Chapter 1

Historical Introduction to the Elementary Particles

This chapter is a kind of "folk history" of elementary particle physics. Its purpose is to provide a sense of how the various particles were first discovered, and how they fit into the overall scheme of things. Along the way some of the fundamental ideas that dominate elementary particle theory are explained. This material should be read quickly, as background to the rest of the book. (As history, the picture presented here is certainly misleading, for it sticks closely to the main track, ignoring the false starts and blind alleys that accompany the development of any science. That's why I call it "folk" history—it's the way particle physicists like to remember the subject—a succession of brilliant insights and heroic triumphs unmarred by foolish mistakes, confusion, and frustration. It wasn't really quite so easy.)

1.1 THE CLASSICAL ERA (1897–1932)

It is always a little artificial to pinpoint such things, but I'd say that elementary particle physics was born in 1897, with J. J. Thomson's discovery of the electron.[1] (It is fashionable to carry the story all the way back to Democritus and the Greek atomists, but apart from a few suggestive words their metaphysical speculations have nothing in common with modern science, and although they may be of modest antiquarian interest, their relevance is infinitesimal.) Thomson knew that *cathode rays* emitted by a hot filament could be deflected by a magnet. This suggested that they carried electric charge; in fact, the direction of the curvature required that the charge be negative. It seemed, therefore, that these were not rays at all, but rather streams of particles. By passing the beam through crossed electric and magnetic fields, and adjusting the field strength until the net deflection was zero, Thomson was able to determine the velocity of the particles (about a tenth the speed of light) as well as their charge-to-mass ratio. (See Fig. 1.1 and Problem 1.1). This ratio turned out to be enormously greater than for any known ion, indicating that either the charge was extremely large or the mass was very small. Indirect evidence pointed to the second conclusion. Thomson called the particles *corpuscles,* and their charge the *electron.* Later the word electron was applied to the particles themselves.

Thomson correctly surmised that these electrons were essential constituents of atoms; however, since atoms as a whole are electrically neutral and very much heavier than electrons, there immediately arose the problem of how the compensating plus charge—and the bulk of the mass—is distributed within an atom. Thomson himself imagined that the electrons were suspended in a heavy, positively charged paste, like (as he put it) the plums in a pudding. But Thomson's model was decisively repudiated by Rutherford's famous scattering experiment, which showed that the positive charge, and most of the mass, was concentrated in a tiny core, or *nucleus,* at the center of the atom. Rutherford demonstrated this by firing a beam of α-particles (ionized helium atoms) into a thin sheet of

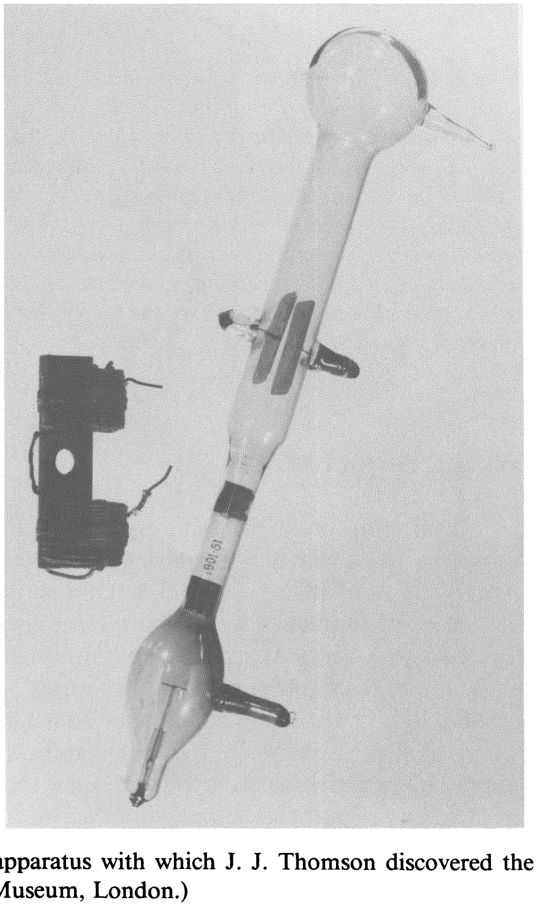


Figure 1.1 The apparatus with which J. J. Thomson discovered the electron. (Photo courtesy Science Museum, London.)

gold foil (see Fig. 1.2). Had the gold atoms consisted of rather diffuse spheres, as Thomson supposed, then all of the α-particles should have been deflected a bit, but none would have been deflected much—any more than a bullet is deflected much when it passes, say, through a bag of sawdust. What in *fact* occurred was that *most* of the α-particles passed through the gold completely undisturbed, but a few of them bounced off at wild angles. Rutherford's conclusion was that the α-particles had encountered something very small, very hard, and very heavy. Evidently the positive charge, and virtually all of the mass, was concentrated at the center, occupying only a tiny fraction of the volume of the atom (the electrons are too light to play any role in the sattering; they are knocked right out of the way by the much heavier α-particles).

The nucleus of the lightest atom (hydrogen) was given the name *proton* by Rutherford. In 1914 Niels Bohr proposed a model for hydrogen consisting of a single electron circling the proton, rather like a planet going around the sun, held in orbit by the mutual attraction of opposite charges. Using a primitive version of the quantum theory, Bohr was able to calculate the spectrum of hydrogen, and the agreement with experiment was nothing short of spectacular. It was natural then to suppose that the nuclei of heavier atoms were composed of two or more protons bound together, supporting a like number of orbiting electrons. Unfortunately, the next heavier atom (helium), although it does indeed carry two electrons, weighs *four* times as much as hydrogen, and lithium (three electrons) is *seven* times the weight of hydrogen, and so it goes. This dilemma

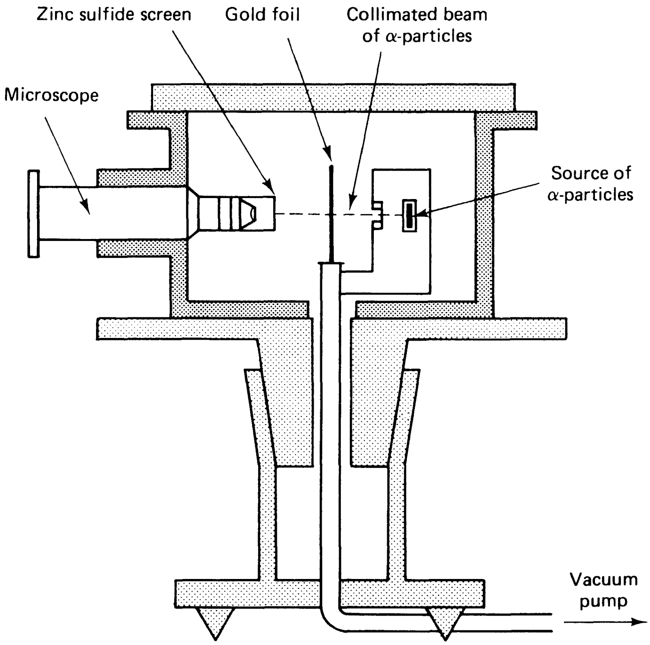


Figure 1.2 Schematic diagram of the apparatus used in the Rutherford scattering experiment. Alpha particles scattered by the gold foil strike a fluorescent screen, giving off a flash of light, which is observed visually through a microscope.

was finally resolved in 1932 with Chadwick's discovery of the neutron—an electrically neutral twin to the proton. The helium nucleus, it turns out, contains two neutrons in addition to the two protons; lithium evidently includes four; and in general the heavier nuclei carry very roughly the same number of neutrons as protons. (The number of neutrons is in fact somewhat flexible: the same atom, chemically speaking, may come in several different *isotopes,* all with the same number of protons, but with varying numbers of neutrons.)

The discovery of the neutron put the final touch on what we might call the *classical period* in elementary particle physics. Never before (and I'm sorry to say never since) has physics offered so simple and satisfying an answer to the question, "What is matter made of?" In 1932 it was all just protons, neutrons, and electrons. But already the seeds were planted for the three great ideas that were to dominate the *middle period* (1930–1960) in particle physics: Yukawa's meson, Dirac's positron, and Pauli's neutrino. Before we come to that, however, I must back up for a moment to introduce the photon.

1.2 THE PHOTON (1900–1924)

In some respects the photon is a very "modern" particle, having more in common with the W and Z (which were not discovered until 1983) than with the classical trio. Moreover, it's hard to say exactly when or by whom the photon was really "discovered," although the essential stages in the process are clear enough. The first contribution was made by Planck in 1900. Planck was attempting to explain the so-called *blackbody spectrum* for the electromagnetic radiation emitted by a hot object. Statistical mechanics, which had proved brilliantly successful in explaining other thermal processes, yielded nonsensical results when applied to electromagnetic fields. In particular, it led to the famous "ultraviolet catastrophe," predicting that the total power radiated should be *infinite.* Planck found that he could escape the ultraviolet catastrophe—and fit the experimental curve—if he assumed that electromagnetic radiation is *quantized,* coming in little "packages" of energy

$$E = h\nu \tag{1.1}$$

where ν is the frequency of the radiation and h is a constant, which Planck adjusted to fit the data. The modern value of Planck's constant is

$$h = 6.626 \times 10^{-27} \text{ erg s} \tag{1.2}$$

Planck did not profess to know *why* the radiation was quantized; he assumed that it was due to a peculiarity in the emission process: For some reason a hot surface only gives off light* in little squirts.

Einstein, in 1905, put forward a far more radical view. He argued that quantization was a feature of the electromagnetic field itself, having nothing to

* In this book the word *light* stands for *electromagnetic radiation,* whether or not it happens to fall in the visible region.

do with the emission mechanism. With this new twist, Einstein adapted Planck's idea, and his formula, to explain the *photoelectric effect:* When electromagnetic radiation strikes a metal surface, electrons come popping out. Einstein suggested that an incoming light quantum hits an electron in the metal, giving up its energy ($h\nu$); the excited electron then breaks through the metal surface, losing in the process an energy w (the so-called *work function* of the material—an empirical constant that depends on the particular metal involved). The electron thus emerges with an energy

$$E \leq h\nu - w \tag{1.3}$$

(It may lose some energy before reaching the surface. That's the reason for using $\leq$, instead of $=$.) Einstein's formula (1.3) is pretty trivial to *derive,* but it carries an extraordinary implication: The maximum electron energy is *independent of the intensity of the light* and depends only on its *color* (frequency). To be sure, a more intense beam will knock out *more* electrons, but their *energies* will be the same.

Unlike Planck's theory, Einstein's theory met a hostile reception, and over the next 20 years he was to wage a lonely battle for the *light quantum.*[2] In saying that electromagnetic radiation is *by its nature* quantized, regardless of the emission mechanism, Einstein came dangerously close to resurrecting the discredited particle theory of light. Newton, of course, had introduced such a *corpuscular* model, but a major achievement of nineteenth-century physics was the decisive repudiation of Newton's idea in favor of the rival wave theory. No one was prepared to see that accomplishment called into question, even when the experiments came down on Einstein's side. In 1916 Millikan completed an exhaustive study of the photoelectric effect and was obliged to report that "Einstein's photoelectric equation . . . appears in every case to predict exactly the observed results. . . . Yet the semicorpuscular theory by which Einstein arrived at his equation seems at present wholly untenable."[3]

What finally settled the issue was an experiment conducted by A. H. Compton in 1923. Compton found that the light scattered from a particle at rest is shifted in wavelength, according the equation

$$\lambda' = \lambda + \lambda_c(1 - \cos\theta) \tag{1.4}$$

where λ is the incident wavelength, λ' is the scattered wavelength, θ is the scattering angle, and

$$\lambda_c = h/mc \tag{1.5}$$

is the so-called *Compton wavelength* of the target particle (mass m). Now, this is *precisely* the formula you get (Problem 3.24) if you treat light as a particle of zero rest mass with energy given by Planck's equation, and apply the laws of conservation of (relativistic) energy and momentum—just as you would for an ordinary elastic collision (Fig. 1.3). That clinched it; here was direct and incontrovertible experimental evidence that light behaves as a particle, on the subatomic scale. We call this particle the *photon* (a name suggested by the chemist Gilbert Lewis, in 1926); the symbol for a photon is γ (from *gamma ray*). How

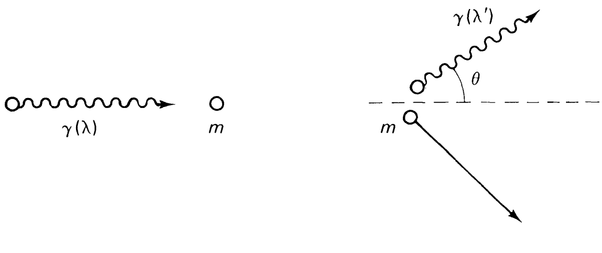


Figure 1.3 Compton scattering. A photon of wavelength λ scatters off a particle, initially at rest, of mass m. The scattered photon carries wavelength λ' given by equation (1.4).

the particle nature of light on this level is to be reconciled with its well-established wave behavior on the macroscopic scale (exhibited in the phenomena of interference and diffraction) is a story I'll leave to the quantum texts.

