the center of momentum frame, we define the quantity s such that

$$\begin{aligned}
s &= (\tilde{p}_1 + \tilde{p}_2)^2 \\
&= (\tilde{p}_1 + \tilde{p}_2)_\mu (\tilde{p}_1 + \tilde{p}_2)^\mu
\end{aligned}$$

Carrying out the multiplication, and exploiting the fact that we are using the center of momentum frame,

$$s = (E_1 + E_2)^2,$$

so

$$s = E_{cm}^2. \tag{2.6}$$

Because we used four-vectors, this is *always* true, regardless of what frame we use.

Example: Beam and Target

If the beam is particle 1 and the target particle 2, we can calculate the energy straightforwardly. The center of mass energy is

$$E_{cm}^2 = (\tilde{p}_1 + \tilde{p}_2)^2$$

where

$$\begin{aligned}
\tilde{p}_1 &= (E_{lab}, \vec{p}_{lab}) \\
\tilde{p}_2 &= (m_2, \vec{0}).
\end{aligned}$$

Then

$$\begin{aligned}
\tilde{p}_1 + \tilde{p}_2 &= (m_2 + E_{lab}, \vec{p}_{lab}) \\
(\tilde{p}_1 + \tilde{p}_2)^2 &= (m_2 + E_{lab})^2 - \vec{p}_{lab} \cdot \vec{p}_{lab} \\
&= m_2^2 + 2m_2 E_{lab} + E_{lab}^2 - p_{lab}^2.
\end{aligned}$$

Then since $E_{lab} = p_{lab}^2 + m_1^2$,

$$\begin{aligned}
E_{cm}^2 &= (\tilde{p}_1 + \tilde{p}_2)^2 = m_2^2 + 2m_2 E_{lab} + p_{lab}^2 + m_1^2 - p_{lab}^2 \\
E_{cm}^2 &= m_2^2 + m_1^2 + 2m_2 E_{lab}.
\end{aligned}$$

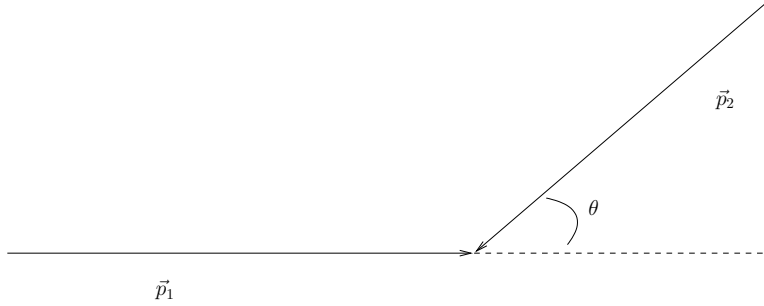


Figure 2.1: Two beams colliding.

Note that in order to double the energy observed in the lab frame, we must *quadruple* it in the center of mass frame. As we shall see in the next example, this can be improved. However, most experiments are done this way rather than with two colliding beams for a number of reasons: The target can be made very large in order to increase the “luminosity” (i.e., increase the number of interactions).

Example: Colliding Beams

TODO...

General Collisions

There is of course another case, that of two colliding beams that are not travelling in exact opposite directions. Now we have

$$\begin{aligned}
 s &= E_{cm}^2 = (\tilde{p}_1 + \tilde{p}_2)^2 \\
 &= [(E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2] \\
 &= E_1^2 + 2E_1E_2 + E_2^2 - p_1^2 - p_2^2 - 2\vec{p}_1 \cdot \vec{p}_2 \\
 &= m_1^2 + m_2^2 + 2E_1E_2 + E_2^2 - 2p_1p_2 \cos(\theta).
 \end{aligned}$$

We can simplify this using $\gamma^2 = (1 - \beta^2)^{-1}$ and $E^2 = p^2 + m^2 = p^2 + E^2/\gamma^2$. Some quick arithmetic shows that

$$p^2 = \beta^2 E^2. \quad (2.7)$$

Finally, we can rewrite the center of momentum energy squared as

$$E_{cm}^2 = m_1^2 + m_2^2 + 2E_1E_2 [1 - \beta_1\beta_2 \cos(\theta)]. \quad (2.8)$$

2.2.4 Other Invariant Variables

In addition to s , there are also several other useful invariants. One is the *momentum transfer* t , defined by

$$t = (\tilde{p}_1 - \tilde{p}_2)^2. \quad (2.9)$$

Chapter 3

Conservation Laws

Often in science, the first method of discrimination between two theories that equally match observations is Occam's razor. Today in physics, this is achieved via symmetry: The theory which is more symmetric is deemed the better one.

But what exactly do we mean by symmetry? Essentially, it means that we expect physics to work the same way under coordinate transformations. For example, if I do an experiment in Austin, I expect to get the same results as the same experiment performed in Moscow. Another way to look at it is the following. For every symmetry in a theory, there exists a corresponding conservation law¹. A conserved quantity, to be precise, is one that does not change with time.

3.1 Conservation Laws and Quantum Mechanics

As we all know, in quantum mechanics, observables are eigenvalues of operators. Say we have some (time independent) operator F acting on the wavefunction Ψ . Then the observable associated with F is its expectation value

$$\langle \Psi | F | \Psi \rangle \equiv \langle F \rangle = \int \Psi^* F \Psi d^3 r.$$

To find if $\langle F \rangle$ is constant in time, we take the time derivative.

$$\begin{aligned} \frac{d\langle F \rangle}{dt} &= \frac{d}{dt} \int \Psi^* F \Psi d^3 r \\ &= \int \frac{d\Psi^*}{dt} F \Psi d^3 r + \int \Psi^* F \frac{d\Psi}{dt} d^3 r. \end{aligned}$$

The Schrödinger equation states

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi, \tag{3.1}$$

and its conjugate,

$$-i\hbar \frac{\partial \Psi^*}{\partial t} = (H\Psi)^* = \Psi^* H. \tag{3.2}$$

¹When we talk about conservation laws, we are usually really only talking about *local* conservation laws. For example, we all know what happens to the teacher with the bicycle wheel on the turntable when he flips the wheel: He turns the opposite way on the turntable in order to conserve angular momentum. However, if he does this again, but this time with a student sitting on the other end of the class simultaneously flipping her wheel the opposite way, the same thing happens as before. Thus, angular momentum is conserved locally.

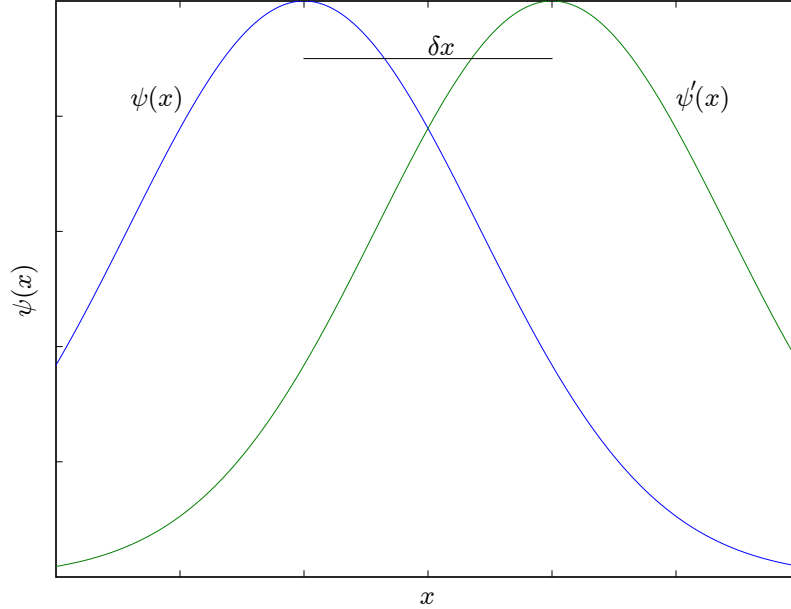


Figure 3.1: A small displacement of the wavefunction $\psi(x)$.

Thus,

$$\frac{d\langle F \rangle}{dt} = \frac{i}{\hbar} \int \Psi^* (HF - FH) \Psi d^3r. \quad (3.3)$$

Therefore, if $[H, F] = 0$, we can say that the quantity measured by F is conserved. Regardless of what operator F actually is, we see from this discussion that *any* time independent operator that commutes with the Hamiltonian measures a conserved quantity.

3.1.1 Translations

For the small change as in Figure 3.1, the new wavefunction is

$$\psi'(x) = \psi(x - \delta x) \sim \psi(x) - \delta x \frac{\partial \psi(x)}{\partial x}.$$

Recall that the momentum operator is given by $p = -i\hbar \partial/\partial x$. Then

$$\psi'(x) = \psi(x) + i\delta x p \psi(x).$$

We define the *generator of infinitesimal translations*² to be

$$D \equiv (1 + i\delta x p). \quad (3.4)$$

For a finite translation, we simply raise D to the N^{th} power:

²This has a fancy name because it comes from group theory. To greatly simplify things, a generalization can apply to other transformations, such as those in time.

$$D^N \rightarrow e^{i\delta xp}. \quad (3.5)$$

So from Equations 3.4 and 3.5, we see that requiring physics to be invariant under spatial translations depends on the p operator. In turn, this implies

$$[D, H] = 0, \quad (3.6)$$

i.e., momentum is conserved. That is, we now have a connection between translational symmetry and conservation of (linear) momentum.

3.1.2 Rotations

The case of a rotational transformation is nearly identical to that of the translational transformation. For some wavefunction ψ and some rotation angle ϕ ,

$$\begin{aligned} \psi'(\phi) &= \psi(\phi - \delta\phi) \\ &\approx \psi(\phi) - \delta\phi \frac{\partial}{\partial\phi} \psi(\phi). \end{aligned}$$

The angular momentum operator³ (along the z axis) is defined as

$$\begin{aligned} J_z &= -(xp_y - yp_x) \\ &= \frac{-i\hbar}{2\pi} \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \\ J_z &= -i\hbar \frac{\partial}{\partial\phi}. \end{aligned} \quad (3.7)$$

This gives us the generator of infinitesimal rotations:

$$R \equiv 1 + \frac{i}{\hbar} J_z \delta\phi, \quad (3.8)$$

and finite rotations:

$$R^N \rightarrow e^{iJ_z \delta\phi/\hbar} \quad (3.9)$$

Again,

$$[R, H] = 0. \quad (3.10)$$

3.2 Other Conservation Laws

In addition to the conservation of linear and angular momentum, there are several other conservation laws listed in Table 3.1.

Gauge invariance symmetry arises from a combination of gauge invariance in electrodynamics and the ability to add phase factors to wave functions without changing anything physical⁴. When put together, we find that charge is conserved.

An *accidental* symmetry is one in which all equations involving a quantity show that it is conserved, yet there is no fundamental reason for it. In other words, there is no explicit equation that states that lepton

³This analysis applies equally well to either orbital or spin angular momentum.

⁴Tacking on $e^{i\phi}$ doesn't matter in the long run since the physical quantity we care about is $|\psi|^2$.

Symmetry	Conservation Law
Additive Symmetries	
Translation	Linear Momentum
Rotation	Angular Momentum
Time Translation	Energy
Gauge Invariance	Charge
<i>Accidental</i>	Lepton Number
<i>Accidental</i>	Baryon Number
Multiplicative Symmetries	
Charge Conjugation	Charge Conjugation
Reflection (Time Reversal)	Parity

Table 3.1: A listing of symmetries and their associated conserved quantities. Note that the conservation of lepton and baryon numbers is *accidental*, i.e., the fact that they are conserved is not fundamental, but instead “just happens.”

number, e.g., *must* be conserved. For the conservation of lepton number, leptons are assigned a lepton number of 1 while antileptons are assigned -1. All other particles have lepton number 0. This conservation law allows the pion decay

$$\pi^+ \rightarrow \mu^+ + \nu_\mu$$

to occur. Note that in terms of lepton number, we have

$$0 \rightarrow (-1) + (+1) = 0.$$

Also note that until recently, it was believed that lepton flavor was conserved. However, with neutrino oscillation and several experiments, this is now known to be false.

Baryon number conservation is similar to that of lepton number. Baryons are assigned baryon number of 1, while antibaryons get -1. All other particles have baryon number 0. This conservation law allows the reaction

$$p \rightarrow e^+ + \nu_e + n.$$

Charge conjugation and parity combine to say that a matter universe behaves in exactly the same way as an antimatter universe.

An additive conservation law is one in which the sum of the conserved quantity on each side of the reaction must be equal. A multiplicative conservation law acts similarly, only it is the product that must be conserved rather than the sum.

3.3 Reflection Invariance

The parity operator P acts on a wavefunction ψ according to

$$P\psi(\vec{r}) = \psi(-\vec{r}). \quad (3.11)$$

If ψ is an eigenfunction of parity, then

$$P\psi(\vec{r}) = \pm\psi(\vec{r}). \quad (3.12)$$

Of course, applying the parity operator again should give us the original wavefunction back.

$$P[P\psi(\vec{r})] = \psi(\vec{r}). \quad (3.13)$$

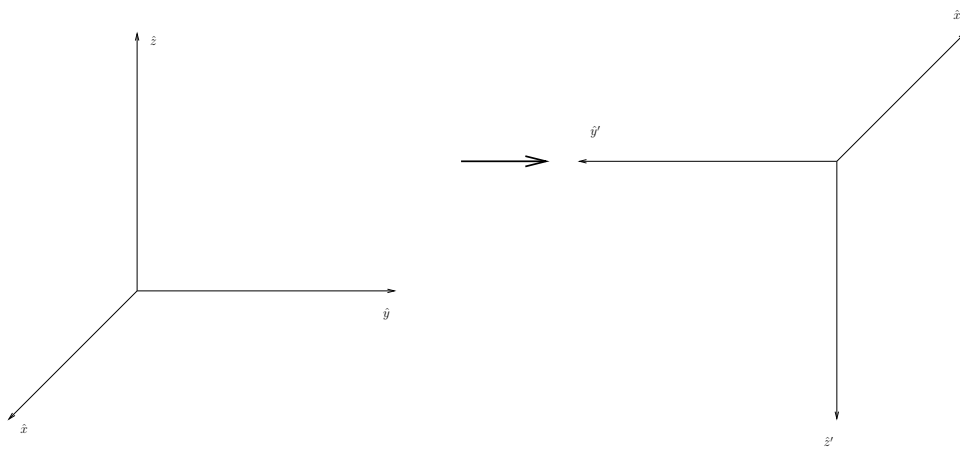


Figure 3.2: Reflecting a coordinate system by direct flipping of all the coordinates.

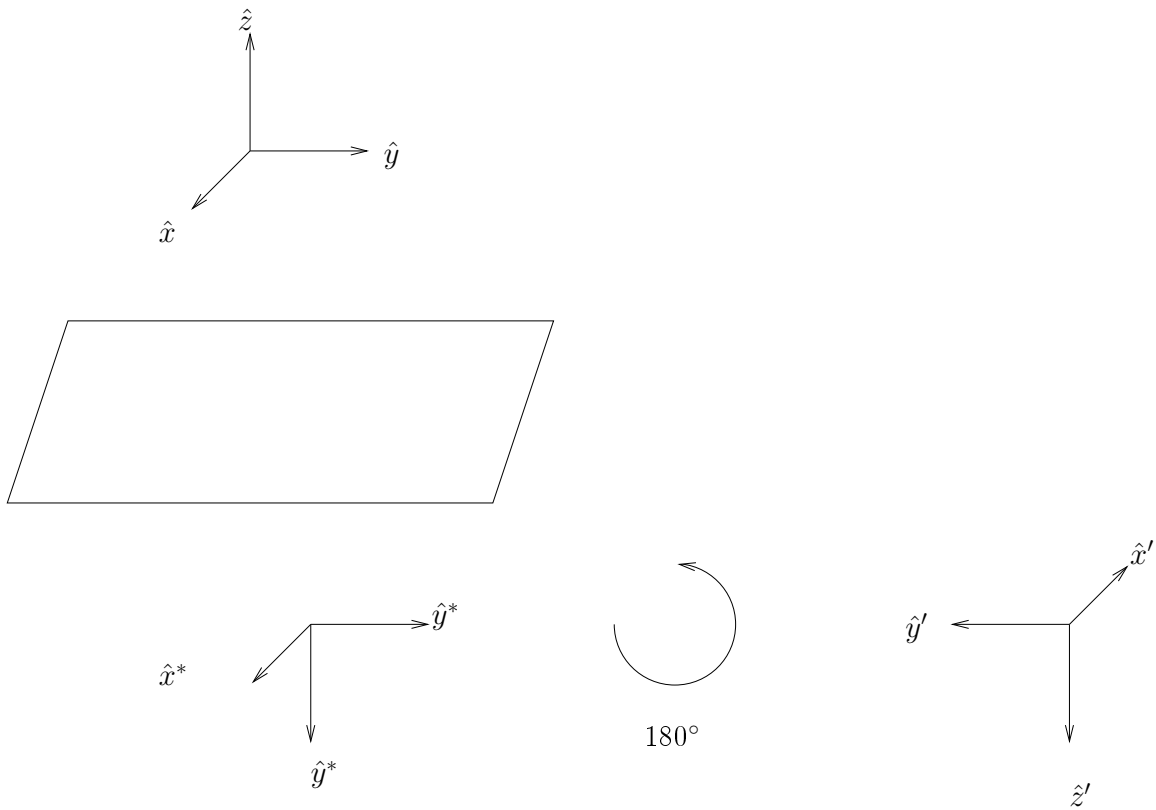


Figure 3.3: Reflecting a coordinate system by first mirroring it, then rotating it by 180 degrees.

Quantities which transform like

$$\vec{\alpha} \rightarrow -\vec{\alpha}$$

we call vectors. Examples include the spatial coordinate \vec{r} , momentum \vec{p} , and electric field \vec{E} . Quantities which transform like

$$\vec{\beta} \rightarrow \vec{\beta}$$

we call *axial* or *pseudovectors*. Examples include⁵ angular momentum \vec{L} and magnetic field \vec{B} .

Other quantities do not change under parity. That is, they transform like

$$q \rightarrow q.$$

Such quantities are known as *scalars*. Examples include charge, scalar potential, etc.

Finally, some quantities transform like

$$\frac{\vec{s} \cdot \vec{p}}{E} \rightarrow -\frac{\vec{s} \cdot \vec{p}}{E}.$$

These are called *pseudoscalars*. The above example is the *helicity*.

3.3.1 Example: Sinusoidal Wavefunctions

Take the wavefunction $\psi = \cos(x)$. Under parity,

$$P\psi(x) = \cos(-x) = \cos(x) = +\psi(x).$$

Because ψ remains the same, we say it is “even” under parity. If instead we have $\psi = \sin(x)$, then

$$P\psi = \sin(-x) = -\sin(x) = -\psi(x).$$

In this case, we say that ψ is “odd” under parity. Not everything needs to be an eigenfunction of parity. For example, say we have $\psi = \cos(x) + \sin(x)$. Then

$$P\psi = \cos(x) - \sin(x) \neq \pm\psi.$$

3.3.2 Example: Hydrogen Atom

The Coulomb potential $V(r)$ has spherical symmetry, thus under parity,

$$PV(\vec{r}) = V(-\vec{r}) = V(\vec{r}).$$

Therefore, we expect the wavefunction to also be parity symmetric. Hydrogen wavefunctions take on the form

$$\begin{aligned} \psi(r, \theta, \phi) &= X(r)Y_l^m(\theta, \phi) \\ &= X(r)\sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}P_l^m[\cos(\theta)]e^{im\phi}. \end{aligned}$$

Note that ϕ only appears in the complex exponential part. That is, it has no effect on the location of the electron. So far, so good for symmetry. The parity operation in spherical coordinates involves transforming θ and ϕ by

⁵The easiest way to visualize this is to use the mirror method of reflection (Figure 3.3). The direction of, say, the current, does not change in the mirror. This even after rotation, the magnetic field still points in the same direction.

$$\begin{aligned}\theta &\rightarrow \pi - \theta \\ \phi &\rightarrow \pi + \phi.\end{aligned}$$

Then the ϕ part of the wavefunction transforms like

$$e^{im\phi} \rightarrow e^{im(\pi+\phi)} = (-1)^m e^{im\phi}.$$

The Legendre polynomial goes to

$$P_l^m[\cos(\theta)] \rightarrow P_l^m[\cos(\pi - \theta)] = (-1)^{l+m} P_l^m[\cos(\theta)].$$

So as a whole, the spherical harmonic transforms as

$$Y_l^m(\theta, \phi) \rightarrow Y_l^m(\pi - \theta, \pi + \phi) = (-1)^l Y_l^m(\theta, \phi).$$

The $(-1)^l$ factor in front then implies that the Y_l^m are eigenstates of parity.

Now for the actual conservation of parity. Say we have a decay from $l = 1$ to $l = 0$. This produces a photon, which then must have (-1) parity!

$$\begin{aligned}{}^*H^{l=1} &\rightarrow H^{l=0} + \gamma \\ P = -1 &\rightarrow (P = 1) \times (P = -1) = -1\end{aligned}$$

It is not clear here exactly how the photon carries parity, but we say it does because we believe in parity conservation⁶.

3.3.3 Example: Pions

Say we have the following reaction:

$$p + p \rightarrow \pi^+ + p + n.$$

First we need to know what the intrinsic parities are. We choose the parity of nucleons⁷ (protons and neutrons) to be $(+1)$. Then $P_{\pi^+} = (+1)$ (remember, parity is a multiplicative conservation law).

The π^+ is actually the “antipion,” while the π^- is the “pion.” The quark composition of these mesons is $(u\bar{d})$ for the former, and $(\bar{u}d)$ for the latter. Thus, we see that $P_{\pi^-} = (-1)$. In general, antiparticles have the opposite parity as particles.

Quarks are spin-1/2, and the π^- has spin-0. So either we must have $s = 1$ with $l = 0$, or $s = 1$ with $l = 1$ such that the total angular momentum $\vec{J} = \vec{L} + \vec{S}$ is zero. This is because we still have parity going like $(-1)^l$ (this is the case in general). In general, for an excited state, there is a factor of $(-1)^l$ tacked on to the end of the intrinsic parity for both individual particles and each side of a reaction overall.

Interestingly, parity alone can tell us the state of the pion. Or, another way to look at it is that the different case is a different particle with a different mass, and so we can see (rather indirectly) the effects of parity.

⁶There are some notable exceptions to the conservation of parity, which we will discuss later.

⁷This is completely arbitrary. The fact that nucleons are baryons, and baryon number is conserved, allows us to choose either ± 1 however we want.

Chapter 4

Scattering

4.1 History

In 1909, Geiger and Marsden performed the first scattering experiments with α particles aimed at a thin foil. At the time, the atom was described by the “plum pudding model” (see Figure 4.1). They would have expected the α ray to go straight through with no scattering, but to their surprise, sometimes (though quite rarely) they would bounce back at extreme angles. This could be explained if the atom is composed of a hard, massive core surrounded by an electron cloud.

4.2 Rutherford’s Nuclear Model

Ernest Rutherford was the first to quantitatively explain the data first observed by Geiger and Marsden. His model consisted of the following.

1. Positive¹ charge is concentrated in the nucleus, which is point like.
2. Electrons surround the nucleus and have little effect on scattering.
3. Target nuclei are very massive so that there is no nuclear recoil.
4. Classical mechanics describes the collision. That is, the α particles were nonrelativistic ($\lambda_{dB} \ll p_\alpha$).
5. *Only* Coulomb interactions are present. Thus, the potential is given by $V(\vec{r}) = Zze^2/(4\pi\epsilon_0 r)$, where Z is the nuclear charge and z is the α particle charge (in units of e).
6. There are no excitations² of the incident α ray or the nucleus; i.e., collisions are completely elastic.

Ultimately, what we want from this model is a way to calculate the probability that a given α particle is scattered into a particular solid angle Ω . Define the *intensity* I as the number of incident particles per cm^2 s; N_T the number of target particles seen by the beam; the scattered flux $C(\theta, \phi)$ as the number of outgoing particles per second, or

$$C(\theta, \phi) = \frac{\# \text{ incident particles}}{\text{cm}^2 \text{ s}} \times \text{probability of scattering into } \Delta\Omega \times \# \text{ possible scatters} \times \Delta\Omega.$$

If we imagine the foil as a lattice of tiny discs, the cross sectional areas determine the probability of a scattering event. Thus

¹Was there really a reason for Rutherford to believe that the nucleus was positively charged? A likely explanation for this is that the electron was known by this point to be negative and light compared to the proton.

²This, of course, is not the way that Rutherford would have phrased it.

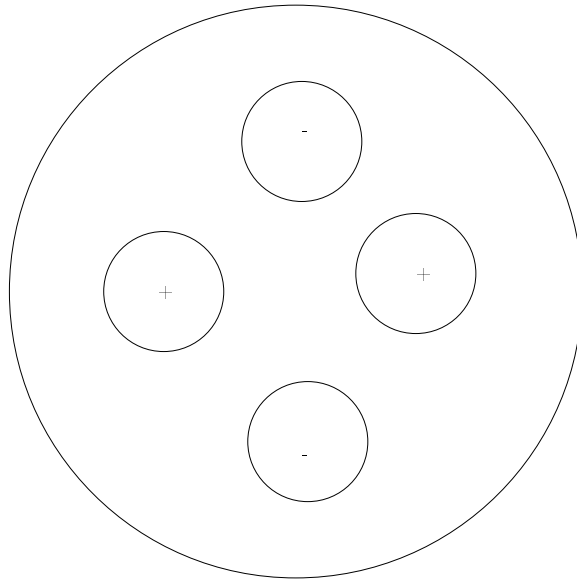


Figure 4.1: In the plum pudding model, the atom consists of protons and electrons next to each other, forming a neutral “blob.”

