9 Particle Theory

9.1 Introduction

In 1948 quantum mechanics entered a new phase. Increasingly precise experimental results required new calculation methods, as the existing methods were hopelessly inadequate to deal with the complications of the theory. Richard Feynman came up with a new method that led to enormous simplifications. The method relied heavily on little drawings, now called Feynman diagrams. For a given situation one would draw a few of these diagrams, and then there were simple rules that provided the calculational answers in connection with them. As these diagrams are moreover very appealing intuitively they have become the universal tools of particle physics.

Historically, the work of Feynman was a tremendous step forward. In itself it did not really add to the theory, but it made working with it practical. It became simple to do calculations. The first domain conquered was the theory of photons and electrons, quantum electrodynamics (QED). That theory had its difficulties, but these difficulties could be overcome using a procedure called renormalization. That is discussed in this Chapter. Using Feynman's techniques this procedure becomes transparent.

Besides the interactions of photons and electrons there are other interactions, notably weak interactions. It took many decades to understand these forces. The renormalization procedure was not sufficient to eliminate all troubles. Progress came with the idea that new forces, new particles, with suitable interactions,

Richard Feynman (1918–1988). The most important contribution of Feynman, in my view, is his introduction of the diagram method named after him, and the theoretical tool, path integrals, that he developed. Truly wonderful work.

Part of the formal theory associated with those diagrams was published before, in French, by Stückelberg in the somewhat inaccessible journal, *Helvetica Physica Acta*. This including the idea that a positron may be viewed as an electron going backwards in time (this is basically the idea of crossing). It is unlikely that Feynman knew of that work, yet when he learned of it he dutifully acknowledged that in his papers. There are some anecdotes associated with that, not necessarily true.

On the evening of the day (in 1965) that Feynman celebrated his Nobel prize he received a telegram during the party: "Send back my notes, please", signed Stückelberg. According to my source (unpublished biography of Stückelberg by Ruth Wenger) the originator of the joke was Gell-Mann. I asked Gell-Mann if he had sent this telegram, but he denied that, adding that it was a nice idea.

When Feynman, after receiving his prize in Stockholm, gave a lecture at CERN, Geneva, he was afterwards introduced to Stückelberg. He asked Stückelberg: "Why did you not draw diagrams?" To which Stückelberg answered: "I had no draughtsman''. Stückelberg, always the perfect gentleman and very conscious of his standing as a baron, apparently felt it below his dignity to draw those simple figures himself.

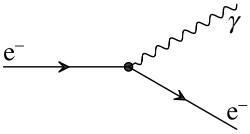
Feynman was a very charming person to talk to, and he was a gifted teacher. Well-known are his textbooks on physics, and he came very much in the public eye in connection with his part in the understanding of the Challenger disaster.

could be introduced such that the theory became manageable, i.e. renormalizable. The famous Higgs particle is one of those particles. From a mathematical playground this became reality when these hypothetical particles (except the Higgs particle) were actually discovered, and moreover were demonstrated to have the requisite properties. It is this work that was honored with the 1979 (Glashow, Salam and Weinberg) and 1999 ('t Hooft and Veltman) Nobel prizes. The work of Glashow, Salam and Weinberg concerned the construction of the actual model, while 't Hooft and Veltman elucidated the mathematical structure, showing that this model was renormalizable. By model we mean here a precise list of particles and their interactions. Without the simplifications due to Feynman's methods that progress would have been unthinkable. Not only in experimental physics but in theoretical physics as well the advance in techniques leads to new developments and insights.

9.2 Feynman Rules

Feynman rules are the main tools of the contemporary particle theorist. These rules incorporate the basic concepts of quantum mechanics; most importantly they can be represented in terms of drawings, diagrams, that have a strong intuitive appeal. A few basic concepts must be understood first to appreciate these drawings.

In Feynman diagrams particles are represented by lines, and interactions between particles by points where these lines join. Such an interaction point is called a vertex. The most obvious example is the interaction of electrons with photons. It is an interaction that we see literally almost permanently: the emission of light by electrons. In Feynman diagram language this interaction is represented in a very simple manner, see figure below. The electron, represented by a line with an arrow, shakes off a photon and moves on. The arrow is not there to indicate the direction of movement, but rather that of the flow of (negative) electric charge. Later on the meaning of the arrow will be changed slightly, but for now this will do.

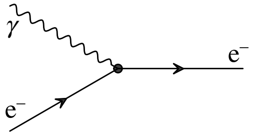


The interaction of electrons with light has been well understood for a long time, and we have a precise quantitative understanding of the physics corresponding to this diagram. However, it must be understood that this simple diagram applies to many situations; the difference is in the possible initial and final configurations. That is typical for quantum mechanics: specify the initial and final configurations and then the theory provides the calculation of the probability for the process to happen.

Thus in practice each diagram must be supplemented with a precise specification of the initial and final state. Very often these initial states are particles coming from accelerators and the final states are the outgoing particles observed in detectors; in other words freely moving particles that collide or emerge from a collision. However, there are other situations. Concerning the above diagram the electron may initially be in a higher orbit in an atom,[a] and fall to a state of lower energy, a lower orbit, thereby emitting a photon. Another example is the emission of radio waves by an emitter. Electrons, moving back and forth in the antenna, shake off photons. In both cases the emerging photons are freely moving particles, but not the electrons, they are tied down in some way.

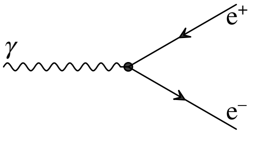
An important lesson that we can draw from this diagram is that particles can be created in an interaction. First the photon was not there, and some time later it came into existence. The opposite happens when a photon hits the eye: the photon is absorbed by an electron which then somehow leads to excitation of a nerve. The diagram corresponding to this process is shown in the figure below.

[a]Electrons can go into a higher orbit due to collisions between atoms or electrons or by absorbing light.



The difference is that here the photon is incoming, not outgoing. We "crossed" the photon line, i.e. we moved the line from outgoing to incoming. "Crossing" is an important property of Feynman diagrams: when moving a line from in to out or vice versa a new diagram results which corresponds to another possible process. This opens up interesting possibilities, especially if we apply crossing to other than photon lines, for example to electron lines. Let us consider the last diagram and apply crossing to the incoming electron line.