Although the photon initially *forced* itself on an unreceptive community of physicists, it eventually found a natural place in quantum field theory, and was to offer a whole new perspective on electromagnetic interactions. In classical electrodynamics, we attribute the electrical repulsion of two electrons, say, to the electric field surrounding them; each electron contributes to the field, and each one responds to the field. But in quantum field theory, the electric field is *quantized* (in the form of photons), and we may picture the interaction as consisting of a stream of photons passing back and forth between the two charges, each electron continually emitting them and continually absorbing them. And the same goes for *any* noncontact force: where classically we interpret "action at a distance" as "mediated" by a *field,* we now say that it is mediated by an *exchange of particles* (the *quanta* of the field). In the case of electrodynamics, the mediator is the photon; for gravity, it is called the *graviton* (though a fully successful quantum theory of gravity has yet to be developed and it may well be centuries before anyone detects a graviton experimentally).

You will see later on how these ideas are implemented in practice, but for now I want to dispel one common misapprehension. When I say that every force is mediated by the exchange of particles, I am *not* speaking of a merely *kinematic* phenomenon. Two ice skaters throwing snowballs back and forth will of course move apart with the succession of recoils; they "repel one another by exchange of snowballs," if you like. But that's *not* what is involved here. For one thing, this mechanism would have a hard time accounting for an *attractive* force. You might think of the mediating particles, rather, as "messengers," and the message can just as well be "come a little closer" as "go away."

I said earlier that in the "classical" picture ordinary matter is made of atoms, in which electrons are held in orbit around a nucleus of protons and neutrons by the electrical attraction of opposite charges. We can now give this model a more sophisticated formulation by attributing the binding force to the exchange of photons between the electrons and the protons in the nucleus. However, for the purposes of atomic physics this is overkill, for in this context quantization of the electromagnetic field produces only minute effects (notably the

Lamb shift and the anomalous magnetic moment of the electron). To excellent approximation we can pretend that the forces are given by Coulomb's law (together with various magnetic dipole couplings). The point is that in a bound state enormous numbers of photons are continually streaming back and forth, so that the "lumpiness" of the field is effectively smoothed out, and classical electrodynamics is a suitable approximation to the truth. But in *most* elementary particle processes, such as the photoelectric effect or Compton scattering, *individual* photons are involved, and quantization can no longer be ignored.

1.3 MESONS (1934–1947)

Now there is one conspicuous problem to which the "classical" model does not address itself at all: What holds the *nucleus* together? After all, the positively charged protons should repel one another violently, packed together as they are in such close proximity. Evidently there must be some other force, more powerful than the force of electrical repulsion, that binds the protons (and neutrons) together; physicists of that less imaginative age called it, simply, the *strong force.* But if there exists such a potent force in nature, why don't we notice it in everyday life? The *fact* is that virtually every force we experience directly, from the contraction of a muscle to the explosion of dynamite is electromagnetic in origin; the only exception, outside a nuclear reactor or an atomic bomb, is gravity. The answer must be that, powerful though it is, the strong force is of very short *range.* (The range of a force is like the arm's reach of a boxer—beyond that distance its influence falls off rapidly to zero. Gravitational and electromagnetic forces have *infinite* range, but the range of the strong force is about the size of the nucleus itself.)*

The first significant theory of the strong force was proposed by Yukawa in 1934. Yukawa assumed that the proton and neutron are attracted to one another by some sort of *field,* just as the electron is attracted to the nucleus by an electric field and the moon to the earth by a gravitational field. This field should properly be quantized, and Yukawa asked the question: What must be the properties of its *quantum*—the particle (analogous to the photon) whose exchange would account for the known features of the strong force? For example, the short range of the force indicated that the mediator would be rather heavy; Yukawa calculated that its mass should be nearly 300 times that of the electron, or about a sixth the mass of a proton. (See Problem 1.2.) Because it fell between the electron and the proton, Yukawa's particle came to be known as the *meson* (meaning "middle-weight"). [In the same spirit the electron is called a *lepton* ("light-weight"), whereas the proton and neutron are *baryons* ("heavy-weight").] Now, Yukawa knew that no such particle had ever been observed in the laboratory, and he therefore assumed his theory was wrong. But at the time a number of systematic studies

* This is a bit of an oversimplification. Typically, the forces go like $e^{-(r/a)}/r^2$, where a is the "range." For Coulomb's law and Newton's law of universal gravitation, $a = \infty$; for the strong force a is about 10^{-13} cm (one fermi).

of cosmic rays were in progress, and by 1937 two separate groups (Anderson and Neddermeyer on the West Coast, and Street and Stevenson on the East) had identified particles matching Yukawa's description. Indeed, the cosmic rays with which you are being bombarded every few seconds as you read this consist primarily of just such middle-weight particles.

For a while everything seemed to be in order. But as more detailed studies of the cosmic ray particles were undertaken, disturbing discrepancies began to appear. They had the wrong lifetime and they seemed to be significantly lighter than Yukawa had predicted; worse still, different mass measurements were not consistent with one another. In 1946 (after a period in which physicists were engaged in a less savory business) decisive experiments were carried out in Rome demonstrating that the cosmic ray particles interacted very weakly with atomic nuclei.[4] If this was really Yukawa's meson, the transmitter of the strong force, the interaction should have been dramatic. The puzzle was finally resolved in 1947, when Powell and his co-workers at Bristol[5] discovered that there are actually *two* middle-weight particles in cosmic rays, which they called π (or "pion") and μ (or "muon"). (Marshak reached the same conclusion simultaneously, on theoretical grounds.[6]) The true Yukawa meson is the π; it is produced copiously in the upper atmosphere, but ordinarily disintegrates long before reaching the ground. (See Problem 3.4.) Powell's group exposed their photographic emulsions on mountain tops (see Fig. 1.4). One of the decay products is the lighter (and longer-lived) μ, and it is primarily muons that one observes at sea level. In the search for Yukawa's meson, then, the muon was simply an imposter, having nothing whatever to do with the strong interactions. In fact, it behaves in every way like a heavier version of the electron and properly belongs in the *lepton* family (though some people to this day call it the "mu-meson" by force of habit).

1.4 ANTIPARTICLES (1930–1956)

Nonrelativistic quantum mechanics was completed in the astonishingly brief period 1923–1926, but the relativistic version proved to be a much thornier problem. The first major achievement was Dirac's discovery, in 1927, of the equation that bears his name. The Dirac equation was supposed to describe free electrons with energy given by the relativistic formula $E^2 - \mathbf{p}^2c^2 = m^2c^4$. But it had a very troubling feature: For every positive-energy solution ($E = +\sqrt{\mathbf{p}^2c^2 + m^2c^4}$) it admitted a corresponding solution with *negative* energy ($E = -\sqrt{\mathbf{p}^2c^2 + m^2c^4}$). This meant, given the natural tendency of every system to evolve in the direction of lower energy, that the electron should "runaway" to increasingly negative states, radiating off an infinite amount of energy in the process. To rescue his equation, Dirac proposed a resolution that made up in brilliance for what it lacked in plausibility: He postulated that the negative energy states are all filled by an infinite "sea" of electrons. Because this sea is always there, and perfectly uniform, it exerts no net force on anything, and we are not normally aware of it. Dirac then invoked the Pauli exclusion principle (which says that no two electrons can occupy the same state), to "explain" why the

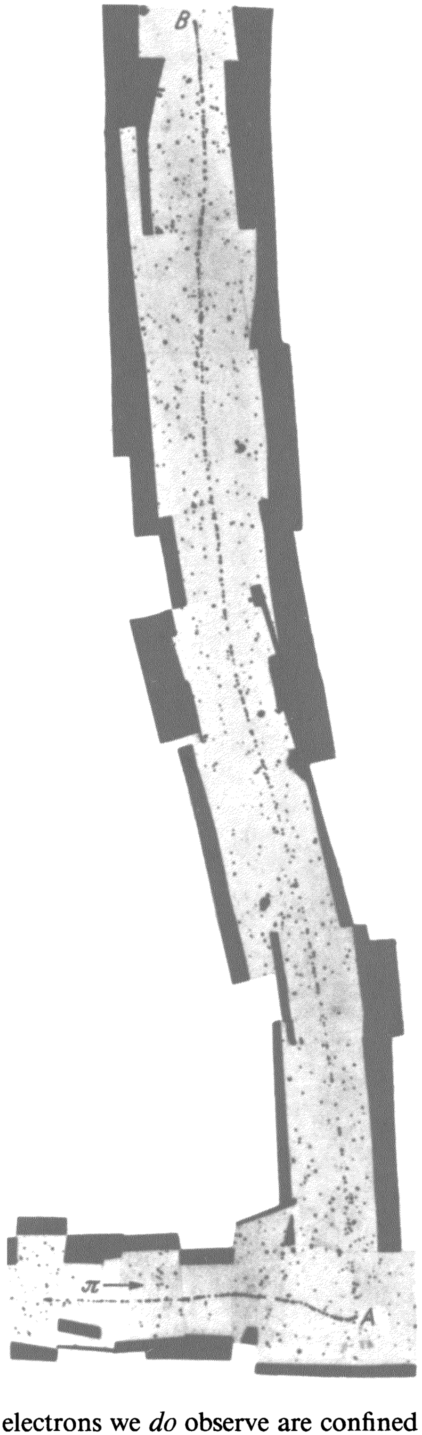


Figure 1.4 One of Powell's earliest pictures showing the track of a pion in a photographic emulsion exposed to cosmic rays at high altitude. The pion (entering from the left) decays into a muon and a neutrino (the latter is electrically neutral, and leaves no track). Reprinted by permission from C. F. Powell, P. H. Fowler, and D. H. Perkins, *The Study of Elementary Particles by the Photographic Method* (New York: Pergamon, 1959). First published in *Nature* **159,** 694 (1947).

electrons we *do* observe are confined to the positive energy states. But if this is true, then what happens when we impart to one of the electrons in the "sea" an energy sufficient to knock it into a positive energy state? The *absence* of the "expected" electron in the sea would be interpreted as a net positive charge in that location, and the absence of its expected negative energy would be seen as a net positive energy. Thus a "hole in the sea" would function as an ordinary particle with *positive* energy and *positive* charge. Dirac at first hoped that these holes might be *protons,* but it was soon apparent that they had to carry the same mass as the electron itself—2000 times too light to be a proton. No such particle was known at the time, and Dirac's theory appeared to be in trouble. What may have seemed a fatal defect in 1930, however, turned into a spectacular triumph in late 1931, with Anderson's discovery of the *positron* (Fig. 1.5), a positively charged twin for the electron, with precisely the attributes Dirac required.[7]

Figure 1.5 The positron. In 1932, Anderson took this photograph of the track left in a cloud chamber by a cosmic ray particle. The chamber was placed in a magnetic field (pointing into the page) which caused the particle to travel in a curve. But was it a negative charge traveling downward, or a positive charge traveling upward? In order to tell, Anderson had placed a lead plate across the center of the chamber (the thick horizontal line in the photograph). A particle passing through the plate slows down, and subsequently moves in a tighter circle. By inspection of the curves, it is clear that this particle traveled upward, and hence must have been positively charged. From the curvature of the track, and from its texture, Anderson was able to show that the mass of the particle was close to that of the electron. (Photo courtesy California Institute of Technology)

Still, many physicists were uncomfortable with the notion that we are awash in an infinite sea of invisible electrons, and in the forties Stuckelberg and Feynman provided a much simpler and more compelling interpretation of the negative-energy states. In the Feynman–Stuckelberg formulation the negative-energy solutions are reexpressed as *positive*-energy states of a *different particle* (the positron); the electron and positron appear on an equal footing, and there is no need for Dirac's "electron sea" or for its mysterious "holes." We'll see in Chapter 7 how this—the modern interpretation—works. Meantime, it turned out that the dualism in Dirac's equation is a profound and universal feature of quantum field theory: For *every* kind of particle there must exist a corresponding *antiparticle,* with the same mass but opposite electric charge. The positron, then, is the *antielectron.* (Actually, it is in principle completely arbitrary which one you call the "particle" and which the "antiparticle"—I could just as well have said that the electron is the antipositron. But since there are a lot of electrons around, and not so many positrons, we tend to think of electrons as "matter" and positrons as "antimatter"). The (negatively charged) antiproton was first observed experimentally at the Berkeley Bevatron in 1955, and the (neutral) antineutron was discovered at the same facility the following year.[8]

The standard notation for antiparticles is an overbar. For example, p denotes the proton and $\bar{p}$ the antiproton; n the neutron and $\bar{n}$ the antineutron. However, in some cases it is more customary simply to specify the charge. Thus most people write e^+ for the positron (not $\bar{e}$) and μ^+ for the antimuon (not $\bar{\mu}$). [But you must not *mix* conventions: $\bar{e}^+$ is ambiguous, like a double negative—the reader doesn't know if you mean the positron or the *anti*positron, (which is to say, the electron).] Some neutral particles are their *own* antiparticles. For example, the photon: $\bar{\gamma} = \gamma$. In fact, you may have been wondering how the antineutron differs physically from the neutron, since both are uncharged. The answer is that neutrons carry other "quantum numbers" besides charge (in particular, baryon number), which change sign for the antiparticle. Moreover, although its *net* charge is zero, the neutron *does* have a charge *structure* (positive at the center and at the edges, negative in between) and a magnetic dipole moment. These, too, have the opposite sign for $\bar{n}$.