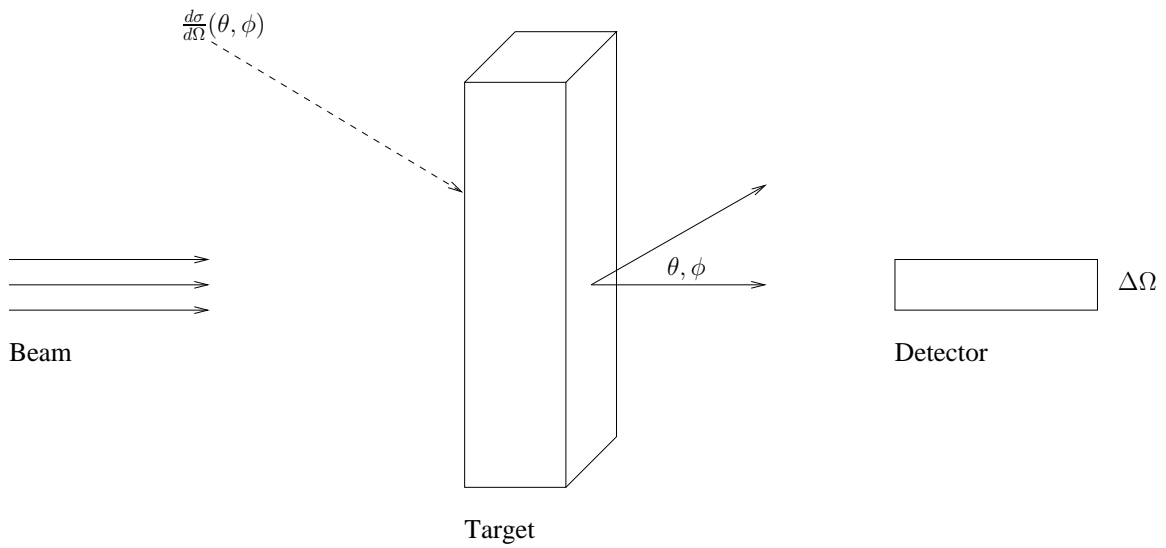


Figure 4.2: A schematic of the setup used to demonstrate Rutherford’s model of the atom.

$$C(\theta, \phi) = I \frac{d\sigma}{d\Omega}(\theta, \phi) N_T \Delta\Omega,$$

meaning that the *differential cross section* is given by

$$\frac{d\sigma}{d\Omega} = \frac{C(\theta, \phi)}{I N_T \Delta\Omega}.$$

Note that everything on the right hand side is a measurable quantity. $C(\theta, \phi)$ is essentially our data, and N_T can be calculated based on the density and thickness of the foil.

What we actually want to find is the *total* cross section, $\sigma = \int (d\sigma/d\Omega) d\Omega$. This is the fraction of the beam scattered, i.e.,

$$\sigma = \frac{\text{“area” of single target} \times N_T}{\text{area of beam}}.$$

So if N_B is the number of particles in the beam,

$$\frac{C}{N_B} = \frac{\sigma N_T}{\text{area of beam}},$$

thus

$$\sigma = \frac{C}{(N_B/A) N_T}. \quad (4.1)$$

Note that $N_B/A = I$.

4.3 Mean Free Path Between Scatters

If n is the number density of target particles, then the *mean free path* is given by

$$\lambda = \frac{1}{n\sigma} \quad (4.2)$$

As a function of distance, the number of particles in the beam drops off like

$$N = N_0 e^{-x/\lambda}$$

and

$$C = N_0 - N = N_0 (1 - e^{-x/\lambda}) \approx \frac{N_0 x}{\lambda} = N_0 n \sigma x,$$

where x in this case is the thickness of the foil. Also, $n\sigma x$ is the probability of some scattering.

4.4 Rutherford’s Calculation

First, we must realize that having just the Coulomb potential

$$V(r) = \frac{Zze^2}{4\pi\epsilon_0 r}$$

means that the probability of a scatter into any angle θ depends on the distance from the target to the beam alone.

The hyperbolic orbit of the α particle implies that we can treat it classical. Its energy is then given by

$$E = \frac{1}{2} m (\dot{r}^2 + r^2 \dot{\theta}^2) + V(r) = \text{constant}. \quad (4.3)$$

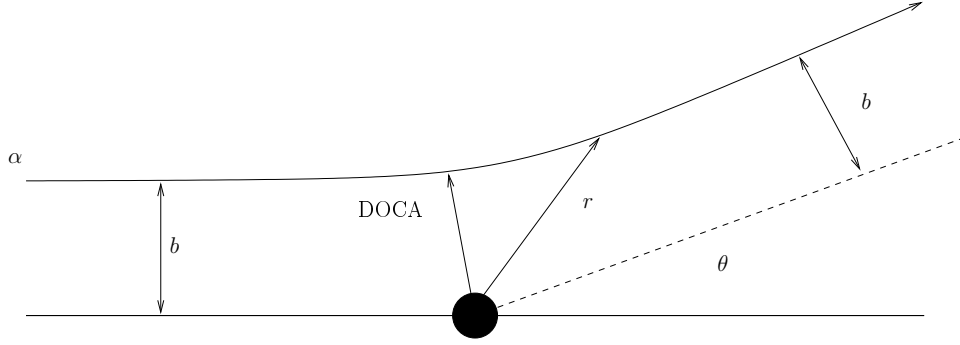


Figure 4.3: The geometry of a Rutherford scattering event. The DOCA is the “distance of closest approach”; b is the “impact parameter”; r is the distance from the scattering center to the α particle at any time.

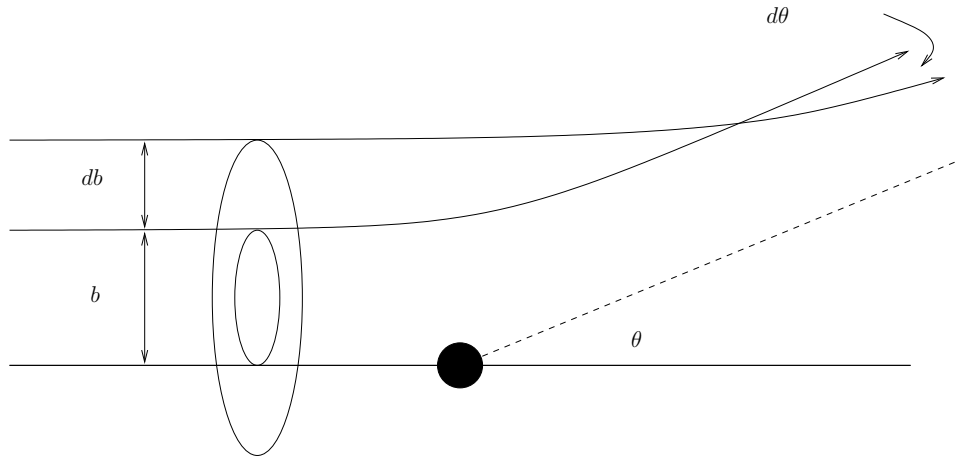


Figure 4.4: An α particle with impact parameter between b and $b + db$ is scattered between angles θ and $\theta + d\theta$.

(Of course, this is only really constant if the nucleus is stationary.) The angular momentum is also conserved and is given by

$$l = mr^2\dot{\theta} = m\sigma b = \text{constant}. \quad (4.4)$$

By carrying out the full calculations, it can be shown that the *impact parameter* b is

$$b = \frac{1}{8\pi\epsilon_0 T} \cot\left(\frac{\theta}{2}\right), \quad (4.5)$$

where T is the kinetic energy.

For a particular scattering event, we can't know the particular b ahead of time. Particles in the range $(b, b + db)$ go into angles of range $(\theta, \theta + d\theta)$ (see Figure 4.4). The number of particles in per second is given by $I2\pi bdb$, where I is the intensity. Similarly, the number of particles out per second is $I(d\sigma/d\Omega)N_T\Omega = I(d\sigma/d\Omega)2\pi \sin(\theta)d\theta$. Because the number in has to equal the number out,

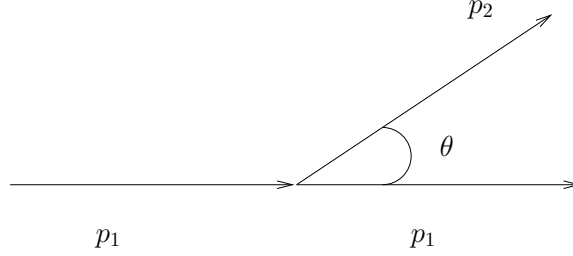


Figure 4.5: The momentum of the α particle before and after the scattering event.

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= \frac{b}{\sin(\theta)} \left| \frac{db}{d\theta} \right| \\ \frac{d\sigma}{d\Omega}(\theta) &= \left(\frac{Zze^2}{16\pi\epsilon_0 T} \right)^2 \frac{1}{\sin^4\left(\frac{\theta}{2}\right)}. \end{aligned} \quad (4.6)$$

This is the *Rutherford (classical) cross section*. Examining this formula, we see that this makes sense logically. It is most likely that an α particle will go straight through (since \sin^4 is maximized for $\pi/2$ radians), and less likely to go at right angles or backwards. Essentially, when θ is large, b is small. In other words, the collision is elastic. The momentum transfer is

$$\vec{q} = \vec{p}_1 - \vec{p}_2 \approx 2p \sin\left(\frac{\theta}{2}\right),$$

where p represents either the magnitude of p_1 or p_2 (see Figure 4.5).

However, this doesn't allow us to probe the *structure* of the nucleus. In order to do this, we need to utilize diffraction: A diffraction pattern³ is related to the "slit" shape by a Fourier transform.

4.5 Mott Scattering

Mott scattering is electron scattering off nuclei. It is basically an "optical" model of Rutherford scattering. Here, the differential cross section is

$$\left(\frac{d\sigma}{d\Omega} \right)_{\text{Mott}} = Z^2 e^2 |f(\theta)|^2, \quad (4.7)$$

where $f(\theta)$ is the Fourier transform part. This would only work if the nucleus were a point particle, but in fact, the nucleus is an extended body, so

$$\frac{d\sigma}{d\Omega} = |F(\theta)|^2 \left(\frac{d\sigma}{d\Omega} \right)_{\text{Mott}}, \quad (4.8)$$

where $F(\theta)$ is called the *form factor*:

$$\begin{aligned} F(\theta) &= \frac{\int \rho(\vec{r}) e^{i\vec{q}\cdot\vec{r}} dV}{\int \rho(\vec{r}) dV} \\ &= \frac{1}{Ze} \int \rho(\vec{r}) e^{i\vec{q}\cdot\vec{r}} dV, \end{aligned} \quad (4.9)$$

³Technically, even with a point particle, a diffraction pattern would occur. However, the first minimum would appear at infinity, meaning that we cannot really say whether diffraction is occurring or not.

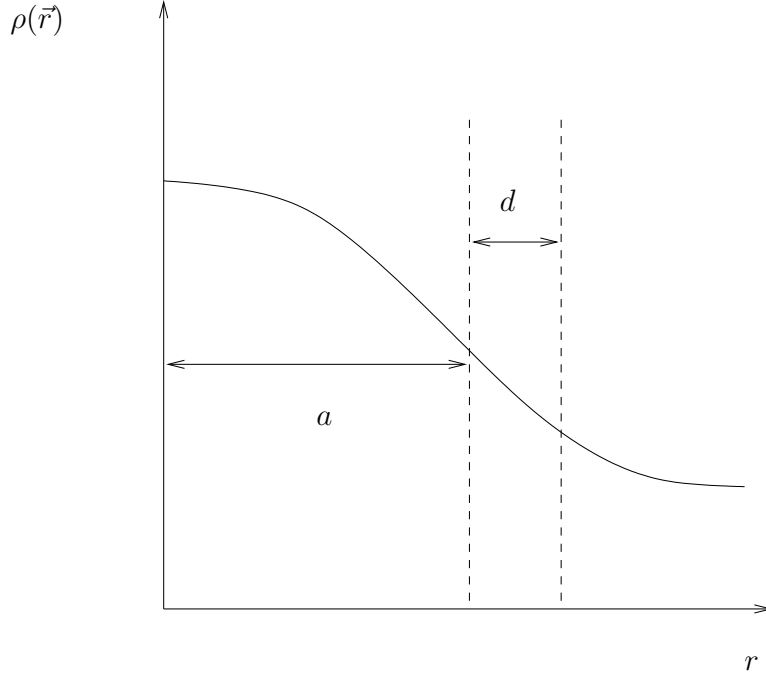


Figure 4.6: The Saxon-Woods nuclear charge distribution; $a \leq 10$ fm.

where \vec{q} is the momentum transfer and ρ is the nuclear charge density distribution. If $\theta \rightarrow 0$, $q \rightarrow 0$, and so $F(0) = 1$. This is just the Fourier transform of the charge distribution. Usually, we rewrite the form factor in terms of $F(q^2)$. While we don't always know $\rho(\vec{r})$, we find it by fitting to the diffraction pattern that we observe. One common ρ is the *Saxon-Woods* distribution:

$$\rho(\vec{r}) = \frac{1}{1 + \exp\left[\frac{r-a}{d}\right]}, \quad (4.10)$$

where a is the nuclear radius and d describes the density fall-off. The number of nucleons is proportional to the volume, which is on the order of a^3 . In turn, a is generally proportional to the cube of the atomic number A ($a \approx 1.2A^{1/3}$).

4.6 Rutherford Scattering Using Time-Dependent Perturbation Theory

For scattering to occur, there must be some interaction. This is represented by the *interaction potential* or *perturbation* in the Hamiltonian:

$$H = H_0 + V. \quad (4.11)$$

Here, H_0 is the “known” Hamiltonian (i.e., the Hamiltonian before perturbing it with V).

Our first step is to expand the wavefunction into a sum of orthonormal basis states, $|u_n\rangle$:

$$\psi(t) = \sum c_n(t) |u_n\rangle e^{-i\omega_n t}. \quad (4.12)$$

The orthonormal states, of course, must satisfy the original Hamiltonian's Schrödinger equation to be of any use to us. That is,

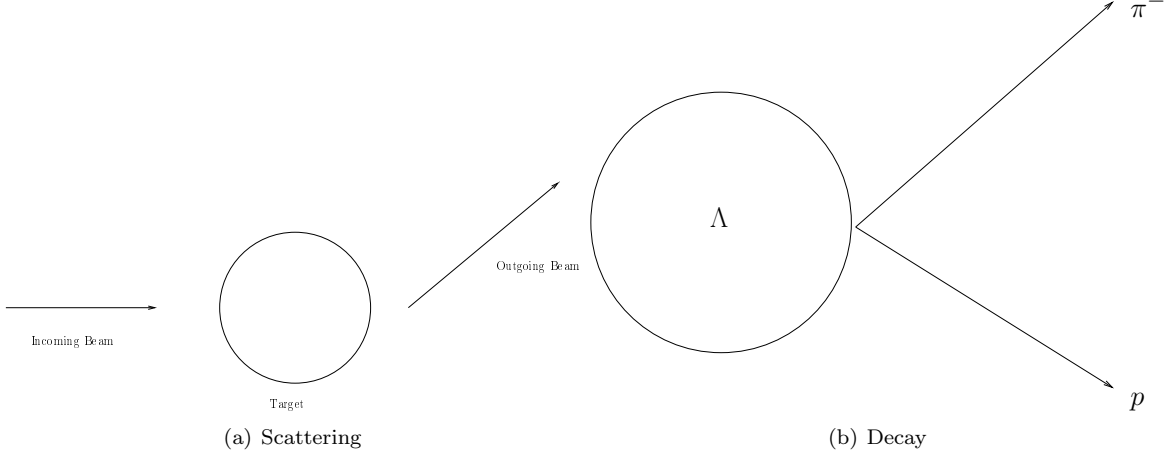


Figure 4.7: Quantum mechanical versions of “Rutherford scattering.” In 4.7(a), this is nothing but an example of a quantum mechanical transition. The trivial case would be sans target. In 4.7(b), this is really no different than a scattering event. The trivial case here would be a Λ particle that does not decay.

$$H_0 |u_n\rangle = E_n |u_n\rangle,$$

and

$$\langle u_j | H_0 | u_k \rangle = E_k \delta_{jk}.$$

We also define these states such that they have no time dependence themselves (i.e., they are stationary states). Next, we write the modified Schrödinger equation as

$$H\psi = (H_0 + V)\psi = i\hbar \frac{\partial \psi}{\partial t}. \quad (4.13)$$

From Equation 4.12, we can write $V\psi$ as

$$\begin{aligned} V\psi &= \sum c_n(t) V |u_n\rangle e^{-i\omega_n t} \\ &= i\hbar \dot{\psi} - H_0 \psi \\ &= i\hbar \sum \dot{c}_n(t) |u_n\rangle e^{-i\omega_n t} + \hbar \sum \omega_n c_n(t) |u_n\rangle e^{-i\omega_n t} - \sum E_n c_n(t) |u_n\rangle e^{-i\omega_n t} \\ V\psi &= i\hbar \sum \dot{c}_n(t) |u_n\rangle e^{-i\omega_n t}. \end{aligned} \quad (4.14)$$

(Note that in the last step, $\hbar\omega_n = E_n$, cancelling out the last two sums). So far we have done nothing but mathematical manipulations. What we really want to know is the probability of a transition to a final state $|u_f\rangle$. To do this, we use the usual “inner product trick” with $\langle u_f |$:

$$\sum c_n(t) \langle u_f | V | u_n \rangle e^{-i\omega_n t} = i\hbar \sum \dot{c}_n \langle u_f | u_n \rangle e^{-i\omega_n t},$$

where of course $\langle u_f | u_n \rangle = \delta_{fn}$. Then

$$\frac{1}{\hbar} \sum c_n(t) \langle u_f | V | u_n \rangle e^{-i\omega_n t} = i\dot{c}_f(t) e^{-i\omega_n t}.$$

Rearranging, we find that

$$i\dot{c}_f(t) = \frac{1}{\hbar} \sum c_n(t) \langle u_f | V | u_n \rangle e^{-i(\omega_n - \omega_f)t}. \quad (4.15)$$

This is the *interaction representation* of the Schrödinger equation. At this point, it is still an exact solution. Note that if $V = 0$, the particle is *always* in the $|u_f\rangle$ state since the right hand side is zero, implying that $c_f(t) \equiv 1$. Unfortunately, we don't know the time dependence of the c_n 's, so we can't really solve this equation exactly. This is where perturbation theory finally comes in.

If V is small, the c_n 's won't change much over time. This in turn means that we can expand the coefficients and drop off higher order terms, in this case using only those of first order. Then if the initial state has $c_i = 1$ with all others 0,

$$\dot{c}_f(t) = -\frac{i}{\hbar} V_{fi} e^{-i(\omega_i - \omega_f)t}, \quad (4.16)$$

where V_{fi} is the matrix element

$$V_{fi} \equiv \langle u_f | V | u_i \rangle. \quad (4.17)$$

With the first order assumptions, we can directly integrate Equation 4.16:

$$c_f(t) = -\frac{i}{\hbar} \int_0^t V_{fi} e^{-i(\omega_i - \omega_f)t'} dt'.$$

The transition probability is given by the modulus squared of c_f , so we find

$$|c_f(t)|^2 = \frac{|V_{fi}|^2}{\hbar^2(\omega_f - \omega_i)^2} 2 \{1 - \cos[(\omega_f - \omega_i)t]\}. \quad (4.18)$$

The transition *rate* is

$$\frac{d}{dt} |c_f(t)|^2 = \frac{2|V_{fi}|^2 \sin[(\omega_f - \omega_i)t]}{\hbar^2(\omega_f - \omega_i)} \equiv w_{if}(t) \quad (4.19)$$

But we must make sure that conservation laws are accounted for. So for the overall transition rate, we must integrate over phase space:

$$w(t) = \frac{2}{\hbar} \int_{\Delta E} |V_{fi}|^2 \frac{\sin[(\omega_f - \omega_i)t]}{\omega_f - \omega_i} \rho_f(E) dE,$$

where $\rho_f(E)$ is the density of states and

$$dE = \hbar d(\omega_f - \omega_i).$$

(E here is not the relativistic energy since we're still discussing a modified Rutherford scattering). Due to the $(\omega_f - \omega_i)^{-1}$ term, there is a sharp peak in the transition rate when $\omega_f = \omega_i$. Because of this peak, without too much loss of generality, we can change the limits of integration to be over all space to utilize

$$\int_{-\infty}^{+\infty} \frac{\sin(ax)}{x} dx = \pi.$$

Then the transition rate is

$$w = \frac{2\pi}{\hbar} |V_{fi}|^2 \rho_f(E_F). \quad (4.20)$$

This is known as *Fermi's golden rule*⁴. This tells us, for example, that an electron decaying into a muon is not very likely because $m_\mu \gg m_e$, meaning there is not really any available phase space.

⁴Fermi's golden rule was actually derived by Dirac, though Fermi referred to it as a "golden rule" in his lectures at the University of Chicago. See, for example, <http://www.physicstoday.org/pt/vol-54/iss-8/p11.html>.

In terms of scattering,

$$dw = \frac{d\sigma}{d\Omega} n_{\text{beam}} v_i.$$

(Unit wise, on the right hand side we have $\text{cm}^2 \times \#/\text{cm}^3 \times \text{cm/s} = 1/\text{s}$). Then to get the Rutherford cross section, we need only the density of states and matrix elements. The density of states is given by (???)

$$\begin{aligned} \rho_f(E_f) &= \frac{dn_f}{dE_f} = \frac{1}{h^3} p_f^2 + dp_f + d\Omega \frac{1}{E_f} \\ &= \frac{p_f^2}{(2\pi)^3} \frac{dp_f}{dE_f} d\Omega. \end{aligned}$$

Choosing plane waves as our basis states, the matrix elements are

$$V_{fi} = \int e^{-i\vec{p}_f \cdot \vec{x}} V(x) e^{-i\vec{p}_i \cdot \vec{x}} d^3x.$$

Thus,

$$\frac{d\sigma}{d\Omega} = \frac{1}{(2\pi)^2} \left| \int V(x) e^{-i\vec{q} \cdot \vec{x}} d^3x \right|^2 \frac{p_f^2}{v_i} \frac{dp_f}{dE_f}. \quad (4.21)$$

For Rutherford scattering, the interaction potential is the Coulomb potential

$$V(\vec{x}) = \frac{k}{x},$$

where $k \equiv (4\pi\epsilon_0)^{-1}$. (Huh? What happened to charge?). Then

$$\begin{aligned} V_{fi} &= \int \frac{k}{x} e^{-i\vec{q} \cdot \vec{x}} d^3x = \int \frac{k}{r} e^{-iqr \cos(\alpha)} r^2 dr d\phi d[\cos(\alpha)] \\ &= 2\pi k \int_0^{+\infty} \frac{dr}{iq} (e^{iqr} - e^{-iqr}) e^{-r/a}, \end{aligned}$$

where α is the angle between the momentum transfer and the point at which we are evaluating and $e^{-r/a}$ is called the *screening factor*. Then

$$V_{fi} = \frac{4\pi k}{q^2 + a^{-2}} \sim \frac{4\pi k}{q^2} \quad (4.22)$$

if $q^2 \gg a^{-2}$, where a is the approximate size of the atom. Assuming still no nuclear recoil, $|p_i| \approx |p_f|$, and

$$q^2 = 4p^2 \sin^2 \left(\frac{\theta}{2} \right).$$

Then the (classical) final state energy is

$$E_f = E^{\text{beam}} + e^{\text{target}} \approx E^{\text{beam}} = \frac{p_f^2}{2m}.$$

This gives us the density of states term:

$$\frac{dp_f}{dE_f} = \frac{m}{p_f}.$$

(m is the mass of the α particle). Then

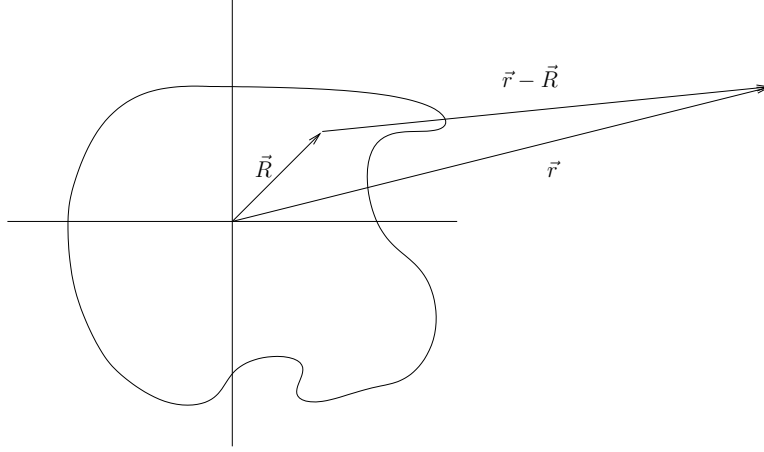


Figure 4.8: An extended body has a different expression for the Coulomb potential. Here, \vec{R} is the source point and \vec{r} is the field point.

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= \frac{1}{(2\pi)^2} \frac{(4\pi^2 k^2)}{16p_i^4 \sin^4(\theta/2)} \frac{p_f^3 m}{v_i p_f} \\ &= \frac{k^2}{16T^2 \sin^4(\theta/a)} \frac{p_f m}{p_i^4 v_i}. \end{aligned}$$

Note that this last part can be written approximately as $m^2/(4p_i^4 \approx 1/(16T^2))$. Finally, we find the collision cross section:

$$\frac{d\sigma}{d\Omega} = \frac{k^2}{16T^2 \sin^4(\theta/2)}.$$

So far, we have stuck with the assumption that the nucleus is a point particle by writing the potential as we did. What we really need to do is to consider an extended body as in Figure 4.8. In this case, we have

$$V(\vec{r}) = \frac{Ze^2}{4\pi\epsilon_0} \int \frac{\rho(\vec{R})d^3R}{|\vec{r} - \vec{R}|}. \quad (4.23)$$

Then computing the matrix elements,

$$\begin{aligned} V_{fi} &= \int V(\vec{r}) e^{i\vec{q}\cdot\vec{r}} d^3r \\ &= \frac{Ze^2}{4\pi\epsilon_0} \int \int \frac{\rho(\vec{R}) e^{i\vec{q}\cdot\vec{r}} d^3r d^3R}{|\vec{r} - \vec{R}|} \\ &= \frac{Ze^2}{4\pi\epsilon_0} \int \rho(\vec{R}) e^{i\vec{q}\cdot\vec{r}} d^3R \int \frac{e^{i\vec{q}\cdot(\vec{r}-\vec{R})}}{|\vec{r} - \vec{R}|} d^3(\vec{r} - \vec{R}) \end{aligned}$$

The first integral is the Fourier transform, i.e., the form factor that was mentioned earlier! Thus

$$\frac{d\sigma}{d\Omega} = \frac{(Z\alpha)^2}{16T^2} \frac{|F(q^2)|^2}{\sin^4(\theta/2)}. \quad (4.24)$$

Chapter 5

The Nucleus

Before we begin, some notation. A nucleus X with N neutrons and Z protons is represented by

A_ZX

where $A = N + Z$ is the total number of nucleons (the atomic *mass number*). For example, ${}^{16}_8\text{O}$ contains 8 neutrons and 8 protons.

5.1 Nuclear Masses

The mass of a nucleus is given by

$$M(Z, A) = ZM_p + (A - Z)M_n - B, \quad (5.1)$$

where B is the *binding energy*. In order for a nucleus to be stable, B must be positive so that the result on the right hand side of Equation 5.1 is less than that on the left hand side (i.e., the sum of the parts is greater than the whole).

However, it is next to impossible to actually be able to calculate B — in fact, no nucleus has an exact solution. However, we do know that the maximum binding energy per nucleon is 8.7 MeV for A around 60. What we need is a way of “estimating” (or rather extracting from data) B .