The result of this electron line crossing is another figure: a photon changes into an electron pair. According to the arrow one of its members has the charge moving in the opposite way, and we observe it as positive charge going out. So this is the rule: when a particle is outgoing, and the arrow points inwards we interpret that as the opposite charge. Thus this particle is now like an electron, except its charge is positive. It is called the positron, the antiparticle of the electron. So our crossing rule gets refined: crossing a line changes a particle into an antiparticle (and vice versa).

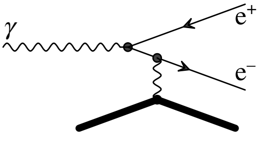


Positrons were experimentally discovered in the 1930s, and today antiparticles are an almost automatically accepted part of particle physics. Some particles are identical with their antiparticles: the photon is an example. Lines corresponding to such a particle carry no arrow and the particle has no charge.

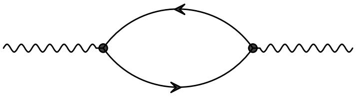
Meanwhile there is another important element to understand. The interactions obey strictly the laws of energy and momentum conservation. An electron in an atom can emit a photon while dropping to a lower energy state. But a free electron cannot emit a photon. Consider an electron at rest, i.e. with zero momentum; it is then in its lowest energy state. If it were to emit a photon of finite energy then an electron with even less energy would be left behind, which is not possible. The same then holds also for a freely moving electron, which one could imagine to be an electron at rest as seen by a moving observer. Since the way processes go should not depend on the frame from which they are observed, especially not whether the observer is moving or not, it follows that if some process is not possible for one observer it should also not occur for any other.

Likewise a photon cannot change in mid-flight into an electron-positron pair, even if it is a high-energy photon. This can be understood by realizing that this high-energy photon appears as a photon of lower energy to another observer moving in the same direction as that photon. A photon always moves with the speed of light, and one can never catch up with it like in the case of a particle with mass; instead, when an observer races in front of the photon he will still see it coming with the speed of light, but it appears red-shifted, i.e. it is perceived as a photon of lower energy. If the observer moves fast enough, the photon energy can for this observer become less than needed to create an electron pair (whose energy at rest is twice the rest mass energy of one electron).

In other circumstances, where another object absorbs or adds some momentum or energy, photon conversion to an electron-positron pair can happen. In collisions with nuclei a high energy photon will in fact readily convert into an electron-positron pair. An observer moving in the same direction as the photon would see a photon of lower energy, but it would then from his point of view collide with a moving nucleus, and there is still enough energy for pair creation. An electron or positron moving through matter may likewise emit a photon, commonly called bremsstrahlung (literally brake-radiation). In Chapter 7 there is a bubble chamber picture of a neutrino event. In that picture one can see several electron-positron pairs, appearing like sea-gulls. These pairs are due to photons coming from bremsstrahlung by an earlier electron or positron. All of these processes involve an extra photon carrying momentum and energy from electron or positron to or from a nucleus; an example is shown in the figure. The fat line represents a nucleus.

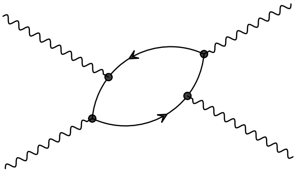


The next point is one of quantum mechanics. Particles can exist with 'inadmissible' energies, provided that this occurs only for a short time. The more inadmissible the energy, the shorter the duration. What we mean here by inadmissible is an energy different from the value that one must normally assign to a particle with a given momentum. For example, an electron at rest has an energy corresponding to its rest mass multiplied by the speed of light squared ($E = mc^2$). An electron with zero energy is hence not possible. Yet quantum mechanics allows the existence of zero energy electrons and even negative energy electrons, or of electrons with inadmissibly large energies (for example a very high energy electron at rest), provided this takes place only for short times. In particular, referring to the discussion above, a photon in flight can momentarily become an electron-positron pair, but very quickly the pair must recombine again into a photon. This possibility is shown in the figure below.



A particle in an inadmissible state of energy and/or momentum is called a **virtual** particle. Because the relation between energy and momentum is not that of a free particle ($E = mc^2$ for a particle of zero momentum) such a particle is said to be "off mass-shell". At this point we may recall Chapter 4 where the concept of mass shell was extensively discussed. Particles off mass-shell, virtual particles, are parts of diagrams, and we may even have some intuitive feeling about them, but we should never make the mistake of treating them as real particles. They occur as intermediate objects in a calculation, in a diagram, but they cannot be observed directly. They are like the photons in the two slit experiment. They move from light source to screen, and one may ask through which slit they pass. That, however, is a senseless question that can never be answered. We are not even sure if those photons actually go through any of the slits. That is the philosophy of quantum mechanics, and you better get used to it. Here we have diagrams and we can make calculations; it is like using wave theory to compute the interference pattern on the screen. But be careful not to think too much of the virtual particles as real objects. Still, within limits, it is helpful to think of them as a variant of the particles that they represent. Consider a virtual particle as a sort of calculational help. It makes you understand processes in a more intuitive way, and that is the path that we shall take.

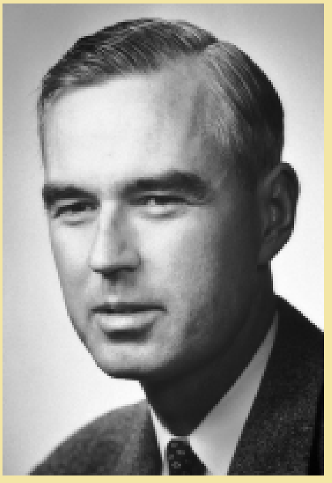
Keeping the above remarks in mind, and using the language of virtual particles, it follows that a photon, for a small fraction of time, can become a virtual electron-positron pair. This is actually of some consequence if we let another photon cross the path of the first one. On the level of diagrams new possibilities arise. If the second photon catches the first one in a dissociated state, it could be absorbed by one of the two virtual particles, to be emitted again by the other. The figure below shows the diagram.



As we know precisely the quantitative value of the photon-electron coupling (the previous diagrams), and also know the quantum-mechanical behaviour of these particles, the probability of the process can be calculated. It would be observed experimentally as the scattering of light by light. You might say that it is still possible to have interactions with virtual particles. The effect is quite small, so you cannot observe it by crossing the beams of two flashlights. Nonetheless, it has been observed. The classical (Maxwell) theory of radiation does not allow such a process. It is a purely quantum-mechanical effect.