There is a general principle in particle physics that goes under the name of *crossing symmetry.* Suppose that a reaction of the form

$$A + B \longrightarrow C + D$$

is known to occur. Any of these particles can be "crossed" over to the other side of the equation, provided it is turned into its antiparticle, and the resulting interaction will also be allowed. For example,

$$\begin{aligned} A &\longrightarrow \bar{B} + C + D \\ A + \bar{C} &\longrightarrow \bar{B} + D \\ \bar{C} + \bar{D} &\longrightarrow \bar{A} + \bar{B} \end{aligned}$$

In addition, the *reverse* reaction occurs $C + D \longrightarrow A + B$, but technically this derives from the principle of *detailed balance,* rather than from crossing symmetry. Indeed, as we shall see, the *calculations* involved in these various reactions are practically identical. We might almost regard them as different manifestations of the same fundamental process. Now, there is one important *caveat* in this: Conservation of energy may veto a reaction that is otherwise permissible. For example, if A weighs less than the sum of B, C, and D, then the decay $A \longrightarrow \bar{B} + C + D$ cannot occur; similarly, if A and C are light, whereas B and D are heavy, then the reaction $A + \bar{C} \longrightarrow \bar{B} + D$ will not take place unless the initial kinetic energy exceeds a certain "threshold" value. So perhaps I should say that the crossed (or reversed) reaction is *dynamically* permissible, but it may or may not be *kinematically* allowed. The power and beauty of crossing symmetry can scarcely be exaggerated. It tells us, for instance, that Compton scattering

$$\gamma + e^- \longrightarrow \gamma + e^-$$

is "really" the same process as pair annihilation

$$e^- + e^+ \longrightarrow \gamma + \gamma$$

although in the laboratory they are completely different phenomena.

The union of special relativity and quantum mechanics, then, leads to a pleasing matter/antimatter symmetry. But this raises a disturbing question: How come *our* world is populated with protons, neutrons, and electrons, instead of antiprotons, antineutrons, and positrons? Matter and antimatter cannot coexist for long—if a particle meets its antiparticle, they annihilate. So maybe it's just a historical accident that in our corner of the universe there happened to be more matter than antimatter, and pair annihilation has eliminated all but a leftover residue of matter. If this is so, then presumably there are other regions of space in which antimatter predominates. Unfortunately, the astronomical evidence is pretty compelling that all of the observable universe is made of ordinary matter. Recently, Wilczek and others have put forward a possible explanation for this cosmic asymmetry. I shall not go into it here, but if you are interested, I recommend Wilczek's article in *Scientific American* (December 1980).

1.5 NEUTRINOS (1930–1962)

For the third strand in the story we return again to the year 1930.[9] A problem had arisen in the study of nuclear beta decay. In beta decay a radioactive nucleus A is transformed into a slightly lighter nucleus B, with the emission of an electron:

$$A \longrightarrow B + e^- \tag{1.6}$$

Conservation of charge requires that B carry one more unit of positive charge than A. [We now realize that the underlying process here is the conversion of a neutron (in A) into a proton (in B), but remember that in 1930 the neutron had not yet been discovered.] Thus the "daughter" nucleus (B) lies one position farther along on the Periodic Table. There are many examples of beta decay: Potassium goes to calcium (${}^{40}_{19}\mathrm{K} \longrightarrow {}^{40}_{20}\mathrm{Ca}$), copper goes to zinc (${}^{64}_{29}\mathrm{Cu} \longrightarrow {}^{64}_{30}\mathrm{Zn}$), tritium goes to helium (${}^{3}_{1}\mathrm{H} \longrightarrow {}^{3}_{2}\mathrm{He}$), and so on. [The upper number is the *atomic*

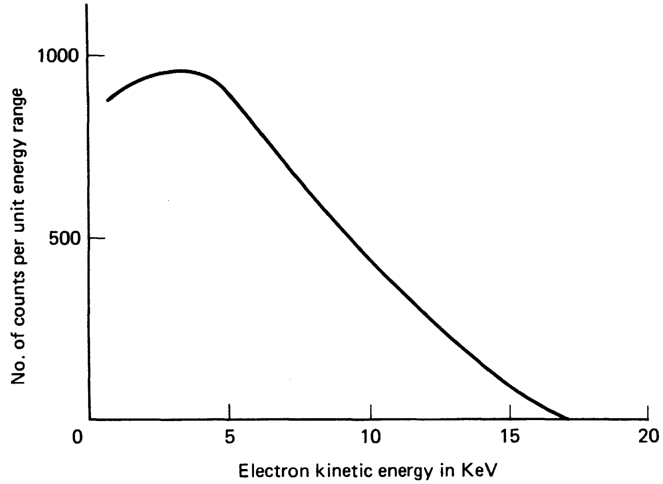


Figure 1.6 The beta decay spectrum of tritium ($^3_1H \longrightarrow {^3_2He}$). (*Source:* G. M. Lewis, *Neutrinos* (London: Wykeham, 1970), p. 30.)

weight (the number of neutrons plus protons) and the lower number is the *atomic number* (the number of protons).]

Now, it is a characteristic of two-body decays such as expression (1.6) that the outgoing energies are kinematically determined, in the center-of-mass frame. Specifically, if the "parent" nucleus (A) is at rest, so that B and e come out back-to-back with equal and opposite momenta, then conservation of energy dictates that the electron energy is

$$E = \left(\frac{m_A^2 - m_B^2 + m_e^2}{2m_A}\right)c^2 \tag{1.7}$$

The derivation of this result will be explained in Chapter 3; for now, the point to notice is that E is *fixed,* once the three masses are specified. But when the experiments are done it is found that the emitted electrons vary considerably in energy. Equation (1.7) only determines the *maximum* electron energy, for a particular beta-decay process (see Fig. 1.6).

This was a most disturbing result. Niels Bohr (not for the first time) was ready to abandon the law of conservation of energy.* Fortunately, Pauli took a more sober view, suggesting that another particle was emitted along with the electron, a silent accomplice that carries off the "missing" energy. It had to be electrically neutral, to conserve charge (and also, of course, to explain why it left no track); Pauli proposed to call it the *neutron.* The whole idea was greeted with some skepticism, and in 1932 Chadwick preempted the name. But in the following year Fermi presented a theory of beta decay that incorporated Pauli's

* It is interesting to note that Bohr was an outspoken critic of Einstein's light quantum (prior to 1924), that he discouraged Dirac's work on the relativistic electron theory (telling him, incorrectly, that Klein and Gordon had already succeeded), that he opposed Pauli's introduction of the neutrino, that he ridiculed Yukawa's theory of the meson, and that he disparaged Feynman's approach to quantum electrodynamics.

particle and proved so brilliantly successful that Pauli's suggestion had to be taken seriously. From the fact that the observed electron energies range up to the value given in equation (1.7) it follows that the new particle is extremely light; as far as we know, its mass is in fact *zero.* Fermi called it the *neutrino.* (For reasons you'll see in a moment, we now call it the *anti*neutrino.) In modern terminology, then, the fundamental beta-decay process is

$$n \rightarrow p^+ + e^- + \bar{\nu} \tag{1.8}$$

(neutron goes to proton plus electron plus antineutrino).

Now, you may have noticed something peculiar about Powell's picture of the disintegrating pion (Fig. 1.4): The muon emerges at about 90° with respect to the original pion direction. (That's not the result of a *collision,* by the way; collisions with atoms in the emulsion account for the dither in the tracks, but they cannot produce an abrupt left turn.) What that kink indicates is that some *other* particle was produced in the decay of the pion, a particle that left no footprints in the emulsion, and hence must have been electrically neutral. It was natural (or at any rate *economical*) to suppose that this was again Pauli's neutrino:

$$\pi \rightarrow \mu + \nu \tag{1.9}$$

A few months after their first paper, Powell's group published an even more striking picture, in which the subsequent decay of the muon is also visible (Fig. 1.7). Now, muon decays had been studied for many years, and it was well established that the charged secondary is an electron. From the figure there is clearly a neutral product as well, and you might guess that it is again a neutrino. However, this time it is *two* neutrinos:

$$\mu \rightarrow e + 2\nu \tag{1.10}$$

How do we know there are *two* of them? Same way as before: We repeat the experiment over and over, each time measuring the energy of the electron. If it always comes out the same, we know there are just two particles in the final state. But if it *varies,* then there must be (at least) three. By 1949* it was clear that the electron energy in muon decay is *not* fixed, and the emission of two neutrinos was the accepted explanation. By contrast, the *muon* energy in *pion* decay is perfectly constant, within experimental uncertainties, confirming that this is a genuine two-body decay.

By 1950, then, there was compelling *theoretical* evidence for the existence of neutrinos, but there was still no direct *experimental* verification. A skeptic might have argued that the neutrino was nothing but a bookkeeping device—a purely hypothetical particle whose only function was to rescue the conservation laws. It left no tracks, it didn't decay; in fact, no one had ever seen a neutrino *do anything.* The reason for this is that neutrinos interact extraordinarily weakly

* Here, and in the original beta-decay problem, conservation of angular momentum also requires a third outgoing particle, quite independently of energy conservation. But the spin assignments were not so clear in the early days, and for most people energy conservation was the compelling argument. In the interests of simplicity, I will keep angular momentum out of the story until Chapter 4.

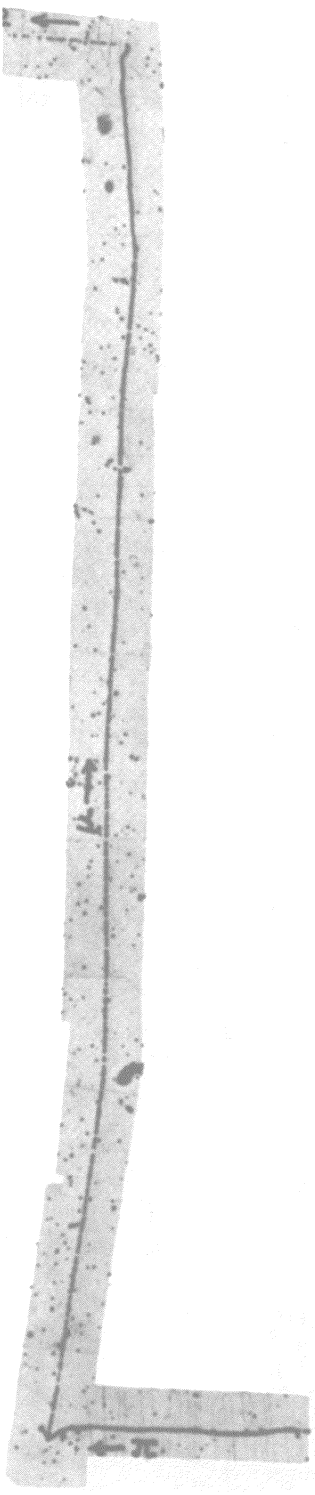


Figure 1.7 Here, a pion decays into a muon (plus a neutrino); the muon subsequently decays into an electron (and two neutrinos). Reprinted by permission from C. F. Powell, P. H. Fowler, and D. H. Perkins, *The Study of Elementary Particles by the Photographic Method* (New York: Pergamon, 1959). First published in *Nature* **163,** 82 (1949).

with matter; a neutrino of moderate energy could easily penetrate a thousand light-years(!) of lead.* To have a chance of detecting one you need an extremely intense source. The decisive experiments were conducted at the Savannah River nuclear reactor in South Carolina, in the mid-fifties. Here Cowan and Reines set up a large tank of water and watched for the "inverse" beta-decay reaction

$$\bar{\nu} + p^+ \rightarrow n + e^+ \tag{1.11}$$

At their detector the antineutrino flux was calculated to be 5×10^{13} particles per square centimeter per second, but even at this fantastic intensity they could only hope for two or three events every hour. On the other hand, they developed an ingenious method for identifying the outgoing positron. Their results provided unambiguous confirmation of the neutrino's existence.[10]

As I mentioned earlier, the particle produced in ordinary beta decay is actually an antineutrino, not a neutrino. Of course, since they're electrically neutral, you might ask—and many people *did*—whether there is any distinction between a neutrino and an antineutrino. The neutral pion, as we shall see, is its *own* antiparticle; so too is the photon. On the hand, the antineutron is definitely *not* the same as a neutron. So we're left in a bit of a quandary: *Is* the neutrino the same as the antineutrino, and if not, what property distinguishes them? In the late fifties, Davis and Harmer put this question to an experimental test.[11] From the positive results of Cowan and Reines, we know that the crossed reaction

$$\nu + n \rightarrow p^+ + e^- \tag{1.12}$$

must also occur, and at about the same rate. Davis looked for the analogous reaction using *anti*neutrinos:

$$\bar{\nu} + n \rightarrow p^+ + e^- \tag{1.13}$$

He found that this reaction does *not* occur, and thus established that the neutrino and antineutrino are distinct particles.

