This chapter introduces the fundamental forces by which elementary particles interact, and the Feynman diagrams we use to represent these interactions. The treatment is entirely qualitative and can be read quickly to get a sense of the “lay of the land.” The quantitative details will come in Chapters 6 through 10.

### 2.1 THE FOUR FORCES

As far as we know, there are just four fundamental forces in nature: strong, electromagnetic, weak, and gravitational. They are listed in the following table in order of decreasing strength:*

<table>
<thead>
<tr>
<th>Force</th>
<th>Strength</th>
<th>Theory</th>
<th>Mediator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong</td>
<td>10</td>
<td>Chromodynamics</td>
<td>Gluon</td>
</tr>
<tr>
<td>Electromagnetic</td>
<td>$10^{-2}$</td>
<td>Electrodynamic</td>
<td>Photon</td>
</tr>
<tr>
<td>Weak</td>
<td>$10^{-13}$</td>
<td>Flavordynamics</td>
<td>Wand Z</td>
</tr>
<tr>
<td>Gravitational</td>
<td>$10^{-42}$</td>
<td>Geometrodynamics</td>
<td>Graviton</td>
</tr>
</tbody>
</table>

To each of these forces there belongs a physical theory. The classical theory of gravity is, of course, Newton’s law of universal gravitation. Its relativistic generalization is Einstein’s general theory of relativity (“geometrodynamics” would be a better term). A completely satisfactory quantum theory of gravity has yet to be worked out; for the moment, most people assume that gravity is simply

* The “strength” of a force is an intrinsically ambiguous notion — after all, it depends on the nature of the source and on how far away you are. So the numbers in this table should not be taken too literally, and (especially in the case of the weak force) you will see quite different figures quoted elsewhere.
too weak to play a significant role in elementary particle physics. The physical theory that describes electromagnetic forces is called electrodynamics. It was given its classical formulation by Maxwell over one hundred years ago. Maxwell's theory was already consistent with special relativity (for which it was, in fact, the main inspiration). The quantum theory of electrodynamics was perfected by Tomonaga, Feynman, and Schwinger in the 1940s. The weak forces, which account for nuclear beta decay (and also, as we have seen, the decay of the pion, the muon, and many of the strange particles) were unknown to classical physics; their theoretical description was given a relativistic quantum formulation right from the start. The first theory of the weak forces was presented by Fermi in 1933; it was refined by Lee and Yang, Feynman and Gell-Mann, and many others, in the fifties, and put into its present form by Glashow, Weinberg, and Salam, in the sixties. For reasons that will appear in due course, the theory of weak interactions is sometimes called *flavordynamics*; in this book I refer to it simply as the Glashow–Weinberg–Salam (GWS) theory. (The GWS model treats weak and electromagnetic interactions as different manifestations of a single electroweak force, and in this sense the four forces reduce to three.) As for the strong forces, beyond the pioneering work of Yukawa in 1934 there really was no theory until the emergence of chromodynamics in the mid-seventies.

Each of these forces is mediated by the exchange of a particle. The gravitational mediator is called the graviton, electromagnetic forces are mediated by the photon, strong forces by the *gluon*, and weak forces by the intermediate vector bosons, $W$ and $Z$. These mediators transmit the force between one quark or lepton and another. In principle, the force of impact between a bat and a baseball is nothing but the combined interaction of the quarks and leptons in one with the quarks and leptons in the other. More to the point, the strong force between two protons, say, which Yukawa took to be a fundamental and irreducible process, must be regarded as a complicated interaction of six quarks. This is clearly not the place to look for simplicity. Rather, we must begin by analyzing the force between one truly elementary particle and another. In this chapter I will show you qualitatively how each of the relevant forces acts on individual quarks and leptons. Subsequent chapters develop the machinery needed to make the theory quantitative.

### 2.2 QUANTUM ELECTRODYNAMICS (QED)

Quantum electrodynamics is the oldest, the simplest, and the most successful of the dynamical theories; the others are self-consciously modeled on it. So I'll begin with a description of QED. *All* electromagnetic phenomena are ultimately reducible to the following elementary process:
This diagram reads: Charged particle e enters, emits (or absorbs) a photon, $\gamma$, and exits.* For the sake of argument, I’ll assume the charged particle is an electron; it could just as well be a quark, or any lepton except a neutrino (the latter is neutral, of course, and does not experience an electromagnetic force).

To describe more complicated processes, we simply patch together two or more replicas of this primitive vertex. Consider, for example, the following:

Here, two electrons enter, a photon passes between them (I need not say which one emits the photon and which one absorbs it; the diagram represents both orderings), and the two then exit. This diagram, then, describes the interaction between two electrons; in the classical theory we would call it the Coulomb repulsion of like charges (if the two are at rest). In QED this process is called Møller scattering; we say that the interaction is “mediated by the exchange of a photon,” for reasons that should now be apparent.

Now, you’re allowed to twist these “Feynman diagrams” around into any topological configuration you like—for example, we could stand the previous picture on its side:

The rule of the game is that a particle line running “backward in time” (as indicated by the arrow) is to be interpreted as the corresponding antiparticle going forward (the photon is its own antiparticle, that’s why I didn’t need an arrow on the photon line). So in this process an electron and a positron annihilate to form a photon, which in turn produces a new electron-positron pair. An electron and a positron went in, an electron and a positron came out (not the same ones, but then, since all electrons are identical, it hardly matters). This represents the interaction of two opposite charges: their Coulomb attraction. In QED this process is called Bhabha scattering. There is a quite different diagram which also contributes:

* In this book time always flows upward; the traditional convention. Particle physicists tend increasingly to let $t$ run horizontally (to the right), but there is no established consensus on the matter.
Both diagrams must be included in the analysis of Bhabha scattering.