5.2 Nucleon Interaction

Nucleons are color neutral, so why should they interact at all? The answer is for the same reason (albeit with charge instead of color) that neutral gas particles can interact with each other — the distribution of color is *not* spherically symmetric. Gluons only “see” colored particles, so as strange as it may seem, they do not mediate the strong force inside the nucleus. Instead, it is the light, spinless π^0 (technically virtual π^0 s) that does this in a way that is analagous to how atoms combine, “exchanging” electrons to bind into a molecule¹.

5.3 The Liquid Drop Model

In order to understand and “predict” the binding energy, we can use (for nuclei with A around 20-60) the *liquid drop model*. Here, we think of the nucleus as a liquid drop because of its roughly uniform charge density among other things.

¹It was originally thought that the exchange of π^0 mesons *was* the strong interaction. This led to the development of the Yukawa potential by Hideki Yukawa. This potential turns out to still be very important and will be discussed later.

The assumptions made in the model are as follows:

1. The nucleus is spherical.
2. The nucleons behave like molecules in a drop.
 - Short range attractive forces.
 - Shorter range repulsive forces.
3. Constant nuclear density.

Using these assumptions, we can write down the *semi-empirical mass formula*:

$$B(Z, A) = a_V A - a_S A^{2/3} - \frac{a_C Z^2}{A^{1/3}} - \frac{a_A (A - 2Z)^2}{A} + \delta(Z, A). \quad (5.2)$$

Given this formula, we can find the a coefficients by plotting data and fitting an equation of this form to it. Each term will be explained in turn below.

An interesting fact: There exist no stable $A = 8$ nuclei. As arbitrary as this seems, it turns out that this is very important to the evolution of the universe.

The Volume Term

The term corresponding to the constant a_V is the *volume* term. It is essentially what the binding energy would be if every nucleon experienced the same amount of force based on being surrounded by neighbor nucleons.

The Surface Term

The a_S term corrects for the fact that the nucleons on the surface of the nucleus do not experience the same force as interior nucleons do (there aren't any forces on the outside of the nucleus).

The Coulomb Term

Protons will obviously repulse each other, hence the a_C term. This works as usual, with the potential being $q^2/(4\pi\epsilon_0 r)$, where $r \sim A^{1/3}$ and $q \sim Z^2$.

The Asymmetry Term

For the most part, a proton can be changed to a neutron in the nucleus without much difference since the strong force is much stronger than the electromagnetic. This is known as *isospin symmetry*. However, the Pauli exclusion principle mandates that there must be some difference, which the a_A term factors in. If we take a proton in the Fermi stack of protons and turn it into a neutron, then this new neutron must go on top of the Fermi stack of neutrons. If the levels of the neutron and proton stacks were equal before, an energy of ΔE must be lost to add the new neutron to the stack (see Figure 5.1). Hence, the binding energy will be less.

The Pairing Term

While there is little difference between, say, proton-proton binding and proton-neutron binding, it is found that like particles bind to each other ever so slightly more. This is known as the *pairing effect*. The $\delta(A, Z)$ in Equation 5.2 corrects for this and is defined as

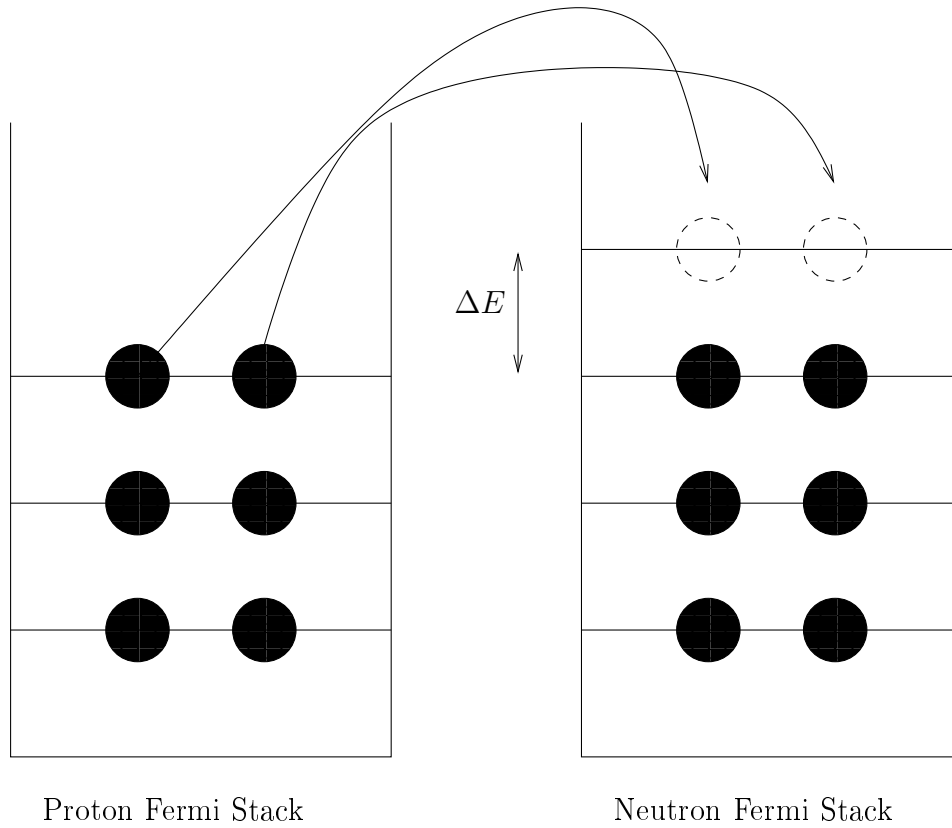


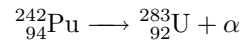
Figure 5.1: If a proton (or two) is removed and becomes a neutron, it must go to the top of the fermi stack of neutrons. If before the transformation, the Fermi stacks were at the same energy level, an energy ΔE must be given up in order to make the change.

$$\delta = \frac{a_P}{A^{1/2}} \times \begin{cases} 0 & \text{if } A \text{ is odd} \\ -1 & \text{if both } N \text{ and } Z \text{ are odd} \\ +1 & \text{if both } N \text{ and } Z \text{ are even} \end{cases} \quad (5.3)$$

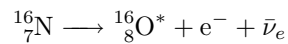
We say the nucleus is odd-odd (oo) if Z and N are both odd and even-even (ee) if both are even. The latter are more stable.

5.4 Nuclear Instability and Decay

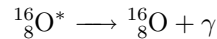
A familiar example of a nuclear decay is



This is alpha decay². Another familiar nuclear decay process is beta decay, e.g.,



Here, the * denotes that the O is in an excited state, thus it will spontaneously decay via



5.4.1 Decay Rates

To get the rate of decay, we use Fermi's golden rule (Equation 4.20). The change in number of a particular nucleus is simply

$$dN(t) = -\omega N(t) dt. \quad (5.4)$$

If we know the initial number $N(0)$ of nuclei, then with ω constant,

$$N(t) = N(0)e^{-\omega t} \quad (5.5)$$

This implies a mean lifetime³ of $\tau = 1/\omega$. The intensity of the decay is

$$\begin{aligned} I(t) &= -\frac{dN(t)}{dt} = N(0)\omega e^{-\omega t} \\ &= I(0)e^{-\omega t}. \end{aligned} \quad (5.6)$$

5.4.2 Units of Measurement

Two units are commonly used in describing nuclear decay. The first is the Curie, defined as

$$1 \text{ Ci} = 3.7 \times 10^{10} \text{ disintegrations/s.}$$

The Curie is based on the nuclear decay of radium. Usually, we deal with quantities that are in the mCi or less.

The other unit frequently used is the Becquerel:

$$1 \text{ Bq} = 1 \text{ disintegration/s.}$$

²The α particle is of course a ${}^4_2\text{He}$ nucleus.

³Note that this is not the "half-life," which is more useful in the sense of human intuition. In terms of mathematics, however, the mean lifetime is a better quantifier.

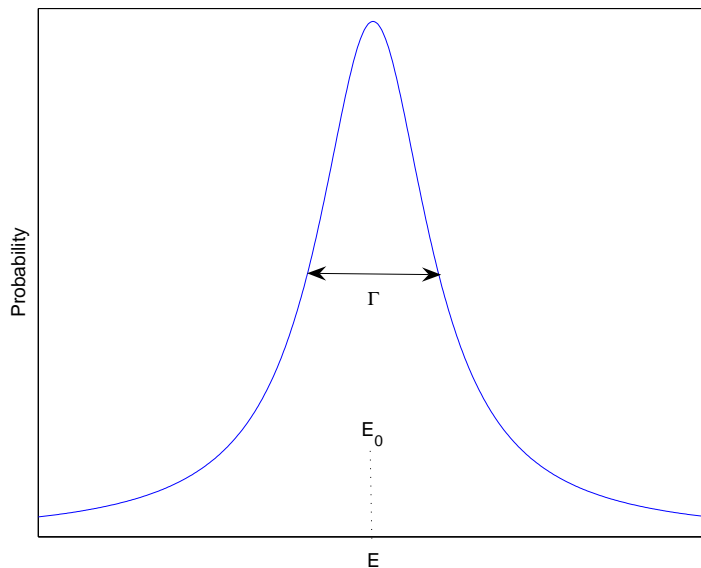
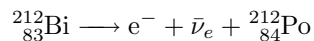


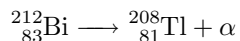
Figure 5.2: Probability of observing energy E in a nuclear decay.

5.4.3 Multiple Decay Branches

Many decays can occur in multiple ways. For example, $^{212}_{83}\text{Bi}$ can take one of the following branches:



or



In such cases, we have

$$\frac{dN(t)}{dt} = -\omega_1 N - \omega_2 N, \quad (5.7)$$

$$N(t) = N(0)e^{-(\omega_1 + \omega_2)t}; \quad (5.8)$$

that is, the decay occurs as if there was one lifetime $\omega^{-1} = (\omega_1 + \omega_2)^{-1}$. The probability of a particular decay path then is $f_1 = \omega_1/\omega$ and $f_2 = \omega_2/\omega$.

By the energy-time uncertainty principle,

$$\Delta E \Delta t \geq \frac{1}{2}, \quad (5.9)$$

so $\Delta E > 1/(2\Delta t)$. The probability of observing an energy E in a nuclear decay is given by

$$P(E) = \frac{1}{(E - E_0)^2 + \Gamma^2/4}, \quad (5.10)$$

where Γ is the full width at half maximum (see Figure 5.2). The result is a Lorentzian (the Fourier transform of $e^{-t/\tau}$).

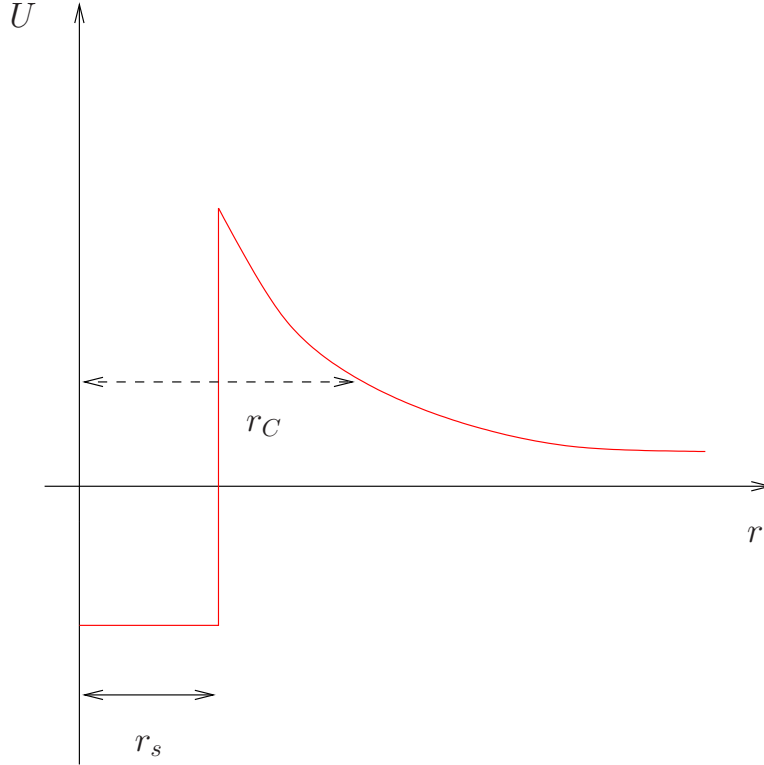
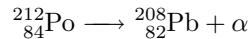


Figure 5.3: The potential experienced by an α particle during alpha decay. The distance r_s is the surface radius, i.e., the size of the nucleus; r_C is the distance an alpha particle must be for it to have as much energy as the Coulomb potential.

Γ is basically ΔE , so we can use it to get the lifetime Δt . The more branches, the higher the rate by Fermi's golden rule, and consequently the shorter the lifetime. In other words, ΔE is broader for shorter lifetimes. This can actually be used to know how many branches there are. For example, using such analysis is how we know that there are three neutrinos!

5.5 Alpha Decay

Alpha decay involves emitting an alpha particle (${}^4\text{He}$). For example, the reaction



Here, $\tau = 43 \times 10^{-7}$ s. However, we cannot use Fermi's golden rule because the α particle inside the nucleus is a strong interaction (i.e., it is not perturbative at all).

Instead, we look at what happens quantum mechanically by using the Schrödinger equation to calculate the decay rate. Doing this obviously requires knowing the potential. Outside the nucleus, we have the normal Coulomb potential. However, inside we have essentially a potential well, which for simplicity we model as a finite square well (see Figure 5.3).

Say the kinetic energy T of the α particle is greater than the potential energy within r_s ($T > U(r_s)$). Then the α particle will leave the nucleus “immediately” (somewhere on the order of 10^{-22} s). On the other hand, if $T < U(r_s)$, the only way out of the nucleus is through tunneling. For this, we need the Schrödinger equation in terms of spherical coordinates:

$$\psi = u(r)Y_l^m(\theta, \phi). \quad (5.11)$$

Starting outside the nucleus, we consider the radial part:

$$-\frac{\hbar^2}{2m} \frac{1}{r} \frac{d^2}{dr^2}(ru) + \left[\frac{2Z_d e^2}{4\pi\epsilon_0 r} + \frac{\hbar^2}{2m} \frac{l(l+1)}{r^2} \right] u = Tu, \quad (5.12)$$

where Z_d is the number of protons in the daughter nucleus (lead in the example given above). It is important to note that in this circumstance, m is *not* the mass of the α particle. Instead, it is the reduced mass:

$$m = \frac{m_\alpha m_d}{m_\alpha + m_d}. \quad (5.13)$$

To make the following calculations easier, we will for the time being assume $l = 0$. Also, we will presume that u is of the form $u = f(r)/r$. Then,

$$\frac{-\hbar^2}{2m} \frac{d^2 f}{dr^2} + \frac{2Z_d e^2}{4\pi\epsilon_0 r} f = Tf.$$

This is remarkable similar to the potential barrier, where the only difference was a constant in the place of the Coulomb potential. From our experience with potential barriers, we had

$$f(r) = \begin{cases} e^{\pm ikr} & T > U \\ e^{\pm \kappa r} & T < U \end{cases}$$

where

$$k^2 = \frac{2m}{\hbar^2}(T - U) \quad \kappa^2 = \frac{2m}{\hbar^2}(U - T)$$

In our case, we don't have a straight barrier, but a Coulomb potential. But this can be represented by a series of small, decreasing barriers, so our solutions take the form $e^{\phi r}$. So, we have

$$\frac{\hbar^2}{2m} \left[\frac{d^2 \phi}{dr^2} + \left(\frac{d\phi}{dr} \right)^2 \right] = \left(\frac{2Z_d e^2}{4\pi\epsilon_0 r} - T \right)$$

, but if the potential is slowly varying, we can ignore the first term on the left,

$$\frac{d\phi}{dr} = \pm \sqrt{\frac{2m}{\hbar^2} \left(\frac{2Z_d e^2}{4\pi\epsilon_0 r} - T \right)}.$$

Thus, this term being ignored, we have the solution

$$\phi(r) = \pm \int \sqrt{\frac{2m}{\hbar^2} \left(\frac{2Z_d e^2}{4\pi\epsilon_0 r} - T \right)} dr. \quad (5.14)$$

When $T > U(r)$, we have

$$f(r) = A \exp \left(i \int_{r_C}^r k(r) dr \right) + B \exp \left(-i \int_{r_C}^r k(r) dr \right),$$

but because this is an incoming wave, $B = 0$. (What is k ??). If $r_s < r < r_C$,

$$f(r) = C \exp \left(\int_r^{r_C} \kappa(r) dr \right) + D \exp \left(\int_r^{r_C} \kappa(r) dr \right).$$



Figure 5.4: The decay on the left has angular momentum, while that on the right does not.

For the same reason as above, $D = 0$, and $\kappa(r)$ is given by

$$\kappa(r) = \sqrt{\frac{2m}{\hbar^2} \left(\frac{2Z_d e^2}{4\pi\epsilon_0 r} - T \right)}. \quad (5.15)$$

Because the wavefunctions need to match at boundaries, the transmission probability is the ratio of $|\psi|^2$ at $r = r_C, r = r_S$:

$$\frac{4\pi r_C^2 |u(r_C)|^2}{4\pi r_S^2 |u(r_S)|^2} = \frac{|f(r_C)|^2}{|f(r_S)|^2} \sim e^{-G},$$

where

$$G = 2 \int_{r_S}^{r_C} \kappa(r) dr.$$

Integrating, we find

$$G = \frac{\pi}{\hbar c} \left(\frac{2Z_d e^2}{4\pi\epsilon_0} \right) \sqrt{\frac{2mc^2}{T}} \mathcal{G}(r_S/r_C),$$

where

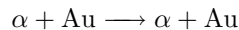
$$\mathcal{G}(r_S/r_C) = \frac{2}{\pi} \left[\arccos \left(\sqrt{\frac{r_S}{r_C}} \right) - \sqrt{\frac{r_S}{r_C} \left(1 - \frac{r_S}{r_C} \right)} \right]$$

Now we consider the total flux of α particles emitted from the nucleus. If $1/\tau_0$ is the number of α particles created at r_s per second, the transmission probability is $\exp(-G)/\tau$ such that $\tau = \tau_0 \exp(-G)$.

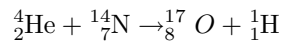
Of course, all of this was for $l = 0$. What if $l > 0$? Then the l part of the wavefunction behaves as a potential barrier because then the nucleus and α particle must separate to conserve angular momentum, which costs energy.

5.6 Nuclear Reactions and Cross Sections

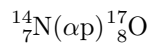
Rutherford observed both *elastic* and *inelastic* scatterings. Rutherford scattering,



is an example of elastic scattering. The inelastic scattering observed⁴ by Rutherford was the reaction



Inelastic scattering differs from elastic in that the products are not the same as the reactants. Another notation for writing the above reaction is

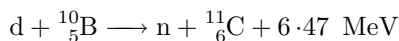


⁴This was the first ever observation of the transmutation of elements — in a sense, alchemy come true!

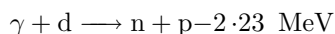
In this notation, the “less important” particles (i.e., the light ones) are put inside the parentheses, with the first particle being on the left hand side of the reaction, and the second being on the right hand side.

In these types of reactions, the typical energy scale is on the order of 1 MeV, so they are above the Coulomb barrier (fusion, in which the reaction occurs via quantum tunneling). The consequence of this is that neutron capture is much easier since.

The *Q-value* is the difference in kinetic energy in a given reaction. For example, the reaction

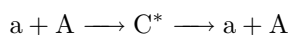


releases energy because it has a *Q*-value greater than zero; the reaction



requires 2.23 MeV in order to occur (in this case, it is needed by the photon). The above reaction is an example of *photodistintegration*.

In some cases, a reaction may result in an excited, intermediate state⁵ being created, which then must decay. The two possibilities here are



or

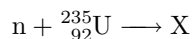


If this is the case, then we want to be able to distinguish between an “actual” intermediate state versus no intermediate state. In other words, how do we know that C^* really ever existed? The typical timescale for Rutherford scattering is on the order of 10^{-21} or 10^{-22} s. If the lifetime of C^* (which we can get from the energy-time uncertainty principle) is longer than the time scale for Rutherford scattering, then we can safely assume that C^* did in fact exist. In fact, we can even use the uncertainty principle to tune the center of mass energy to hit a resonance in C^* production (i.e., the cross section can be made incredibly large)!

There can of course also be multiple channels of decay, just as in other types of decay mentioned previously.

5.7 Fission, Chain Reactions, and Moderation

Nuclear fission is the general process by which a nucleus splits into two daughter nuclei. One such fission reaction of particular interest is that of the fission of uranium-235:



(where X represents the products which, aside from including free neutrons, are unimportant to the reaction). In nature, ${}^{238}\text{U}$ is more common, but it is not so easy to get it to fission. This is because neutrons naturally are too fast to allow for a significant cross section.

In order to get a sustained chain reaction, there must be a *moderator* which slows down the neutrons via elastic scatterings. With the neutrons then slowed, the cross section becomes large enough for capture by the ${}^{235}\text{U}$ nuclei, thus beginning the reaction.

5.8 Nuclear Fusion

Unlike in the case of fission, there exists no spontaneous fusion in nature. This is because now the Coulomb barrier must be overcome, which is why tunneling is vital to fusion in the sun. Fusion in the sun consists of combining a deuteron with tritium:

⁵This intermediate state is sometimes referred to as a *compound nucleus*.

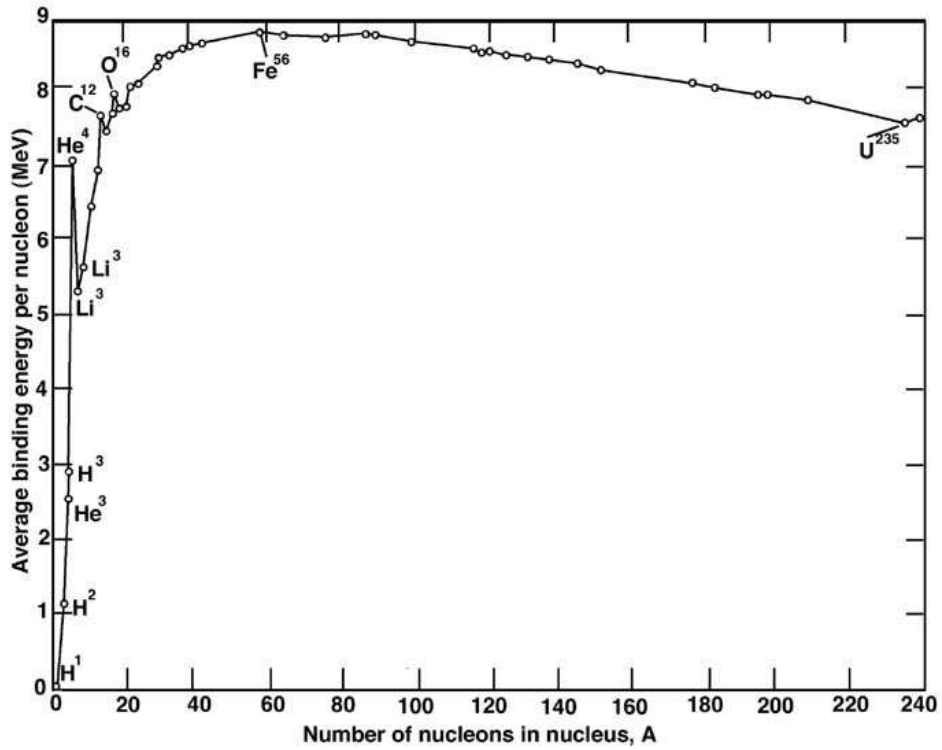
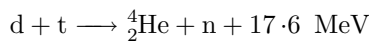
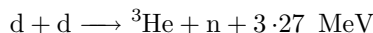


Figure 5.5: Average binding energy per nucleon. Fusion occurs for lighter elements because the curve shifts upwards for low A . The curve peaks around iron, which is why no fusion can take place beyond this point. On the far right of the curve lies plutonium. Because the curve shifts downwards, fission will release energy. Image courtesy Wikipedia.



While the probability to tunnel through the Coulomb barrier is fairly small, the sheer amount of hydrogen in the sun makes the overall probability quite significant.

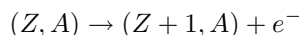
Scientists trying to achieve sustainable fusion sometimes try the reaction



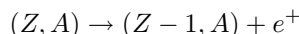
Note that the amount of energy released in this reaction is far lower than that in the sun! One reason why fusion has alluded scientists for so long is that incredibly high temperatures are needed in order to get the fusible material at high enough energies for tunneling to be probabilistically significant⁶.

5.9 Beta Decay

Beta decay is the process⁷



or

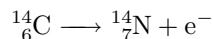


In the center of momentum frame, the available energy is

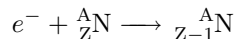
$$E_{\text{avail}} = {}^A_Z M - {}^A_{Z-1} M - m_e.$$

If the only extranuclear decay product were the electron (or positron), then every time beta decay occurred, there would only be *one* possible energy for the electron—that is, its spectrum would be a delta function. However, when the experiments were performed c. 1920, it was found that the spectrum was continuous, with the maximum value of energy at the predicted value (Figure 5.6).

This was obviously troubling, and Bohr, et al. even suggested that perhaps energy is only conserved on average in order to make sense of the data. Needless to say, many physicists of the day felt very uncomfortable with this idea. Another problem involved the decay



${}^{14}\text{C}$ is a spin-0 particle, ${}^{14}\text{N}$ has spin-1, and e^- is a fermion of spin-1/2 — clearly there is no way to combine these spins such that angular momentum is conserved. Physicists of the day were even more aghast that angular momentum might not be conserved than they were about the possibility of energy not being conserved. Another “violation” of energy conservation was found in electron capture:



5.9.1 The Neutrino

Wolfgang Pauli’s solution was to introduce a new particle, later named the *neutrino* by Fermi. Being spin-1/2, the neutrino allows us to hold onto both above conservation laws. With the neutrino, the real beta decay reactions are

⁶The only way the hydrogen bomb works is to trigger the fusion explosion with a fission reaction to get it hot enough to start fusing.

⁷Yes, I’m leaving out the neutrinos for now. I’ll get to that in a moment.