The effect just described is somewhat similar to the so-called tunneling effect, well known to students of quantum mechanics. A particle may cross an energy barrier even if it has not enough energy to go over the top. An electron could cross a mountain even if it had not enough energy to get to the top. It may "tunnel" through. The tunnel, however, should not be too long; the probability for this to happen goes down very quickly as a larger distance must be covered. Here again there is the question. If the electron is initially on one side of the tunnel and finally at the other side, it seems only natural to say that "it has passed through the mountain". Such a statement is however beyond the limits of quantum mechanics. There is no way to establish if the electron actually ever was halfway in the mountain. The electron is there a virtual electron. The moment that you try to locate it (analogous to establishing through which slit a photon passes) the effect disappears. In an intuitive sense the electron passed through the mountain, and you may use that picture to devise experiments, such as sending other electrons from other directions, having them influence one another inside the mountain. In other words, in some sense interactions between virtual particles are quite possible. What you observe, however, are the initial and final configurations, never the intermediate virtual particles.

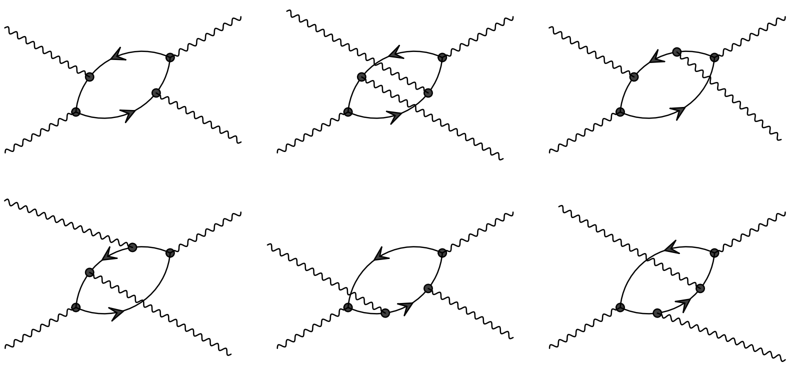
One more effect must be discussed, namely interference. Light interferes with itself, and this is a property of particles as well. The way this works is that for a given situation there may be



Willis Lamb (1913). In a splendid series of experiments Lamb found a discrepancy in the spectrum of hydrogen, as compared to the theoretical predictions of quantum theory. The discrepancy, now called the Lamb shift, was reported at a conference on Shelter Island (near Long Island) in June 1947. Kramers, successor to Lorentz in Leiden, the Netherlands, had been aware of the possibility of such effects, and lectured about his insights. The participants then recognized that the effect was due to higher order effects of quantum field theory, not taken into account up to then because people did not know how to handle the infinities of that theory. In his lecture Kramers also came up with the idea of renormalizability. The classic example is that of a small ball moving through water; a thin layer of water attaches to that ball and moves with it, thereby effectively increasing its mass. Likewise, the electron mass supposedly derives partly from the energy contained in the electric field around that electron. Kramers suggested that one must clearly separate the “bare” mass (the mass of the electron not including the contribution of the field) and the physical mass, that what you see experimentally. The bare mass is in fact not observable, and one simply chooses it in such a way that after addition of the field energy the observed mass value comes out. That the calculation of the field energy produces infinity is regrettable, but by choosing a bare mass of minus that same infinity (plus something extra) the correct experimental result can be reproduced. That is the kind of thing theorists do: sweeping infinities under the rug, smuggling them away. Ugly as it is, the theorists present at Shelter Island (including Feynman) followed Kramers’ suggestion, and produced a calculation of the Lamb shift that agreed with the experimental results of Lamb.

more than one way to go from a given initial state to a given final state. These different possibilities may interfere, either constructively or destructively. That is certainly something in which elementary particles seem to differ from billiard balls or cannon balls. In actual fact, the laws of quantum mechanics apply equally well to macroscopic objects; the point is that for the latter the effects become too small to be observable. Imagine a machine gun firing point-like bullets at two slits; the interference pattern on the screen would be incredibly small, the distance between the top in the middle and the adjacent peaks would be something like 10^{-37} m. That is much, much smaller than the size of a nucleus!

In calculations with particles the theorist draws as many diagrams as applicable (i.e. diagrams with the same initial and final configuration), writes down the corresponding mathematical expressions and sums them up. The different possibilities may add up or subtract, i.e. interfere. Only from this sum total can the probability of the process happening be calculated (effectively by squaring it). For example, to compute light-by-light scattering one must consider six diagrams, and combine their contributions. The figure below shows the possibilities.



All of these diagrams correspond to contributions of possibly different sign, and these contributions interfere. After taking all these contributions together the result must be squared and that is then a real probability. Never, but never could one say that one or the other possibility as represented by the different diagrams was what actually happened. Again, that would be like asking through which slit the photon passed.

9.3 Infinities

Where life becomes difficult is implicit in these diagrams. Not only must they be summed over all the different configurations, but over all the different energy-momentum values of the virtual particles as well. Consider for example again the temporary transformation of a photon into an electron-positron pair. The virtual electron or positron of this pair can have an infinite range of possible energies, including also negative energies. For example, the electron may be very energetic, while the positron would have very negative energy. The particles are then very far off mass-shell. The total energy must of course be equal to the energy of the photon, energy conservation being strictly enforced by Nature.

In calculating a process one must sum over all the possibilities. One must take all virtual configurations into account, no matter how much off mass-shell. That leads often to a hard calculation. Moreover, sometimes the summation gives an infinite result. It is a question of magnitude of contributions. If the configurations with an electron-positron pair of very high energy (very high negative energy for one of them and very high positive energy for the other) keep on contributing as much as configurations with low energy electron-positron pairs, then there is simply no end to the summation. The central question then is to what extent configurations containing virtual particles very far off mass-shell keep on contributing in a sizable manner as compared to contributions of configurations very nearly on mass-shell. One may put it in the following way. Normally contributions are smaller, damped, as the particles are more off mass-shell. The crucial thing is the amount of damping. If there is no or too little damping one is in trouble. **This, in a nutshell, is the problem of infinities in quantum field theory**.

Let us try to give an example for the case of sound. Imagine yourself standing in a crowd, hearing the conversations around you. Normally conversations held by people at some distance will not bother you very much because the volume of sound decreases with the distance. Thus there is a damping factor associated with sound generated at some distance. Imagine now that there would be no such damping factor, that a conversation far away would be heard by you as strongly as a conversation nearby. That would be horrible. You would hear a conversation between two Chinese in Beijing and two Russians in Moscow as strongly as a nearby discussion. Clearly, you would go mad. The noise would be unbearable.