Using just two vertices we can also construct the following diagrams, describing, respectively, pair annihilation, \( e^- + e^+ \rightarrow \gamma + \gamma \); pair production, \( \gamma + \gamma \rightarrow e^- + e^+ \); and Compton scattering, \( e^- + \gamma \rightarrow e^- + \gamma \):

[Notice that Bhabha and Maller scattering are related by crossing symmetry (see Section 1.4); as are the three processes shown here. In terms of Feynman diagrams, crossing symmetry corresponds to twisting or rotating the figure.] If we allow more vertices, the possibilities rapidly proliferate; for example, with four vertices we obtain, among others, the following diagrams:

In each of these figures two electrons went in and two electrons came out. They too describe the repulsion of like charges (Maller scattering). The “innards” of the diagram are irrelevant as far as the observed process is concerned. Internal lines (those which begin and end within the diagram) represent particles that are not observed—indeed, that cannot be observed without entirely changing the process. We call them “virtual” particles. Only the external lines (those which
2.2 QUANTUM ELECTRODYNAMICS (QED)

enter or leave the diagram) represent “real” (observable) particles. The external lines, then, tell you what physical process is occurring; the internal lines describe the mechanism involved.

Please understand: these Feynman diagrams are purely symbolic; they do not represent particle trajectories (as you might see them in, say, a bubble chamber photograph). The vertical dimension is time, and horizontal spacings do not correspond to physical separations. For instance, in Bhabha scattering the electron and positron are attracted, not repelled (as the diverging lines might seem to suggest). All the diagram says is: “Once there was an electron and a positron; they exchanged a photon; then there was an electron and a positron again.” Each Feynman diagram actually stands for a particular number, which can be calculated using the so-called Feynman rules (you’ll learn how to do this in Chapter 6). Suppose you want to analyze a certain physical process (say, Möller scattering). First you draw all the diagrams that have the appropriate external lines (the one with two vertices, all the ones with four vertices, and so on), then you evaluate the contribution of each diagram, using the Feynman rules, and add it all up. The sum total of all Feynman diagrams with the given external lines represents the actual physical process. Of course, there’s a problem here: there are infinitely many Feynman diagrams for any particular reaction! Fortunately, each vertex within a diagram introduces a factor of $\alpha = (e^2/\hbar c) = \alpha$, the fine structure constant. Because this is such a small number, diagrams with more and more vertices contribute less and less to the final result, and, depending on the accuracy you need, may be ignored. In fact, in QED it is rare to see a calculation that includes diagrams with more than four vertices. The answers are only approximate, to be sure, but when the approximation is valid to six significant digits, only the most fastidious are likely to complain.

The Feynman rules enforce conservation of energy and momentum at each vertex, and hence for the diagram as a whole. It follows that the primitive QED vertex by itself does not represent a possible physical process. We can draw the diagram, but calculation would assign to it the number zero. The reason is purely kinematical: $e^- e^- + y$ would violate conservation of energy. (In the center-of-mass frame the electron is initially at rest, so its energy is $mc^2$. It cannot decay into a photon plus a recoiling electron because the latter alone would require an energy greater than $mc^2$.) Nor, for instance, is $e^- + e^+ y$ kinematically possible, although it is easy enough to draw the diagram:

![Feynman Diagram](image)

In the center-of-mass system the electron and positron enter symmetrically with equal and opposite velocities, so the total momentum before the collision is obviously zero. But the final momentum cannot be zero, since photons always
travel at the speed of light; an electron-positron pair can annihilate to make two photons, but not one. Within a larger diagram, however, these figures are perfectly acceptable, because, although energy and momentum must be conserved at each vertex, a virtual particle does not carry the same mass as the corresponding free particle. In fact, a virtual particle can have any mass—whatever the conservation laws require.* In the business, we say that virtual particles do not lie on their mass shell. External lines, by contrast, represent real particles, and these do carry the “correct” mass.

[Actually, the physical distinction between real and virtual particles is not quite as sharp as I have implied. If a photon is emitted on Alpha Centauri, and absorbed in your eye, it is technically a virtual photon, I suppose. However, in general, the farther a virtual particle is from its mass shell the shorter it lives, so a photon from a distant star would have to be extremely close to its “correct” mass; it would have to be very close to “real.” As a calculational matter, you would get essentially the same answer if you treated the process as two separate events (emission of a real photon by star, followed by absorption of a real photon by eye). You might say that a real particle is a virtual particle which lasts long enough that we don’t care to inquire how it was produced, or how it is eventually absorbed.]

2.3 QUANTUM CHROMODYNAMICS (QCD)

In chromodynamics, color plays the role of charge, and the fundamental process (analogous to $e^- e^- + \gamma$) is quark $\rightarrow$ quark-plus-gluon (since leptons do not carry color, they do not participate in the strong interactions):

As before, we combine two or more such “primitive vertices” to represent more complicated processes. For example, the force between two quarks (which is responsible in the first instance for binding quarks together to make baryons, and indirectly for holding the neutrons and protons together to form a nucleus) is described in lowest order by the diagram:

* In special relativity, the energy $E$, momentum, $p$, and mass $m$ of a free particle are related by the equation $E^2 - p^2c^2 = mc^4$. But for a virtual particle $E^2 - p^2c^2$ can take on any value. Many authors interpret this to mean that virtual processes violate conservation of energy (see Problem 1.2). Personally, I consider this misleading, at best. Energy is always conserved.
We say that the force between two quarks is “mediated” by the exchange of gluons.

At this level chromodynamics is very similar to electrodynamics. However, there are also important differences, most conspicuously, the fact that whereas there is only one kind of electric charge (it can be positive or negative, to be sure, but a single number suffices to characterize the charge of a particle), there are three kinds of color (red, green, and blue). In the process $q \rightarrow q + g$, the color of the quark (but not its flavor) may change. For example, a blue up-quark may convert into a red up-quark. Since color (like charge) is always conserved, this means that the gluon must carry away the difference—in this instance, one unit of blueness and minus one unit of redness:

Gluons, then, are “bicolored,” carrying one positive unit of color and one negative unit. There are evidently $3 \times 3 = 9$ possibilities here, and you might expect there to be 9 kinds of gluons. For technical reasons, which we’ll come to in Chapter 9, there are actually only 8.

Since the gluons themselves carry color (unlike the photon, which is electrically neutral), they couple directly to other gluons, and hence in addition to the fundamental quark-gluon vertex, we also have primitive gluon-gluon vertices, in fact, two kinds: three gluon vertices and four gluon vertices:

This direct gluon–gluon coupling makes chromodynamics a lot more complicated than electrodynamics, but also far richer, allowing, for instance, the possibility of glueballs (bound states of interacting gluons, with no quarks on the scene at all).