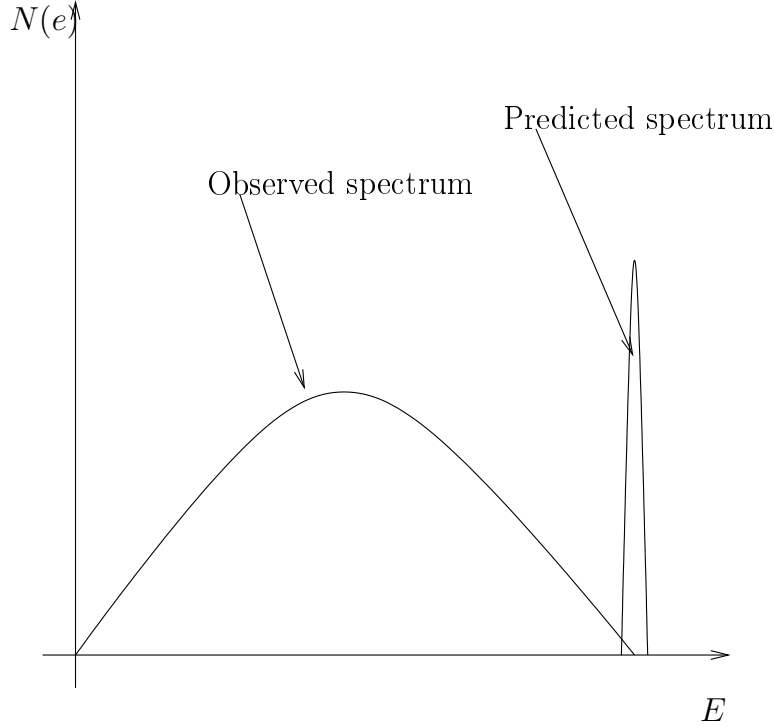
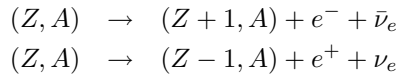


Figure 5.6: The beta decay spectrum.



To explain why the neutrino had not been detected in beta decay, Pauli proposed that it must be neutral and largely noninteracting, in addition to either massless or of very little mass in order not to make E_{avail} too low.

5.9.2 Fermi's Calculation

Fermi set out to calculate the rate for beta decay under the assumption that the neutrino really exists. Recall his golden rule:

$$\omega = \frac{2\pi}{\hbar} G_F^2 |M_{if}|^2 \frac{dN_f}{dE_f},$$

The M matrix elements are somewhat vague, however. Fermi reasoned that for the total angular momentum of the leptons being zero, $|M|^2 \sim 1$. This is the *Fermi transition*. If the total angular momentum of the leptons is 1, then $|M|^2 \sim 3$. This is the *Gamow-Teller transition*⁸. We begin by defining the momenta and energies for each particle involved in the decay as shown in Figure 5.8.

We begin by noticing that the order of magnitude for the final state energy E_f is about 1 MeV, as is $p_e c$. The nucleus (proton) on the other hand, is heavy. Making some classical approximations, $T \sim p^2 c^2 / (2m_p c^2) = p^2 / (2m_p) \sim 10^{-3}$ MeV; thus we can safely ignore the kinetic energy of the proton. If at this

⁸The easiest way to explain why the factors of 1 and 3 is by saying that it is related to the spin degeneracy.

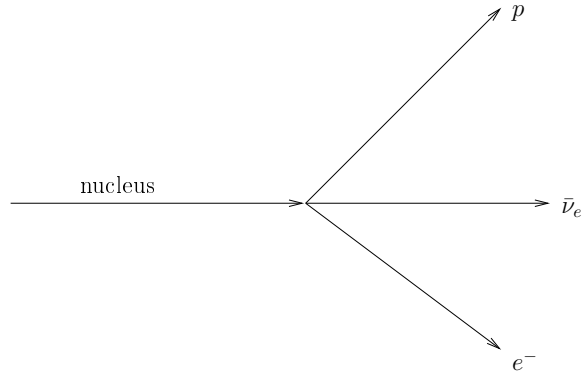


Figure 5.7: The “four-Fermi interaction,” also known as the contact interaction. There is really no difference between a neutron splitting up into a proton, antineutrino, and an electron or an electron being absorbed by a proton to form a neutron and a neutrino.

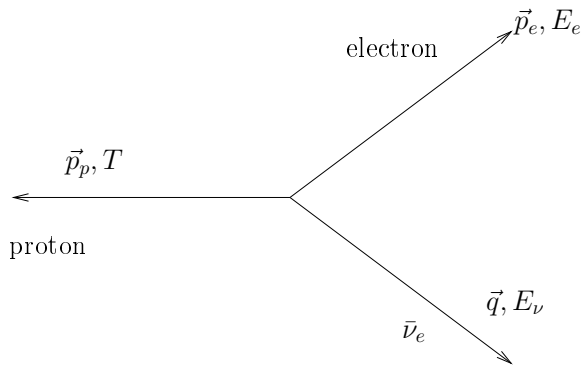


Figure 5.8: The notation we use for Fermi’s calculation of the rate of beta decay.

point, we make the further assumption⁹ that $m_\nu = 0$, then the energy of the neutrino is given by simply $E_\nu = qc$. Conservation of energy then states

$$qc = E_f - E_e; \quad (5.16)$$

conservation of momentum tells us

$$\vec{p}_p + \vec{q} + \vec{p}_e = 0 \quad (5.17)$$

since we are in the parent nucleus's rest frame.

Using a density of states calculation, the number of final states available is given by

$$\frac{4\pi p_e^2 dp_e}{(2\pi)^2 \hbar^3},$$

while that available to the antineutrino is given by

$$\frac{4\pi q^2 dq}{(2\pi)^3 \hbar^3}.$$

Then the total number of final states is given by

$$dN = \frac{(4\pi)^2}{(2\pi)^6 \hbar^6} p_e^2 q^2 dp_e dq.$$

For a given p_e and E_e , $q = (E_f - E_e)/c$ with $dq = dE_f/c$. So we have

$$\frac{dN}{dE_f} = \frac{1}{(4\pi^4) \hbar^6 c^3} p_e^2 (E_f - E_e)^2 dp_e.$$

From here, one could calculate the transition rate. Instead, we look at the spectrum of electron energies, which is¹⁰

$$N(p_e) dp_e \propto p_e^2 (E_f - E_e)^2 dp_e. \quad (5.18)$$

If we did the same calculation with the exception that the neutrino is *not* massless, we would get

$$N(p_e) dp_e \propto p_e^2 (E_f - E_e)^2 \sqrt{1 - \left(\frac{m_\nu c^2}{E_f - E_e}\right)^2} dp_e. \quad (5.19)$$

Comparing these two formulae can in theory allow us to measure the mass of the neutrino (see Figure 5.9).

The *total* rate of beta decay¹¹ we get from first integrating N :

$$N = \int_0^{E_f} E_e^2 (E_f - E_e)^2 dE_e = \frac{E_f^5}{30},$$

so the rate is

$$\omega = \frac{G_F^2 |M_{if}|^2}{60\pi^3 (\hbar c)^6 \hbar} E_f^5. \quad (5.20)$$

With this prediction, the existence of neutrinos could now begin to be experimentally verified. Fermi asserted that “inverse beta decay” (or electron capture),

⁹Indeed, the neutrino is not massless, as we now believe, though it is incredibly small, so the following results are still approximately valid.

¹⁰This is incredibly sloppy notation, but everyone does it. Sorry.

¹¹As it turns out, Fermi's golden rule is really a *differential* rate, even though it's not written that way.

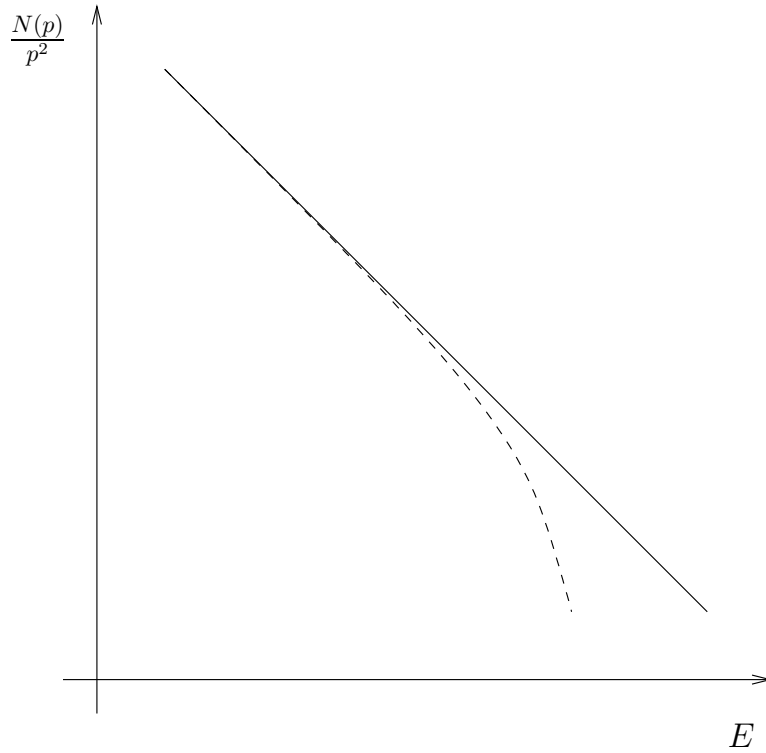
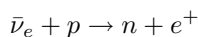


Figure 5.9: The Kurie plot. The solid line represents a massless neutrino; the dashed line is for a massive neutrino. The distance between these two curves is directly proportional to the neutrino mass. While in principle, this method of detecting the neutrino mass would work, it is enormously difficult experimentally due to the upper limit on the neutrino mass being around only a few eV (the lower limit is around 0.05 eV!!!).



is the same thing on fundamental grounds as beta decay. Given the rate in Equation 5.20, the cross section can be computed. We shall examine the more generic process $a + b \rightarrow c + d$. The cross section per particle is

$$\frac{d\sigma}{d\Omega} = \frac{\omega}{v_i} = \frac{2\pi}{\hbar} \frac{|M_{if}|^2}{v_i} \left(\frac{\rho_f}{(2\pi\hbar)^3} \frac{dp_f}{dE_f} \right),$$

where v_f and v_i are the relative final and initial velocities in the center of momentum frame. It is easy to show that $dp_f/dE_f = v_f^{-1}$. Then the cross section becomes (ignoring spin)

$$\frac{d\sigma}{d\Omega} = \frac{1}{4\pi^2\hbar^4} |M_{if}|^2 \frac{p_f^2}{v_i v_f}.$$

If the final state particles have spins s_c and s_d , then the number of states available is increased by a factor

$$g_f = (2s_c + 1)(2s_d + 1);$$

and for the initial state,

$$g_i = (2s_a + 1)(2s_b + 1).$$

For the spins, we average over the initial and final states, thereby multiplying by g_f/g_i , so we get the cross section

$$\sigma \sim \frac{d\sigma}{d\Omega} \Delta\Omega = \frac{G^2}{\pi} |M_{if}|^2 \frac{p^2}{v_i v_f}. \quad (5.21)$$

Here we see that G^2 has units E^{-4} . Since $p^2 \sim E^2$, and E^{-2} has units of area (which we can see from the Heisenberg uncertainty principle), we see that the units are correct for the above expression.

Finally, we can take Fermi's rate for beta decay and calculate the cross section needed to detect Pauli's "undetectable" neutrino — we can detect it because if a neutrino is a decay product, it *must* interact with matter (see inverse beta decay)! We find that $\sigma \sim 10^{-43} \text{ E}^2 \text{ cm}^2$. This means that for a 1 MeV antineutrino, the mean free path through lead is about 50 light years! Thus detecting the neutrino is no small task since the probability for interaction is so low and thus requires a huge flux of neutrinos.

Experimentally, it was far easier initially to detect antineutrinos, because with the process of electron capture, the resulting neutron and positron can be detected separately, giving a unique signature for the antineutrino. This was first accomplished by Reines and Cowan in 1953 outside of a nuclear reactor.

5.9.3 Parity Nonconservation

In the 1950s, there were thought to be two particles, called the τ and θ , that seemed exactly the same in every way except that the τ decayed into three pions and the θ into two. That is, while the τ and θ were of the same mass and most quantum numbers, they did not share the same parity. Today, we know that these particles are one and the same (now known as the K^+ meson)

In 1956–1957, Wu and colleagues observed the pure Gamow-Teller transition



Figure 5.10 outlines the experimental setup. If parity were conserved, then the intensity of electrons at each detector should be the same. Instead, they found that the intensity was given by

$$I(\theta) = 1 + \alpha \frac{\vec{\sigma} \cdot \vec{p}}{E}, \quad (5.22)$$

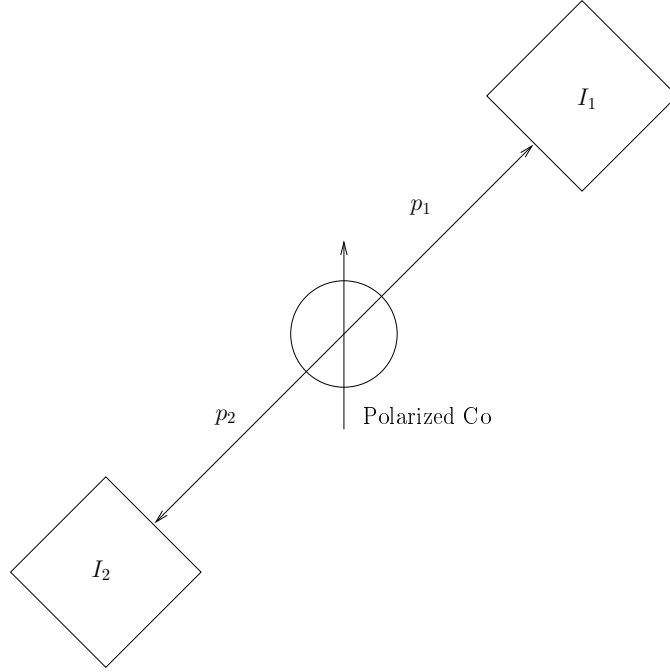


Figure 5.10: A schematic of Wu's experiment that showed the nonconservation of parity in weak interactions. Performing a three-dimensional reflection on everything shows that if parity were conserved, we could switch I_1 , I_2 , p_1 , and p_2 and still get the same results.

where $\vec{\sigma}$ is the electron (positron) spinor and \vec{p} the electron (positron) momentum, and $\alpha = \pm 1$ (plus for the positron, minus for the electron). This intensity can in no way conserve parity since 1 is a scalar, whereas the *helicity* (the fraction in Equation 5.22) is a pseudo scalar. That is, it *cannot* be an eigenstate of parity, and hence parity is clearly violated in weak interactions! Equation 5.22 can be rewritten as

$$I(\theta) = 1 + \alpha \frac{v}{c} \cos(\theta). \quad (5.23)$$

Then if I^+ represents spin aligned with motion and I^- spin antialigned with motion, the longitudinal polarization is

$$P = \frac{I^+ - I^-}{I^+ + I^-}. \quad (5.24)$$

We can see that the faster the electron goes, the more the spin is aligned with the direction of motion.

Furthermore, it can be shown that for a massless neutrino, the neutrino is *always* left-handed and the antineutrino is *always* right-handed.

Chapter 6

Angular Momentum and Raising and Lowering Operators

Angular momentum is incredibly important in quantum mechanics. In some sense, different spins can be different particles, such as in the case of the neutrino (left-handed) and antineutrino (right-handed). However, angular momentum, and namely spin, is important because the methods and mathematics we use can be generalized to other concepts, as well.

6.1 Commutators

Recall that

$$[J_x, J_y] = iJ_z, \quad (6.1)$$

and similarly for all cyclic permutations thereof. This means that an eigenstate of J_z cannot simultaneously be an eigenstate of J_x , for example. However, the angular momentum squared operator, $J^2 = J_x^2 + J_y^2 + J_z^2$ is a scalar, and therefore *does* commute with each of the angular momentum components:

$$[J^2, J_n] = 0. \quad (6.2)$$

Then if we choose to measure J_z in addition to J^2 ,

$$J^2 |j m\rangle = j(j+1) |j m\rangle \quad (6.3)$$

$$J_z |j m\rangle = m |j m\rangle, \quad (6.4)$$

with $m = -j, j+1, \dots, j-1, j$ and j either an integer or half-integer.

6.2 Raising and Lowering Operators

It is often useful to define the raising and lower operators,

$$J_{\pm} = J_x \pm iJ_y. \quad (6.5)$$

These operators can be used in combinations to form J_x and J_y ; e.g., $J_x = (J_+ + J_-)/2$. Evaluating the commutator of J_{\pm} with J_z ,

$$\begin{aligned}
[J_z, J_\pm] &= [J_z, J_x \pm iJ_y] \\
&= i(J_y \mp iJ_x) \\
&= \pm(J_x + iJ_y) \\
[J_z, J_\pm] &= \pm J_\pm
\end{aligned} \tag{6.6}$$

By using this commutator on the state $|j m\rangle$, we see that J_\pm moves the state up or down in J_z by one. That is, $J_\pm |j m\rangle$ returns the state with the eigenvalue of J_z of $m \pm 1$. Now, we would hope that J_+ operating on a state with maximum m would in effect “destroy” the state, and similarly for J_- — we can’t have a state higher (or lower) than the maximum (or minimum) allowed! In order to verify this, we need to find what the constant is when operating with J_\pm . That is, given

$$J_\pm = C_\pm^j |j m \pm 1\rangle,$$

we want to find the C values.

Say we repeatedly apply J_- to the state of maximum $m = m_{\max}$ n times:

$$J_-^n |j m_{\max}\rangle = C |j m_{\max} - n\rangle.$$

If the next application of J_- would then lead to zero on the right hand side, $m_{\max} - n = m_{\min} = -m_{\max}$. That is, the number of times we can apply J_- starting from the m_{\max} state is

$$n = 2m_{\max}. \tag{6.7}$$

Since n is an integer, we also verify that m is either integral or half-integral.

Now to get the coefficients that we wanted, we begin with the orthonormality of the states, namely $\langle j m | j m \rangle = 1$. Then,

$$\langle j m | J_\pm^\dagger J_\pm | j m \rangle = \langle j m \pm 1 | C_\pm^* C_\pm | j m \pm 1 \rangle = C_\pm^* C_\pm,$$

where the dagger indicates the Hermitian conjugate and the asterisk the complex conjugate. But we also have that

$$J_\pm^\dagger J_\pm = J_\mp J_\pm = J_x^2 + J_y^2 \mp i[J_x, J_y] = J^2 - J_z^2 \pm J_z.$$

If we then combine this with what is above, and temporarily denote the eigenvalue of J^2 with ξ , then we find that

$$C_\pm = \sqrt{\xi - m^2 \pm m}. \tag{6.8}$$

Since $J_+ |j m\rangle = 0$, we see that $j \equiv m_{\max}$, and so we now see that

$$J_\pm |j m\rangle = \sqrt{j(j+1) - m(m \pm 1)} |j m \pm 1\rangle. \tag{6.9}$$

6.3 Addition of Angular Momenta

Say we have more than one particle. If, say, we have two, with associated momenta \vec{J}_1 and \vec{J}_2 , then the total angular momentum for the system is

$$\vec{J} = \vec{J}_1 + \vec{J}_2. \tag{6.10}$$

Now imagine a state $|j m\rangle$ which is a superposition of the states $|j_1 m_1\rangle$ and $|j_2 m_2\rangle$. One possible configuration is

$$|j\ m\rangle = |1\ 1\rangle = \left| \frac{1}{2}\ \frac{1}{2} \right\rangle_1 \left| \frac{1}{2}\ \frac{1}{2} \right\rangle_2.$$

To get the state $|1\ 0\rangle$, one needs only apply the lowering operator:

$$J_- |1\ 1\rangle = (J_{-1} + J_{-2}) \left| \frac{1}{2}\ \frac{1}{2} \right\rangle_1 \left| \frac{1}{2}\ \frac{1}{2} \right\rangle_2,$$

but J_{-1} only operates on state 1, while J_{-2} only operates on state 2. Then we find

$$\sqrt{2}|1\ 0\rangle = \left| \frac{1}{2}\ \frac{-1}{2} \right\rangle_1 \left| \frac{1}{2}\ \frac{1}{2} \right\rangle_2 + \left| \frac{1}{2}\ \frac{1}{2} \right\rangle_1 \left| \frac{1}{2}\ \frac{-1}{2} \right\rangle_2.$$

Applying J_- once more, we get

$$|1\ -1\rangle = \left| \frac{1}{2}\ \frac{-1}{2} \right\rangle_1 \left| \frac{1}{2}\ \frac{-1}{2} \right\rangle_2.$$

These three states corresponding to $m = \{1, 0, -1\}$ are referred to as the *triplet* states, while the $|0\ 0\rangle$ state is the *singlet* state¹.

¹The singlet state cannot be obtained via raising and lowering operators. Instead, we can construct the singlet state by finding a state which is orthogonal to the others.

Chapter 7

Charge Conjugation and Time Reversal Symmetries

At this juncture, we come to some other symmetries and conservation laws that were alluded to in Chapter 3.

7.1 Charge Conjugation

The charge conjugation operator C reverses the charge of a particle; e.g., $C|e^- \rangle \rightarrow |e^+ \rangle$. Immediately, we notice that C effectively is changing matter to antimatter. Furthermore, we see that a charged particle can never be an eigenstate of charge conjugation — only a neutral particle that is its own antiparticle can be a charge conjugation eigenstate¹. One such particle is the neutral pion, π^0 . Operating with C ,

$$C|\pi^0 \rangle = \eta|\pi^0 \rangle.$$

As with parity, we must get the same result if we operate on the result with C , namely

$$C^2|\pi^0 \rangle = \eta^2|\pi^0 \rangle = |\pi^0 \rangle$$

implies that $\eta = \pm 1$.

Note that $C|p \rangle \rightarrow |\bar{p} \rangle$ changes the baryon number from 1 to -1. This again demonstrates that the effect of C is to change matter to antimatter.

Like P , C is multiplicative and is conserved in electromagnetic and strong interactions, but not necessarily in weak interactions. To get the C eigenvalue of the neutral pion, we look at

$$\pi^0 \rightarrow 2\gamma$$

A rigorous derivation would require the use of field theory, so we improvise here. Recall from electrodynamics that the *vector potential* is given by

$$\vec{A}(\vec{x}) = \int \frac{\vec{j}(\vec{x}')d^3x'}{|\vec{x} - \vec{x}'|}. \tag{7.1}$$

This becomes an operator that creates photons:

$$\tilde{A}|0 \rangle = |\gamma \rangle \tilde{A}^\dagger |\gamma \rangle = |0 \rangle$$

¹For example, there exists an antineutron, since a neutron is constituted of three quarks. That is, the neutron is not an eigenstate of charge conjugation, despite being neutral. This is another hint that C changes particles to antiparticles, since it has an effect on neutral particles as well.

Then using C on \vec{A} results in a sign change due to the current density \vec{j} :

$$C\vec{A}|0\rangle = -\vec{A}|0\rangle = -|\gamma\rangle$$

Thus, we conclude that $C|\gamma\rangle = -|\gamma\rangle$, i.e., the charge conjugation number of the photon is $C_\gamma = -1$. For n photons, we have $C = (-1)^n$. Therefore, $C_{\pi^0} = 1$.

7.1.1 Positronium

As an example, we consider the case of positronium, the bound state of an electron and positron. Because neither particle is an eigenstate of charge conjugation, we can say that they are in fact not two distinct particles, but rather different states of the *same* particle. We can change between these states by using C , much like the raising and lowering operators for angular momentum (see Chapter 6). Treating these particles as different states of the same particle makes them indistinguishable particles. Because the electron and positron are fermions, the overall wavefunction must be odd, but it now has three parts instead of only a spin and space part:

$$\psi_{e^+e^-} = \phi(\vec{r}_1 - \vec{r}_2)\alpha(\text{spin})\chi(\text{charge}). \quad (7.2)$$

Odd symmetry implies that under interchange ($\vec{r}_1 - \vec{r}_2 \rightarrow \vec{r}_2 - \vec{r}_1$), we must have $\psi_{e^+e^-} \rightarrow \psi_{e^-e^+}$. Clearly, the spatial part of the wavefunction has $(-1)^l$ symmetry.

For spin, positronium can either have $s = 0$ or $s = 1$. Spin-1 implies even symmetric under interchange, while spin-0 implies odd symmetric. Then, the symmetry of the spin part of the wavefunction is $(-1)^{s+1}$. We then have for positronium

$$(-1)^l(-1)^{s+1}S_{\text{charge}} = -1.$$

It can be shown that $S_{\text{charge}} = (-1)^{l+s}$, which makes the interchange equivalent to

$$C\psi_{e^+e^-} = (-1)^{l+s}\psi_{e^+e^-}. \quad (7.3)$$

Thus, when positronium decays, the number of resultant photons is dependent on $l + s$.

7.2 CP Conservation and Violation

As an example, take beta decay. Let state $|a\rangle$ be that illustrated in Figure 7.1(a). Under parity, $P|a\rangle$ results in a left-handed antineutrino (Figure 7.1(b)), which does not occur in nature. As for charge conjugation, $C|a\rangle$, we get a right-handed neutrino (Figure 7.1(c)). But if we operate with *both* charge conjugation and parity, namely $CP|a\rangle$ (Figure 7.1(d)), we *should* find a conserved quantity in even the weak interaction. As it turns out, even CP violation can occur, and was first discovered in 1964 by James Cronin and Val Fitch.

7.3 Time Reversal Invariance

In classical mechanics, every process is inherently invariant under time reversal. Yet, in thermodynamics, we see many examples of processes that are inherently irreversible. The quintessential example is the process of scrambling eggs. While we never experience eggs unscrambling and jumping back into the shell, there is nothing on a *fundamental* level that prevents this — it is just so highly unlikely that it will not happen.

The following is a list of what happens to some familiar constructs under time reversal T :

1. $x \rightarrow x$.
2. $\vec{p} = m\frac{d\vec{x}}{dt} \rightarrow -\vec{p}$.
3. $\vec{L} = \vec{r} \times \vec{p} \rightarrow -\vec{L}$.

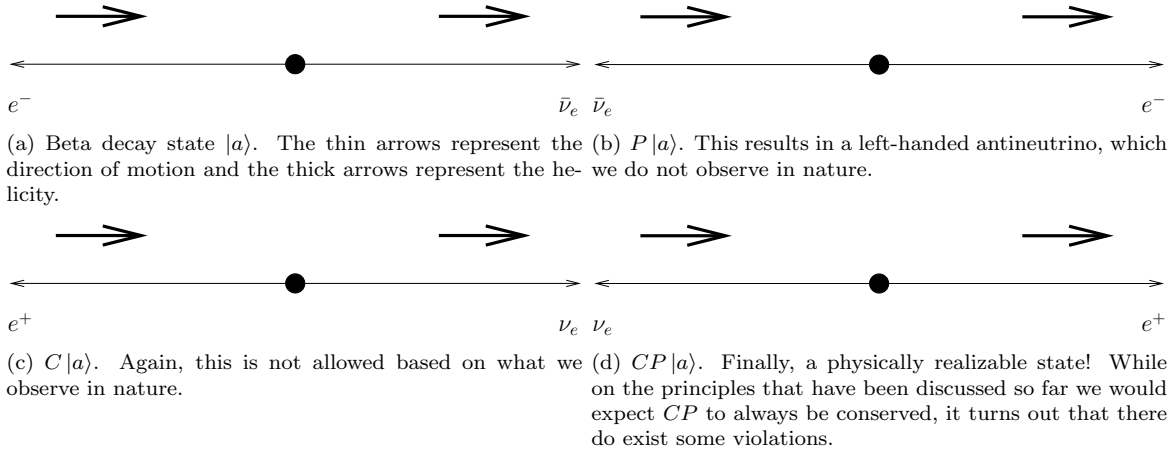
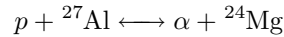


Figure 7.1: Beta decay under parity.