Actually, in a large room filled with people, the sound volume would go down with distance, but there is the opposite effect of there being more people in a larger circle around you. That would more or less overcome the distance damping effect. In a large room full with people the total amount of noise may be very large indeed, and you may have to shout to carry on a conversation yourself! And if the room were infinitely large you might go deaf.

Several factors affect the occurrence of the infinities mentioned. To begin with, the more a particle is away from its mass shell the shorter is the time it is allowed to exist in that state. Consequently there is normally a damping factor associated with the occurrence of any virtual particle. This damping is stronger as the particle is more virtual. Furthermore, the damping is also a function of the intrinsic properties of the particle (more about that below). Another factor is the behaviour of the vertices, i.e. of the coupling, as a function of the energies of the particles involved. By and large these couplings have no strong energy dependence, although there are exceptions.

A difficult point is the behaviour of virtual particles as function of their intrinsic properties. The main property in this respect is the "spin" of the particle. One of the very surprising discoveries in the domain of quantum physics was the discovery that particles have an intrinsic angular momentum, as if they were spinning around an axis. For a body of some size, like a billiard ball, that is easy to imagine. But for a particle that for all we know has no size that is very hard to imagine. Yet it is there, and every elementary particle has a very definite spin as this intrinsic angular momentum is called (it may be zero). When particles are created or absorbed the interaction is always such that angular momentum is conserved. If a spinning particle enters an interaction then the angular momentum is preserved throughout, and it appears in the final state either in the form of other spinning particles, or else through non-spinning particles that revolve around each other, or both. All this is quite complicated, but fortunately we need only a few facts related to this. Spin is measured in terms of a specific basic unit, and spin is always a multiple of $\frac{1}{2}$ in terms of that unit.

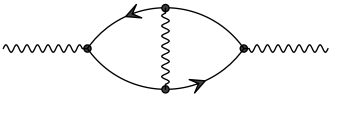
As it happens, no elementary particle observed to date is of the spin zero variety. The so far hypothetical Higgs particle has spin zero. Most particles have spin $\frac{1}{2}$, and the remainder have spin 1, except for the graviton (the particle responsible for gravitational interactions similarly to the photon in electromagnetism) that has spin 2. Here now is the important property relevant to our discussion about virtual particles: as their spin becomes higher, virtual particles are less damped at higher energy. Particles of spin 1 are barely damped at high energy in their contributions to a virtual process. Particles of spin 2 are even worse: the quantum theory of gravitation is in a very poor shape. Quantum field theory for particles of spin 1 (with the exception of the photon) was not part of our understanding of Nature up to 1971. No one knew how to handle the virtual contributions. They invariably led to infinite sums.

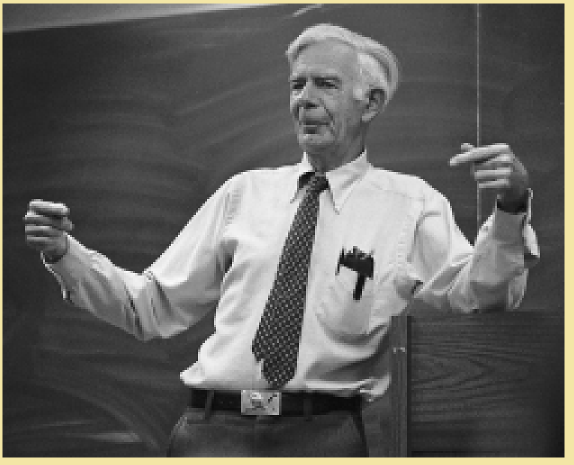
Even if the photon has spin 1, and thus has not much of a damping factor associated with it, there is still effective damping due to the way that different diagrams tend to compensate each other. As a consequence the theory of electrons interacting with photons became manageable, using the renormalization technique discussed below. In 1948 Feynman, Tomonaga and Schwinger worked out the theory for which they received the Nobel prize in 1965. However, the weak interactions, involving other spin 1 particles, remained intractable.

What changed the situation for weak interactions was the discovery that the worst effects in individual diagrams can, theoretically, be cured by introducing new interactions and particles (and hence new diagrams) in such a way that in the sum total the bad parts cancel out. Thus new particles were introduced into the theory. That will be discussed at length later on. For the photon a similar mechanism, involving cancellations between different diagrams but without the introduction of new particles, was, as mentioned above, partly understood since 1948. Using the method of renormalization quantum electrodynamics produced finite numerical results that could be compared with experiment. One of the most important results is the magnitude of the magnetic moment of the electron. Any charged particle with spin usually has a magnetic moment, which one might consider as a consequence of the spinning charge. The predictions of quantum electrodynamics for the magnitude of this magnetic moment have been verified to a truly fantastic degree of accuracy. But before discussing this we must first fill in some gaps, and explain about perturbation theory.

9.4 Perturbation Theory

As we have pointed out, one must sum up all possibilities when considering any process. That includes summing over all energy/momentum distributions of the virtual particles. However, also additional emissions/absorptions of virtual particles must be taken into account. The figure shows an example: the virtual electron emits a photon which is absorbed by the virtual positron (or the other way around).





Horace R. (Dick) Crane (1907). Crane, a professor at the University of Michigan, discovered a very sensitive method to measure the magnetic moment of the electron. In certain units, the value of this magnetic moment (usually called g) is 2 if no radiative corrections are included. When an electron moves through a magnetic field its magnetic moment (pointing in the same direction as its spin) will not change relative to the direction of motion along the trajectory if indeed this g is exactly 2. However, any deviation of the magnetic moment from that value will cause the electron spin to rotate. By measuring the amount of rotation an accurate measurement of the difference between the actual value and the value 2 can be achieved. In this way one measures pure and simple the quantum corrections that cause the magnetic moment to have a value different from 2. This method is called the $g - 2$ method. The theoretical calculations and the various experiments on this magnetic moment have occupied many physicists in the course of time.

Crane is a modest man who is not given to advocate his own achievements. Things being what they are in this business he did not always get the proper credit for his wonderful idea. He did get the US presidential medal, a very high US distinction.

On retirement he decided to help the Ann Arbor Hands-on museum (an initiative of Cynthia Yao, wife of a theorist at the University of Michigan). He managed to keep most items exhibited in working order, which is a must for survival of any museum of this type. How Crane did it is a mystery to me, because visitors, often including whole bus loads from schools, are not particularly careful.

In his spare time Crane grows orchids.

Richard Garwin (1928, left) and **Valentine Telegdi** (1922). These two physicists did pioneering experiments (each a different one, with other collaborators) on parity violation after the theoretical analysis of Lee and Yang. They showed that the muons in the decay of the pion ($\pi \to \mu$ + neutrino) were polarized. This implies parity violation in pion decay.