Another difference between chromodynamics and electrodynamics is the size of the coupling constant. Remember that each vertex in QED introduces a factor of $\alpha = \frac{1}{137}$, and the smallness of this number means that we need only consider Feynman diagrams with a small number of vertices. Experimentally, the corresponding coupling constant for the strong forces, $\alpha_s$, as determined, say, from the force between two protons, is greater than 1, and the bigness of this number plagued particle physics for decades. For instead of contributing less and less, the more complex diagrams contribute more and more, and Feynman’s procedure, which worked so well in QED, is apparently worthless. One
of the great triumphs of quantum chromodynamics (QCD) was the discovery that in this theory the number that plays the role of coupling “constant” is in fact not constant at all, but depends on the separation distance between the interacting particles (we call it a “running” coupling constant). Although at the relatively large distances characteristic of nuclear physics it is big at very short distances (less than the size of a proton) it becomes quite small. This phenomenon is known as asymptotic freedom; it means that within a proton or a pion, say, the quarks rattle around without interacting much. Just such behavior was found experimentally in the deep inelastic scattering experiments. From a theoretical point of view, the discovery of asymptotic freedom rescued the Feynman calculus as a legitimate tool for QCD, in the high-energy regime.

Even in electrodynamics, the effective coupling depends somewhat on how far you are from the source. This can be understood qualitatively as follows. Picture first a positive point charge \( q \) embedded in a dielectric medium (i.e., a substance whose molecules become polarized in the presence of an electric field). The negative end of each molecular dipole will be attracted toward \( q \), and the positive end repelled away, as shown in Figure 2.1. As a result, the particle acquires a “halo” of negative charge, which partially cancels its field. In the
presence of the dielectric, then, the effective charge of any particle is somewhat reduced:

\[ q_{\text{eff}} = q / \varepsilon \]  

(2.1)

(The factor \( \varepsilon \) by which the field is reduced is called the dielectric constant of the material; it is a measure of the ease with which the substance can be polarized.*) Of course, if you are in closer than the nearest molecule, then there is no such screening, and you “see” the full charge \( q \). Thus if you were to make a graph of the effective charge, as a function of distance, it would look something like Figure 2.2. The effective charge increases at very small distances.

Now, it so happens that in quantum electrodynamics the vacuum itself behaves like a dielectric; it sprouts positron-electron pairs, as shown in Feynman diagrams such as these:

The virtual electron in each “bubble” is attracted toward \( q \), and the virtual positron is repelled away; the resulting vacuum polarization partially screens the charge and reduces its field. Once again, however, if you get too close to \( q \), the screening disappears. What plays the role of the “intermolecular spacing” in this case is the Compton wavelength of the electron, \( \lambda_c = h / mc = 2.43 \times 10^{-10} \text{ cm} \). For distances smaller than this the effective charge increases, just as it did in Figure 2.2. Notice that the unscreened (“close-up”) charge, which you might regard as the “true” charge of the particle, is not what we measure in any ordinary experiment, since we are seldom working at such minute separation distances. [An exception is the Lamb shift—a tiny perturbation in the spectrum of hydrogen—in which the influence of vacuum polarization (or rather, its absence at short distances) is clearly discernible.] What we have always called “the charge of the electron” is actually the fully screened effective charge.

So much for electrodynamics. The same thing happens in QCD, but with an important added ingredient. Not only do we have the quark-quark-gluon vertex (which, by itself, would again lead to an increasing coupling strength at short distances), but now there are also the direct gluon-gluon vertices. So in addition to the diagrams analogous to vacuum polarization in QED, we must now also include gluon loops, such as these:
It is not clear \textit{a priori} what influence these diagrams will have on the story\textsuperscript{3} as it turns out, their effect is the opposite: There occurs alund of competition between the quark polarization diagrams (which drive $\alpha_s$ up at short distances) and gluon polarization (which drives it down). Since the former depends on the number of quarks in the theory (hence on the number of flavors, $f$), whereas the latter depends on the number of gluons (hence on the number of colors, $n$), the winner in the competition depends on the relative number of flavors and colors. The critical parameter turns out to be

$$a = 2f - 11n$$

(2.2)

If this number is positive, then, as in QED, the effective coupling increases at short distances; if it is negative, the coupling decreases. In the Standard Model, $f = 6$ and $n = 3$, so $a = -21$, and the QCD coupling decreases at short distances. Qualitatively, this is the origin of asymptotic freedom.

The final distinction between QED and QCD is that whereas many particles carry electric charge, no naturally occurring particles carry color. Experimentally, it seems that quarks are confined in colorless packages of two (mesons) and three (baryons). As a consequence, the processes we actually observe in the laboratory are necessarily indirect and complicated manifestations of chromodynamics. It is as though our only access to electrodynamics came from the van der Waals forces between neutral molecules. For example, the (strong) force between two protons involves (among many others) the following diagram:
2.4 WEAK INTERACTIONS

There is no particular name for the “stuff” that produces weak forces, in the sense that electric charge produces electromagnetic forces and color produces strong forces. Some people call it “weak charge.” Whatever word you use, all quarks and all leptons carry it. (Leptons have no color, so they do not participate in the strong interactions; neutrinos have no charge, so they experience no electromagnetic forces; but all of them join in the weak interactions.) There are two kinds of weak interactions: charged (mediated by the W’s and neutral (mediated by the Z). The theory is cleaner for leptons than it is for quarks, so let us begin with the leptons.