4. Spin: $\vec{\sigma} \rightarrow -\vec{\sigma}$.
5. $\vec{E} = -\nabla\phi \rightarrow \vec{E}$.
6. $\vec{B} \rightarrow -\vec{B}$.
7. Electric dipole: $\vec{\sigma} \cdot \vec{E} \rightarrow -\vec{\sigma} \cdot \vec{E}$.
8. Magnetic dipole: $\vec{\sigma} \cdot \vec{B} \rightarrow \vec{\sigma} \cdot \vec{B}$.

T is conserved under the electromagnetic and strong interactions. For example, the reaction



happens in both directions at approximately the same rate (we compare the rates by calculating the cross section for each). Like P and CP , T is not always conserved in weak interactions.

Unlike CP symmetry, which is occasionally broken, it is believed that CPT symmetry is *always* conserved. In fact, the CPT theorem² states that if CPT symmetry is violated, then Lorentz invariance ultimately fails. Although nearly all mainstream physicists believe in CPT symmetry, there are some ways to test for violations. Perhaps the most common is to check the masses of a particle and its antiparticle. If they differ even slightly, then CPT symmetry would be violated, and modern physics would be setup for a complete upheaval.

²The CPT theorem was first put forth by John Bell, the same physicist behind the famous Bell's inequality.

Chapter 8

Rotations, SU(2), and Isospin

8.1 Rotations

Recall that we can write a rotation as

$$R_z |j\ m\rangle = e^{-i\theta J_z} |j\ m\rangle = e^{-i\theta m} |j\ m\rangle.$$

For a spin-1/2 system, we can write

$$\begin{aligned} \left| \frac{1}{2}, \frac{1}{2} \right\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ \left| \frac{1}{2}, -\frac{1}{2} \right\rangle &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned}$$

The i^{th} component of angular momentum, J_i is then given by $J_i = (1/2)\sigma_i$, where σ_i denotes the i^{th} Pauli spin matrix. These are defined as

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{8.1}$$

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{8.2}$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{8.3}$$

The general rotational operator R is unitary and associative, meaning that it belongs to the mathematical group SU(2), which is the group of all 2×2 unitary matrices. The “S” is for “special,” and refers to the fact that the determinant of the Pauli spin matrices (in this context) is 1.

8.1.1 An Example

Here we make an example of applying a rotation along the y axis.

$$\begin{aligned} R_y(\theta) |j, m\rangle &= e^{-i\theta J_y} |j, m\rangle \\ &= \left(1 - i\theta J_y + \frac{1}{2}(-i\theta)^2 + \dots \right) \times |j, m\rangle. \end{aligned}$$

If we carry out the Taylor expansion in full detail, we find that we get sines and cosines, leaving

$$\begin{aligned} R_y(\theta) |j, m\rangle &= \left[\cos\left(\frac{\theta}{2}\right) - iJ_y \sin\left(\frac{\theta}{2}\right) \right] \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta/2) & \sin(\theta/2) \\ -\sin(\theta/2) & \cos(\theta/2) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \end{aligned}$$

This is, of course, the familiar rotation matrix that we all know and love. Here, a represents the amplitude of being in the $|\uparrow\rangle$ state, while $b = (1 - a^2)^{1/2}$ is the amplitude for being in the $|\downarrow\rangle$ state. In a general rotational coordinate system, the amplitude of being in the $|\rightarrow\rangle$ state is

$$a \cos(\theta/2) + b \sin(\theta/2).$$

Some notation that is sometimes used:

$$\left\langle \frac{1}{2}, \frac{1}{2} \left| R_y(\theta) \right| \frac{1}{2}, \frac{1}{2} \right\rangle = d_{\frac{1}{2}, \frac{1}{2}}^{\frac{1}{2}}(\theta).$$

8.2 Isospin

Recall from the liquid drop model (Section 5.3) that the strong interaction appears (largely) not to distinguish between protons and neutrons. Heisenberg proposed in 1932 that the proton and neutron could then be treated as different states of one “nucleon.” Note that these states are *not* basis vectors, but just state vectors:

$$\begin{aligned} p &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ n &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

This treatment of one nucleon is known as *isotopic spin*, or more frequently, *isospin*. Isospin belongs to the SU(2) group, and the strong interaction (roughly) conserves isospin. Then

$$\begin{aligned} I_3 &= \frac{1}{2} \text{ (proton)} \\ I_3 &= -\frac{1}{2} \text{ (neutron)} \end{aligned}$$

The strong interaction conserves the *total* I , not just I_3 (just as is the case for angular momentum). A general nucleon may be represented as a superposition of proton and neutron states,

$$|\psi\rangle = a|p\rangle + b|n\rangle.$$

We can even rotate in isospin space: $I_i = (1/2)\tau_i$ (where for historical reasons, $\tau_i = \sigma_i$).

Today, we know why it is that isospin is “conserved.” Recall that the proton is $|uud\rangle$ and the neutron is $|udd\rangle$. This means that changing a proton into a neutron (or vice versa) is equivalent to changing an up quark to a down quark (or vice versa). In order for that to be the case, then $m_p \approx m_n$, except the Coulomb interaction would affect the proton mass. We would expect that if $m_d = m_u$, the proton would be heavier due to this, but as it turns out, the mass of the down quark is slightly larger than that of the up quark by about 1.3 MeV. Thus, $m_n > m_p$, but only by a small amount. So, while isospin is not “really” conserved, it can still be useful.

Because isospin belongs to $SU(2)$, we can convert a proton to a neutron (and vice versa) with raising and lowering operators, just like we can change angular momenta states with raising and lowering operators. Mathematically, the scattering of, say, two protons is equivalent to the scattering of, say, two spin up electrons. Protons (and neutrons) can only have *one* component of isospin.

Since isospin is due to the mass difference between the up and down quarks being about 1%, apparently other u, d combinations have isospin symmetry, as well. For example, the pions:

$$\begin{aligned}\pi^+ &= u\bar{d} & I_3 &= +1 \\ \pi^- &= d\bar{u} & I_3 &= -1 \\ \pi^0 &= 1/\sqrt{2}(d\bar{d} - u\bar{u}) & I_3 &= 0\end{aligned}$$

The first two pions are each of mass approximately 140 MeV, while the neutral pion is about 135 MeV. Analogously to triplet spin states, we see from the above that the pion is an $I = 1$ isospin multiplet.

8.2.1 Example: Deuteron

When treating the nucleons as identical particles, the total nuclear wavefunction now becomes

$$\psi_{\text{total}} = \phi(\text{space})\alpha(\text{spin})\chi(\text{isospin}). \quad (8.4)$$

The deuteron has $J = 1, l = 0$. Since we are treating the neutron and proton as identical¹ fermions, ψ_{total} must be antisymmetric. Then in terms of isospin, we may have either one of the triplet states, namely

$$\begin{aligned}|I = 1, I_3 = 1\rangle &= |p\rangle_1 |p\rangle_2 \\ |I = 1, I_3 = 0\rangle &= \frac{1}{\sqrt{2}}(|p\rangle_1 |n\rangle_2 + |n\rangle_1 |p\rangle_2) \\ |I = 1, I_3 = -1\rangle &= |n\rangle_1 |n\rangle_2,\end{aligned}$$

or the singlet state,

$$|I = 0, I_3 = 0\rangle = \frac{1}{\sqrt{2}}(|p\rangle_1 |n\rangle_2 - |n\rangle_1 |p\rangle_2).$$

The space part of ψ_{total} , ϕ , goes like $(-1)^l = 1$, so it is symmetric. Similarly, we know that the spin part, α , is symmetric, so χ *must* be antisymmetric. That is, it must be in the singlet state. We say that the deuteron is an *isosinglet*. In $SU(2)$ notation, this is written schematically as

$$2 \otimes 2 = 1 \oplus 3.$$

It can be shown that isospin and charge are related by

$$\frac{Q}{e} = I_3 + \frac{B}{2}, \quad (8.5)$$

where B is the baryon number. This applies to all baryon and meson combinations of up and down quarks and antiquarks.

¹Note that what we really are doing is treating the particles as identical as far as the strong interaction is concerned. Clearly, via the Coulomb interaction, we can distinguish between the proton and the neutron. However, since the strong force dominates the electromagnetic force in the nucleus, this is a valid treatment given the u and d mass discussion earlier.

8.2.2 π^\pm -Nucleon Scattering

Assuming only isospin matters in the strong interaction (i.e., ignoring the Coulomb interaction), we now examine (charged) pion scattering from nucleons². Some possibilities of this include

$$\begin{aligned}\pi^+ p &\rightarrow \pi^+ p \\ \pi^- n &\rightarrow \pi^- n \\ \pi^+ n &\rightarrow \pi^+ n \\ \pi^- p &\rightarrow \pi^- p \\ \pi^- p &\rightarrow \pi^0 n \\ \pi^+ n &\rightarrow \pi^0 p\end{aligned}$$

Each of these possibilities entails an $I = 3/2$, $I_3 = 3/2$ system on the left hand side. We can get the right hand side I_3 values via Equation 8.5. Because isospin is conserved in strong interactions, then the right hand side must also have $I = 3/2$. Schematically, we say H_q is the amplitude to go from $I = q$ to $I = q$ ³.

To get the cross section for a particular scattering event, we need to calculate

$$\sigma \sim |\langle \psi_f | H_{3/2} + H_{1/2} | \psi_i \rangle|^2,$$

except all that we care about is the isospin part, χ , of the wavefunction. So, e.g., say we are interested in $|\pi^+ p\rangle = |3/2, 3/2\rangle$. Then it follows that

$$\begin{aligned}\sigma_{\pi^+ p \rightarrow \pi^+ p} &\sim |\langle \pi^+ p | H_{3/2} | \pi^+ p \rangle|^2 \\ &= \left| \left\langle \frac{3}{2}, \frac{3}{2} \left| H_{3/2} \right| \frac{3}{2}, \frac{3}{2} \right\rangle \right|^2 \\ &= |H_{3/2}|^2\end{aligned}$$

In the case of the $\pi^- n \rightarrow \pi^- n$ interaction, we have similarly

$$\begin{aligned}\sigma_{\pi^- n \rightarrow \pi^- n} &\sim |\langle \pi^- n | H_{3/2} | \pi^- n \rangle|^2 \\ &= \left| \left\langle \frac{3}{2}, -\frac{3}{2} \left| H_{3/2} \right| \frac{3}{2}, -\frac{3}{2} \right\rangle \right|^2 \\ &= |H_{3/2}|^2.\end{aligned}$$

That is, the above two cross sections are the same when only isospin matters.

But what if we are interested in one of the other scatterings? Say, for example, we are interested in comparing $\pi^- p \rightarrow \pi^- p$ and $\pi^- p \rightarrow \pi^0 n$. In the first case, $|\psi_i\rangle = |\psi_f\rangle = |3/2, -1/2\rangle$. If we wish to expand this, we get

$$\begin{aligned}\left| \frac{3}{2}, -\frac{1}{2} \right\rangle &= a |1, -1\rangle \left| \frac{1}{2}, \frac{1}{2} \right\rangle + b |1, 0\rangle \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \\ &= a |\pi^- p\rangle + b |\pi^0 n\rangle.\end{aligned}$$

We also have

²There is really no reason not to include here the π^0 as an incident particle; we are just choosing to focus on the π^\pm mesons specifically as a particular example of the usefulness of isospin.

³these ‘‘amplitudes’’ are basically the same thing as matrix elements. Note that the matrix from which these H values come is diagonalized.

$$\begin{aligned} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle &= c |1, -1\rangle \left| \frac{1}{2}, \frac{1}{2} \right\rangle + d |1, 0\rangle \left| \frac{1}{2}, \frac{1}{2} \right\rangle \\ &= c |\pi^- p\rangle + d |\pi^0 n\rangle. \end{aligned}$$

(???) Now, it must be that $c = b$ and $d = -a$. From a table of Clebsch-Gordan coefficients, we find that $a = \sqrt{1/3}$ and $b = \sqrt{2/3}$. Now we can solve for $|1, 1\rangle |1/2, -1/2\rangle$ as a superposition of $I = 3/2$ and $I = 1/2$ states:

$$\begin{aligned} |\pi^- p\rangle &= |1, -1\rangle \left| \frac{1}{2}, \frac{1}{2} \right\rangle \\ &= \sqrt{\frac{1}{3}} \left| \frac{3}{2}, -\frac{1}{2} \right\rangle - \sqrt{\frac{2}{3}} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle. \end{aligned}$$

Now we have enough information to calculate the cross sections for both reactions. For $\pi^- p \rightarrow \pi^0 n$, we have

$$\begin{aligned} \sigma_{\pi^- p \rightarrow \pi^0 n} &\sim \left| \langle \pi^- p | H_{3/2} + H_{1/2} | \pi^0 n \rangle \right|^2 \\ &= \left| \left(\sqrt{13} \left\langle \frac{3}{2}, -\frac{1}{2} \right| - \sqrt{\frac{2}{3}} \left\langle \frac{1}{2}, -\frac{1}{2} \right| \right) | H_{3/2} + H_{1/2} | \left(\sqrt{\frac{2}{3}} \left| \frac{3}{2}, -\frac{1}{2} \right\rangle + \sqrt{\frac{1}{3}} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \right) \right|^2 \\ &= \left| \sqrt{\frac{2}{3}} (H_{3/2} - H_{1/2}) \right|^2. \end{aligned}$$

Similarly,

$$\sigma_{\pi^- p \rightarrow \pi^- p} \sim \left| \frac{1}{3} H_{3/2} + \frac{2}{3} H_{1/2} \right|^2.$$

Now, all we would need to know to calculate the actual (relative) cross sections are the H values. As it turns out, we can find these from the first two cross sections we found since they only depend on $H_{3/2}$.

Chapter 9

Relativistic Quantum Mechanics

9.1 Lagrangians and the Universe

The ultimate goal of physics is to be able to write a Lagrangian, L , which describes everything¹. In practice, it is often easier to deal with the *Lagrangian density*, \mathcal{L} , rather than the Lagrangian itself. Take, for example, waves on a stretched string (see Figure 9.1). In this case, L and \mathcal{L} are related by

$$L = T - U = \int_0^l \mathcal{L} dx. \quad (9.1)$$

\mathcal{L} is a function of the continuous variable $\phi(\vec{x}, t)$:

$$\mathcal{L} = \frac{1}{2}\rho \left(\frac{\partial\phi}{\partial t}\right)^2 - \frac{1}{2}F \left(\frac{\partial\phi}{\partial x}\right)^2, \quad (9.2)$$

where ρ is the mass per unit length and F is the tension on the string. We can see that $\mathcal{L} = \mathcal{L}(\dot{\phi}, \phi')$, where the dot represents a time derivative, and the prime a space derivative. To get the equations of motion, we must solve Lagrange's equation,

$$\frac{d}{dt} \left(\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\right) - \frac{\partial}{\partial x} \left(\frac{\partial\mathcal{L}}{\partial\phi'}\right) = 0. \quad (9.3)$$

Solving for this particular example, we end up with the wave equation:

$$\rho \frac{\partial^2\phi}{\partial x^2} - F \frac{\partial^2\phi}{\partial x^2} = 0. \quad (9.4)$$

While we do not have the ultimate Lagrangian, we know some of the properties it should have, enumerated as follows:

1. It must reproduce known equations of motion (and field equations, such as Maxwell's equations).
2. It must be invariant under known symmetries (e.g., Lorentz invariance, rotation, translation, etc.).
3. It may contain other hypothesized symmetries and even "accidental" symmetries (such as lepton number conservation).
4. It must include all degrees of freedom (e.g., spin).
5. It must agree with experiment and make *testable* predictions.

¹The only such Lagrangian that we have exactly today is that for the vacuum.

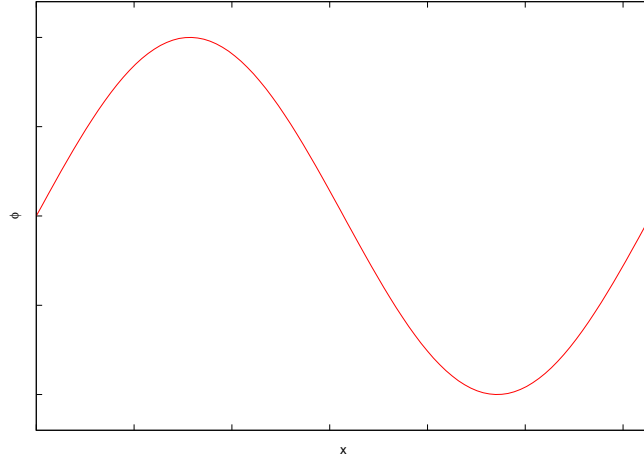


Figure 9.1: A stretched string gives rise to the wave equation. One can obtain this result either by Newtonian mechanics or Lagrangian mechanics.

Usually, it would also be helpful to start with a particular equation of motion so that we know where we are going. A particularly good starting point is the Schrödinger equation. But the Schrödinger equation is missing some pieces of physics that we know about today, namely relativity and spin. . . .

9.2 Relativistic Quantum Mechanics

There are a number of possible approaches to adding relativity to the Schrödinger equation. The simplest is to start by making everything a four-vector. Recall from the heuristic argument to get the Schrödinger equation, we begin by setting $E - p^2/(2m) = 0$ and take as axioms that $E = i\hbar\partial/\partial t$ and $p = i\hbar\nabla$. Putting these into our assumption, we wind up with

$$i\frac{\partial\psi}{\partial t} + \frac{1}{2m}\nabla^2\psi = 0.$$

The “predictions” of the Schrödinger equation are the probability density $\rho = |\psi|^2$ such that $\int |\psi|^2 d^3x = 1$.

It is also often useful to find the change, or *flux*, of this probability density; that is, we want the *probability current*:

$$-\frac{\partial}{\partial t} \int_V \rho dV = \int_S \vec{j} \cdot \hat{n} dS = \int_V \nabla \cdot \vec{j} dV. \quad (9.5)$$

From this, we get the *continuity equation*,

$$\frac{\partial\rho}{\partial t} + \nabla \cdot \vec{j} = 0. \quad (9.6)$$

In the case of the Schrödinger equation, we have that the probability current is

$$\vec{j} = -\frac{i}{2m} (\psi^*\nabla\psi - \psi\nabla\psi^*). \quad (9.7)$$

The continuity equation associated with the Schrödinger wavefunction is then

$$\frac{\partial\psi^*\psi}{\partial t} + \nabla \cdot \left[\frac{1}{2mi} (\psi^*\nabla\psi - \psi\nabla\psi^*) \right] = 0. \quad (9.8)$$

For example, consider the wavefunction $\psi = N \exp[i(\vec{p} \cdot \vec{x} - Et)]$. Then $\rho = |N|^2$ and $\vec{j} = (\vec{p}/m)|N|^2$. (At this point, we also note the future relevance of the four-vector (ρ, \vec{j}) .)

Now we can add relativity to quantum mechanics. From Einstein, $E^2 = |\vec{p}|^2 + m^2$. Using the energy and angular momentum operators, we get the *Klein-Gordon equation*,

$$-\frac{\partial^2 \phi}{\partial t^2} + \nabla^2 \phi = m^2 \phi. \quad (9.9)$$

In order for the Klein-Gordon equation to be valid quantum mechanically, it must have a *conserved current*, i.e., have a \vec{j} that is conserved under the continuity equation. In order to do this, we multiply by $i\phi^*$ and the complex conjugate Klein-Gordon equation by $-i\phi$ and add to get

$$\frac{\partial}{\partial t} \left[i \left(\phi^* \frac{\partial \phi}{\partial t} - \phi \frac{\partial \phi^*}{\partial t} \right) \right] + \nabla \cdot [-i(\phi^* \nabla \phi - \phi \nabla \phi^*)] = 0,$$

where we note that the first term is ρ and the second term is \vec{j} in Equation 9.6. That is, the four-current for the Klein-Gordon equation is defined as

$$j^\mu = (\rho, \vec{j}) = i(\phi^* \partial^\mu \phi - \phi \partial^\mu \phi^*), \quad (9.10)$$

where $\partial^\mu = \partial/\partial x^\mu$ and $x^0 = t, x^1 = x, \dots$

While this does add in relativity to the Schrödinger equation, it does have some shortcomings. First, the energy eigenvalues are $E = \pm \sqrt{|\vec{p}|^2 + m^2}$, which means there can be *negative* energies. These lead to *negative* probabilities, which does not make sense mathematically. Secondly, the Klein-Gordon equation does not account for spin². In short, while it is an improvement, the Klein-Gordon equation is by no means a complete description.

To explain the negative energy possibilities, Dirac proposed that perhaps these negative energies weren't unphysical at all. Rather, they exist in the universe, but they are already filled up, and so they can be ignored. Today, we instead go along with Feynman's interpretation that a negative energy particle is actually an *antiparticle* moving forwards in time. Recall the Klein-Gordon conserved current is given by Equation 9.10. Then for an electron,

$$j^\mu(e^-) = -2e|N|^2(E, \vec{p}),$$

while for the positron (e^+), the sign changes on the p . Thus, the emission of a positron is entirely equivalent to the absorption of an electron. The Lagrangian that gives rise to the Klein-Gordon equation is given by

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi. \quad (9.11)$$

The first term is the *kinetic energy term* and the second is the *mass term*.

9.3 The Dirac Equation

9.3.1 Motivation

Due to the problems of the Klein-Gordon equation, Dirac set out to include spin into a relativistically correct "Schrödinger equation." A particle with spin³ has two states:

$$\psi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$$

This state ψ is called a *spinor* and is a four component column vector. The ϕ_1 component represents a spin up particle/antiparticle while ϕ_2 represents a spin down particle/antiparticle.

²Consequently, the Klein-Gordon equation can be used to explain spinless particles, such as the π^0

³Spin-1/2 is really all that is considered here.

We want our new, relativistically (and spin) correct equation to be first order like the Schrödinger equation. So, we start with

$$H\psi = (\vec{\alpha} \cdot \vec{p} + \beta m) \psi. \quad (9.12)$$

Of course, we wish to preserve the special relativistic relation $E^2 = p^2 + m^2$, which means that

$$\begin{aligned} H^2\psi &= (p^2 + m^2) \psi \\ &= (\alpha_i p_i + \beta m) (\alpha_j p_j + \beta m) \psi \\ &= [\alpha_i^2 p_i^2 + (\alpha_i \alpha_j + \alpha_j \alpha_i) p_i p_j + (\alpha_i \beta + \beta \alpha_i) p_i m + \beta^2 m^2] \psi. \end{aligned}$$

In order for the equality to hold, it must be true then that the middle two terms are zero, which in turn tells us that $\alpha_1, \alpha_2, \alpha_3$, and β all anticommute and that $\alpha_1^2 = \alpha_2^2 = \alpha_3^2 = \beta^2 = 1$. Because of the anticommutation relations, we are also lead inexorably to realize that the α s and β *must be matrices*. Additionally, they must be hermitian, traceless, have even dimensionality, and have eigenvalues of ± 1 . These matrices are not unique, but the minimal dimensionality that satisfies the above requirements is 4×4 . Noting that the Pauli matrices are defined by

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (9.13)$$

$$\sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad (9.14)$$

$$\sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (9.15)$$

One convenient representation is the *chiral representation*, which uses the matrices

$$\alpha^i = \begin{pmatrix} -\sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}, \quad (9.16)$$

$$\beta = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}. \quad (9.17)$$

Other representations are often used, though this choice will turn out to be useful in that it shows effects related to the neutrino helicity discussed in Chapter 5.

9.3.2 Dirac Gamma Matrices and the Covariant Form

Next⁴, we define the *Dirac gamma matrices* to be

$$\gamma^\mu = (\beta, \beta\alpha^i). \quad (9.18)$$

(I could provide some examples here of what, say, γ^1 looks like, but it seems to be a waste of time — it is just basic matrix multiplication, after all.) By multiplying Equation 9.12 by β from the left, we find the *covariant form*⁵ of the Dirac equation,

$$(i\gamma^\mu \partial_\mu - m) \psi = 0. \quad (9.19)$$

If we instead write this out more formally and in less compact notation, we have

⁴The amount of notation here becomes voluminous. Bear with me.

⁵Note that $\partial^\mu \equiv (\partial/\partial t, -\nabla)$. The minus sign in front of the gradient operator is motivated by our definition of the four-momentum, $p^\mu \equiv (E, \vec{p}) = (i\partial/\partial t, -i\nabla) \equiv i\partial^\mu$. That is, the negative sign goes with the “index up” derivative operator because of the negative sign in the momentum operator.

$$[i(\gamma^0\partial_0 - \gamma^1\partial_1 - \gamma^2\partial_2 - \gamma^3\partial_3) - m]\psi = 0.$$

When we put in the wavefunction⁶ $\psi = (\psi_1, \psi_2, \psi_3, \psi_4)$, we get an incredibly complicated matrix equation that gives us a set of four coupled partial differential equations. It can be shown, however, that these equations end up forming a diagonalized matrix. If we let $\psi_L \equiv (\psi_1, \psi_2)$ and $\psi_R \equiv (\psi_3, \psi_4)$ (the choice of names will become apparent later), then we can rewrite the Dirac equation yet again as

$$i \begin{pmatrix} 0 & \frac{\partial}{\partial t} + \vec{\sigma} \cdot \nabla \\ \frac{\partial}{\partial t} - \vec{\sigma} \cdot \nabla & 0 \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = m \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}. \quad (9.20)$$

This equation thus gives the two coupled differential equations

$$\begin{aligned} i\sigma^0\partial_0\psi_L - i\sigma^i\partial_i\psi_L &= m\psi_R \\ i\sigma^0\partial_0\psi_R + i\sigma^i\partial_i\psi_R &= m\psi_L. \end{aligned}$$

As for the solutions, we already know them! Note that in all that we have done, there has been no mention of any potential — we are dealing only with a free particle. Thus, we should get plane wave solutions:

$$\begin{aligned} \psi_L &= u_L e^{i(\vec{p}\cdot\vec{r} - Et)} \\ \psi_R &= u_R e^{i(\vec{p}\cdot\vec{r} - Et)}, \end{aligned}$$

where the u 's are unknown (at this point) functions. In the center of momentum frame, $E = \pm m$, and so we have for our differential equations

$$\begin{aligned} i\frac{\partial\psi_L}{\partial t} &= m\psi_R \\ i\frac{\partial\psi_R}{\partial t} &= m\psi_L. \end{aligned}$$

In the case where $E = m$, we then have $\psi_L = u \exp(-imt)$ and $\psi_R = u \exp(-imt)$ (???) where $u = (u_1, u_2) = u_1(1, 0) + u_2(0, 1)$. The possible solutions, sans normalization, are then

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

For $p \neq 0$, we define

$$\chi^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \chi^{(2)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Then,

$$u^{(s)} = N \begin{pmatrix} \chi^{(s)} \\ \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi^{(s)} \end{pmatrix},$$

where N is a normalization factor.