Concerning $g - 2$, after the original proposal of Crane, physicists were quick to realize that the method could be very suitable to measure the anomalous magnetic moment of the muon. The muon decays in a way that depends on its spin orientation (also this indicates parity violation), and by recording the decay products the direction of the spin (and thus the magnetic moment) could be measured relatively simply. At high energy the muon lives long enough to make it traverse a substantial distance.

On the theoretical side the necessary theory (not the quantum corrections) was first formulated by Ken Case, a theorist of the University of Michigan. A subsequent classic paper by Bargmann, Michel and Telegdi became the standard concerning the treatment of spin.

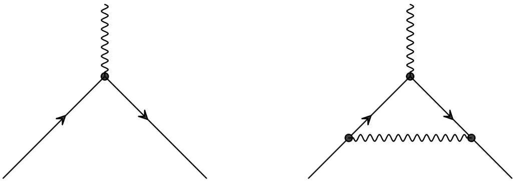
At CERN a group of physicists mounted the first muon $g - 2$ experiment, in 1959. In 1961 a result was published (Charpak, Farley, Garwin, Muller, Sens, Telegdi, Zichichi) with an accuracy of 1.9%, agreeing perfectly with theory. Since then, using muon storage rings at CERN and Brookhaven, this accuracy has improved greatly.

Telegdi is a long-time friend of mine who helped me enormously with this book. Of Hungarian origin, he became a stateless person at the end of World War II and traveled around with a self-made passport.

Here we have another complication of quantum theory: there is no end to this chain. One can exchange one photon, two photons, whatever number, and to boot any of these photons can momentarily become an electron-positron pair, etc. It looks hopeless. How can one calculate all these diagrams?

As luck has it, there is in many cases no need to consider all of these possibilities. The reason is that there is a factor associated with any vertex and that factor, at least for quantum electrodynamics, is quite small: the electric charge. The emission or absorption of a photon by an electron (or positron) is proportional to the electric charge of the electron. Indeed, if the electron had no charge it would not interact with the electromagnetic field. For this reason, a diagram as shown above with an additional photon exchanged between electron and positron gives a contribution that is down by a factor e^2, where $-e$ is the electric charge of the electron. In practice there are some additional factors, and the relevant dimensionless quantity is what physicists call the fine-structure constant, $\alpha = e^2/4\pi\hbar c$. Numerically $\alpha \approx \frac{1}{137}$, so that a diagram with an extra photon exchange indicates a contribution of the order of 1% as compared to that of the diagram without that photon exchange. So if we restrict ourselves for a given process to diagrams with the least number of vertices we may expect an answer that is accurate to 1%. And if that were not enough we can include diagrams with two more vertices and get an accuracy of 0.01% (i.e., 1 part in 10^4).

Here we see a fact of field theory: it is **perturbation theory**. Rarely, in fact never, can we compute things exactly but we can approximate them to any desired precision. That is of course true assuming we can find our way though the maze of summations (over energy/momentum distributions) that arises when considering a diagram with many virtual particles. The calculation of the magnetic moment of the electron is in practice perhaps the most advanced example. The electron, possessing spin, has like any rotating charged object a magnetic moment. In other words, the electron has not only a charge, but it also acts as a tiny magnet as well. That part of the interaction of an electron with the electromagnetic field is also subject to quantum corrections, and the figure below shows the lowest order diagram and a next order (in α) diagram.

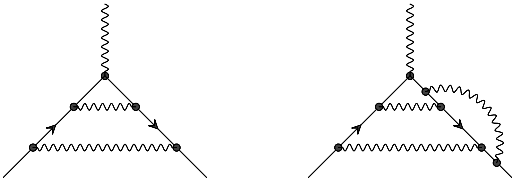


In suitable units the magnetic moment of the electron, disregarding quantum corrections, is 1 (in another context units are often chosen such that it is 2). The second and higher order contributions alter that magnetic moment by a tiny amount; to give an idea about the accuracy achieved we quote here the theoretical result for this anomalous magnetic moment (including fourth and sixth order contributions as well):

$$a_e = 0.5\left(\frac{\alpha}{\pi}\right) - 0.328478965\left(\frac{\alpha}{\pi}\right)^2 + 1.181241456\left(\frac{\alpha}{\pi}\right)^3 - 1.4\left(\frac{\alpha}{\pi}\right)^4$$
$$= 0.001159652201$$

as compared to the experimental value 0.001159652188. Note that $\alpha/\pi \approx 0.00232$. The error margins in both the theoretical and experimental values are of the order of the difference between the values quoted here. In other words, the agreement is excellent. Also the anomalous magnetic moment of the muon has been computed with great precision: $a_\mu = 0.00116591849$ (there is an uncertainty of 70 in the last four decimals), to be compared with the most recent experimental value $a_\mu = 0.00116592030$ (uncertainty 80). The sophistication involved in both theory and experiment is mind boggling. The calculation of the coefficient of α^3 has taken some 20 years, involving some 72 diagrams, while the calculation of the α^4 term (891 diagrams) has been done mainly by numerical approximation methods, using up years of super-computer time. Any experiment achieving an accuracy of one part in a thousand is already difficult, let alone the experiment relevant here, having an accuracy of order of one part in 10^6. The most spectacular experiment for the electron is based on measurements performed on a single electron, caught in an electromagnetic trap (Dehmelt, Nobel prize 1989). For the muon things are slightly different, because it happens that with this accuracy, for the muon, diagrams involving a W or a Z^0 or even quarks must be included. That makes it even more interesting. At the time of this writing the latest measurement of the magnetic moment of the muon, quoted above, was done at Brookhaven using a muon storage ring.

Here a remark concerning the way theorists talk about these things. They usually classify the subsequent orders of perturbation theory by means of loops. The lowest order diagram is called a tree diagram (no loop), the next order diagrams have one-loop, the next order two loops etc. The figures below show examples of a two-loop and a three-loop diagram.



The indeed amazing agreement between theory and experiment, involving these very complicated quantum effects, must be seen as a strong support for the theoretical insights as well as for the validity of perturbation theory. Many theorists would have liked a formulation of the theory not involving approximations, but so far perturbation theory is all we have. In certain instances one has been able to sum up the contributions of some classes of diagrams to all orders, but we do not have any general non-perturbative version of the theory. This is a fact of life. Let us be happy with the notion that we know at least the basic ingredients that form the basis of the quantum mechanical calculations.