2.4.1 Leptons

The fundamental charged vertex looks like this:

\[ \begin{align*}
\text{\textbullet$\bar{u}$ & \textbullet$u$ & \textbullet$d$ & \longrightarrow \\
\longrightarrow & \\
\text{\textbullet$u$ & \textbullet$\bar{d}$ & \textbullet$d$ & \longrightarrow} \end{align*} \]

\[ \begin{align*}
\text{\textbullet$\bar{u}$ & \textbullet$\bar{d}$ & \textbullet$d$ & \longrightarrow} \end{align*} \]

\[ \begin{align*}
\text{\textbullet$\bar{u}$ & \textbullet$\bar{d}$ & \textbullet$d$ & \longrightarrow} \end{align*} \]
A negative lepton (it could be $e^-$, $\mu^-$, or $\tau^-$) converts into the corresponding neutrino, with emission of a $W^-$ (or absorption of a $W^+$): $l^- \rightarrow \nu_l + W^-$. As always, we combine the primitive vertices together to generate more complicated reactions. For example, the process $\mu^- + \nu_e \rightarrow e^- + \nu_\mu$ would be represented by the diagram:

![Diagram](image)

Such a neutrino-muon scattering event would be hard to set up in the laboratory, but with a slight twist essentially the same diagram describes the decay of the muon, $\mu^- e^- + \nu_\mu + \bar{\nu}_e$, which happens all the time:

![Diagram](image)

(Technically, this is only the lowest-order contribution to muon decay, but in weak interaction theory one almost never needs to consider higher-order corrections.)

The fundamental neutral vertex is:

![Diagram](image)

In this case $l$ can be any lepton (including neutrinos). The $Z$ mediates such processes as neutrino-electron scattering ($\nu_\mu + e^- \rightarrow \nu_\mu + e^-$):

![Diagram](image)

Although charged weak processes were recognized from the start (beta decay itself is a charged process), the theoretical possibility of neutral weak processes was not appreciated until 1958. The Glashow–Weinberg–Salam (GWS) model

* This implies, of course that $l^- \rightarrow \bar{\nu}_l + W^+$ is also an allowed vertex.
includes neutral weak processes as essential ingredients, and their existence was confirmed experimentally at CERN in 1973. The reason it took so long for neutral weak processes to be discovered is twofold: (1) nobody was looking for them and (2) they tend to be masked by much stronger electromagnetic effects. For example, the Z can be exchanged between two electrons, but so can the photon:

Presumably there is a minute correction to Coulomb’s law that’s attributable to the first diagram, but the photon-mediated process overwhelmingly dominates. Experiments at DESY (in Hamburg) studied the reaction $e^- + e^+ \to \mu^- + \mu^+$ at very high energy and found unmistakable evidence of a contribution from the Z. But to observe a pure neutral weak interaction one has to go to neutrino scattering, in which there is no competing electromagnetic mechanism, and neutrino experiments are notoriously difficult.

### 2.4.2 Quarks

Notice that the leptonic weak vertices connect members of the same generation: $e^-$ converts to $\nu$, (with emission of $W^-$), or $\mu^- \to \mu^-$ (emitting a Z), but $e^-$ never goes to $\nu$ nor $\mu^-$ to $\nu$. In this way the theory enforces the conservation of electron number, muon number, and tau number. It is tempting to suppose that the same rule applies to the quarks, so that the fundamental charged vertex is:

A quark with charge $-\frac{1}{3}$ (which is to say: $d$, $s$, or $b$) converts into the corresponding quark with charge $+\frac{2}{3}$ ($u$, $c$, or $t$, respectively), with the emission of a $W^-$. The outgoing quark comes the same color as the ingoing one, but a different flavor. It’s not that the W comes off the “missing” flavor—after all, the W must be capable of coupling to leptons, which have no flavor; rather, flavor is simply not conserved in weak interactions. (Because quark flavor typically changes at a weak vertex, as quark color changes at a strong vertex, weak interaction theory is sometimes called “flavordynamics.”)

The far end of the Wline can couple to leptons (a “semileptonic” process), or to other quarks (a purely hadronic process). The most important semileptonic process is undoubtedly $d + \nu_e \to u + e$: 

\[ d + \nu_e \to u + e \]
Because of quark confinement, this process would never occur in nature as it stands. However, turned on its side, and with the $\bar{u}$ and $d$ bound together (by the strong force), this diagram represents a possible decay of the pion, $\pi^- \rightarrow e^- + \nu_e$:

(For reasons to be discussed later, the more common decay is actually $\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$, but the diagram is the same: with $e$ replaced by $\mu$.) Moreover, essentially the same diagram accounts for the beta decay of the neutron ($n \rightarrow p^+ + e^- + \bar{\nu}_e$):

Thus, apart from strong interaction contamination (in the form of the “spectator” $u$ and $d$ quarks), the decay of the neutron is identical in structure to the decay of the muon, and closely related to the decay of the pion. In the days before the quark model, these appeared to be three very different processes.
2.4 WEAK INTERACTIONS

Eliminating the electron-neutrino vertex in favor of a second quark vertex we obtain a purely hadronic weak interaction, $\Delta^0 \rightarrow p^+ + \pi^-:$

Actually, this particular decay also proceeds by the strong interaction:

The weak mechanism is an immeasurably small contribution. We'll see more realistic examples of nonleptonic weak interactions in a moment.

The fundamental neutral vertex for leptons $(l^- l + Z)$ leaves the lepton species unchanged; again, it is natural to suppose that the same applies to quarks:

* The $\Delta^0$ has the same quark content as the neutron, but this decay is not possible for neutrons because they are not heavy enough to make a proton and a pion.
This leads to neutrino-scattering processes such as $\nu_\mu + p \rightarrow \nu_\mu + p$:

$Z^0$ exchange also makes a tiny contribution to the electron-proton force within an atom. As before, this contribution is masked by the dominant electromagnetic force, but it is detectable in certain carefully chosen atomic transitions.