⁶This *is* a column vector, I just am sick of taking up so much space. A comma, in my notation, represents a “new line.”

9.3.3 Conserved Current

Define the *adjoint spinor*

$$\bar{\psi} = \psi^\dagger \gamma^0. \quad (9.21)$$

Then the adjoint of Equation 9.19 is

$$i\partial_\mu \bar{\psi} \gamma^\mu + m\bar{\psi} = 0. \quad (9.22)$$

Then one can show that there exists a continuity equation for the Dirac equation, just as there is for the Schrödinger and Klein-Gordon equations:

$$\bar{\psi} \gamma^\mu \partial_\mu \psi + \partial_\mu \bar{\psi} \gamma^\mu \psi + \partial_\mu (\bar{\psi} \gamma^\mu \psi) = 0. \quad (9.23)$$

(??? Check on plus sign). The *conserved current* is then

$$j^\mu = \bar{\psi} \gamma^\mu \psi, \quad (9.24)$$

with

$$\rho = j^0 = \bar{\psi} \gamma^0 \psi = \psi^\dagger \psi = \sum_{j=1}^4 |\psi_j|^2. \quad (9.25)$$

In order to get the “real” current, $\pm e$ needs to be multiplied by Equation 9.24, where the + is for the positron, and the – for the electron.

9.3.4 Other Conserved Currents and γ^5

In addition to Equation 9.24, there are a number of other *bilinears* in $\bar{\psi}, \psi$, and γ^μ that could potentially be currents. First, we define⁷

$$\gamma^5 \equiv i\gamma^0 \gamma^1 \gamma^2 \gamma^3. \quad (9.26)$$

Some of the bilinears are then

$$\begin{aligned} \bar{\psi} \psi &\rightarrow \text{scalar} \\ i\bar{\psi} \gamma^5 \psi &\rightarrow \text{pseudoscalar} \\ \bar{\psi} \gamma^\mu \psi &\rightarrow \text{vector} \\ \bar{\psi} \gamma^5 \gamma^\mu \psi &\rightarrow \text{axialvector} \end{aligned}$$

If we manipulate equations involving ψ and γ^5 for long enough, we discover an interesting property. Namely,

$$\frac{1}{2} (1 - \gamma^5) \psi = \begin{pmatrix} \sigma^0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \psi^L \\ \psi^R \end{pmatrix} = \begin{pmatrix} \psi^L \\ 0 \end{pmatrix}.$$

Similarly,

$$\frac{1}{2} (1 + \gamma^5) \psi = \begin{pmatrix} 0 & 0 \\ 0 & \sigma^0 \end{pmatrix} \begin{pmatrix} \psi^L \\ \psi^R \end{pmatrix} = \begin{pmatrix} 0 \\ \psi^R \end{pmatrix},$$

where σ_0 is the identity matrix. That is, γ^5 projects out one part of the spinor. Thus any current with γ^5 only works on either ψ_L or ψ_R . If this seems awfully familiar, it should be — this is reminiscent of beta

⁷Where'd the 4 go? Originally, the (0, 1, 2, 3) indices were written as (1, 2, 3, 4). Evidently, when that was switched, no one bothered to make the 5 a 4.

decay, where we only observe left-handed neutrinos and right-handed antineutrinos! At the very least, we can say that γ^5 is certainly related (at least mathematically) with the weak interaction.

We now consider a specific example. The covariant form of the Dirac equation, when multiplied out, reads

$$i\gamma^\mu \partial_\mu \psi = m\psi.$$

But using $\psi = (\psi_L, \psi_R)$, we can write these as two coupled differential equations:

$$\begin{aligned} i\left(\frac{\partial}{\partial t} + \vec{\sigma} \cdot \nabla\right) \psi_L &= m\psi_R \\ i\left(\frac{\partial}{\partial t} - \vec{\sigma} \cdot \nabla\right) \psi_R &= m\psi_L \end{aligned}$$

When $m = 0$, then we can rewrite these yet again to be

$$\begin{aligned} i(\partial_0 + \sigma^i \partial_i) \psi_L &= 0 \\ i(\partial_0 - \sigma^i \partial_i) \psi_R &= 0 \end{aligned}$$

Immediately, we can guess the solutions to these equations. Presuming that the motion is only in the z direction, then

$$\begin{aligned} \psi_L &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{i(pz - Et)}, \\ \psi_R &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{i(pz - Et)}. \end{aligned}$$

The first solution is the left-handed helicity case, while the second is the right-handed case. That is,

$$\begin{aligned} \vec{\sigma} \cdot \hat{p} \psi_L &= -\psi_L, \\ \vec{\sigma} \cdot \hat{p} \psi_R &= \psi_R. \end{aligned}$$

So massless particles have no choice but to be in a positive or negative helicity state; massive particles are in a superposition of helicity states. Consequently, this is why the neutrino was long believed to be massless — since we never seem to observe right-handed neutrinos (or left-handed antineutrinos), physicists took this to be an indication of zero mass.

If a particle does have mass, then we wish to find the probability of finding ψ_L (say) in the right handed helicity state. Because the particle is massive, $v < c$. The closer v is to c , the less likely it is to find a “normally left-handed” particle in a right-handed state. We saw an example of this with the electron in beta decay — the higher the electron energy, the more polarized they were. More concretely, the probability of being left-handed for the ψ_L state is proportional to $(1 + v/c)/2$.

As for the parity operation on either ψ_L or ψ_R , it can easily be shown that for $\vec{r} \rightarrow -\vec{r}$, $\psi_L(\vec{r}) \rightarrow \psi_R(\vec{r})$.

9.3.5 The Dirac Equation Lagrangian

As we did for the Klein-Gordon equation, we can also write down the Lagrangian density for the Dirac equation. Again, we use the covariant form to see that

$$\begin{aligned} \mathcal{L} &= \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \\ &= i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi. \end{aligned} \tag{9.27}$$

As before, the first term is the kinetic energy term and the second is the mass term. The Lagrange equation then becomes

$$\frac{\partial \mathcal{L}}{\partial \bar{\psi}} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \bar{\psi})} \right) = 0. \quad (9.28)$$

Say we have a *scalar* (spinless) particle interacting with a spin-1/2 particle. The Lagrangian for this interaction will then combine that for the Klein-Gordon equation (Equation 9.11) and that for the Dirac equation (Equation 9.27):

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi + i \bar{\psi} \gamma^\mu \partial_\mu \psi - m \bar{\psi} \psi.$$

However, unfortunately with our current model, these particles will not actually interact. In order for an interaction to occur, we need some sort of current that includes both ϕ and ψ — essentially, this is how new theories are built. That is, they are built by starting with a hypothetical Lagrangian and determining its consequences.

Chapter 10

Quantum Electrodynamics

10.1 Maxwell's Equations

Thus far in all our dealings with the electromagnetic interaction, we have treated it as a classical field theory, via Maxwell's equations:

$$\begin{aligned}\nabla \cdot \vec{E} &= \rho, \\ \nabla \cdot \vec{B} &= 0, \\ \nabla \times \vec{E} &= -\frac{\partial \vec{B}}{\partial t}, \\ \nabla \times \vec{B} &= \frac{\partial \vec{E}}{\partial t} + \vec{j}.\end{aligned}$$

Maxwell's equations also satisfy the continuity equation¹. In order to relativize electromagnetism, we need to find a valid four-vector to use. As alluded to before, we can use the potentials, so that

$$A^\mu = (\phi, \vec{A}), \quad (10.1)$$

where

$$\vec{B} = \nabla \times \vec{A}, \quad (10.2)$$

$$\vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t}. \quad (10.3)$$

The *electromagnetic field tensor*² is given by

$$\begin{aligned}F^{\mu\nu} &= \partial^\mu A^\nu - \partial^\nu A^\mu \\ &= \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}.\end{aligned} \quad (10.4)$$

We can then use this to rewrite Maxwell's equations in another form. The homogenous equation is

¹We often times will use the Lorentz covariant form of the continuity equation, $\partial_\mu J^\mu = 0$.

²This is sometimes referred to as the *Maxwell stress tensor*.

$$\partial^\lambda F^{\mu\nu} + \partial^\mu F^{\lambda\nu} + \partial^\nu F^{\lambda\mu} = 0, \quad (10.5)$$

while the inhomogenous equation³ is

$$\partial_\mu F^{\mu\nu} = J^\nu \quad (10.6)$$

10.2 Electromagnetic Lagrangian

Next, we wish to find the Lagrangian associated with the electromagnetic field. Recall from the stretched string that we found a quadratic equation in the continuous variable ϕ (see Section 9.1). The electromagnetic field is (thus far) also continuous, and so we expect a similar Lagrangian. In fact, we find

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - J^\mu A^\mu. \quad (10.7)$$

The first term represents the homogenous part, while the second term represents the inhomogenous part.

We know that the potential energy for the electrostatic field is $U = q\phi$. Or, relativistically, we would have $q \rightarrow J^\mu$ and $\phi \rightarrow A^\mu$. As with the stretched string, we expect to get a wave equation if we solve for the equations of motion using Lagrange's equation.

10.3 Symmetries

Recall *gauge invariance* from classical electrodynamics, namely

$$\begin{aligned} A^\mu &\rightarrow A^\mu + \partial^\mu \chi \\ &= \left(\phi + \frac{\partial \chi}{\partial t}, \vec{A} - \nabla \chi \right). \end{aligned}$$

Making this transformation changes nothing because the field tensor then contains

$$\partial^\mu (\partial^\nu \chi) - \partial^\nu (\partial^\mu \chi) = 0.$$

In this *gauge transformation*, $\partial^\mu \chi$ serves as the *gauge*. Thus, we see that the homogenous term of \mathcal{L} will be invariant under this gauge.

For the inhomogenous piece, the change in \mathcal{L} (the change in action) is

$$\Delta s = - \int J_\mu \partial^\mu \chi d^4 x.$$

Using integration by parts, we find yet another continuity equation:

$$\Delta s = \int \partial^\mu J_\mu \chi d^4 x = 0.$$

This means that charge is conserved! We have finally demonstrated that *gauge invariance leads to charge conservation*.

³Note then that, e.g., $\partial_\mu F^{0\nu} = \rho$ is the same as writing the "old" equation $\nabla \cdot \vec{E} = \rho$, except that it saves a bit of ink!

10.4 Quantization of the Electromagnetic Field

In order to add electromagnetism to the Dirac equation, we will need to do the following:

1. Quantize the electromagnetic field.
2. “Fieldify” the spinors. This is sometimes called the *second quantization*.

We begin by looking at the wave equation from the Lagrangian:

$$\nabla^2 \vec{A} - \frac{\partial^2 \vec{A}}{\partial t^2} = 0. \quad (10.8)$$

The solution for one particular momentum \vec{k} and polarization h is

$$\vec{A} = a_h(\vec{k})\epsilon_h(\vec{k})e^{i(\vec{k}\cdot\vec{x}-\omega t)} + a_h^*(\vec{k})\epsilon_h^*(\vec{k})e^{-i(\vec{k}\cdot\vec{x}-\omega t)}, \quad (10.9)$$

where $\epsilon_h(\vec{k})$ is the polarization vector. The general solution (in a small volume V) is the sum over all possible momenta (a Fourier series):

$$\vec{A} = \sum_{h,\vec{k}} \frac{1}{\sqrt{\omega V}} \left[a_h(\vec{k})\epsilon_h(\vec{k})e^{i(\vec{k}\cdot\vec{x}-\omega t)} + a_h^*(\vec{k})\epsilon_h^*(\vec{k})e^{-i(\vec{k}\cdot\vec{x}-\omega t)} \right]. \quad (10.10)$$

From this solution, we could now calculate \vec{E} and \vec{B} .

In order to find the Hamiltonian, recall from electrodynamics that, in the appropriate units,

$$H = \frac{1}{2} \int (E^2 + B^2) d^3x.$$

Putting Equation 10.10 into the above, we find that

$$H = \sum_{h,\vec{k}} \omega \left[a_h^*(\vec{k})a_h(\vec{k}) + a_h(\vec{k})a_h^*(\vec{k}) \right]. \quad (10.11)$$

This is eerily reminiscent of the harmonic oscillator Hamiltonian. Then, by analogy, *we assume that the a_h 's are operators*. Then there will be a commutation relation

$$\left[a_h(\vec{k}), a_{h'}^\dagger(\vec{k}') \right] = \delta_{hh'}\delta_{\vec{k}\vec{k}'}, \quad (10.12)$$

because we want them to be an orthonormal set (we use the daggers now in order to be more general). This tells us that

$$a_h(\vec{k})a_h^\dagger(\vec{k}) - a_h^\dagger(\vec{k})a_h(\vec{k}) = 1.$$

The Hamiltonian then becomes

$$H = \sum_{h,\vec{k}} \omega \left[a_h^\dagger(\vec{k})a_h(\vec{k}) + \frac{1}{2} \right].$$

However, because of the sum, the zero-point energy (the 1/2 term) is *infinite*, and so we ignore it⁴. Thus, we write the quantized electromagnetic Hamiltonian as

$$H = \sum_{h,\vec{k}} \omega a_h^\dagger(\vec{k})a_h(\vec{k}) \quad (10.13)$$

⁴Somehow, this isn't a very satisfying justification. In fact, some argue that this zero point energy is the source (or perhaps one of the sources) of the dark energy that is accelerating the expansion of the universe.

Say we operate on a particular state $|n\rangle$ with frequency ω and helicity h . Then the summation goes away (we're only dealing with one state) and we have

$$H|n\rangle = \omega a_h^\dagger(\vec{k}) a_h(\vec{k}) |n\rangle = n\omega |n\rangle.$$

Thus far, no real surprises. Now if we operate with $H a_h^\dagger(\vec{k})$, we see that

$$\begin{aligned} H a_h^\dagger(\vec{k}) |n\rangle &= \omega a_h^\dagger(\vec{k}) a_h^\dagger(\vec{k}) a_h(\vec{k}) |n\rangle \\ &= \omega a_h^\dagger(\vec{k}) \left[a_h^\dagger(\vec{k}) a_h(\vec{k}) + 1 \right] |n\rangle \\ &= \omega a_h^\dagger(\vec{k}) |n\rangle + \omega a_h^\dagger(\vec{k}) a_h^\dagger(\vec{k}) a_h(\vec{k}) |n\rangle \\ &= \omega a_h^\dagger(\vec{k}) (1+n) |n\rangle. \end{aligned}$$

So apparently $a_h^\dagger(\vec{k})$ raises the energy by one unit. Thus, we say that a_h^\dagger is a *photon creation operator*. Similarly, we can show that a_h is a *photon annihilation operator*. From all of this, we see that electromagnetic fields are linear combinations of creation and annihilation operators so that

$$\langle n | H | n \rangle = 0,$$

because H changes the state $|n\rangle$. That is, the electromagnetic field is a superposition of many states of photons, or, equivalently, a superposition of waves with different momenta. The creation and annihilation operators act according to

$$a_h^\dagger(\vec{k}) |n\rangle = \sqrt{n+1} |n+1\rangle \quad (10.14)$$

$$a_h(\vec{k}) |n\rangle = \sqrt{n} |n-1\rangle. \quad (10.15)$$

Similarly, it can be shown that $\langle n' | n \rangle = \delta_{nn'}$.

Now we must make fermions (by way of ψ and $\bar{\psi}$) into fields. This is done completely analogously to the quantization of the electromagnetic field, so we simply start by saying

$$\psi(x, t) = \sum_{s, \vec{p}} \left[\tilde{b}_s(\vec{p}) u_s(\vec{p}) e^{i(\vec{p}\cdot\vec{x} - \omega t)} + \tilde{d}_s^\dagger(\vec{p}) v_s(\vec{p}) e^{-i(\vec{p}\cdot\vec{x} - \omega t)} \right].$$

This implies that \tilde{b}_s destroys $E > 0$ electrons, while the \tilde{d}_s^\dagger operator creates $E > 0$ positrons. In $\bar{\psi}$, we will have \tilde{d}_s destroy $E > 0$ positrons and \tilde{b}_s^\dagger create $E > 0$ electrons.

Now, we interpret $|\psi|^2$ as a *particle density* rather than a probability density. Finally, we can put everything together, letting $J^\mu = j^\mu = -e\bar{\psi}\gamma^\mu\psi$ (an interaction between the Fermi field and the electromagnetic field):

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - e\bar{\psi}\gamma^\mu\psi A_\mu. \quad (10.16)$$

We call $(1/4)F^{\mu\nu}F_{\mu\nu}$ the *dynamic term* for the electromagnetic field.

10.5 Local Gauge Invariance

Previously, we showed that the gauge invariance of Maxwell's equations necessitated the conservation of charge (see Section 10.3). However, what we really want to have is *local gauge invariance*. That is, we demand that charge should be conserved *locally*. If an electron is present somewhere in the vicinity of my desk, then I expect that electron (or at least another series of particles with net charge $-e$) to be still be present in the vicinity of my desk in the next nanosecond.

In the presence of an electromagnetic field, the Dirac equation changes according to $E \rightarrow E - q\phi$ and $\vec{p} \rightarrow \vec{p} - q\vec{A}$ so that we have

$$[\gamma^\mu (\partial_\mu - qA_\mu) - m]\psi = 0.$$

If we apply a gauge transformation to A_μ , we find that $\psi \rightarrow \psi' = \psi \exp(-iq\chi)$ (where q replaces e , the charge). Note that while the *wavefunction* changes, the actual *physics* does not — the phase factor disappears in the end when we go to make a prediction by squaring the magnitude of ψ .

Note also that this is quite reminiscent of the generator of infinitesimal rotations in SU(2) (see Chapter 8). This similarity means that we have another group associated with the wavefunction's transformation, namely the U(1) group⁵. Generally, a U(1) gauge transformation has the form $\psi \rightarrow \psi' = \psi \exp(-i\alpha)$.

Transforming ψ under U(1), we can see the consequences of a local gauge transformation by having α depend on position⁶: $\alpha = q\chi(\vec{x})$. Then it can be shown that in order to keep \mathcal{L} invariant, it is *required* that the field A_μ transform like $A_\mu \rightarrow A_\mu + \partial_\mu\chi$; this is *precisely* gauge invariance in electrodynamics! That is, *local U(1) gauge invariance requires the existence of photons!* For this reason, we call the electromagnetic field a *U(1) gauge field*.

Furthermore, one may ask why the Lagrangian does not take the form

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}c^2A_\mu A^\mu - J^\mu A_\mu.$$

The reason: The second term, which is the mass term (cf. the $m^2\phi^*\phi$ term in the Klein-Gordon equation), would require photons to have mass. *But this Lagrangian is not locally gauge invariant.* Hence, photons also *have to be massless*⁷.

10.6 Interactions

Now that we have a Lagrangian that describes more than just a free particle, we want to be able to calculate cross sections (e.g., for the simple electron scattering interaction, $e^- + e^- \rightarrow e^- + e^-$) using Fermi's golden rule (Equation 4.20). The interaction Hamiltonian is given by

$$H_{int} = \int \vec{j} \cdot \vec{A} d^3x. \quad (10.17)$$

The Fermi fields are ψ and $\bar{\psi}$, where

$$\psi = \sum_p \frac{1}{\sqrt{2EV}} \sum_s \left[b_s(\vec{p}) u_s(\vec{p}) e^{i(\vec{p}\cdot\vec{x} - Et)} + d_s^\dagger(\vec{p}) v_s(\vec{p}) e^{-i(\vec{p}\cdot\vec{x} - Et)} \right],$$

and $\bar{\psi}$ is the same except with primed momenta and spin, and complex conjugates taken.. Then Equation 10.17 becomes

⁵There is no 'S' because a scalar isn't "special." That is, the determinant is meaningless when dealing with a one component quantity. More precisely, the only member of "SU(1)" would be the number 1.

⁶Remember, the tilde indicates that the quantity is a four-vector.

⁷This can be shown in the following way. We know that

$$\frac{\partial^2 A}{\partial t^2} - \nabla^2 A^2 + c^2 A = 0.$$

It can then easily demonstrated that

$$\frac{1}{2}(A^\mu - \partial^\mu\chi)(A_\mu - \partial_\mu\chi) \neq \frac{1}{2}c^2 A^\mu A_\mu.$$

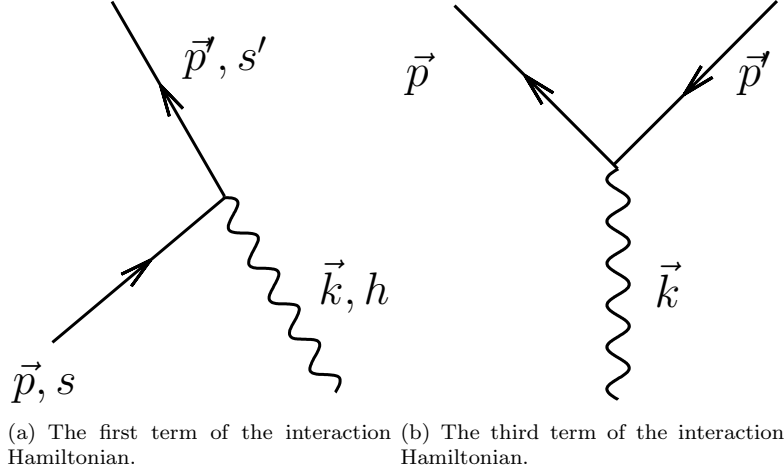


Figure 10.1: First order terms of the interaction Hamiltonian. These do not conserve momentum, and thus we must go to second order. Time moves forward in the upwards direction.

$$\vec{j} = \sum_{p,p'} \sum_{s,s'} \frac{1}{\sqrt{2EV}} \frac{1}{\sqrt{2E'V}} \left[b_{s'}^\dagger(\vec{p}') u_{s'}^*(\vec{p}') e^{-i(\vec{p}' \cdot \vec{x} - E't)} + d_{s'}(\vec{p}') v_{s'}^*(\vec{p}') e^{i(\vec{p}' \cdot \vec{x} - E't)} \gamma^0 \gamma^\mu \right] \\ \times \left[b_s(\vec{p}) u_s(\vec{p}) e^{i(\vec{p} \cdot \vec{x} - Et)} + d_s^\dagger(\vec{p}) v_s(\vec{p}) e^{-i(\vec{p} \cdot \vec{x} - Et)} \right]$$

We will also need the vector potential \vec{A} :

$$\vec{A} = \sum_{h,\vec{k}} \left\{ \frac{1}{\sqrt{\omega V}} \left[a_h(\vec{k}) \epsilon_h(\vec{k}) e^{i(\vec{k} \cdot \vec{x} - \omega t)} + a_h^\dagger \epsilon_h^*(\vec{k}) e^{-i(\vec{k} \cdot \vec{x} - \omega t)} \right] \right\}.$$

The only thing remaining is to “just” write down the interaction Hamiltonian H_{int} . However, this turns out to be an absolute mess of eight very large terms. Rather than write it down, notice that in each term, there will be two creation operators and one annihilation operator, or two annihilation operators and one creation operator. Furthermore, the ϵ terms with the spinors (u_s , v_s) indicate an interaction depending on the spin of the fermions and photons — this is not entirely unexpected, as particles with a magnetic moment should behave differently than those without. There are also integrals that are delta functions when normalized (from the exponentials in the above formulae) — this indicates the conservation of momenta of the particles.

Rather than use the mess that results, we often keep track of everything with diagrams (Figure 10.1). However, we see from these diagrams that the terms of H_{int} cannot conserve both energy and momentum simultaneously! What went wrong? Remember that when we derived Fermi’s golden rule, we had $d\sigma = (|m|^2 dQ)/F$ where dQ is Lorentz invariant phase space:

$$dQ = (2\pi)^4 \delta^4(p_C + p_D - p_A - p_B) \frac{d^3 p_C}{(2\pi)^3 2E_C} \frac{d^3 p_D}{(2\pi)^3 2E_D},$$

for the interaction $A + B \rightarrow C + D$, and F is the flux:

$$F = (\vec{v}_A - \vec{v}_B) \cdot 2E_A \cdot 2E_B.$$

The only thing we really need (and the only thing that really changes for each interaction) is $|m|^2$. However, when we derived the golden rule, we only did it to first order. The diagrams in Figure 10.1 are thus incorrect

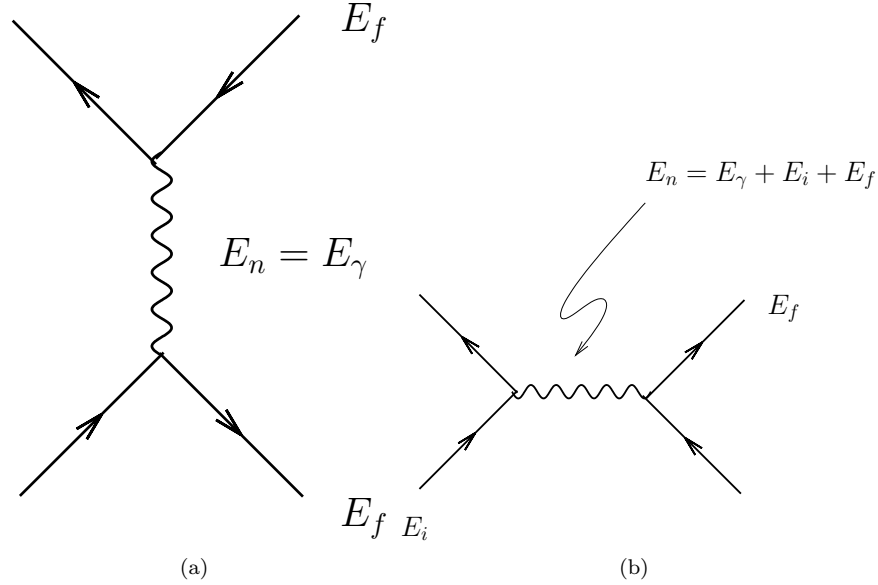


Figure 10.2: The two possible Feynman diagrams for electron–positron scattering.

because first order is not enough in this case! We really need to go to second order, where Fermi’s golden rule adds the term

$$T_{fi} = -2\pi i \sum_{n \neq i} \frac{V_{fn} V_{ni}}{E_i - E_n + i\epsilon} \delta(E_f - E_i). \quad (10.18)$$

Putting all this together, we get valid, momentum conserving interactions. For the interaction $e^- + e^+ \rightarrow e^- + e^+$, we have the diagrams as illustrated in Figure 10.2. There, we would have

$$\begin{aligned} m &\sim V_{fn} \frac{1}{E_i - E_\gamma} V_{ni} + V_{fn} \frac{1}{E_f - 2E_i - E_\gamma} V_{ni} \\ &= V_{fn} \frac{2E_\gamma}{E_i^2 - E_\gamma^2} V_{ni}. \end{aligned}$$

But $E_i^2 = (\tilde{p}_A^2 + \tilde{p}_B^2) + (\tilde{p}_A^2 + \tilde{p}_B^2)$ and $E_\gamma^2 = m_\gamma^2 + p^2$, where $\vec{p} = \vec{p}_A + \vec{p}_B$. So

$$\frac{1}{E_i^2 - E_\gamma^2} = \frac{1}{(\tilde{p}_A + \tilde{p}_B) - m_\gamma^2} = \frac{1}{q^2},$$

where q is the momentum transfer. We call this q^{-2} piece the *propagator* — it corresponds to the internal line in the second order *tree-level diagram*.