Davis's result was not unexpected. In fact, back in 1953 Konopinski and Mahmoud[12] had introduced a beautifully simple rule for determining which reactions [such as (1.12)] will work, and which [like (1.13)] will not. In effect,† they assigned a *lepton number* $L = +1$ to the electron, the muon, and the neutrino, and $L = -1$ to the positron, the positive muon, and the antineutrino (all other particles are given a lepton number of zero). They then proposed the law of conservation of lepton number (analogous to the law of conservation of charge): In any physical process, the sum of the lepton numbers before must equal the sum of the lepton numbers after. Thus the Cowan–Reines reaction (1.11) is allowed ($L = -1$ before and after), but the Davis reaction (1.13) is forbidden (on the left $L = -1$, on the right $L = +1$). [It was in anticipation of this rule that I called the beta-decay particle, in expression (1.8), an *anti*neutrino.] In

* That's a comforting realization when you learn that hundreds of billions of neutrinos pass through every square inch of your body per second, night and day, coming from the sun (they hit you from below, at night, having passed right through the earth).

† Konopinski and Mahmoud (ref. 12) did not use this terminology, and they got the muon assignments wrong. But never mind, the essential idea was there.

view of the conservation of lepton number, the charged pion decays (1.9) should actually be written

$$\begin{aligned} \pi^- &\longrightarrow \mu^- + \bar{\nu} \\ \pi^+ &\longrightarrow \mu^+ + \nu \end{aligned} \tag{1.14}$$

and the muon decays (1.10) are really

$$\begin{aligned} \mu^- &\longrightarrow e^- + \nu + \bar{\nu} \\ \mu^+ &\longrightarrow e^+ + \nu + \bar{\nu} \end{aligned} \tag{1.15}$$

What property distinguishes the neutrino from the antineutrino, then? The cleanest answer is: *lepton number*—it's $+1$ for the neutrino and -1 for the antineutrino. These numbers are experimentally determinable, just as electric charge is, by watching how the particle in question interacts with others. (As we shall see, they also differ in their *helicity:* the neutrino is "left-handed" whereas the antineutrino is "right-handed." But this is a technical matter best saved for later.)

There is a final twist to the neutrino story. Experimentally, the decay of a muon into an electron plus a *photon* is never observed:

$$\mu^- \not\rightarrow e^- + \gamma \tag{1.16}$$

and yet this process is consistent with conservation of charge and conservation of the lepton number. Now, there's a very reliable rule of thumb in particle physics (generally attributed to Richard Feynman) which says that whatever is not expressly *forbidden* is *mandatory.* The absence of $\mu \longrightarrow e + \gamma$ suggests a law of conservation of "mu-ness"; but then how are we to explain the observed decays $\mu \longrightarrow e + \nu + \bar{\nu}$? The answer occurred to a number of people in the late fifties and early sixties:[13] Suppose there are two different *kinds* of neutrino—one associated with the electron (ν_e) and one with the muon (ν_μ). If we assign a *muon number* $L_\mu = +1$ to μ^- and ν_μ, and $L_\mu = -1$ to μ^+ and $\bar{\nu}_\mu$, and at the same time an *electron number* $L_e = +1$ to e^- and ν_e, and $L_e = -1$ to e^+ and $\bar{\nu}_e$, and refine the conservation of lepton number into two separate laws—conservation of electron number and conservation of muon number—we can then account for *all* the allowed and forbidden processes. Neutron beta decay becomes

$$n \longrightarrow p^+ + e^- + \bar{\nu}_e \tag{1.17}$$

the pion decays are

$$\begin{aligned} \pi^- &\longrightarrow \mu^- + \bar{\nu}_\mu \\ \pi^+ &\longrightarrow \mu^+ + \nu_\mu \end{aligned} \tag{1.18}$$

and the muon decays take the form

$$\begin{aligned} \mu^- &\longrightarrow e^- + \bar{\nu}_e + \nu_\mu \\ \mu^+ &\longrightarrow e^+ + \nu_e + \bar{\nu}_\mu \end{aligned} \tag{1.19}$$

I said earlier that when pion decay was first analyzed it was "natural" and "economical" to assume that the outgoing neutral particle was the same as in beta decay, and that's quite true: It *was* natural, and it *was* economical, but it was *wrong*.

The first experimental test of the two-neutrino hypothesis (and the separate conservation of electron and muon number) was conducted at Brookhaven in 1962.[14] Using about 10^{14} antineutrinos from π^- decay, Lederman, Schwartz, Steinberger, and their collaborators identified 29 instances of the expected reaction

$$\bar{\nu}_\mu + p^+ \longrightarrow \mu^+ + n \tag{1.20}$$

and no cases of the forbidden process

$$\bar{\nu}_\mu + p^+ \longrightarrow e^+ + n \tag{1.21}$$

With only one kind of neutrino the second reaction would be just as common as the first. (Incidentally, this experiment presented truly monumental shielding problems. Steel from a dismantled warship was stacked up 44 feet thick, to make sure that nothing except neutrinos got through to the target.)

By 1962, then, the lepton family had grown to eight: the electron, the muon, their respective neutrinos, and the corresponding antiparticles (Table 1.1). The leptons are characterized by the fact that they do not participate in strong interactions. For the next 14 years things were pretty quiet, as far as the leptons go, so this is a good place to pause and let the strongly interacting particles—the mesons and baryons, known collectively as the *hadrons*—catch up.

1.6 STRANGE PARTICLES (1947–1960)

For a brief period in 1947 it was possible to believe that the major problems of elementary particle physics were solved. After a lengthy detour in pursuit of the muon, Yukawa's meson (the π) had finally been apprehended. Dirac's positron had been found, and Pauli's neutrino, although still at large (and, as we have

TABLE 1.1 THE LEPTON FAMILY, 1962–1976

	Lepton number	Electron number	Muon number
Leptons			
e^-	1	1	0
ν_e	1	1	0
μ^-	1	0	1
ν_μ	1	0	1
Antileptons			
e^+	−1	−1	0
$\bar{\nu}_e$	−1	−1	0
μ^+	−1	0	−1
$\bar{\nu}_\mu$	−1	0	−1

seen, still capable of making mischief), was basically under control. The role of the muon was something of a puzzle ("Who ordered *that?*" Rabi asked); it seemed quite unnecessary in the overall scheme of things. On the whole, however, it looked in 1947 as though the job of elementary particle physics was essentially done.

But this comfortable state did not last long. In December of that year Rochester and Butler[15] published the cloud chamber photograph shown in Figure 1.8. Cosmic ray particles enter from the upper left and strike a lead plate, producing a neutral particle, whose presence is revealed when it decays into two charged secondaries, forming the upside-down "V" in the lower right. Detailed analysis shows that these charged particles are in fact a π^+ and a π^-. Here, then, was a new neutral particle with at least twice the mass of the pion; we call it the K^0 ("kaon"):

$$K^0 \rightarrow \pi^+ + \pi^- \tag{1.22}$$

In 1949, Powell published the photograph reproduced in Figure 1.9, showing the decay of a charged kaon:

$$K^+ \rightarrow \pi^+ + \pi^+ + \pi^- \tag{1.23}$$

(The K^0 was first known as the V^0 and later as the θ^0; the K^+ was originally called the τ^+. Their identification as neutral and charged versions of the same basic particle was not completely settled until 1956—but that's another story, to which we shall return in Chapter 4.) The kaons behave in some respects like heavy pions, and so the *meson* family was extended to include them. In due course, many more mesons were discovered—the η, the ϕ, the ω, the ρ's, and so on.

Meanwhile, in 1950 another neutral "V" particle was found, this time by Anderson's group at Cal Tech. The photographs were similar to Rochester's (Fig. 1.8), but this time the products were a p^+ and a π^-. Evidently this particle is substantially heavier than the proton; we call it the Λ:

$$\Lambda \rightarrow p^+ + \pi^- \tag{1.24}$$

The lambda belongs with the proton and the neutron in the *baryon* family. To appreciate this, we must go back for a moment to 1938. The question had arisen, "Why is the proton stable?" Why, for example, doesn't it decay into a positron and a photon:

$$p^+ \rightarrow e^+ + \gamma \tag{1.25}$$

Needless to say, it would be unpleasant for us if this reaction were common (all atoms would disintegrate), and yet it does not violate any law known in 1938. (Actually, this particular process does violate conservation of lepton number, but that law was not recognized, remember, until 1953.) Stückelberg[16] proposed to account for the stability of the proton by asserting a law of conservation of baryon number: Assign to all baryons (which in 1938 meant the proton and the neutron) a "baryon number" $A = +1$, and to the antibaryons ($\bar{p}$ and $\bar{n}$)

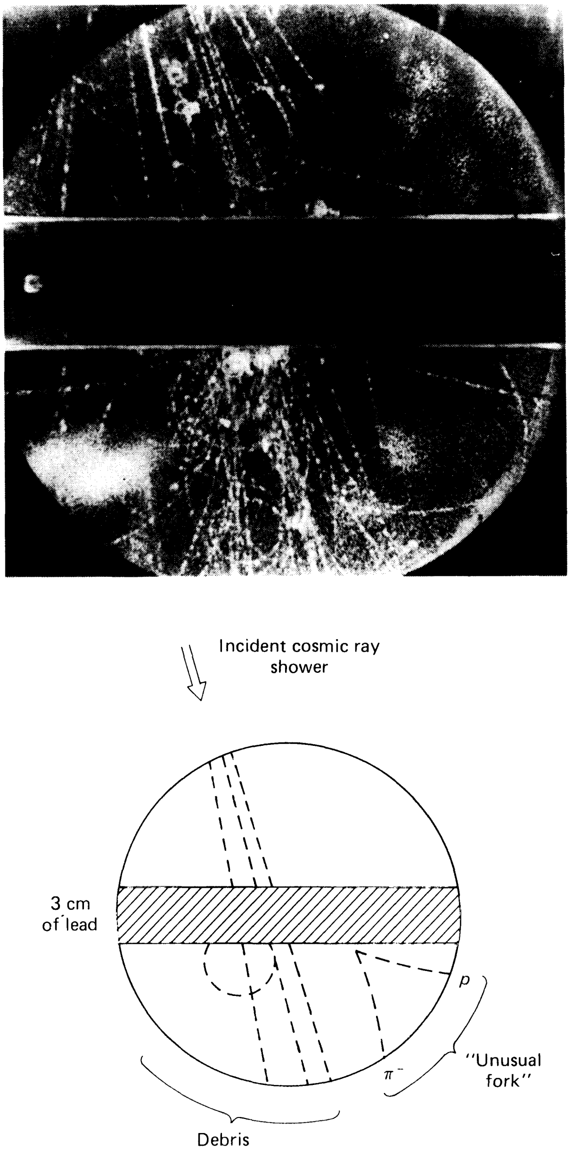


Figure 1.8 The first strange particle. Cosmic rays strike a lead plate, producing a K^0, which subsequently decays into a pair of charged pions. (Photo courtesy of Prof. G. D. Rochester. Reprinted by permission from *Nature* **160,** 855. Copyright © 1947, Macmillan Journals Limited.)