So far, it’s all pretty simple: The quarks mimic the leptons, as far as the weak interactions are concerned. The only difference is that because of the confining property of the strong force, there are generally spectator quarks present, which go along for the ride. Sad to say, this picture is a little too simple. For as long as the fundamental quark vertex is allowed to operate only within each generation, we can never hope to account for strangeness-changing weak interactions, such as the decay of the lambda $(\Lambda \rightarrow p^+ + \pi^-)$ or the omega-minus $(\Omega^- \rightarrow A + K^-)$, which involve the conversion of a strange quark into an up-quark:

The solution to this dilemma was suggested by Cabibbo in 1963, applied to neutral processes by Glashow, Illiopoulos, and Maiani (GIM) in 1970, and extended to three generations by Kobayashi and Maskawa (KM) in 1973.* The essential idea is that the quark generations are “skewed,” for the purposes of weak interactions. Instead of

* The Cabibbo/GIM/KM mechanism will be discussed more fully in Chapter 10.
2.4 WEAK INTERACTIONS

\[
\begin{pmatrix}
  u \\
  d
\end{pmatrix}, \quad
\begin{pmatrix}
  c \\
  s
\end{pmatrix}, \quad
\begin{pmatrix}
  t \\
  b
\end{pmatrix}
\]

(2.3)

the weak force couples the pairs

\[
\begin{pmatrix}
  u' \\
  d'
\end{pmatrix}, \quad
\begin{pmatrix}
  c' \\
  s'
\end{pmatrix}, \quad
\begin{pmatrix}
  t' \\
  b'
\end{pmatrix}
\]

(2.4)

where \(d', s',\) and \(b'\) are linear combinations of the physical quarks \(d, s,\) and \(b:\)

\[
\begin{pmatrix}
  d' \\
  s' \\
  b'
\end{pmatrix} = \begin{pmatrix}
  V_{ud} & V_{ts} & V_{tb} \\
  V_{cs} & V_{ud} & V_{cb} \\
  V_{sb} & V_{cs} & V_{ub}
\end{pmatrix}
\begin{pmatrix}
  d \\
  s \\
  b
\end{pmatrix}
\]

(2.5)

If this 3 X 3 Kobayashi–Maskawa matrix were the unit matrix, then \(d', s',\) and \(b'\) would be the same as \(d, s,\) and \(b,\) and no “cross-generational” transitions could occur. “Upness-plus-downness” would be absolutely conserved (just as the electron number is); “strangeness-plus-charm” would be conserved (like muon number); and so would “topness-plus-bottomness” (like tau number). But it’s not the unit matrix (although it’s pretty close); experimentally, the magnitudes of the matrix elements are:

\[
\begin{pmatrix}
  0.9705 \text{ to } 0.9770 & 0.21 \text{ to } 0.24 & 0. \text{ to } 0.014 \\
  0.21 \text{ to } 0.24 & 0.971 \text{ to } 0.973 & 0.036 \text{ to } 0.070 \\
  0. \text{ to } 0.024 & 0.036 \text{ to } 0.069 & 0.997 \text{ to } 0.999
\end{pmatrix}
\]

(2.6)

\(V_{ud}\) measures the coupling of \(u\) to \(d,\) \(V_{us}\) the coupling of \(u\) to \(s,\) and so on. The fact that the latter is nonzero is what permits strangeness-changing processes, such as the decay of the \(A\) and the \(\Omega^-\), to occur.

2.4.3 Weak and Electromagnetic Couplings of \(W\) and \(Z\)

There are also direct couplings of \(W\) and \(Z\) to one another, in GWS theory (just as there are direct gluon-gluon couplings in QCD):

Moreover, because the \(W\) is charged, it couples to the photon:
Although these interactions are critical for the internal consistency of the theory, as we shall see in Chapter 11, they are of limited practical importance at this point in time (see Problem 2.6).

2.5 DECAYS AND CONSERVATION LAWS

One of the most striking general properties of elementary particles is their tendency to disintegrate; we might almost call it a universal principle that every particle decays into lighter particles, unless prevented from doing so by some conservation law. The neutrinos and the photon are stable (having zero mass, there is nothing lighter for them to decay into); the electron is stable (it’s the lightest charged particle, so conservation of charge prevents its decay); and the proton is presumably stable (it’s the lightest baryon, and the conservation of baryon number saves it). By the same token, the positron and the antiproton are stable. But apart from these, all particles spontaneously disintegrate, even the neutron, although it becomes stable in the protective environment of many atomic nuclei. In practice, our world is populated mainly by protons, neutrons, electrons, photons, and neutrinos; more exotic things are created now and then (by collisions) but they do not last long. Each unstable species has a characteristic mean lifetime, for the muon it’s $2.2 \times 10^{-6}$ sec; for the $\pi^+$ it’s $2.6 \times 10^{-8}$ sec; for the $\pi^0$ it’s $8.3 \times 10^{-17}$ sec. Most particles exhibit several different decay modes; $64\%$ of all $K^+$s, for example, decay into $\mu^+ + \nu_\mu$, but $21\%$ go to $\pi^+ + \pi^0$, $6\%$ to $\pi^+ + \pi^+ + \pi^-$, $5\%$ to $(e^+ + \nu_e + \pi^0)$, and so on. One of the goals of elementary particle theory is to calculate these lifetimes and branching ratios.

A given decay is governed by one of the three fundamental forces; $\Delta^+ \rightarrow p^+ \pi^+$, for example, is a strong decay; $\pi^0 \rightarrow \gamma + \gamma$ is electromagnetic; and $\Sigma^- \rightarrow n + \pi^-$ is weak. How can we tell? Well, if a photon comes out, the process is certainly electromagnetic, and if a neutrino emerges, the process is certainly weak. But if neither a photon nor a neutrino is present, it’s a little harder to say. For example, $\Sigma^- \rightarrow n + \pi^-$ is weak, but $A^- \rightarrow n + \pi^-$ is strong. I’ll show you in a moment how to figure that out, but first I want to mention the most dramatic experimental difference between strong, electromagnetic, and weak decays: A typical strong decay involves a lifetime around $10^{-23}$ sec, a typical electromagnetic decay takes about $10^{-16}$ sec, and weak decay times range from around $10^{-13}$ sec (for the $\tau$) up to 15 min (for the neutron). For a given type of interaction, the decay generally proceeds more rapidly the larger the mass difference between the original particle and the decay products, just as a ball rolls faster down a steeper hill. There are exceptions: $\pi^+ \rightarrow \mu^+ + \nu_\mu$, for example, is faster by a factor of $10^4$ than $\pi^+ \rightarrow e^+ + \nu_e$, but such cases demand special explanations. It is this kinematic effect that accounts for the enormous range in weak interaction lifetimes. In particular, the proton and electron together are so