10.7 Feynman Diagrams and Feynman Rules

Nobel prize winner Julian Schwinger actually managed to meticulously do all the above calculations by hand. Richard Feynman (also a Nobel Prize winner) also started out doing calculations like this, but quickly found himself going mad. In order to keep track of all the terms, he drew little diagrams. Finally, it occurred to him that he could just *start* with the diagrams as a mathematical tool to represent the underlying physics.

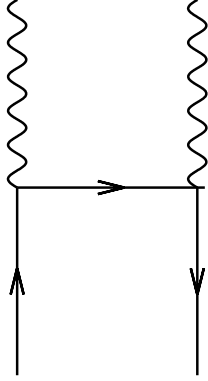


Figure 10.3: A tree level diagram with a fermionic propagator. The propagator has $E^2 \neq p^2 + m^2$. We say that it is *off mass shell* because of this.

We must always begin by drawing all *topologically distinct* diagrams for a particular process. For example, take *Bhabha scattering*, $e^-e^+ \rightarrow e^-e^+$, shown in Figure 10.4. In Rutherford scattering, $e^-e^- \rightarrow e^-e^-$, there is only one topologically distinct diagram that may be drawn.

Each (tree level) Feynman diagram has three basic pieces:

1. External lines.
2. Internal line (the propagator).
3. Two vertices.

In order for these diagrams to be useful, we follow the *Feynman rules* in order to calculate the matrix element m . They are (for what we have done thus far) as follows ($m = 1 \times$ (factors) from these rules):

1. **External Lines.**

- (a) External spinless boson $\rightarrow 1$
- (b) Spin-1/2 fermion $\rightarrow u_s(p)$ on the right if the initial state; $\bar{u}_s(p)$ on the left if the final state.
- (c) Spin-1/2 antifermion $\rightarrow \bar{v}_s(p)$ on the left if the initial state; $v_s(p)$ on the right if the final state.

2. **Vertices.**

- (a) $-iq\gamma^\mu$ in between the above.

3. **Internal Lines (Propagator).**

- (a) Photon \rightarrow

$$\frac{-ig_{\mu\nu}}{q^2 + i\epsilon}.$$

- (b) Fermion \rightarrow

$$\frac{i(\gamma^\mu q_\mu + m)}{q^2 - m^2 + i\epsilon}$$

When put together, we will have $m \sim$ (current) \times (propagator) \times (current).

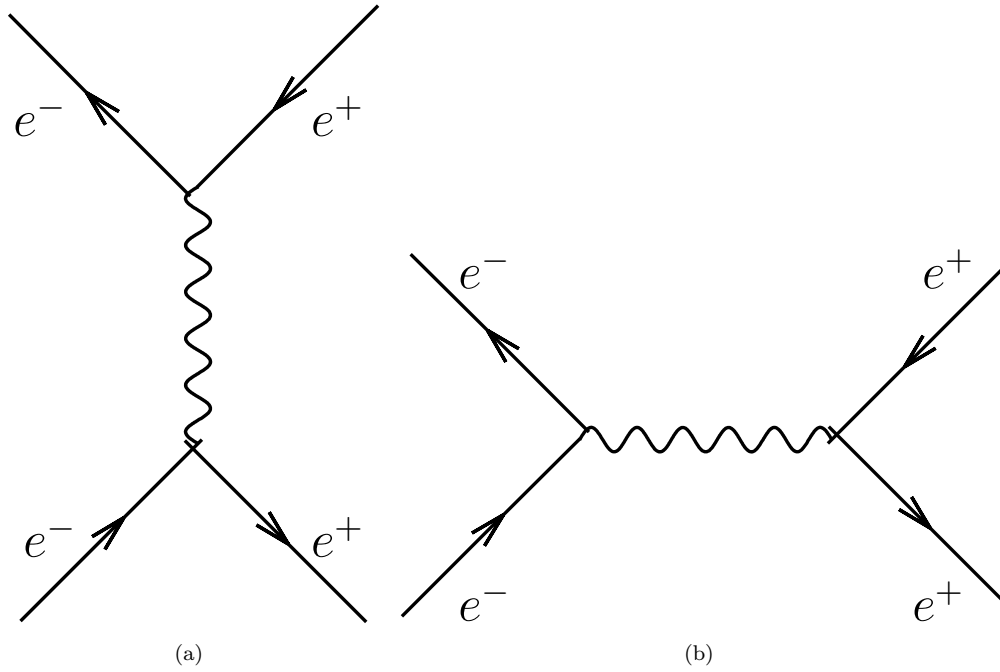


Figure 10.4: The two topologically distinct Feynman diagrams for Bhabha scattering. In all the diagrams that we draw, the forward flow of time is in the upwards direction (“time flows upwards”).

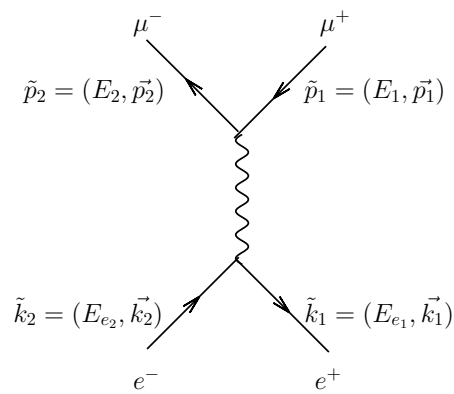
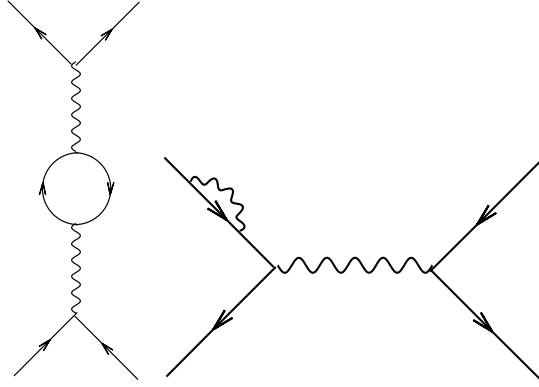


Figure 10.5: The setup for the example.



(a) A fourth order Feynman diagram. (b) Another fourth order Feynman diagram. The extra photon is known as a *self energy*.

Figure 10.6: Higher order Feynman diagrams.

Example: $e^+e^- \rightarrow \mu^+\mu^-$

See Figure 10.5. After applying the Feynman rules we have

$$m = \bar{v}_{s_1}^e(\vec{k}_1)(+ie\gamma^\rho)u_{s_2}^e\left(\frac{-ig_{\rho\tau}}{q^2 + i\epsilon}\right)\bar{u}_{s_2}^\mu(\vec{p}_2)(-ie\gamma^\tau)v_{s_1}^\mu.$$

Here, we use ρ and τ instead of the usual μ and ν to label the indices to sum over in order to avoid having multiple μ 's that mean different things. Note that m is an amplitude. This means that when another diagram is topologically distinct, the matrix elements must be added and then squared — thus we can get diagrams interfering with each other, and potentially even canceling each other out! Also, note the $1/q^2$ dependence. This means that as the momentum transfer increases, the amplitude decreases.

It is often useful to use $|\bar{m}|^2$, which is the average m over all spins,

$$|\bar{m}|^2 = \frac{1}{4} \sum_{s_1, s_2} mm^\dagger.$$

To get the differential cross section, there are several tricks involving $\text{tr}(\gamma)$. Eventually, we find that

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4s} [1 + \cos^2(\theta)], \tag{10.19}$$

where $s = (\vec{p}_1 + \vec{p}_2)^2$ is the total center of mass energy and α is the finestructure constant. The total cross section is then

$$\sigma = \frac{4\pi\alpha^2}{3s}.$$

10.7.1 Higher Order Diagrams

There can of course be higher order diagrams. However, because of the α dependence, each higher term gets smaller (remember, $\alpha \approx 1/137$).

We say that the lines in the diagrams are the *bare charges*. However, the vacuum is constantly pair producing and annihilating, thus polarizing the vacuum and subsequently shielding the charge. However, there are effectively an infinity of possible diagrams. We can solve this problem by saying that the lines are

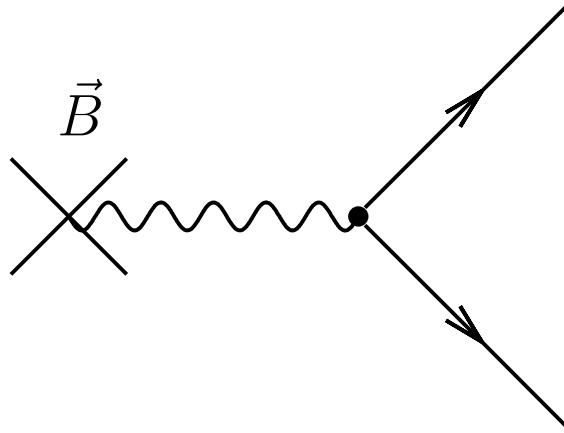


Figure 10.7: An electron moves in the presence of a magnetic field.

whatever charge we actually measure. This means that *the finestructure constant actually increases with energy*. It takes the form

$$\alpha(q^2) = \frac{\alpha(\mu^2)}{\left[1 - \frac{1}{\pi}\alpha(\mu^2) \ln\left(\frac{q^2}{\mu^2}\right)\right]}, \quad (10.20)$$

where μ is some scaling factor. Unfortunately, this means that the electromagnetic interaction becomes *infinitely large at high enough energies*. Nevertheless, QED remains one of the most precise theories ever formulated.

10.8 Magnetic Fields

Say an electron is moving in a magnetic field (see Figure 10.7). Because the electron has a magnetic moment, there can also be photons between the lepton lines in the figure. Recall that the *Bohr magneton* is given by

$$\mu_B = \frac{e\hbar}{2mc}. \quad (10.21)$$

Then the magnetic moment on the electron is given by

$$\mu = g\mu_B s, \quad (10.22)$$

where $g = 2$ for the “normal” (first order) magnetic moment. The deviation from the theoretical value of 2 can be measured in the electron (though usually muons are used for several reasons):

$$\left(\frac{g-2}{2}\right)^{\text{th}} = 0.5\left(\frac{\alpha}{\pi}\right) - 0.328\left(\frac{\alpha}{\pi}\right)^2 + \dots = 115965230 \pm 10 \times 10^{-11}. \quad (10.23)$$

This is the highest precision prediction ever made with the standard model and holds up well to experiment.

Chapter 11

The Weak Interaction

We now return our attention to the weak interaction, although with a much more sophisticated set of tools now. Unfortunately, unlike the electromagnetic case, there is no classical theory to guide us; instead we only have experiment.

Recall that in 1957, Wu, et al. found that only negative helicity electrons participate in beta decay (parity violation). From Dirac, the probability of being in a negative helicity state when in the left-handed chiral state is given by

$$\frac{1}{2} \left(1 + \frac{v}{c} \right),$$

while that for the positive helicity state exchanges the plus sign for a minus. The degree of polarization is then just the sum:

$$-\frac{1}{2} \left(1 + \frac{v}{c} \right) + \frac{1}{2} \left(1 - \frac{v}{c} \right) = -\frac{v}{c}.$$

11.1 Pion Decay

We know that neutrinos are (or once were believed to be) always left-handed and that leptons are universal, with the only difference being their masses:

$$\begin{aligned} m_e &= 0.511 \text{ MeV}, \\ m_\mu &= 106 \text{ MeV}, \\ m_\tau &= 1777 \text{ MeV}. \end{aligned}$$

We will look at the decay of the π^- (recall that $m_\pi \approx 140 \text{ MeV}$), in the following channels:

$$\begin{aligned} \pi^- &\rightarrow e^- \bar{\nu}_e \\ \pi^- &\rightarrow \mu^- \bar{\nu}_\mu. \end{aligned}$$

From Fermi's golden rule, the decay rate is

$$w = |m|^2 \frac{dN}{dE_0}.$$

Then because the electron is more massive, there is more available phase space, and we would expect given this that the pion would decay into an electron and electron antineutrino far more often. However, when we observe the branching ratio, we find that

$$\frac{B(\pi^- \rightarrow e^- \bar{\nu}_e)}{B(\pi^- \rightarrow \mu^- \bar{\nu}_\mu)} = 10^{-4}. \quad (11.1)$$

The reason for this apparent “mistake” is the following. We know that the π^- has 0 spin, and the neutrinos and leptons each have spin-1/2. Then to conserve angular momentum, the spins of the neutrino and lepton must be in opposite direction. Similarly, to conserve momentum in the pion’s rest frame, the neutrino and electron must travel with equal and opposite momenta. Antineutrinos are always right-handed, but because the spins have to be opposite, the lepton must *also* be right-handed!

Of course, the probability of being right-handed increases as mass increases. For the electron, Wu found that

$$\frac{1}{2} \left(1 - \frac{v_e}{c}\right) = \frac{m_e^2}{m_\pi^2 + m_e^2} = 10^{-5},$$

whereas for the muon this same probability calculation yields 0.4. Nature choosing the heavier muon rather than the lighter electron due to helicity considerations is known as *helicity suppression*.

11.2 The Weak Lagrangian

Clearly, something about the weak interaction prefers left-handed particles and right-handed antiparticles. Somehow, we need to include the helicity in the Lagrangian, then. We begin by writing the (approximated) Lagrangian densities for the above pion decay process. The free particle Lagrangians are given by

$$\mathcal{L}_{\text{free}}^\pi = \partial_\mu \phi_\pi^\dagger \partial^\mu \phi_\pi - m_\pi^2 \phi_\pi^\dagger \phi_\pi, \quad (11.2)$$

$$\mathcal{L}_{\text{free}}^{\text{leptons}} = \bar{\psi}_e (\gamma^\mu i \partial_\mu - m_e) \psi_e + \bar{\psi}_\mu (\gamma^\mu i \partial_\mu - m_\mu) \psi_\mu, \quad (11.3)$$

$$\mathcal{L}_{\text{free}}^\nu = \bar{\psi}_{\nu_e} \gamma^\mu i \partial_\mu \frac{1}{2} (1 - \gamma^5) \psi_{\nu_e} + \bar{\psi}_{\nu_\mu} \gamma^\mu i \partial_\mu \frac{1}{2} (1 - \gamma^5) \psi_{\nu_\mu} \quad (11.4)$$

Note that while in principle there should be a term for the tauon in the lepton Lagrangian, we neglect it since it is far too massive for a pion to be able to decay into a tauon in the first place. Note the presence of the projection operator $(1/2) \times (1 - \gamma^5)$ to pick out only the left-handed neutrino state.

The interaction term is given by

$$\mathcal{L}_{\text{int}} = \alpha_\pi [j^\mu \partial_\mu \phi_\pi + j^{\mu\dagger} \partial_\mu \phi_\pi^\dagger], \quad (11.5)$$

where the second term is for the π^+ , α_π is some coupling term, and

$$j^\mu = \bar{\psi}_e \gamma^\mu \frac{1}{2} (1 - \gamma^5) \psi_{\nu_e} + \bar{\psi}_\mu \gamma^\mu \frac{1}{2} (1 - \gamma^5) \psi_{\nu_\mu}. \quad (11.6)$$

Note that the interaction term is just (current) \times (current) and has no propagator. If we look closely at two possible interactions, we notice that there may or may not be a *charge exchange* (see Figure 11.1 and Figure 11.2). In this case, the interaction Lagrangian becomes

$$\mathcal{L}_{\text{int}} = -2\sqrt{2}G_F j_\mu^\dagger j^\mu = \left[\bar{\psi}_e \gamma^\mu \frac{1}{2} (1 - \gamma^5) \psi_{\nu_e} \right] \left[\bar{\psi}_{\nu_e} \gamma_\mu \frac{1}{2} (1 - \gamma^5) \psi_e \right] (-2\sqrt{2}G_F). \quad (11.7)$$

In the case of the neutral current interaction, we will have

$$\mathcal{L}_{\text{int}} = \frac{G_F}{\sqrt{2}} [(j_{\text{neutral}})^\mu (j_{\text{neutral}})_\mu],$$

but we do not yet know what this current is. We do, however, know the form that it will take, namely

$$(j_{\text{neutral}})^\mu = \bar{\psi}_{\nu_\mu} \gamma^\mu \frac{1}{2} (1 - \gamma^5) \psi_{\nu_\mu} + \bar{\psi}_e \gamma^\mu (C_V - C_A \gamma^5) \psi_e.$$

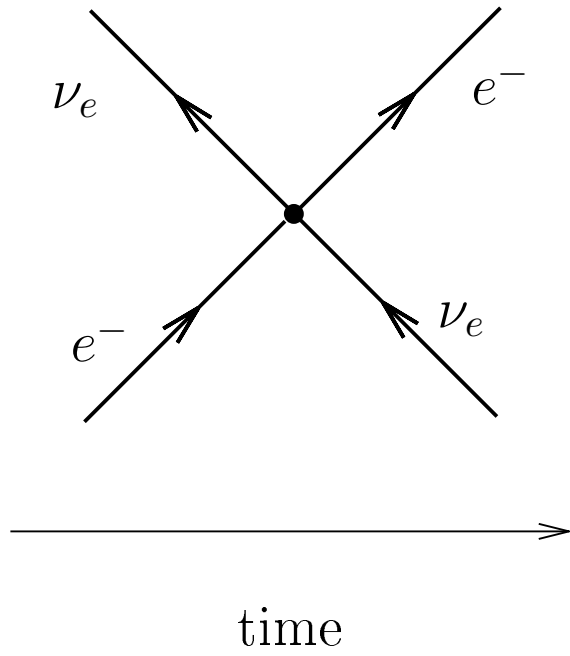


Figure 11.1: A charged current interaction.

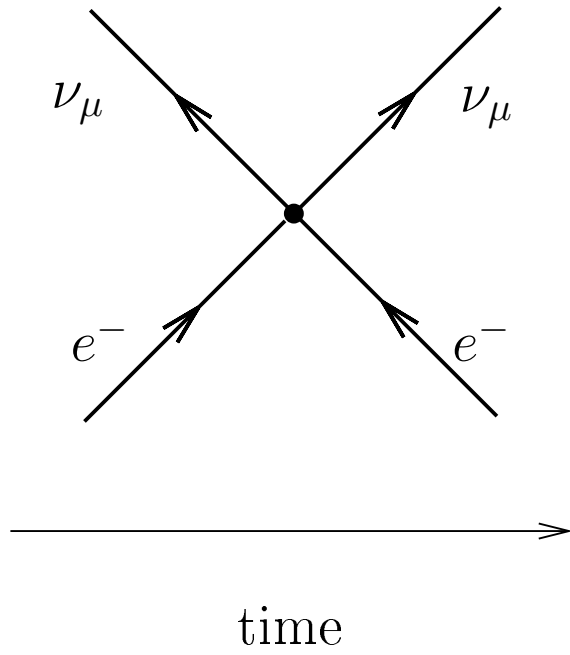


Figure 11.2: A neutral current interaction.

11.3 Cross Sections

We now wish to examine the cross sections for both the neutral current weak interaction (e.g., $\nu_e + e^- \rightarrow e^- + \nu_e$) and for the charged current weak interaction (e.g., $\nu_\mu + e^- \rightarrow \nu_\mu + e^-$). It can be shown that

$$\frac{d\sigma}{dq^2} = \frac{G_F^2}{\pi},$$

making the *total* cross section

$$\sigma_{\text{tot}}(\nu_e) = \frac{G_F^2}{\pi} s,$$

where as usual, $s = (\tilde{p}_e + \tilde{p}_{\nu_e})^2 \approx 2m_e E_{\text{incident}}^{\nu_e}$. We immediately see a problem with this prediction — as the energy increases, we eventually have an *infinite* cross section! Yet this is not the only problem. To discover the other one, we must revisit the calculation of cross sections in the first place.

11.3.1 Calculating Cross Sections — Revisited

Recall the concept of compound nuclei or *resonances* first discussed in Section 5.6. The lifetime of the resonant state is given by $\tau = w^{-1}$, where w is the rate as calculated by Fermi's golden rule,

$$w = \frac{2\pi}{\hbar} |M_{if}|^2 \rho_f.$$

By the uncertainty principle, the width of the outgoing energy distribution is inversely proportional to the lifetime:

$$\Gamma = \frac{\hbar}{\tau} = \hbar w = 2\pi |M_{if}|^2 \int \rho_f d\Omega.$$

The total number of particles in state A decays exponentially, namely

$$N_A(t) = N_A(0)e^{-t/\tau} = N_A(0)e^{-\Gamma t/\hbar},$$

where we note that if there are many channels, $\Gamma = \sum_i \Gamma_i$. The wavefunction for the decaying state is simply

$$\psi(t) = \psi(0)e^{-i\omega_R t} e^{-t/2\tau},$$

where $\omega_R = E_R/\hbar$, the “R” meaning “resonance,” and $E_R = m_R$ in the center of momentum frame. In units of $\hbar = c = 1$, we can then write the wavefunction as

$$\psi(t) = \psi(0)e^{-t(iE_R + 1/2)}.$$

Setting $\omega = E/\hbar$ and performing a Fourier transform, we have

$$g(\omega) = \int_0^{+\infty} \psi(t)e^{i\omega t} dt,$$

and

$$\chi(E) = \int \psi(t)e^{iEt} dt = \psi(0) \int \exp \left\{ -t \left[\frac{\Gamma}{2} + i(E_R - E) \right] \right\} dt.$$

For K some constant, we have

$$\chi(E) = \frac{K}{(E - E_R) - i\Gamma/2},$$

implying that the cross section is

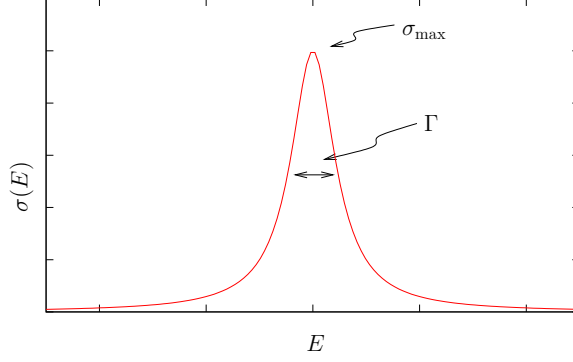


Figure 11.3: The Breit-Wigner resonance curve.

$$\sigma(E) = \sigma_{\max} \frac{\Gamma^2/4}{(E - E_R)^2 + \Gamma^2/4}. \quad (11.8)$$

This is known as the *Briet-Wigner resonance curve* (Figure 11.3). The only task we have left is finding the maximum cross section.

For l an angular momentum (such that $l\hbar = pb$), p a linear momentum, and b the impact parameter, we have for particles in the region $l \rightarrow l + 1$

$$\sigma = \pi(b_{l+1}^2 - b_l^2) = \pi\lambda(2l + 1),$$

where $\lambda = \hbar/p$ is the de Broglie wavelength. Then assuming pure scattering and no absorption,

$$\sigma_{\max} = 4\pi\lambda^2(2l + 1),$$

so in the Breit-Wigner case, we have

$$\sigma_{a+b \rightarrow c+d}^{\text{BW}} = \frac{4\pi\lambda^2(2J + 1)}{(2s_a + 1)(2s_b + 1)} = \frac{\Gamma^2/4}{(E - E_R)^2 + \Gamma^2/4}.$$

However, this is not Lorentz invariant. It can be shown that to make this expression Lorentz invariant, we need

$$\sigma(s) = \sigma_{\max} \frac{M_0^2 \Gamma^2}{(s - M_0)^2 + \Gamma^2 M_0^2},$$

where M_0 is the mass of the resonance. Effectively, when we see this shape experimentally, we then know that the resonance existed. If the shape does not appear, then there was no resonance.

11.3.2 Four-Point Interaction

We now can calculate the maximum cross section for the $\nu_e + e \rightarrow \nu_e + e$ interaction. For point-like scattering (i.e., $l = 0$), it is given by

$$\sigma_{\max} = \pi\lambda^2 \frac{2l + 1}{2s + 1} = \frac{\pi\lambda^2}{2},$$

since both particles are spin-1/2. A calculation from Fermi's golden rule says that

$$\sigma_{\text{tot}}(\nu_e) = \frac{G_F^2 s}{\pi} = \frac{4G_F^2 p^*}{\pi} > \frac{\pi}{2(p^*)^2},$$

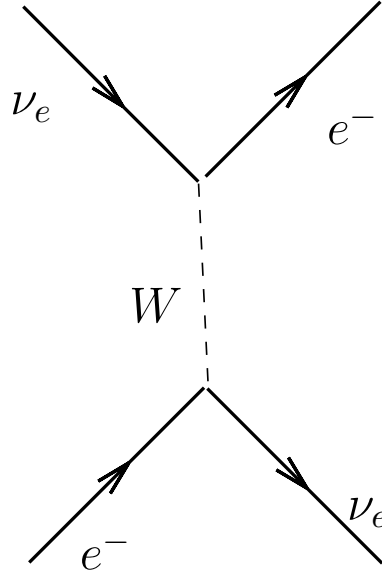


Figure 11.4: A Feynman diagram representing a charged current interaction. In this figure, time flows to the right. The propagator contributes a factor of $1/(1 + q^2/M_W^2)$.

where the starred quantities are in the center of momentum frame. But this tells us that at some point, we have more coming out than what we have going in. This is a *violation of unitarity*.