$A = -1$; then the total baryon number is conserved in any physical process. Thus, neutron beta decay ($n \longrightarrow p^+ + e^- + \bar{\nu}_e$) is allowed ($A = 1$ before and after), and so also is the reaction in which the antiproton was first observed:

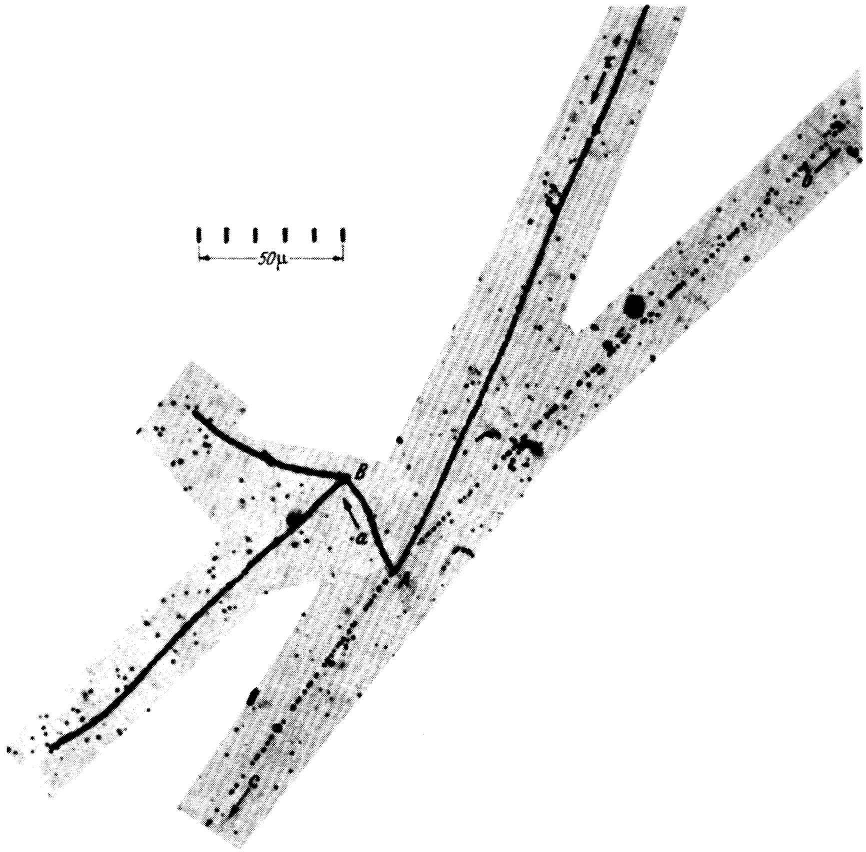


Figure 1.9 K^+, entering from above, decays at A: $K^+ \rightarrow \pi^+ + \pi^+ + \pi^-$. (The π^- subsequently causes a nuclear disintegration at B). [Reprinted by permission from C. F. Powell, P. H. Fowler, and D. H. Perkins, *The Study of Elementary Particles by the Photographic Method* (New York: Pergamon, 1959). First published in *Rep. Prog. Phys.* **13,** 384 (1950).]

$$p + p \rightarrow p + p + p + \bar{p} \tag{1.26}$$

($A = 2$ on both sides). But the proton, as the lightest baryon, has nowhere to go; conservation of the baryon number guarantees its absolute stability.* If we are to retain the conservation of baryon number in the light of reaction (1.24), the lambda must be assigned to the baryon family. Over the next few years many more heavy baryons were discovered—the Σ's, the Ξ's, and the Δ's, and so on. [By the way: unlike leptons and baryons, there is *no* conservation of mesons. In pion decay ($\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$) a meson disappears, and in lambda decay ($\Lambda \rightarrow p^+ + \pi^-$) a meson is created.]

* Recent "grand unified" theories allow for a minute violation of baryon number conservation, and in these theories the proton is *not* absolutely stable. See the article by S. Weinberg in *Scientific American,* June 1981. The experimental situation is discussed by J. M. LoSecco *et al., Scientific American,* June 1985.

It is some measure of the surprise with which these new heavy baryons and mesons were greeted that they came to be known collectively as "strange" particles. In 1952 the first of the modern particle accelerators (the Brookhaven Cosmotron) began operating, and soon it was possible to produce strange particles in the laboratory (before this the only source had been cosmic rays) . . . and with this, the rate of proliferation increased. Willis Lamb began his Nobel Prize acceptance speech in 1955 with the words

> When the Nobel Prizes were first awarded in 1901, physicists knew something of just two objects which are now called "elementary particles": the electron and the proton. A deluge of other "elementary" particles appeared after 1930; neutron, neutrino, μ meson, π meson, heavier mesons, and various hyperons. I have heard it said that "the finder of a new elementary particle used to be rewarded by a Nobel Prize, but such a discovery now ought to be punished by a \$10,000 fine". [Source: Les Prix Nobel 1955, The Nobel Foundation, Stockholm.]

Not only were the new particles unexpected; there is a more technical sense in which they seemed "strange": They are *produced* copiously (on a time scale of about 10^{-23} sec), but they *decay* relatively slowly (typically about 10^{-10} sec). This suggested to Pais[17] and others that the mechanism involved in their production is entirely different from that which governs their disintegration. In modern language, the strange particles are *produced* by the *strong* force (the same one that holds the nucleus together), but they *decay* by the *weak* force (the one that accounts for beta decay and all other neutrino processes). The details of Pais's scheme required that the strange particles be produced in *pairs.* The experimental evidence for this was far from clear at that time, but in 1953 Gell-Mann[18] and Nishijima[19] found a beautifully simple, and, as it developed stunningly successful, way to implement and improve Pais's idea. They assigned to each particle a new property (Gell-Mann called it "strangeness") that (like charge, lepton number, and baryon number) is conserved in any strong interaction, but (*un*like those others) is *not* conserved in a weak interaction. In a pion-proton collision, for example, we might produce *two* strange particles:

$$\begin{aligned} \pi^- + p^+ &\to K^+ + \Sigma^- \\ &\to K^0 + \Sigma^0 \\ &\to K^0 + \Lambda \end{aligned} \tag{1.27}$$

Here the K's carry strangeness $S = +1$, the Σ's and the Λ have $S = -1$, and the "ordinary" particles—π, p, and n—have $S = 0$. But we never produce just *one* strange particle:

$$\begin{aligned} \pi^- + p^+ &\not\to \pi^+ + \Sigma^- \\ &\not\to \pi^0 + \Lambda \\ &\not\to K^0 + n \end{aligned} \tag{1.28}$$

On the other hand, when these particles *decay,* strangeness is *not* conserved:

$$\begin{aligned} \Lambda &\to p^+ + \pi^- \\ \Sigma^+ &\to p^+ + \pi^0 \\ &\to n + \pi^+ \end{aligned} \tag{1.29}$$

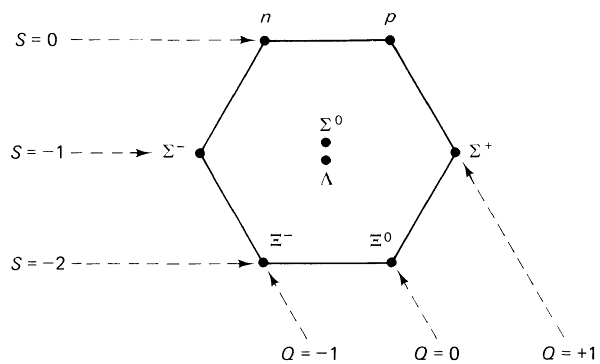
for these are *weak* processes, which do not respect conservation of strangeness.

There is some arbitrariness in the assignment of strangeness numbers, obviously. We could just as well have given $S = +1$ to the Σ's and the Λ, and $S = -1$ to K^+ and K^0; in fact, in retrospect it would have been a little nicer that way. [In exactly the same sense, Benjamin Franklin's original convention for plus and minus charge was perfectly arbitrary at the time, and unfortunate in retrospect since it made the current-carrying particle (the electron) negative.] The significant point is that there exists a consistent assignment of strangeness numbers to all the hadrons (baryons and mesons) that accounts for the observed strong processes and "explains" why the others do not occur. (The leptons and the photon don't experience strong forces at all, so strangeness does not apply to them.)

The garden which seemed so tidy in 1947 had grown into a jungle by 1960, and hadron physics could only be described as chaos. The plethora of strongly interacting particles was divided into two great families—the baryons and the mesons—and the members of each family were distinguished by charge, strangeness, and mass; but beyond that there was no rhyme or reason to it all. This predicament reminded many physicists of the situation in chemistry a century earlier, in the days before the Periodic Table, when scores of elements had been identified, but there was no underlying order or system. In 1960 the elementary particles awaited their own "Periodic Table."[20]

1.7 THE EIGHTFOLD WAY (1961–1964)

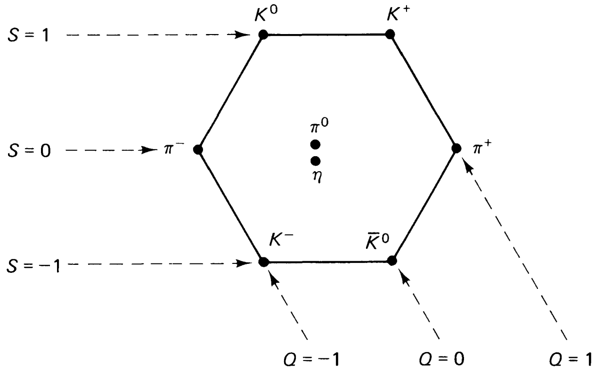
The Mendeleev of elementary particle physics was Murray Gell-Mann, who introduced the so-called *Eightfold Way* in 1961.[21] (Essentially the same scheme was proposed independently by Ne'eman.) The Eightfold Way arranged the baryons and mesons into weird geometrical patterns, according to their charge and strangeness. The eight lightest baryons fit into a hexagonal array, with two particles at the center:



The Baryon Octet

This group is known as the *baryon octet.* Notice that particles of like charge lie along the downward-sloping *diagonal* lines: $Q = +1$ (in units of the proton charge) for the proton and the Σ^+; $Q = 0$ for the neutron, the lambda, the Σ^0, and the Ξ^0; $Q = -1$ for the Σ^- and the Ξ^-. *Horizontal* lines associate particles of like *strangeness:* $S = 0$ for the proton and neutron, $S = -1$ for the middle line and $S = -2$ for the two Ξ's.

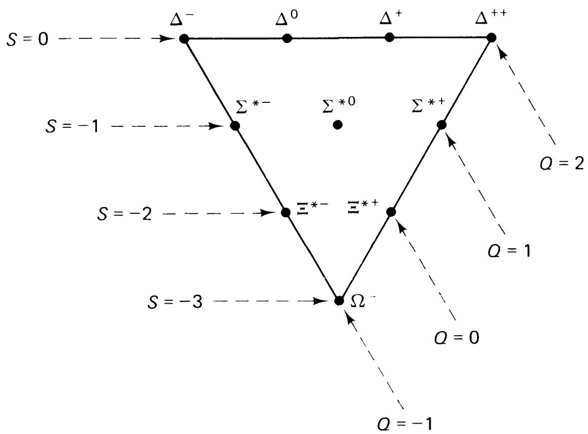
The eight lightest mesons fill a similar hexagonal pattern, forming the (*pseudo-scalar*) *meson octet:*



The Meson Octet

Once again, diagonal lines determine charge, and horizontals determine strangeness; but this time the top line has $S = 1$, the middle line $S = 0$, and the bottom line $S = -1$. (This discrepancy is a historical accident; Gell-Mann could just as well have assigned $S = 1$ to the proton and neutron, $S = 0$ to the Σ's and the Λ, and $S = -1$ to the Ξ's. In 1953 he had no reason to prefer that choice, and it seemed most natural to give the familiar particles—proton, neutron, and pion—a strangeness of zero. After 1961 a new term—*hypercharge*—was introduced, which was equal to S for the mesons and to $S + 1$ for the baryons. But later developments showed that strangeness was the better quantity after all, and the word "hypercharge" has now been taken over for a quite different purpose.)