In my own scientific life this wonderful agreement between theory and experiment has played an important role. It made me deeply conscious of the fact that diagrammatic methods and perturbation theory worked very well, and this stimulated me to continue using these techniques even in the dark times in the middle sixties when false gods were dominating particle theory. Now, of course, with the Standard Model we can apply these methods all over the place.

9.5 Renormalizability

As indicated earlier, there is a problem related to the non-convergence of the summations over all possible distributions of energy/momentum of the virtual particles. In certain cases these sums do not converge, i.e. the result is infinite. That stopped progress for quite some time, until, in about 1948, the idea of **renormalization**, due to Kramers, solved the problem at least on a practical level. The subject theory was quantum electrodynamics, and it was noted that the infinities occurred only in some well defined instances. For example, in the calculation of the magnetic moment of the electron discussed above they did not occur. But the reader will realize that these very same diagrams, which alter the magnetic properties of the electron, will equally well alter the electric charge of the electron, as they simply affect the way a photon interacts with the electron. That change of the electric charge turns out to be infinite. Here then is the big idea: the electric charge as actually observed is the sum total of the basic electric charge (as occurring in the tree diagram) plus the contributions of all higher order diagrams. But we have no idea how large that charge (the basic charge) is without the corrections due to these higher order diagrams.[b] So, let us choose the value of the basic charge such that the total charge comes out equal to the experimentally observed value. In other words, we give the basic

[b]Such corrections are called radiative corrections.

charge a value that has an infinite part as well, but opposite to that part of the higher order corrections, making the sum come out equal to the observed value! Speaking of dirty tricks!

This trick, hard to swallow at first, works very well indeed. The basic fact is that the infinities occur only in conjunction with the free parameters of the theory. A free parameter is a parameter for which there is no theoretical prediction. Any theory has some of these parameters. For example, Newton's theory of gravitation has the gravitational constant. It fixes the strength of the gravitational force. You can determine it by working out one case, for example the orbit of the earth. Basically, the constant is determined from experiment. There is no theory that says how big Newton's gravitational constant should be. It is a free parameter, to be determined from experiment.

The electric charge of an electron is also a free parameter. It is an input to the theory, not something that we can compute. Another such parameter is the mass of the electron. It is not known from any basic principle, and its value must be obtained by measurement. That gives us an opportunity to hide an infinity. Experimentally one observes the basic value of the parameter, but theoretically, what is observed is some input value plus the contribution of many diagrams. The important thing here is that what we observe experimentally includes contributions of all kinds of diagrams. At that point one says: whatever the contribution of diagrams, it goes together with the basic value and the only thing we know is the combination, observed experimentally. Thus if the diagrams give an infinite contribution let us make the basic value also infinite, but with the opposite sign so that the combination comes out to the experimentally observed result. Nonsense minus nonsense gives something ok.

This scheme for getting rid of infinities is called renormalization. It is by itself far from satisfactory. No one thinks that the basic quantities are actually infinite. Rather we believe that the theory is imperfect, but that this imperfection can be isolated and at least for the moment swept under the rug. The miracle is that for quantum electrodynamics all infinities can be absorbed into the available free parameters. So, apart from these infinite corrections to free parameters, everything else (such as the magnetic moment of the electron quoted above) is finite, and insofar as checked, agrees well with experiment.

So here we are. We have a theory imperfect on several counts. First, the theory is only perturbative. Second, infinities occur, even if they can be isolated and hidden. In spite of these imperfections, all this leaves us with a scheme that makes accurate predictions that can be compared with experimental results.

There are also theories such that infinities occur not only in conjunction with free parameters. Such theories cannot make solid predictions. They are called **non-renormalizable**. For a long time theories involving vector particles (spin 1) other than the photon were thought to be of that type, and as such useless. This picture has changed, and we now also have renormalizable theories involving spin 1 particles. These theories are called **gauge theories**.[c] Strictly speaking the name gauge theory refers not to renormalizability but to some mathematical property, which in the case of spin 1 particles leads to a renormalizable theory. There are gauge theories that are not renormalizable, gravitation (involving a spin 2 particle) being one of them. The name **Yang-Mills** theories (named after the inventors, C. N. Yang and R. Mills) refers more narrowly to gauge theories with particles of spin 1.

As a matter of fact, almost all interactions seen in experiments are of the renormalizable type, gravitation being the exception. Quantum effects in gravitation have so far not been understood. That casts a shadow on that theory and its consequences, like for example black holes.

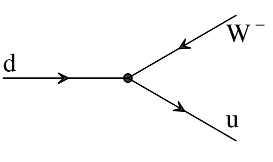
[c]Historically the name gauge was introduced in another context, namely gravitation. It referred at the time to a transformation, a gauge transformation, that changed the length scale. There are similar transformations in the theories that we discuss, but they do not scale length. Nonetheless, the name has stuck.

9.6 Weak Interactions

Weak interactions constitute a different type of interactions, one that does not produce long-range forces such as those in electrodynamics. A photon has zero mass, and can hence have arbitrarily low energy. For this reason a virtual photon of zero energy and small momentum is only a tiny bit "off mass-shell", and little damping is associated with the exchange of such a virtual photon. It is this type of zero energy, low momentum photons that are responsible for long-range electromagnetic interactions. For the same reason the graviton will also give rise to a long-range force. These, however, exhaust the list of long-range forces that we experience in daily life. The weak interactions have a very short range; let us discuss them in some detail.

Weak interactions made their entry into physics through the discovery, by Becquerel in 1896, of β-radioactivity. Experimentally and theoretically it took a really very long time before these interactions were understood, even on a purely phenomenological level. Since it is not our purpose to present here the history of the subject, we shall straightaway describe things as they are understood today.

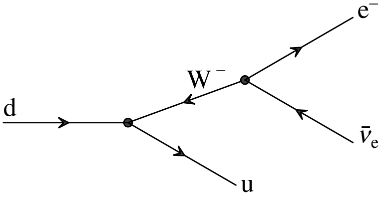
Consider the most fundamental nuclear β-decay, that of a neutron into a proton and an electron (plus an antineutrino). The neutron contains two down quarks and one up quark (denoted by d and u respectively), the proton one d and two u quarks. As a first step, one of the d quarks in the neutron decays into a u quark and a negatively charged vector boson, denoted by W^-. The figure below shows the diagram representing this decay.