In other words, *Fermi's four-point interaction is an incorrect picture of the weak interaction*. Until $p^* \sim 300$ GeV, however, Fermi's picture remains reasonably accurate. As it so happens, we would not have this violation of unitarity if the coupling constant G_F did not have units of (1/Energy)! What we really need is a dimensionless coupling constant akin to α in QED.

In turn, this implies we need a propagator.

11.4 Charged and Neutral Currents

For the charged current, the propagator is the W boson, which takes on either a +1 or -1 charge depending on “which way” it is “going.” Since we have no way of really knowing in which direction it is going, we usually just leave off the arrows on the Feynman diagram representation of the charged current (see Figure 11.4).

We know that the weak interaction should be very short range since Fermi's theory worked up until this point. This implies that the W bosons must be massive in order to last only a short time, since $\Delta E \Delta t \sim \hbar/2$. Now the propagator term is

$$\frac{1}{1 + q^2/M_W^2}, \quad (11.9)$$

so the cross section goes like

$$\sigma \sim \frac{G_F^2 M_W^2}{\pi}.$$

However, now the coupling constant is not Fermi's G_F , but rather

$$g^2 \approx \frac{G_F}{M_W^2},$$

where

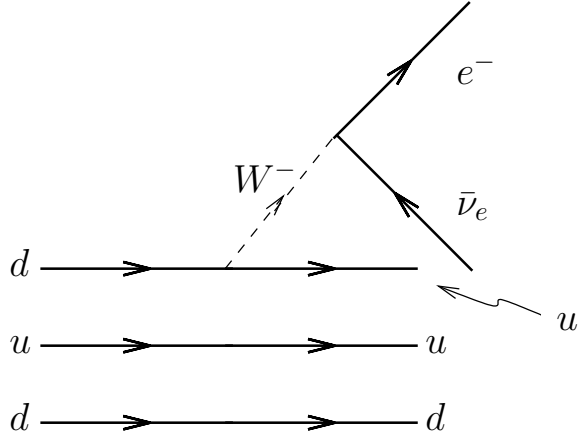


Figure 11.5: Beta decay with the charged current in terms of the constituent quarks. The two quarks that do not participate in the interaction are called *spectator quarks*. Time proceeds to the right.

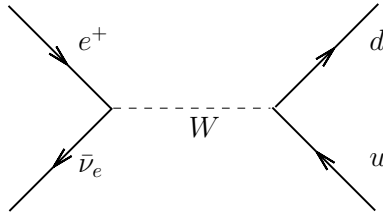


Figure 11.6: Inverse beta decay. Time increases upwards.

$$g = e \sin(\theta_W).$$

11.4.1 Beta Decay Revisited

We now reexamine beta decay, namely

$$n \rightarrow p + e^- + \bar{\nu}_e.$$

Now that we are a little more sophisticated, we know that the neutron is composed of two down quarks and one up quark, whereas the proton is two up quarks and one down quark. This means that effectively, one quark is changing from a down to an up (Figure 11.5). Note that by adding the W boson, we have added a degree of freedom which gets around the problem of unitarity violation.

The process of “inverse beta decay,”

$$\bar{\nu}_e + p \rightarrow n + e^+,$$

is also very similar under our new context and is pictured in Figure 11.6.

11.4.2 Neutral Current

An example of a neutral current interaction is that of neutrino-electron scattering:

$$\nu_e + e^- \rightarrow \nu_e + e^-,$$

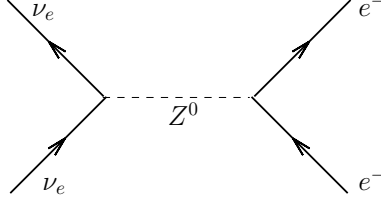


Figure 11.7: Neutrino-electron scattering, an example of a neutral current interaction.

pictured in Figure 11.7. Note a similarity between the Z^0 boson and the photon: Neither can change the charge of anything. Additionally, we note that we never see photon-photon scattering (since the photon couples to charged particles), nor will we see Z^0 - Z^0 scattering. However, both the photon and the Z^0 will couple to the W^\pm since it is charged.

11.5 The Electroweak Interaction

Recall that we “got” photons from gauge invariance symmetry arguments. We would then expect that similar arguments would show a symmetry involved with the weak gauge bosons.

Recall the concept of isospin (Chapter 8). Here, the $SU(2)$ group of Pauli matrices served as rotation generators. Similarly, Weinberg conjectured that there is a left-handed doublet of *weak isospin*:

$$\vec{L} = \begin{pmatrix} \nu_{e,L} \\ e_L \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(1 - \gamma^5) \psi_{\nu_e} \\ \frac{1}{2}(1 - \gamma^5) \psi_{e^-} \end{pmatrix}. \quad (11.10)$$

(There are also similar doublets for the μ and τ leptons.) Because at the time the neutrino was still thought to be a massless particle, there is also a right-handed weak isosinglet:

$$R = (e_R). \quad (11.11)$$

The charged current now becomes

$$j_\alpha = \vec{L}_e \gamma_\alpha \tau_- L_e, \quad (11.12)$$

where τ_- is the $SU(2)$ weak isospin lowering operator.

One may now ask what the coupling constant is. We call it the *weak hypercharge* Y :

$$Y = 2(Q - I_3), \quad (11.13)$$

where Q is the electromagnetic charge and I_3 is the third component of weak isospin. We then see that for both $\nu_{e,L}$ and e_L , $Y = -1$.

Weinberg further postulated that at some point in the distant past, there must have been a good symmetry between the neutrino and the electron, even though today, one has mass and the other doesn't; one has charge and the other is neutral. Perhaps the electron was at one point massless as well.

To address this, it was proposed that the W^\pm , Z^0 , and γ bosons were once contained in a weak isospin triplet,

$$\begin{pmatrix} W_1 \\ W_2 \\ W_3 \end{pmatrix}, \quad (11.14)$$

and a singlet,

$$(B_\mu). \quad (11.15)$$

All of these bosons are massless. In the Lagrangian, the kinetic term changes from $\partial_\mu\psi$ to

$$D_\mu\vec{L} = \left[\partial_\mu + ig_2\vec{\tau} \cdot \vec{W}_\mu + i\left(\frac{g_1}{2}\right) B_\mu \right] \vec{L}, \quad (11.16)$$

where by convention, we write the Pauli matrices as $\vec{\tau}$ instead of the usual $\vec{\sigma}$. This then means that the Lagrangian,

$$\mathcal{L} = \vec{\bar{L}}\gamma^\mu iD_\mu\vec{L} + \psi_{eR}\gamma^\mu iD_\mu\psi_{eR} \quad (11.17)$$

is invariant under the $U(1) \times SU(2)$ symmetric rotation

$$\vec{L}' = U\vec{L} = \exp\left(\frac{-i}{2}g\vec{\tau} \cdot \vec{\Lambda}\right)\vec{L}, \quad (11.18)$$

where g is either g_1 or g_2 and $\vec{\Lambda}$ is the isospin vector.

So far, we have treated everything as massless, yet we know that at least the electron has mass. To reconcile this, we consider the analogy of a dipole in a hot ferromagnet. Above the Curie temperature, the ferromagnet exhibits complete isotropy; that is, the dipole experiences zero net force everywhere. As the ferromagnet cools down, however, domains begin to form in order to try to minimize energy. This *spontaneous symmetry breaking* causes the dipole to experience a force, which gives it an “effective mass.” Because each dipole within the ferromagnet interacts with every other dipole, there must be a wave (and hence a particle) to communicate the motion. So not only has spontaneous symmetry breaking introduced mass, it has also introduced a new particle.

Similarly, in the early universe, it is postulated that spontaneous symmetry breaking gives particles mass (except the photon and gluon, which must have zero mass in order to have $U(1)$ gauge invariance) and makes a new particle, the spinless *Higgs boson* whose coupling constant is mass. Because of this, we say that *the Higgs field generates mass*. The Higgs field has the Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}\mu^2\phi^2 - \frac{1}{4}\lambda\phi^4. \quad (11.19)$$

Note that the last term means that a Higgs bosons can interact with each other.

Since $L = T - V$, we find that V_{\min} occurs when $\phi = \phi_{\min}$, or $\partial\phi/\text{partial}V = 0$. That is,

$$\phi(\mu^2 + \lambda\phi^2) = 0. \quad (11.20)$$

This isn't quite right. If $\mu^2 > 0$, then the potential looks just like that in the case of a hot magnet, as shown in Figure 11.8. But what if μ^2 has nothing to do with mass and instead is just a constant required in the Lagrangian? If $\mu^2 < 0$, then we have the so-called *Mexican hat potential* (Figure 11.9), leading to two minima. With two minima, the vacuum would have to choose one for the Higgs boson to reside. It is postulated that when symmetry was broken, the positive minimum was chosen, and since the Higgs is then in a potential well, it is given mass.

“And here a miracle happens...” Having ϕ real and spoiling the $SU(2)$ symmetry by making μ^2 a parameter, W_1 becomes the W^+ boson, W_2 becomes the W^- , and the photon and Z^0 are formed by a linear combination of the others:

$$\begin{aligned} A_\mu &= \sin(\theta_W)B_\mu + \cos(\theta_W)W_3 \\ Z^0 &= -\cos(\theta_W)B_\mu + \sin(\theta_W)W_3, \end{aligned}$$

and only A_μ remains massless. We call the mixing angle θ_W the *Weinberg* or *weak mixing angle*. It is related to the weak interaction strength constant g by

$$g = e \sin(\theta_W),$$

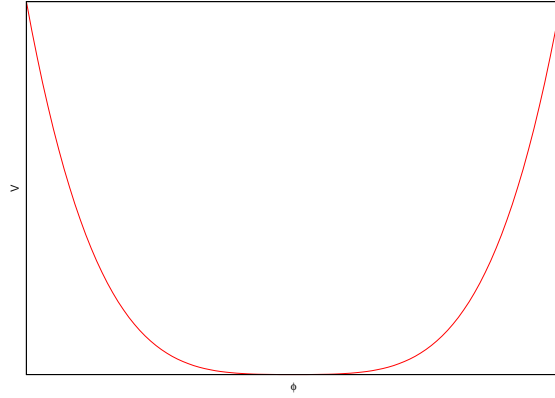


Figure 11.8: The Higgs potential for $\mu^2 > 0$. This case is analogous to the isotropic “hot magnet” example.

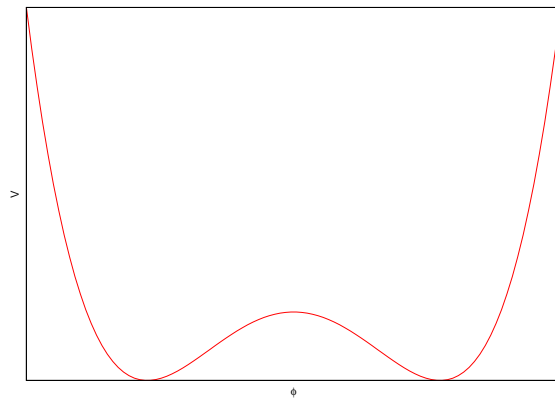


Figure 11.9: The Higgs potential for $\mu^2 < 0$ (the *Mexican hat potential*). Notice there are now two minima, $-V$ and V . In reality, this is actually a three dimensional potential.

and to the masses of the W and Z by

$$\sin(\theta_W) = \frac{M_Z}{M_W}.$$

The ratio of the masses of the Z and W have been observed, and fit this relation very precisely, giving credence to the *electroweak theory*.

Chapter 12

Quarks and Leptons

12.1 Quarks

Just like with the electrons and neutrinos, quarks are present in doublets and singlets. The left-handed doublet (for the u and d quarks) is

$$\begin{pmatrix} u_L \\ d_L \end{pmatrix},$$

while the two right-handed singlets are

$$\begin{pmatrix} u_R \\ d_R \end{pmatrix}.$$

Just like the leptons, quarks have weak isospin. The same process is followed in which there is symmetry breaking and then the Higgs mechanism provides the mass for the quarks.

12.1.1 Meson Decay

Consider the decay of the positively charged pion $\pi^+ \rightarrow \mu^+ \nu_\mu$, where, recall that $m_\pi \sim 140$ MeV. The lifetime of the pion is $\tau \approx 2.6 \times 10^{-8}$ s, and we can now explain the decay in terms of the weak current (Figure 12.1).

Another, stranger meson decay involves the kaon, $K^+ \rightarrow \mu^+ \nu_\mu$, which has a lifetime of $\tau \approx 1.28 \times 10^{-8}$ s and $m_K \sim 493$ MeV. What's so strange about this? Basically, it is that the lifetime of the K^+ is not much shorter than that of the π^+ , despite being considerably more massive! How strange.

Which brings us to the strange quark. Kaon decay led scientists to discover it, and now we know that the quark content of a K^+ is $u\bar{s}$. Due to *strangeness*, the decay of the kaon has a lifetime comparable to that of the pion. This new quantum number means that something different is happening when a u and \bar{s} interact than when, say, a u and \bar{d} interact.

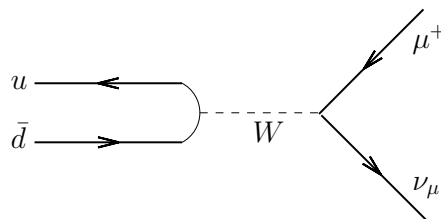


Figure 12.1: Pion decay in the language of quarks and the weak interaction.

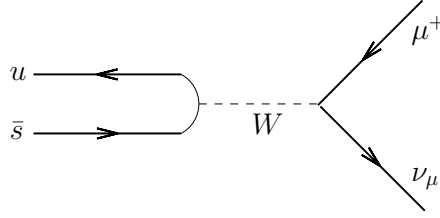


Figure 12.2: Kaon decay. The strange antiquark leads to the decay rate of the kaon being similar to that of the pion.

But how do we go between generations? The interaction term looked like $\bar{L}\gamma\tau\vec{L}$. It could be that there is a new “X” boson, but this turns out not to be the case. Instead, it turns out that the quarks we see are not the “real” quarks, but instead are superpositions of the “real” ones. The “real” quarks are eigenstates of the strong interaction, whereas we observe

$$\begin{pmatrix} u \\ d' \end{pmatrix} \quad \begin{pmatrix} c \\ s' \end{pmatrix},$$

where

$$d' = d \cos(\theta_C) + s \sin(\theta_C) \quad (12.1)$$

$$s' = -d \sin(\theta_C) + s \cos(\theta_C), \quad (12.2)$$

and $\theta_C \approx 12^\circ$ is the *Cabbibo angle*¹. This mixing angle means that the vertex term in our Feynman diagrams (e.g., Figure 12.2) go from having a factor of G_F to having a factor of $G_F \sin(\theta_C)$. There is an associated *mixing matrix*, given by the familiar

$$\begin{pmatrix} \cos(\theta_C) & \sin(\theta_C) \\ -\sin(\theta_C) & \cos(\theta_C) \end{pmatrix} \quad (12.3)$$

In the case of muon decay, say, we still use G_F because there are no quarks involved. Unlike the strong interaction, the weak force interacts with the superpositions of the “real” quarks.

Now imagine neutral kaon decay, $K^0 = \bar{s}d \rightarrow \mu^+\mu^-$ (Figure 12.3). Here, we would get a matrix element of $m = \cos(\theta_C) \sin(\theta_C)$. However, experimentally, we find that this is an incredibly small branching ratio. Glashow, Iliopoulos and Maiani (GIM) proposed that we don’t often see this decay because it is cancelled by another process which involves a virtual *charmed quark* instead of a virtual up quark (Figure 12.4). In this case, $m = -\cos(\theta_C) \sin(\theta_C)$, meaning that the matrix elements cancel². We call this the *GIM mechanism*.

12.2 Meson Oscillation

Mixing can lead to *oscillation of mesons*. For example, take the neutral kaon, $K^0 = \bar{s}d$ and its antiparticle, $\bar{K}^0 = d\bar{s}$ (see Figure 12.5). We can make both K^0 and \bar{K}^0 via

$$n + p \rightarrow \Lambda + \bar{K}^0 + p$$

and

$$n + p \rightarrow K^+ + K^0 + n + n,$$

¹Because it is so small, a d' is almost a d and a s' is almost a s

²They do not precisely cancel due to the difference in c and u masses, but they come very close to doing so. The fact that they don’t exactly is why we can observe neutral kaon decay at all.

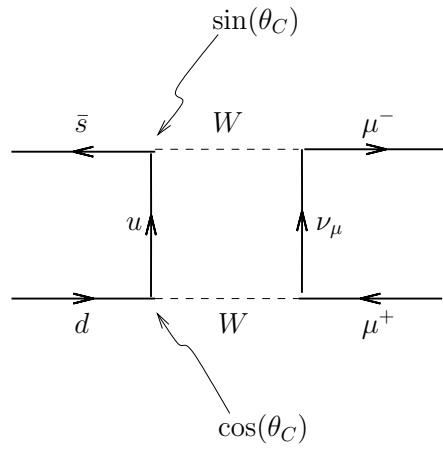


Figure 12.3: Neutral kaon decay by way of an up quark.

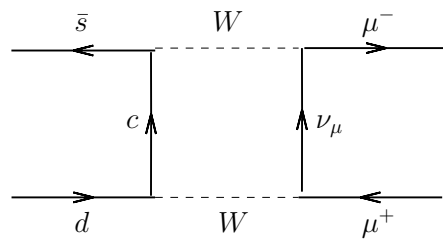


Figure 12.4: Neutral kaon decay by way of a charmed quark.

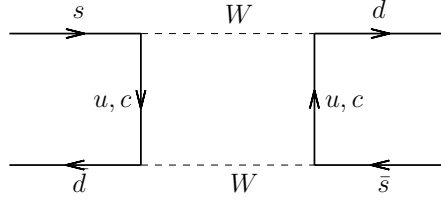


Figure 12.5: Neutral kaon oscillation. Note that these are all strong interactions, which we will talk about in somewhat more depth later.

noting that the last process does not happen as frequently. Recall also that $\Lambda = uds$. If we wait around long enough and check, we can find that a beam of what was initially \bar{K}^0 has become a beam of mostly K^0 ! We can see this by examining the processes

$$\bar{K}^0 + p \rightarrow K^+ + n$$

and

$$K^0 + p \rightarrow \Lambda + \pi^+.$$

12.3 CP Violation

Of course, without being in the presence of protons, kaons can decay by themselves. Such processes are

$$K_S^0 \rightarrow \pi^+\pi^- \text{ or } \pi^0\pi^0$$

and

$$K_L^0 \rightarrow \pi^0\pi^+\pi^-.$$

The “S” stands for “short” since the lifetime is $\tau_S \approx 0.9 \times 10^{-10}$ s compared to the “long” $\tau_L \approx 1.7 \times 10^{-7}$ s. It can be shown that

$$\begin{aligned} K_S &= \frac{1}{\sqrt{2}} (K^0 + \bar{K}^0) \\ K_L &= \frac{1}{\sqrt{2}} (K^0 - \bar{K}^0), \end{aligned}$$

and these are the eigenstates of the weak interaction. If we wait long enough, there should only be K_L decays left since all the K_S “components” would have decayed already.

It can be easily demonstrated that under CP

$$\begin{aligned} CP|K_S\rangle &= +1|K_S\rangle \\ CP|K_L\rangle &= -1|K_L\rangle, \end{aligned}$$

so that K_L and K_S are also eigenstates of CP . Recall from Chapter 5 that parity and charge conjugation symmetry alone are not always conserved, but (usually) together they are. Thus, with the $\pi^-\pi^+$ decay, $CP = +1$ and with $\pi^0\pi^+\pi^-$, $CP = -1$.

In 1964, Fitch and Cronin examined these kaon decays. In particular, they decided to see if K_L would decay into $\pi^+\pi^-$ (which should not happen if CP is a good symmetry). They found that with a fraction of $\epsilon = 2 \times 10^{-3}$, the decay *does* happen, which violates CP ! Actually, the conclusion that was drawn was that

$$\begin{aligned}
K_L &= \frac{1}{\sqrt{1+|\epsilon|^2}} (K_1 + \epsilon K_2) \\
K_S &= \frac{1}{\sqrt{1+|\epsilon|^2}} (K_1 - \epsilon K_2)
\end{aligned}$$

, where K_1 and K_2 are the CP eigenstates: $CP(K_1) = -1$, $CP(K_2) = +1$.

In order to explain CP violation, Kobayashi and Maskawa proposed a third quark generation and introduced the CKM³ matrix:

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \quad (12.4)$$

where $c_{ij} \equiv \cos(\theta_{ij})$, $s_{ij} \equiv \sin(\theta_{ij})$, and δ is the (complex) CP phase. Without the third quark generation, there could be no CP violation.

Furthermore, the mixing is required in the above matrix because the Lagrangian has no intergenerational interaction terms. Some researchers tried to invent a new interaction to explain what was going on, but in the end, the BaBar experiment was able to confirm that the mixing approach was the correct one.

While the CKM matrix demonstrates the origin of CP violation in the experimental realm, the size of the violation is not enough to explain the matter-antimatter asymmetry in the universe.

12.4 Mixing in the Lepton Sector

If neutrinos were massless, then there would be no mixing. Given that a lepton l_α^+ can be accompanied by *any* neutrino ν_i . The neutrino state emitted together with l_α^+ is

$$|\nu_\alpha\rangle = \sum_i U_{\alpha i}^* |\nu_i\rangle,$$

where $|\nu_i\rangle$ are the *mass eigenstates* of the neutrinos and U is the unitary *leptonic mixing matrix*. This means that each mass eigenstate is a superposition of the flavor (or weak) eigenstates. This leads to neutrino oscillations.

It can be shown that the $m\bar{\psi}\psi$ term in the Lagrangian then will become $m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L)$. If $\nu_i = \bar{\nu}_i$, i.e., the neutrino is *Majorana*, then $L \sim m_\mu\bar{\nu}_R$. Alternatively, there could just be fields that do not interact with anything, and so would be completely undetectable!

The Majorana case allows for two new CP -violating phases, which *could* explain the matter-antimatter asymmetry in the universe.

12.5 The Strong Interaction

The strong force couples to *color*, of which there are three: (R)ed, (G)reen, and (B)lue. We know there are three colors by looking at the interactions involving $e^+e^- \rightarrow q\bar{q}$ (Figure 12.6). The rate is proportional to the quark charges:

$$R \sim \sum e_i^2,$$

so, e.g., for the charmed quark,

$$R(\sqrt{s} < 3 \text{ GeV}) \sim \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{2}{3}\right)^2 = \frac{2}{3},$$

³The C is for Cabbibo.

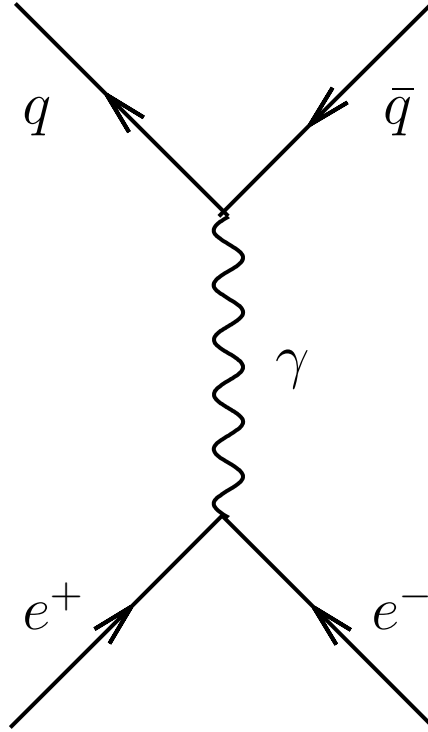


Figure 12.6: An interaction used to determine that there are three colors. Here, time moves to the right.

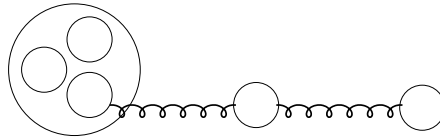


Figure 12.7: Quark jets.

and (ignoring the top quark),

$$R(\sqrt{s} < 10GeV) \sim \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{2}{3}\right)^2 = \frac{11}{9}.$$

Comparing these calculations to the data shows that they are off by exactly a factor of 3, implying the existence of the three colors⁴. All the particles we observe in nature are color-neutral (leptons intrinsically, and mesons and quarks by design).

12.5.1 Finding Quarks

To find quarks, we do what we did when probing the structure of the nucleus — we use diffraction techniques (with, e.g., $e^- + p \rightarrow e^- + p$), focusing particularly on the far away maxima, since these correspond to higher momentum transfer. In such cases, we can get jets of quarks (Figure 12.7).

⁴Actually, there are 6, when you count the anticolors.

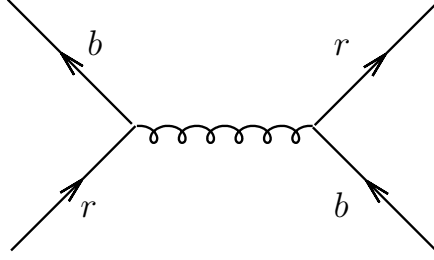


Figure 12.8: A strong interaction Feynman diagram. Each vertex carries a factor of α_S .

12.5.2 Quantum Chromodynamics

The strong force is mediated by gluons, which also carry color, and so can interact with each other. Gluons are massless and spinless, and come in 8 varieties: $r\bar{b}$, $r\bar{g}$, $g\bar{r}$, $g\bar{b}$, $b\bar{g}$, $1/\sqrt{2}(r\bar{r} - b\bar{b})$, and $1/\sqrt{6}(r\bar{r} - b\bar{b} - 2g\bar{g})$, hence *quantum chromodynamics* is $SU(3)$.

Quark interactions carry factors of the strong coupling constant α_S at each vertex in their Feynman diagrams (e.g., Figure ??).

12.5.3 Strong Potential

The strong potential is

$$V = \frac{4}{3} \frac{\alpha_S}{r} + kr. \quad (12.5)$$

Note that α_S goes down with energy, compared to the electromagnetic finestructure constant which goes up with higher energy. This weakness at higher energies is referred to as *asymptotic freedom*.

12.6 A Summary of the Standard Model

Now we have a more complete picture of the standard model, and in terms of group theory, it is $U(1) \times SU(2) \times SU(3)$. However, there are 26 (or so) parameters that are required to make predictions with the standard model:

- $\alpha_S, \alpha_W, \alpha_{EM}$
- CKM matrix elements with 3 mixing angles, 1 phase
- 12 fermion masses
- Masses of W and Z bosons
- $\sin(\theta_W)$
- MNS (neutrino mixing matrix) with 3 angles and 1 phase
- Higgs mass

In addition to all of these parameters, there are also “optional” parameters. If the neutrino turns out to be Majorana, then there are also 2 new phases.

Despite this large number of parameters, there has to date been no single, confirmed experiment that disagrees with the standard model.