Hexagons were not the only figures allowed by the Eightfold Way; there was also, for example, a triangular array, incorporating 10 heavier baryons—the *baryon decuplet:*



The Baryon Decuplet

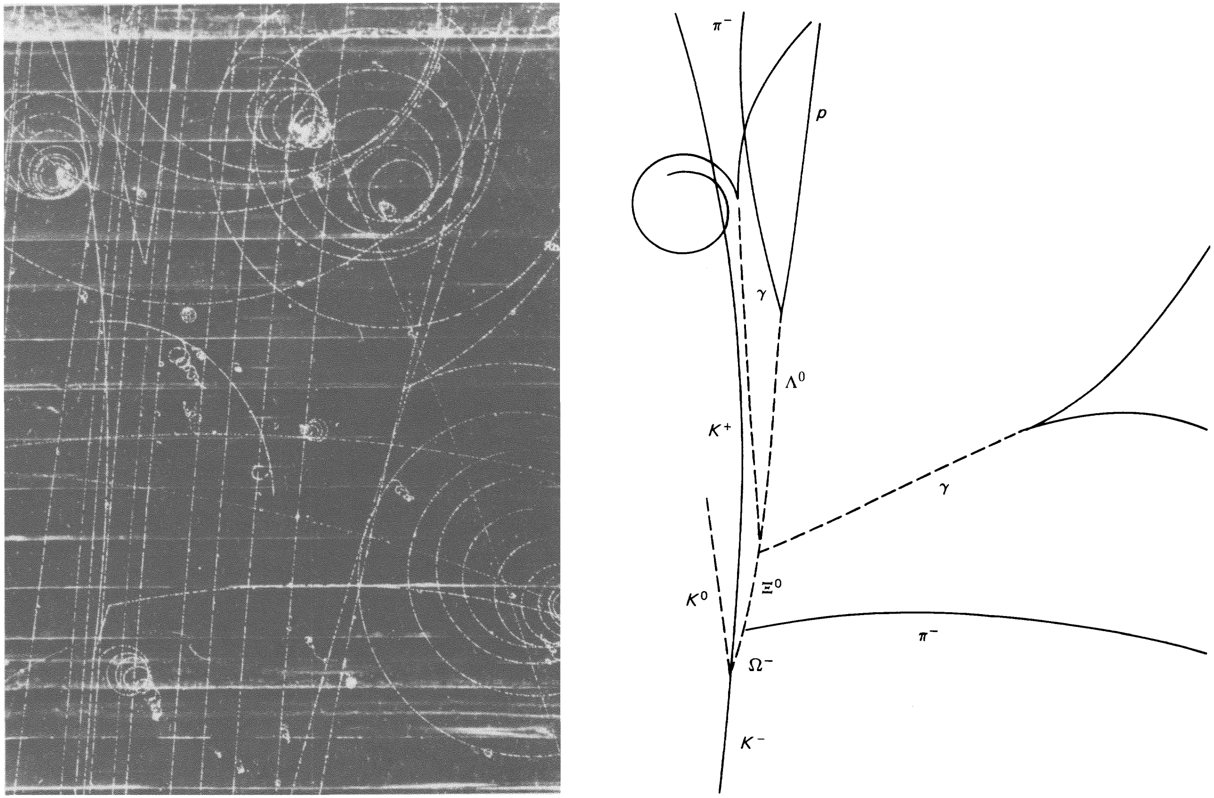


Figure 1.10 The discovery of the Ω^-. The actual bubble chamber photograph is shown on the left; a line diagram of the relevant tracks on the right. (Photo courtesy Brookhaven National Laboratory.)

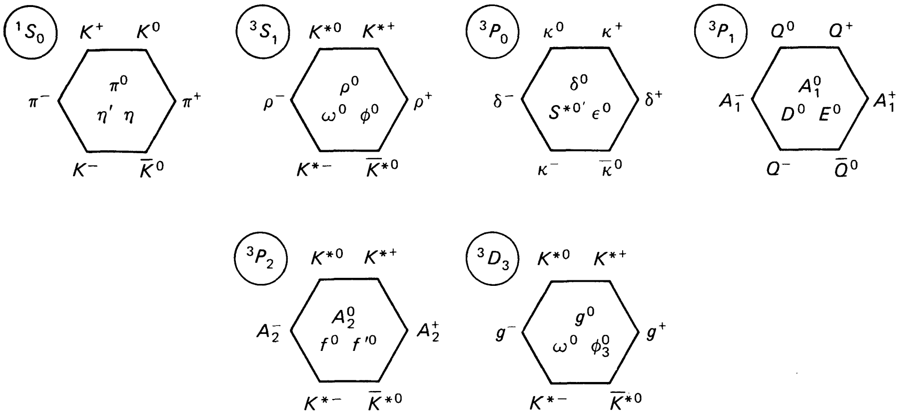


Figure 1.11 Established meson nonets. Obviously, we are running out of letters. It is customary to distinguish different particles represented by the same letter by indicating the mass parenthetically (in MeV/c^2), thus $K^*(892)$, $K^*(1430)$, $K^*(1650)$, and so on. In this figure the supermultiplets are labeled in spectroscopic notation (see Chap. 5). At present, there are no complete baryon supermultiplets beyond the octet and decuplet, although there are many partially filled diagrams.

Now, as Gell-Mann was fitting these particles into the decuplet, an absolutely lovely thing happened. Nine of the particles were known experimentally, but at that time the tenth particle—the one at the very bottom, with a charge of -1 and strangeness -3—was missing: No particle with these properties had ever been detected in the laboratory.[22] Gell-Mann boldly predicted that such a particle would be found, and told the experimentalists exactly how to produce it. Moreover, he calculated its mass—as you can for yourself, in Problem 1.6—and its lifetime, Problem 1.8—and sure enough, in 1964 the famous *omega-minus* particle was discovered,[23] precisely as Gell-Mann had predicted (see Fig. 1.10).

Since the discovery of the omega-minus (Ω^-), no one has seriously doubted that the Eightfold Way is correct.* Over the next 10 years, every new hadron found a place in one of the Eightfold Way *supermultiplets.* Some of these are shown in Figure 1.11. (This is not to say there were no false alarms; particles have a way of appearing and then *dis*appearing. Of the 26 mesons listed on a standard table in 1963, 19 were later found to be spurious!) In addition to the baryon octet, decuplet, and so on, there exist of course an *anti*baryon octet, decuplet, etc., with opposite charge and opposite strangeness. However, in the case of the mesons, the antiparticles lie in the *same supermultiplet* as the corresponding particles, in the diametrically opposite positions. Thus the antiparticle

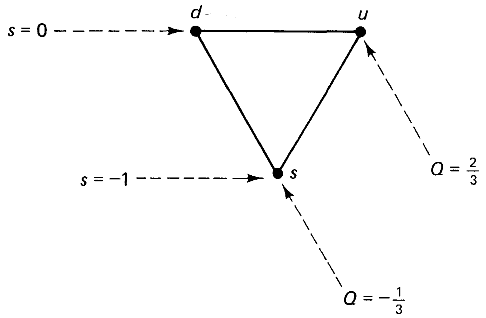
* A similar thing happened in the case of the Periodic Table. There were three famous "holes" (missing elements) on Mendeleev's chart, and he predicted that new elements would be discovered to fill in the gaps. Like Gell-Mann, he confidently described their properties, and within 20 years all three—gallium, scandium, and germanium—were found.

of the pi-plus is the pi-minus, the anti-K-minus is the K-plus, and so on (the pi-zero and the eta are their *own* antiparticles).

Classification is the first stage in the development of any science. The Eightfold Way did more than merely classify the hadrons, but its real importance lies in the organizational structure it provided. I think it's fair to say that the Eightfold Way initiated the modern era in particle physics.

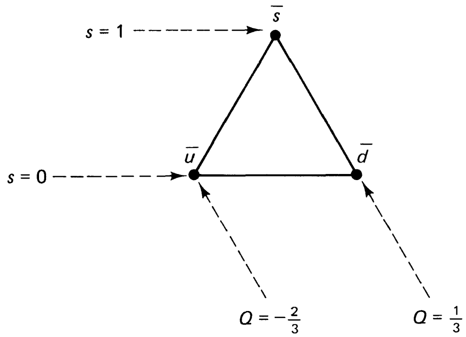
1.8 THE QUARK MODEL (1964)

But the very success of the Eightfold Way begs the question: *Why* do the hadrons fit into these curious patterns? The Periodic Table had to wait many years for quantum mechanics and the Pauli exclusion principle to provide its explanation. An understanding of the Eightfold Way, however, came already in 1964, when Gell-Mann and Zweig independently proposed that all hadrons are in fact composed of even more elementary constituents, which Gell-Mann called *quarks*.[24] The quarks come in three types (or "flavors"), forming a triangular "Eightfold-Way" pattern:



The Quarks

The u (for "up") quark carries a charge of $\frac{2}{3}$ and a strangeness of zero; the d ("down") quark carries a charge of $-\frac{1}{3}$ and $S = 0$; the s (originally "sideways", but now more commonly "strange") quark has $Q = -\frac{1}{3}$ and $S = -1$. To each quark (q) there corresponds an *anti*quark ($\bar{q}$), with the opposite charge and strangeness:



The Antiquarks

The quark model asserts that

1. Every baryon is composed of three quarks (and every *anti*baryon is composed of three *anti*quarks).
2. Every meson is composed of a quark and an antiquark.

With these two rules it is a matter of elementary arithmetic to construct the baryon decuplet and the meson octet. All we need to do is list the combinations of three quarks (or quark–antiquark pairs), and add up their charge and strangeness:

THE BARYON DECUPLET

qqq	Q	S	Baryon
uuu	2	0	Δ^{++}
uud	1	0	Δ^{+}
udd	0	0	Δ^{0}
ddd	−1	0	Δ^{-}
uus	1	−1	Σ^{*+}
uds	0	−1	Σ^{*0}
dds	−1	−1	Σ^{*-}
uss	0	−2	Ξ^{*0}
dss	−1	−2	Ξ^{*-}
sss	−1	−3	Ω^{-}

Notice that there are 10 combinations of three quarks. Three u's, for instance, at $Q = \frac{2}{3}$ each, yield a total charge of $+2$, and a strangeness of zero. This is the Δ^{++} particle. Continuing down the table, we find all the members of the decuplet ending with the Ω^-, which is evidently made of three s quarks.

A similar enumeration of the quark–antiquark combinations yields the meson table:

THE MESON NONET

$q\bar{q}$	Q	S	Meson
$u\bar{u}$	0	0	π^{0}
$u\bar{d}$	1	0	π^{+}
$d\bar{u}$	−1	0	π^{-}
$d\bar{d}$	0	0	η
$u\bar{s}$	1	1	K^{+}
$d\bar{s}$	0	1	K^{0}
$s\bar{u}$	−1	−1	K^{-}
$s\bar{d}$	0	−1	$\bar{K}^{0}$
$s\bar{s}$	0	0	??

But wait! There are *nine* combinations here, and only eight particles in the meson octet. The quark model requires that there be a third meson (in addition to the π^0 and the η) with $Q = 0$ and $S = 0$. As it turns out, just such a particle had already been found experimentally—the η'. In the Eightfold Way the η' had been classified as a *singlet*, all by itself. According to the quark model it properly belongs with the other eight mesons to form a *meson nonet*. (Actually, since $u\bar{u}$, $d\bar{d}$, and $s\bar{s}$ all have $Q = 0$ and $S = 0$, it is not possible to say, on the basis of anything we have done so far, which is the π^0, which the η, and which the η'. But never mind, the point is that there are *three* mesons with $Q = S = 0$.) By the way, the *anti*mesons automatically fall in the same supermultiplet as the mesons: $u\bar{d}$ is the antiparticle of $d\bar{u}$, and vice versa.

You may have noticed that I avoided talking about the baryon *octet*—and it is far from obvious how we are going to get *eight* baryons by putting together three quarks. In truth, the procedure is perfectly straightforward, but it does call for some facility in handling spins, and I would rather save it until Chapter 5. For now, I'll just tantalize you with the mysterious observation that if you take the decuplet and knock off the three corners (where the quarks are identical—*uuu*, *ddd*, and *sss*), and double the center (where all three are different—*uds*), you obtain precisely the eight states in the baryon octet. So the same set of quarks can account for the octet; it's just that some combinations do not appear at all, and one appears twice.

Indeed, all the Eightfold Way supermultiplets emerge in a natural way from the quark model. Of course, the same combination of quarks can go to make a number of different particles: The delta-plus and the proton are both composed of two *u*'s and a *d*; the pi-plus and the rho-plus are both $u\bar{d}$; and so on. Just as the hydrogen atom (electron plus proton) has many different energy levels, so a given collection of quarks can bind together in many different ways. But whereas the various energy levels in the electron/proton system are relatively close together (the spacings are typically several electron volts, in an atom whose rest energy is nearly 10^9 electron volts), so that we naturally think of them all as "hydrogen," the energy spacings for different states of a bound quark system are very large, and we normally regard them as distinct particles. Thus we can, in principle, construct an infinite number of hadrons out of only three quarks. Notice, however, that *some* things are absolutely excluded in the quark model: For example, a baryon with $S = 0$ and $Q = -2$; no combination of the three quarks can produce these numbers. Nor can there be a *meson* with a charge of $+2$ (like the Δ^{++} baryon) or a strangeness of -3 (like the Ω^-). For a long time there were major experimental searches for these so-called "exotic" particles; their discovery would be devastating for the quark model, but none has ever been found (see Problem 1.11).

The quark model *does*, however, suffer from one profound embarrassment: In spite of the most diligent search over a period of 20 years, no one has ever seen an individual quark. Now, if a proton is really made out of three quarks, you'd think that if you hit one hard enough, the quarks ought to come popping out. Nor would they be hard to recognize, carrying as they do the conspicuous label of fractional charge; an ordinary Millikan oil drop experiment would clinch the identification. Moreover, at least one of the quarks should be absolutely stable; what could it decay into, since there is no lighter particle with fractional charge? So quarks ought to be *easy* to produce, *easy* to identify, and *easy* to store, and yet, no one has ever found one.