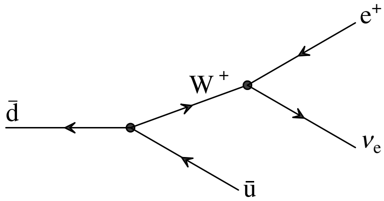
It contains one of the basic vertices of weak interactions. The associated coupling constant ("weak charge"), usually denoted by

g, is somewhat larger than the corresponding one for electromagnetism. Experiment shows that $\alpha_w = g^2/4\pi\hbar c = 1/32$. At this point we have a notational problem, because all particles in this reaction are charged (in terms of a unit such that the charge of the electron is -1, the charges are $-\frac{1}{3}$, $+\frac{2}{3}$ and -1 for d, u and W^-, respectively), and the arrow no longer represents the flow of negative electric charge. Instead it will be used to distinguish particles and antiparticles, where an arrow pointing opposite to the flow of energy indicates an antiparticle. Here there is a choice: is the W^- a particle or an antiparticle? Historically, there was the regrettable mistake of having defined the charge of the electron as negative. We shall not do that here for the W, and define the W^- to be the antiparticle. That is why the arrow in the W-line points inwards.

The W is very massive (80.3 GeV, as compared to the proton mass, 0.938 GeV), and given the low mass of the d quark (about 10 MeV = 0.010 GeV) it must be very virtual (way off mass-shell) in the actual process of d decay. In a second step it must quickly transform into an electron and an antineutrino, an interaction which is another basic vertex of the theory. The figure shows the complete diagram for d decay. Let us note here that an antineutrino is not the same as a neutrino, despite the fact that the neutrino has zero electrical charge.



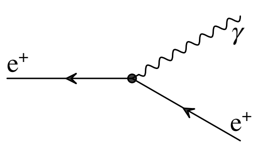
As noted before, the existence of a negatively charged W^- (which we will take to be an antiparticle) implies the existence of a particle with opposite charge, the W^+.



This W^+ would, for example, be involved in the decay of an antineutron into an antiproton, a positron (anti-electron) and a neutrino; that reaction is simply the same reaction as the one discussed above, with all particles replaced by their antiparticles (reversal of all the arrows).

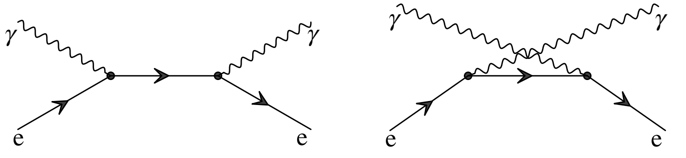
The W has spin 1 and its interactions generally lead, due to the absence of damping at high energies, to a non-renormalizable theory as we discussed before. We shall discuss this point in a systematic way, showing how the situation can be salvaged, and how new particles and interactions must be introduced in order to achieve a tolerable high-energy behaviour.

Let us point out that we have quietly introduced another property of Feynman diagrams. First there was the crossing property discussed before. Lines may be moved from in to out or vice versa, and that gives new possible processes. The point introduced above is this: if in a given diagram all arrows are reversed then another process results, and this new process can also occur in nature. For example, the very first diagram in this Chapter, an electron emitting a photon, when treated this way becomes a positron emitting a photon, see figure below. Once more: the arrow has nothing to do with the movement of the particle; the fact that the arrow points opposite to the movement of the particle (the flow of energy) means that we are dealing with an antiparticle.



9.7 Compton Scattering

Exploring high-energy behaviour can conveniently be done by considering certain simple processes. One must estimate the energy dependence of a process where particles of very high energy are scattered. The simplest and most important example is Compton scattering, that is the scattering of a photon incident on an electron. In lowest order only electromagnetic interactions are of relevance here, and the figure shows the two diagrams that contribute at the lowest level.

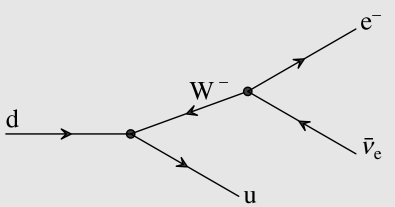


Strictly speaking much of what follows below is true only if photons have non-zero mass, but we shall ignore this subtle point for the sake of simplicity.

The central issue is the behaviour of the theory as dependent on the energy of the particles. It is a necessary property that the theoretical behaviour must not be such that the probability of the process to occur increases if the energy of the incoming particle (in this case the photon) increases. Then that probability would for sufficiently high energy become larger than one, which is nonsense and certainly not acceptable for a physical theory.

We now state a general property of the diagrams shown. Each of these diagrams alone would produce a bad theory. Taking into account only one of the two diagrams, the probability for the process to happen will increase indefinitely if the energy of the photon increases. And that is unacceptable. However, in this case the two diagrams combined produce an acceptable result, the probability is constant if the energy of the photon goes up. The two diagrams compensate each other. This is the wonderful thing: diagrams may compensate each other's bad behaviour. And indeed, that is what does the trick for quantum electrodynamics.

The high-energy behaviour of diagrams as shown above and similar diagrams with different particles can be guessed as follows. We are not explaining anything here, just stating the rules. An incoming or outgoing photon (vector particle, spin 1) contributes a factor proportional to the energy E of that photon. A virtual photon contributes no energy dependence, i.e. it must be counted as a constant. A virtual electron, or generally a virtual spin $\frac{1}{2}$ particle behaves as $1/E$. An incoming or outgoing electron (spin $\frac{1}{2}$ particle) must be counted as $\sqrt{E}$. A virtual scalar particle (spin 0) must be counted as $1/E^2$, an incoming or outgoing spin 0 particle contributes a constant. It must be noted that in special cases the energy dependence might be different from the dependence that one would deduce by counting with these rules; whether or not it does depends on details of the actual couplings which sometimes compensate the aforementioned energy dependence related to the magnitude of the spin. A case in point, one that we shall meet in the following, is the coupling of a vector boson to two real (i.e. not virtual) particles. In the case of gauge theories the vertices are always such that spin effects are neutralized in that instance, i.e. for a (possibly virtual) vector boson coupling to two real particles one can ignore the spin effects for this vector boson. An example is the decay of the down quark, corresponding to a diagram that we will show again.