* The lifetime $\tau$ is related to the half-life $t_{1/2}$ by the formula $t_{1/2} = (\ln 2) \tau = 0.693 \tau$. The half-life is the time it takes for half the particles in a large sample to disintegrate (see Ch. 6, Sect. 6.1).
close to the neutron’s mass that the decay $n \rightarrow p^+ + e^- + \bar{\nu}_e$ barely makes it at all, and the lifetime of the neutron is greater by far than that of any other unstable particle. Experimentally, then, there is a vast separation in lifetime between strong and electromagnetic decays (a factor of about 10 million), and again between electromagnetic and weak decays (a factor of at least a thousand). Indeed, particle physicists are so used to thinking in terms of $10^{-23}$ sec as the “normal” unit of time that the handbooks generally classify anything with a lifetime greater than $10^{-19}$ sec or so as a “stable” particle.*

Now, what about the conservation laws which, as I say, permit certain reactions and forbid others? To begin with there are the purely kinematic conservation laws—conservation of energy and momentum (which we shall study in Chapter 3) and conservation of angular momentum (which comes in Chapter 4). The fact that a particle cannot spontaneously decay into particles heavier than itself is actually a consequence of conservation of energy (although it may seem so “obvious” as to require no explanation at all). The kinematic conservation laws apply to all interactions—strong, electromagnetic, weak, and for that matter anything else that may come along in the future—since they derive from special relativity itself. However, our concern right now is with the dynamical conservation laws that govern the three relevant interactions. Ten years ago I would simply have stated them as empirical rules coming from experiment, which you just have to memorize. It is in that spirit that we encountered them in Chapter 1. But now that we have a workable model for each of the basic forces, it becomes a question of examining the fundamental vertices:

Since all physical processes are obtained by sticking these together in elaborate combinations, anything that is conserved at each vertex must be conserved for the reactions as a whole. So, what do we have?

1. Conservation of charge: All three interactions, of course, conserve electric charge. In the case of the weak interactions the lepton (or quark) that comes out may not have the same charge as the one that went in, but if so, the difference is carried away by the $W$.

* Incidentally, $10^{-23}$ sec is about the time it takes a light signal to cross a hadron (diameter $= 10^{-14}$ m). You obviously cannot determine the lifetime of such a particle by measuring the length of its track [as we did for the $\Omega^-$ in Problem 1.8(b)]. Instead, you make a histogram of mass measurements, and invoke the uncertainty principle: $AE \Delta t = \hbar$. Here $AE = (\Delta m)c^2$, and $\Delta t = \tau$, so we get

$$\tau = \frac{\hbar}{(\Delta m)c^2}$$

Thus the spread in mass is a measure of the particle’s lifetime.
2. Conservation of color: The electromagnetic and weak interactions do not affect color. At a strong vertex the quark color does change, but the difference is carried off by the gluon. (The direct gluon-gluon couplings also conserve color.) However, since naturally occurring particles are always colorless, the observable manifestation of color conservation is pretty trivial: zero in, zero out.

3. Conservation of baryon number: In all the primitive vertices, if a quark goes in, a quark comes out, so the total number of quarks present is a constant. In this arithmetic we count antiquarks as negative, so that, for example, at the vertex \( q + \bar{q} \rightarrow g \) the quark number is zero before and zero after. Of course, we never see individual quarks, only baryons (with quark number 3), antibaryons (quark number \(-3\)), and mesons (quark number zero). So, in practice, it is more convenient to speak of the conservation of baryon number (\( A = 1 \) for baryons, \( A = -1 \) for antibaryons, and \( A = 0 \) for everything else). The baryon number is just \( \frac{1}{3} \) the quark number. Notice that there is no analogous conservation of meson number; since mesons carry zero quark number, a given collision or decay can produce as many mesons as it likes, consistent with conservation of energy.

4. Electron number, muon number, and tau number: The strong forces do not touch leptons at all; in an electromagnetic interaction the same particle comes out (accompanied by a photon) as went in; and the weak interactions only mix together leptons from the same generation. So, the electron number, muon number, and tau number are all conserved. If it weren’t for Cabibbo mixing, there would be a similar conservation of generation type for quarks (upness-plus-downness, strangeness-plus-charm, and beauty-plus-truth), but the fact that the generations are skewed in the weak interactions spoils things, and there is no hadronic analog to conservation of the individual lepton numbers.

5. Approximate conservation of flavor: So far, all the conservation laws we have considered are absolute, in the sense that they hold for all three interactions, as presently understood. An observed violation of any of them would be big news, calling for a major overhaul in our view of the subatomic world. But what about quark flavor? Flavor is conserved at a strong or electromagnetic vertex, but not at a weak vertex, where an up quark may turn into a down quark or a strange quark, with nothing at all picking up the lost upness or supplying the “gained” downness or strangeness. Because the weak forces are so weak, we say that the various flavors are approximately conserved. In fact, as you may remember, it was precisely this approximate conservation that led Gell-Mann to introduce the notion of strangeness in the first place. He “explained” the fact that strange particles are always produced in pairs:

\[
\pi^-(d\bar{u}) + p^+(uud) \rightarrow K^+(u\bar{s}) + \Sigma^-(dds)
\]  

(2.7)

for instance, but

\[
\pi^-(d\bar{u}) + p^+(uud) \nrightarrow \pi^+(ud\bar{d}) + \Sigma^-(dds)
\]  

(2.8)

by arguing that the latter violates conservation of strangeness. (Actually, this is a possible weak interaction, but it will never be seen in the laboratory, because it must compete against enormously more probable strong processes that do
2.5 DECAYS AND CONSERVATION LAWS

conserve strangeness.) In decays, however, the nonconservation of strangeness is very conspicuous, because for many particles this is the only way they can decay; there is no competition from strong or electromagnetic processes. The $\Lambda$, for instance, is the lightest strange baryon; if it is to decay, it must go to $n$ (or $p$) plus something. But the lightest strange meson is the $K$, and $n$ (or $p$) plus $K$ weighs substantially more than the $\Lambda$. If the $\Lambda$ decays at all (and it does, as we know: $\Lambda \rightarrow p^+ + \pi^-$ 64\% of the time, and $\Lambda \rightarrow n + \pi^0$ 36\% of the time), then strangeness cannot be conserved, and the reaction must proceed by the weak interaction. By contrast, the $\Delta^0$ (with a strangeness of zero) can go to $p^+ + \pi^-$ or $n + \pi$ by the strong interaction, and its lifetime is accordingly much shorter.