The failure of experiments to produce isolated quarks occasioned widespread skepticism about the quark model in the late sixties and early seventies. Those who clung to the model tried to conceal their disappointment by introducing the notion of *quark confinement:* perhaps, for reasons not yet understood, quarks are *absolutely confined* within baryons and mesons, so that no matter how hard you try, you cannot get them out. Of course, this doesn't explain anything, it just gives a name to our frustration. But at least it poses sharply what has become a crucial theoretical problem for the eighties: to discover the mechanism responsible for quark confinement. There are some indications that the solution may be at hand.[25]

Even if all quarks are stuck inside hadrons, this does not mean they are inaccessible to experimental study. One can probe the inside of a proton in much the same way as Rutherford probed the inside of an atom—by firing something into it. Such experiments were carried out in the late sixties using high-energy electrons at the Stanford Linear Accelerator Center (SLAC). They were repeated in the early seventies using neutrino beams at CERN, and later still using protons. The results of these so-called "deep inelastic scattering" experiments were strikingly reminiscent of Rutherford's (Fig. 1.12): *Most* of the incident particles pass right through, whereas a small number bounce back sharply. This means that the charge of the proton is concentrated in small lumps, just as Rutherford's results indicated that the positive charge in an atom is concentrated at the nucleus.[26] However, in the case of the proton the evidence suggests *three* lumps,

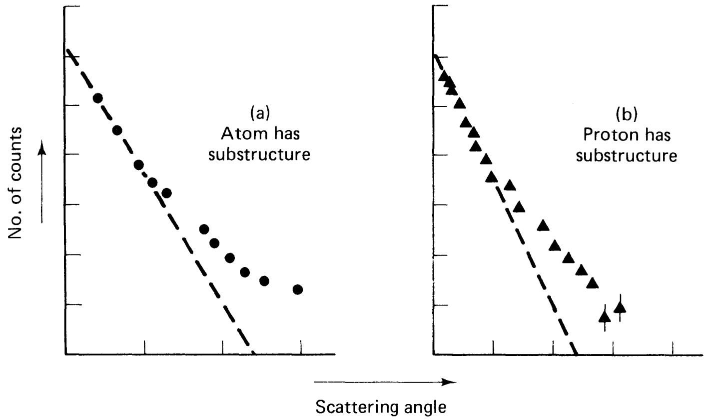


Figure 1.12 (a) In Rutherford scattering the number of particles deflected through large angles indicates that the atom has internal structure (a nucleus). (b) In deep inelastic scattering the number of particles deflected through large angles indicates that the proton has internal structure (quarks). The dashed lines show what you would expect if the positive charge were uniformly distributed over the volume of (a) the atom, (b) the proton. [*Source:* F. Halzen and A. D. Martin, *Quarks and Leptons* (New York: Wiley, 1984), p. 17. Copyright © John Wiley & Sons, Inc. Reprinted by permission.]

instead of *one*. This is strong support for the quark model, obviously, but still not conclusive.

Finally, there was a theoretical objection to the quark model: It appears to violate the Pauli exclusion principle. In Pauli's original formulation the exclusion principle stated that no two electrons can occupy the same state. However, it was later realized that the same rule applies to all particles of half-integer spin (the proof of this is one of the most important achievements of quantum field theory). In particular, the exclusion principle should apply to quarks, which, as we shall see, must carry spin $\frac{1}{2}$. Now the Δ^{++}, for instance, is supposed to consist of three identical u quarks in the same state; it (and also the Δ^- and the Ω^-) appear to be inconsistent with the Pauli principle. In 1964, O. W. Greenberg proposed a way out of this dilemma:[27] He suggested that quarks not only come in three *flavors* (u, d, and s) but each of these also comes in three *colors* ("red," "green," and "blue," say). To make a baryon, we simply take one quark of each color, then the three u's in Δ^{++} are no longer identical (one's red, one's green, and one's blue). Since the exclusion principle only applies to *identical* particles, the problem evaporates.

The color hypothesis sounds like sleight of hand, and many people initially considered it the last gasp of the quark model. As it turned out, the introduction of color was one of the most fruitful ideas of our time. I need hardly say that the term "color" here has absolutely no connection with the ordinary meaning of the word. Redness, blueness, and greenness are simply *labels* used to denote three new properties that, in addition to charge and strangeness, the quarks possess. A *red* quark carries one unit of redness, zero blueness, and zero greenness; its antiparticle carries *minus* one unit of redness, and so on. We could just as well call these quantities X-ness, Y-ness, and Z-ness, for instance. However, the color terminology has one especially nice feature: It suggests a delightfully simple characterization of the particular quark combinations that are found in nature.

All naturally occurring particles are colorless.

By "colorless" I mean that *either* the total amount of each color is zero *or* all three colors are present in equal amounts. (The latter case mimics the optical fact that light beams of three primary colors combine to make white.) This clever rule "explains" (if that's the word for it) why you can't make a particle out of *two* quarks, or *four* quarks, and for that matter why *individual* quarks do not occur in nature. The only colorless combinations you can make are $q\bar{q}$ (the mesons), qqq (the baryons), and $\bar{q}\bar{q}\bar{q}$ (the antibaryons). (You could have *six* quarks, of course, but we would interpret that as a bound state of two baryons.)

1.9 THE NOVEMBER REVOLUTION AND ITS AFTERMATH (1974–1983)

The decade from 1964 to 1974 was a barren time for elementary particle physics. The quark model, which had seemed so promising at the beginning, was in an uncomfortable state of limbo by the end. It had had some striking successes: It neatly explained the Eightfold Way, and correctly predicted the lumpy structure of the proton. But it had two conspicuous defects: the experimental absence of free quarks and inconsistency with the Pauli principle. Those who liked the model papered over these failures with what seemed at the time to be rather transparent rationalizations: the idea of quark confinement and the color hypothesis. But I think it is safe to say that by 1974 most elementary particle physicists felt queasy, at best, about the quark model. The lumps inside the proton were called *partons,* and it was unfashionable to identify them explicitly with quarks.

Curiously enough, what rescued the quark model was not the discovery of free quarks, or an explanation of quark confinement, or confirmation of the color hypothesis, but something entirely different and (almost)[28] completely unexpected: the discovery of the psi meson. The ψ was first observed at Brookhaven by a group under C. C. Ting, in the summer of 1974. But Ting wanted to check his results before announcing them publicly, and the discovery remained an astonishingly well-kept secret until the weekend of November 10–11, when the new particle was discovered independently by Burton Richter's group at SLAC. The two teams then published simultaneously,[29] Ting naming the particle *J,* and Richter calling it ψ. The J/ψ was an electrically neutral, extremely heavy meson—more than three times the weight of a proton (the original notion that mesons are "middle-weight" and baryons "heavy-weight" had long since gone by the boards). But what made this particle so unusual was its extraordinarily long lifetime. For the ψ lasted fully 10^{-20} seconds before disintegrating. Now, 10^{-20} seconds may not impress you as a particularly long time, but you must understand that the *typical* lifetimes for hadrons in this mass range are on the order of 10^{-23} seconds. So the ψ has a lifetime about a thousand times longer than any comparable particle. It's as though someone came upon an isolated village in Peru or the Caucasus where people live to be 70,000 years old. That wouldn't just be some actuarial anomaly; it would be a sign of fundamentally new biology at work. And so it was with the ψ: its long lifetime, to those who understood, spoke of fundamentally new physics. For good reason, the events precipitated by the discovery of the ψ came to be known as the *November Revolution.*[30]

In the months that followed, the true nature of the ψ meson was the subject of lively debate, but the explanation that won was provided by the quark model. It is now universally accepted that the ψ represents a bound state of a new (fourth) quark, the c (for *charm*) and its antiquark: $\psi = (c\bar{c})$. Actually, the idea of a fourth flavor, and even the whimsical name, had been introduced many years earlier, by Bjorken and Glashow.[31] Indeed, there was an intriguing parallel between the leptons and the quarks:

Leptons: e, ν_e, μ, ν_μ

Quarks: d, u, s

If all mesons and baryons are made out of quarks, these two families are left as

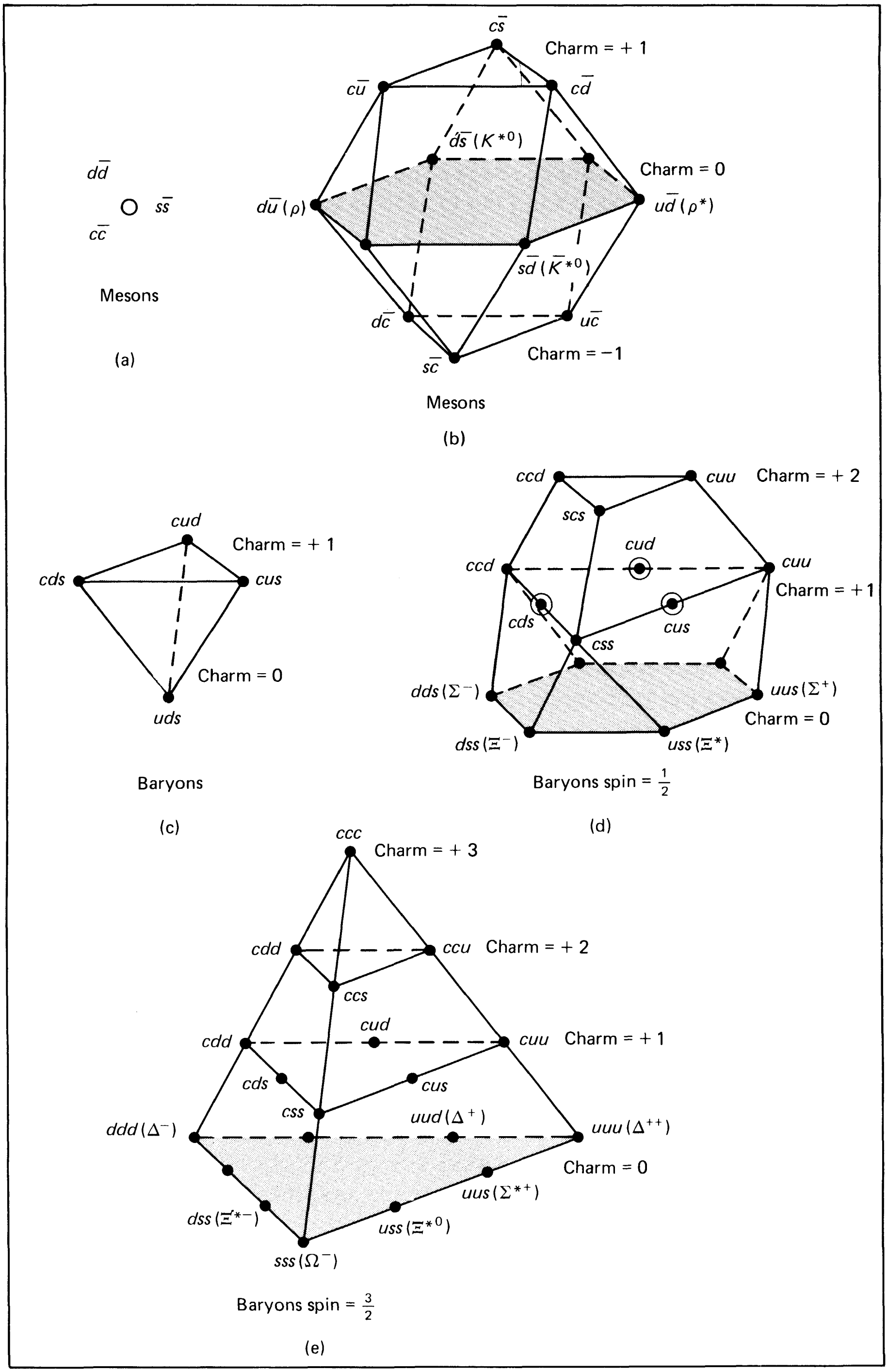


Figure 1.13 Supermultiplets constructed with four quarks. (From "Quarks with Color and Flavor," by S. Glashow. Copyright © Oct. 1975 by Scientific American, Inc. All rights reserved.)

the *truly* fundamental particles. But why *four* leptons and only *three* quarks? Wouldn't it be nicer if there were four of each? Later, Glashow, Iliopoulos, and Maiani[32] offered more compelling technical reasons for wanting a fourth quark, but the simple idea of a parallel between quarks and leptons is another of those farfetched speculations that turned out to have more substance than their authors could have imagined.