6. The OZI Rule: Finally, I must tell you about one very peculiar case that has been on my conscience since Chapter 1. I have in mind the decay of the psi, which, you will recall, is a bound state of the charmed quark and its antiquark: $\psi = c\bar{c}$. The $\psi$ has an anomalously long lifetime ($\sim10^{-20}$ sec); the question is, why? It has nothing to do with conservation of charm; the net charm of the psi is zero. The $\psi$ lifetime is short enough so that the decay is clearly due to the strong interactions. But why is it a thousand times slower than a strong decay “ought” to be? The explanation (if you call it that) goes back to an old observation by Okubo, Zweig, and Iizuka, known as the “OZI rule.” These authors were puzzled by the fact that the $\phi$ meson (whose quark content, $s\bar{s}$, makes it the strange analog to the $\psi$) decays much more often into two $K$'s than into three $\pi$'s (the two pion decay is forbidden for other reasons, which we will come to in Chapter 4), in spite of the fact that the three pion decay is energetically favored (the mass of two $K$'s is 990 MeV/c$^2$, three $\pi$'s weigh only 415 MeV/c$^2$). In Figure 2.4, we see that the three-pion diagram can be cut in two by snipping

![Figure 2.4](image-url)  

**Figure 2.4** The OZI rule: If the diagram can be cut in two by slicing only gluon lines (and not cutting open any external particles), the process is suppressed.
only gluon lines. The OZI rule states that such processes are “suppressed.” Not absolutely forbidden, mind you, for the decay $\phi \rightarrow 3\pi$ does in fact occur, but far less likely than one would otherwise have supposed. The OZI rule is related to asymptotic freedom, in the following sense: In an OZI-suppressed diagram the gluons must be “hard” (high energy), since they carry the energy necessary to make the hadrons into which they fragment. But asymptotic freedom says that gluons couple weakly at high energies (short ranges). By contrast, in OZI-allowed processes the gluons are typically “soft” (low energy), and in this regime the coupling is strong. Qualitatively, at least, this accounts for the OZI rule. (The quantitative details will have to await a more complete understanding of QCD.)

But what does all this have to do with the $\psi$? Well, presumably the same rule applies, suppressing $\psi \rightarrow 3\pi$, and leaving the decay into two charmed $D$ mesons (analogous to the $K$, but with the charmed quarks in place of the strange quarks) as the favored route. Only there’s a new twist in the $\psi$ system, for the $D$’s turn out to be too heavy: A pair of $D$’s weighs more than the $\psi$, so the decay $\psi \rightarrow D^+ + D^-$ (or $D^0 + \bar{D}^0$) is kinematically forbidden, while $\psi \rightarrow 3\pi$ is OZI suppressed, and it is to this happy combination that the $\psi$ owes its unusual longevity.

### 2.6 UNIFICATION SCHEMES

At one time electricity and magnetism were two distinct subjects, the one dealing with pith balls, batteries, and lightning; the other with lodestones, bar magnets, and the North Pole. But in 1820 Oersted noticed that an electric current could deflect a magnetic compass needle, and 10 years later Faraday discovered that a moving magnet could generate an electric current in a loop of wire. By the time Maxwell put the whole theory together in its final form, electricity and magnetism were properly regarded as two aspects of a single subject: electromagnetism.

Einstein dreamed of going a step further, combining gravity with electrodynamics in a single unified field theory. Although this program was not successful, a similar vision inspired Glashow, Weinberg, and Salam to join the weak and electromagnetic forces. Their theory starts out with four massless mediators, but, as it develops, three of them acquire mass (by the so-called Higgs mechanism), becoming the $W$’s and the $Z$, while one remains massless: the photon. Although experimentally a reaction mediated by $W$ or $Z$ is quite different from one mediated by the $\gamma$, if the GWS theory is right they are all manifestations of a single electroweak interaction. The relative weakness of the weak force is attributable to the enormous mass of the intermediate vector bosons; its intrinsic strength is in fact somewhat greater than that of the electromagnetic force, as we shall see in Chapter 10.

Beginning in the early seventies, many people have been working on the obvious next step: combining the strong force (in the form of chromodynamics) with the electroweak force (à la GWS). Several different schemes for implementing this grand unification are now on the table, and although it is too soon to draw
any definitive conclusions, some of the early results are promising. You will recall that the strong coupling constant $\alpha_s$, decreases at short distances (which is to say, for very high-energy collisions). So too does the weak coupling $\alpha_w$, but at a slower rate. Meanwhile, the electromagnetic coupling constant, $\alpha_e$, which is the smallest of the three, increases. Could it be that they all converge to a common limiting value, at extremely high energy? (See Fig. 2.5.) Such is the promise of the grand unified theories (GUTs). Indeed, from the functional form of the running coupling constants it is possible to estimate the energy at which this unification occurs: around $10^{15}\text{ GeV}$. This is, of course, astronomically higher than any currently accessible energy (remember, the mass of the $Z$ is $90\text{ GeV}/c^2$). Nevertheless, it is an exciting idea, for it means that the observed difference in strength among the three interactions is an “accident” resulting from the fact that we are obliged to work at low energies, where the unity of the forces is obscured. If we could just get in close enough to see the “true” strong, electric, and weak charges, without any of the screening effects of vacuum polarization, we would find that they are all equal. How nice!

Another suggestion of the GUTs is that the proton may be unstable, although its half-life is fantastically long (at least $10^{20}$ times the age of the universe). It has often been remarked that conservation of charge and color are in a sense more “fundamental” than the conservation of baryon number and lepton number, because charge is the “source” for electrodynamics, and color for chromodynamics. If these quantities were not conserved, QED and QCD would have to be completely reformulated. But baryon number and lepton number do not function as sources for any interaction, and their conservation has no deep dynamical significance. In the grand unified theories new interactions are contemplated, permitting decays such as

$$p^+ \rightarrow e^+ + \pi^0 \quad \text{or} \quad p^+ \rightarrow \bar{\nu}_u + \pi^+ \quad (2.9)$$

in which baryon number and lepton number change. Several major experiments are now under way to search for these proton decays. So far, the results are negative.

If grand unification works, all of elementary particle physics will be reduced to the action of a single force. The final step, then, will be to bring in gravity, vindicating at last Einstein’s dream. Indeed, many theorists are already working
on this, the ultimate unification. But it is probably safe to say that a detailed theory is still years off—after all, we hardly know how to carry out the most rudimentary calculations in chromodynamics, and here we are speculating about a theory two generations more sophisticated!

REFERENCES AND NOTES

1. Consistent etymology would call for geusidynamics, from the Greek word for “flavor”; see Phys. Today (April 1981), p. 74. M. Gaillard suggests asthenodynamics, from the Greek word for weak.

2. See, for example, E. M. Purcell, Electricity and Magnetism, 2d Ed. (New York: McGraw-Hill, 1985), Sec. 10.1.

3. C. Quigg, in Sci. Am. (April 1985) gives a qualitative interpretation of the effect of gluon polarization, but I do not find it entirely persuasive. Quigg’s article is an outstanding and accessible introduction to the current state of elementary particle physics.


PROBLEMS

2.1. Calculate the ratio of the gravitational attraction to the electrical repulsion between two stationary electrons. (Do I need to tell you how far apart they are?)

2.2. Sketch the lowest-order Feynman diagram representing Delbruck scattering: \( y + y \rightarrow y + y \). (This process, the scattering of light by light, has no analog in classical electrodynamics.)

2.3. Draw all the fourth-order (four vertex) diagrams for Compton scattering. (There are 17 of them; disconnected diagrams don’t count.)

2.4. Determine the mass of the virtual photon in each of the lowest-order diagrams for Bhabha scattering (assume the electron and positron are at rest). What is its velocity? (Note that these answers would be impossible for real photons.)

2.5. (a) Which decay do you think would be more likely, \( \Xi^- \rightarrow \Lambda + \pi^- \) or \( \Xi^- \rightarrow n + \pi^- \)? Explain your answer, and confirm it by looking up the experimental data.
(b) Which decay of the $D^0(c\bar{u})$ meson is more likely,

\[ D^0 \rightarrow K^- + \pi^+ , \quad D^0 \rightarrow \pi^- + \pi^+ , \quad \text{or} \quad D^0 \rightarrow K^+ + \pi \]

Which is least likely? Draw the Feynman diagrams, explain your answer and check the experimental data. (One of the successful predictions of the Cabibbo/GIM/KM model was that charmed mesons should decay preferentially into strange mesons, even though energetically the $2\pi$ mode is favored.)

(c) How about the “beautiful” ($B$) mesons? Should they go to the $D$’s, $K$’s, or $a$’s? How about “truthful” mesons?

2.6. Draw all the lowest-order diagrams contributing to the process $e^+ + e^- \rightarrow W^+ + \nu$. [One of them involves the direct coupling of $Z$ to $W$’s and another the coupling of $y$ to $W$’s, so if a positron-electron collider is ever built with sufficient energy to make two $W$’s, these interactions will be directly observable.]

2.7. Examine the following processes, and state for each one whether it is possible or impossible, according to the Standard Model (which does not include GUTs, with their potential violation of the conservation of lepton number and baryon number). In the former case, state which interaction is responsible — strong, electromagnetic, or weak; in the latter case cite a conservation law that prevents it from occurring. (Following the usual custom, I will not indicate the charge when it is unambiguous, while $y$, $\Lambda$, and $n$ are neutral; $p$ is positive, $e$ is negative; etc.)

(a) $p + \bar{p} \rightarrow \pi^+ + \pi^0$
(b) $\eta \rightarrow \gamma + \gamma$
(c) $\Sigma^0 \rightarrow \Lambda + \pi^0$
(d) $\Sigma^- \rightarrow n + a^-$
(e) $e^+ + e^- \rightarrow \mu^+ + p^-$
(f) $\mu^- \rightarrow e^- + \nu_e$
(g) $\Lambda \rightarrow p + \pi^0$
(h) $\tau^+ \rightarrow n + e^+$
(i) $e^+ + e^- \rightarrow \nu_e + \pi^0$
(j) $p + p \rightarrow \Sigma^+ + n + K^0 + \pi^+ + \pi^0$
(k) $p \rightarrow e^+ + \gamma$
(l) $p + p \rightarrow p + p + p + \bar{p}$
(m) $n + \bar{n} \rightarrow \pi^+ + \pi^- + \pi^0$
(n) $\pi^+ + n \rightarrow \pi^- + p$
(o) $K^- \rightarrow \pi^- + \pi^0$
(p) $\Sigma^+ + n \rightarrow \Sigma^- + p$
(q) $\Sigma^0 \rightarrow \Lambda + \gamma$
(r) $\Xi^- \rightarrow \Lambda + \pi^-
(s) $\Xi^0 \rightarrow p + \pi^-
(t) $\pi^- + p \rightarrow \Lambda + K^0$
(u) $\pi^0 \rightarrow \gamma + \gamma$
(v) $\Sigma^- \rightarrow n + e + \nu_e$

2.8. Some decays involve two (or even all three) different forces. Draw possible Feynman diagrams for the following processes:

(a) $K^+ \rightarrow \mu^+ + \nu_\mu + \gamma$
(b) $\Sigma^+ \rightarrow p + \gamma$

What interactions are involved? (Both these decays have been observed, by the way.)

2.9. The upsilon meson, $b\bar{b}$, is the bottom-quark analog to the $\psi$, $c\bar{c}$. Its mass is 9460 MeV/c$^2$, and its lifetime is $1.5 \times 10^{-20}$ sec. From this information, what can you say about the mass of the $B$ meson, $u\bar{b}$? (The observed mass is 5270 MeV/c$^2$.)

2.10. The $\psi'$ meson, at 3685 MeV/c$^2$, has the same quark content as the $\psi$ (i.e., $c\bar{c}$). Its principal decay mode is $\psi' \rightarrow \psi + \pi^+ + \pi^-$. Is this a strong interaction? Is it OZI-suppressed? What lifetime would you expect for the $\psi'$? (The observed value is $3 \times 10^{-21}$ sec.)