So when the ψ was discovered, the quark model was ready and waiting with an explanation. Moreover, it was an explanation pregnant with implications. For if a fourth quark exists, there should be all kinds of new baryons and mesons, carrying various amounts of charm. Some of these are shown in Figure 1.13; you can work out the possibilities for yourself (Problems 1.14 and 1.15). Notice that the ψ itself carries no *net* charm, for if the c is assigned a charm of $+1$, then $\bar{c}$ will have a charm of -1; the charm of the ψ is, if you will, "hidden." To confirm the charm hypothesis it was important to produce a particle with "naked" (or "bare") charm.[33] The first evidence for charmed baryons ($\Lambda_c^+ = udc$ and possibly $\Sigma_c^{++} = uuc$) appeared already in 1975 (Fig. 1.14);[34] the first charmed mesons ($D^0 = c\bar{u}$ and $D^+ = c\bar{d}$) were found in 1976,[35] and the charmed strange meson ($F^+ = c\bar{s}$) in 1977.[36] (The F meson was recently renamed D_s. There is also some evidence for *usc* and *ssc*.) With these discoveries the interpretation of the ψ as $c\bar{c}$ was established beyond reasonable doubt. More important, the quark model itself was put back on its feet.

However, the story does not end there, for in 1975 a new *lepton* was discovered,[37] spoiling Glashow's symmetry. This new particle (the tau) presumably has its own neutrino, so we are up to six leptons, and only four quarks. But don't dispair, because two years later a new heavy meson (the *upsilon*) was discovered,[38] and quickly recognized as the carrier of a fifth quark, b (for *beauty,* or *bottom,* depending on your taste): $\Upsilon = b\bar{b}$. Immediately the search began for mesons and hadrons exhibiting "naked beauty" (or "bare bottom"). (I'm sorry. *I* didn't invent this terminology. In a way, its silliness is a reminder of how wary people were of taking the quark model seriously, in the early days.) The first *beautiful* baryon, $\Lambda_b = udb$, may have been observed in 1981[39] (the claim is hotly contested[40]); the first beautiful mesons ($B^0 = b\bar{d}$ and $B^- = b\bar{u}$) were found in 1983.[41] At this point it doesn't take much imagination to predict that a sixth quark will eventually be found; it already has a name: t (for *truth,* of course, or *top*). If and when the t quark is discovered (there were some indications in the summer of 1984 that it may have been seen at CERN), Glashow's symmetry will be restored, with six leptons and six quarks. And there (knock on wood) the proliferation stops.

1.10 INTERMEDIATE VECTOR BOSONS (1983)

In his original theory of beta decay (1933) Fermi treated the process as a contact interaction, occurring at a single point, and therefore requiring no mediating

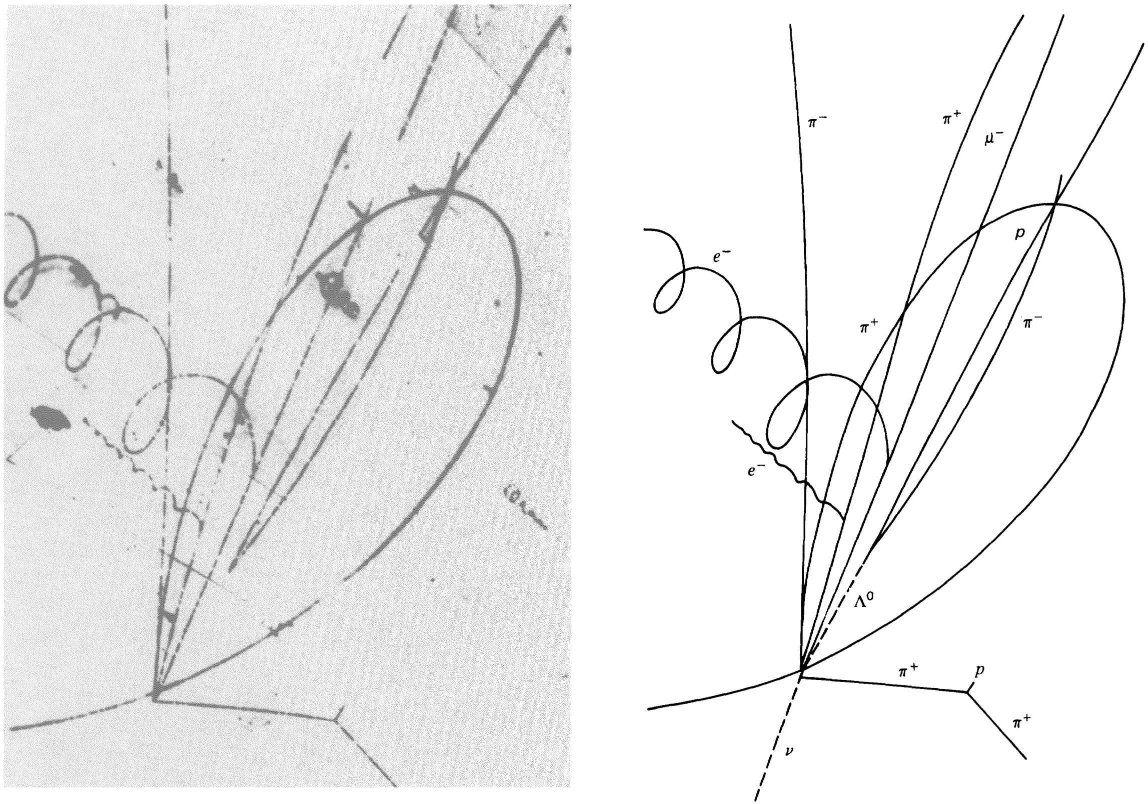


Figure 1.14 The charmed baryon. The probable interpretation of this event is $\nu_\mu + p \longrightarrow \Lambda_c^+ + \mu^- + \pi^+ + \pi^-$. The charmed baryon decays ($\Lambda_c^+ \longrightarrow \Lambda + \pi^+$) too soon to leave a track, but the subsequent decay of the Λ is clearly visible. (Photo courtesy of N. P. Samios, Brookhaven National Laboratory.)

particle. As it happens, the weak force (which is responsible for beta decay) *is* of extremely short range, so that Fermi's model was not far from the truth, and yields excellent approximate results at low energies. However, it was widely recognized that this approach was bound to fail at high energies, and would eventually have to be supplanted with a theory in which the interaction was mediated by the exchange of some particle. The mediator came to be known by the prosaic name *intermediate vector boson.* The challenge for theorists was to predict the properties of the intermediate vector boson, and for experimentalists, to produce one in the laboratory. You may recall that Yukawa, faced with the analogous problem for the strong force, was able to estimate the mass of the pion in terms of the range of the force, which he took to be roughly the same as the size of a nucleus. But we have no corresponding way to measure the range of the weak force; there are no "weak bound states" whose size would inform us—the weak force is simply too feeble to bind particles together. For many years predictions of the intermediate vector boson mass were little more than educated guesses (the "education" coming largely from the failure of experiments at progressively higher energies to detect the particle). By 1962 it was known that the mass had to be at least half the proton mass; 10 years later the experimental lower limit had grown to 2.5 proton masses.

But it was not until the emergence of the electroweak theory of Glashow, Weinberg, and Salam that a really firm prediction of the mass was possible. In this theory there are in fact *three* intermediate vector bosons, two of them charged ($W^{\pm}$) and one neutral (Z). Their masses were calculated to be[42]

$$M_W = 82 \pm 2 \text{ GeV}/c^2, \qquad M_Z = 92 \pm 2 \text{ GeV}/c^2 \tag{1.30}$$

In the late seventies, CERN began construction of a proton–antiproton collider designed specifically to produce these extremely heavy particles (bear in mind that the mass of the proton is 0.94 GeV/c^2, so we're talking about something nearly 100 times as heavy). In January 1983 the discovery of the W (at 81 ± 5 GeV/c^2) was reported by Carlo Rubbia's group,[43] and five months later the same team announced discovery of the Z (at 95 ± 3 GeV/c^2).[44] These experiments represent an extraordinary technical triumph,[45] and they were of fundamental importance in confirming a crucial aspect of the Standard Model, to which the physics community was by that time heavily committed (and for which a Nobel Prize had already been awarded). Unlike the strange particles or the ψ, however, the intermediate vector bosons were long awaited and universally expected, so the general reaction was a sigh of relief, not shock or surprise.

1.11 THE STANDARD MODEL (1978–?)

In the current view, then, all matter is made out of three kinds of elementary particles: leptons, quarks, and mediators. There are six leptons, classified according to their charge (Q), electron number (L_e), muon number (L_μ), and tau number (L_τ). They fall naturally into three *families* (or *generations*):

LEPTON CLASSIFICATION

	l	Q	L_e	L_μ	L_τ
First generation	e	−1	1	0	0
	ν_e	0	1	0	0
Second generation	μ	−1	0	1	0
	ν_μ	0	0	1	0
Third generation	τ	−1	0	0	1
	ν_τ	0	0	0	1

There are also six antileptons, with all the signs reversed. The positron, for example, carries a charge of +1 and an electron number −1. So there are really 12 leptons, all told.

Similarly, there are six "flavors" of quarks, which are classified according to charge, strangeness (S), charm (C), beauty (B), and truth (T). [For consistency, I suppose we should include "upness" (U) and "downness" (D), although these terms are seldom used. They are redundant, inasmuch as the only quark with $S = C = B = T = 0$ and $Q = \frac{2}{3}$, for instance, is the up quark, so it is not necessary to specify $U = 1$ and $D = 0$ as well.] The quarks, too, fall into three generations:

QUARK CLASSIFICATION

	q	Q	D	U	S	C	B	T
First generation	d	$-\frac{1}{3}$	−1	0	0	0	0	0
	u	$\frac{2}{3}$	0	1	0	0	0	0
Second generation	s	$-\frac{1}{3}$	0	0	−1	0	0	0
	c	$\frac{2}{3}$	0	0	0	1	0	0
Third generation	b	$-\frac{1}{3}$	0	0	0	0	−1	0
	t	$\frac{2}{3}$	0	0	0	0	0	1

Again, all signs would be reversed on the table of antiquarks. Meanwhile, each quark and antiquark comes in three colors, so there are 36 of them in all.

Finally, every interaction has its mediators: the photon for the electromagnetic force, two W's and a Z for the weak force, the graviton (presumably) for gravity, . . . but what about the strong force? In Yukawa's original theory (1934) the mediator of strong forces was the pion, but with the discovery of heavy mesons this simple picture could not stand; protons and neutrons could now exchange rho's and eta's and K's and phi's and all the rest of them. The

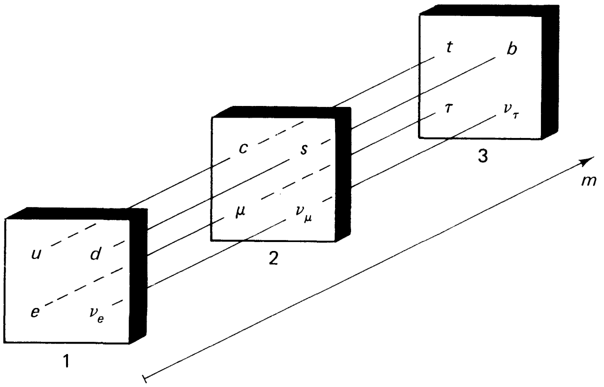


Figure 1.15 The three generations of quarks and leptons, in order of increasing mass.

quark model brought an even more radical revision, for if protons, neutrons, and mesons are complicated composite structures, there is no reason to believe their interaction *should* be simple. To study the strong force at the fundamental level, one should look, rather, at the interaction between individual quarks. So the question becomes: What particle is exchanged between two quarks, in a strong process? This mediator is called the *gluon,* and in the Standard Model there are eight of them. As we shall see, the gluons themselves carry color, and therefore (like the quarks) should not exist as isolated particles. We can hope to detect gluons only within hadrons, or in colorless combinations with other gluons (*glueballs*). Nevertheless, there is substantial indirect experimental evidence for the existence of gluons: The deep inelastic scattering experiments showed that roughly half the momentum of a proton is carried by electrically neutral constituents, presumably gluons; the *jet* structure characteristic of proton scattering at high energies can be explained in terms of the disintegration of quarks and gluons in flight;[46] and glueballs may conceivably have been observed.[47] But no one would say that the experimental evidence is really *compelling,* at this stage.

This is all adding up to an embarrassingly large number of supposedly "elementary" particles: 12 leptons, 36 quarks, 12 mediators (I won't count the graviton, since gravity is not included in the Standard Model). And, as we shall see later, the Glashow-Weinberg-Salam theory calls for at least one *Higgs* particle, so we have a minimum of 61 particles to contend with. Informed by our experience first with atoms and later with hadrons, many people have suggested that some, at least, of these 61 must be composites of more elementary subparticles (see Problem 1.17).[48] Such speculations lie beyond the Standard Model and outside the scope of this book. Personally, I do not think the large number of "elementary" particles in the Standard Model is by itself alarming, for they are tightly interrelated. The eight gluons, for example, are identical except for their colors, and the second and third generations mimic the first (Fig. 1.16). In the next chapter we shall see how this structure leads to the first systematic and comprehensive theory of elementary particle dynamics.