Here the virtual W^- is coupled to real particles on both ends. In that case the energy dependence relating to that virtual W is as that for a scalar particle, i.e. as $1/E^2$. For the readers' convenience, we summarize the various factors given above in the following table.

spin	in/out	virtual	ends*
0	1	$1/E^2$	$1/E^2$
$\frac{1}{2}$	$\sqrt{E}$	$1/E$	$1/E$
1	E	1	$1/E^2$

*Behaviour of virtual particles with real particles attached at both ends and coupled according to a gauge theory

In a renormalizable theory the probability of a process to occur must, as a function of energy, either decrease or at worst tend to a constant value. Even on a purely intuitive level a probability increasing indefinitely as the energy of the incident particle increases is hard to accept. Note that the probabilities (cross sections) are obtained, as mentioned earlier, by squaring the contributions of the diagrams. Counting the expected behaviour for the diagrams shown above for Compton scattering we arrive at a result increasing with energy, in fact as E^2 (so the cross section would go as E^4). This then is unacceptable. However, the second diagram shows also a leading dependence proportional to E^2 but it turns out that it has the opposite sign, and the sum of the two actually behaves as a constant. A somewhat simplistic explanation for this difference in sign is that the intermediate (virtual) electron in the first diagram has a large positive energy, in the second a large negative energy. Thus the factor $1/E$ for the intermediate electron has the opposite sign for the two diagrams.

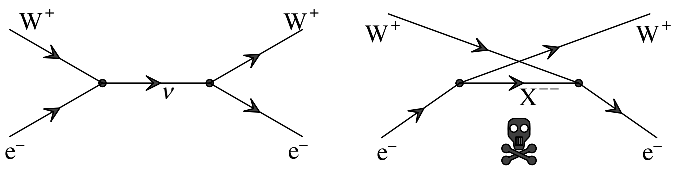
Cancellations of bad behaviour between diagrams is the idea behind gauge theories, of which quantum electrodynamics is the simplest example. Individual diagrams give unacceptable energy behaviour, but everything is arranged in such a way that in the end the bad behaviour cancels. It is a complicated game of cancellations, requiring from time to time the introduction of new particles. Experiment must then verify the existence of those new hypothetical particles having the desired properties. Let us see if we can make this work for weak interactions that involve charged vector bosons.

9.8 Neutral Vector Bosons

In the following we shall ignore electromagnetic interactions, giving rise to diagrams containing a virtual photon in some of the situations discussed below. They are not essential to the reasoning.

Let us first examine W^+ scattering off an electron. The corresponding lowest order diagram is the first one in the figure below. There is a virtual neutrino mediating this process. The behaviour at high energy is bad.

> The high energy behaviour guessed by power counting as specified above is bad, namely as E^2 (E for each of the W's, $\sqrt{E}$ for each of the electrons and $1/E$ for the intermediate neutrino). As the W's are connected to vertices of which one of the particles (the neutrino) is virtual there are no special compensating effects.

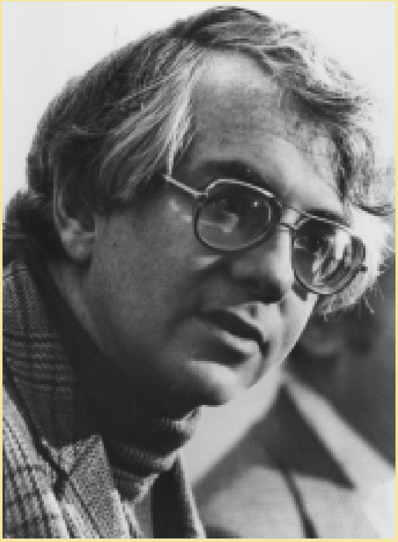


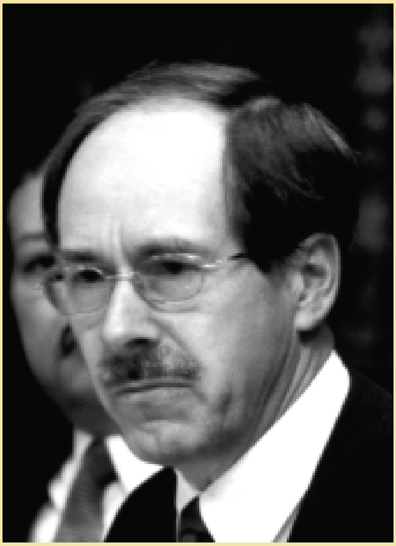
Recalling the case of Compton scattering we might think that the situation can be cured by another process, the one shown in the second diagram in the figure above. Since now the incoming electron emits a positively charged W, the intermediate particle (named X^{--} in the diagram) cannot be a neutrino because charge conservation (a law that holds rigorously) requires the intermediate particle to have charge -2. But no such particle is known. The diagram does not exist. That is why we put a skull and bones below it. What now?

Steven Weinberg (1933) and **Martinus Veltman** (1931). Weinberg and I do not see eye to eye on certain issues. The picture above, taken at the occasion of the 1999 Nobel week, shows Weinberg offering me some explanations that I found difficult to swallow. In 1967 he wrote his most famous paper (for which he was awarded the 1979 physics Nobel prize, together with Glashow and Salam), but up to 1971, at which time the mathematical consistency of his model became clear (after the Amsterdam conference) it was largely ignored. It did not help that in this paper only one experimental consequence was mentioned, followed by the remark that it should not be taken very seriously. Indeed, nobody did. Weinberg refers to the period 1967–1971 as the period that his paper lay dormant. It is now, I believe, the most cited paper in particle physics, followed by the paper of Kobayashi and Maskawa (see Chapter 3).

In 1972 at a conference at Fermilab, Ben Lee, reporting on theory in a session entitled "Perspectives on theory of weak interactions", pulled Weinberg's paper out of obscurity. Ben Lee's talk was very important, as it explained many of the facets of gauge theories to a large audience; not long after that neutral currents (a consequence of the existence of the Z^0) were established by a neutrino experiment at CERN using the gigantic French bubble chamber Gargamelle.

Weinberg's paper contains one of the ingredients that made the Standard Model what it is today. He invoked Higgs forces that we now know to be necessary for mathematical consistency, adding them to the model proposed earlier by Glashow.





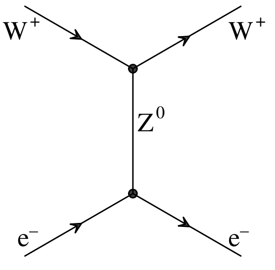
Sheldon Glashow (1932) and **Gerardus 't Hooft** (1946). Glashow is certainly one of the main contributors to the Standard Model. His 1963 paper specified the interactions between leptons and vector bosons (W and Z^0), thereby introducing the Z^0 and moreover the mixing of the photon with the Z^0. That paper was Weinberg's starting point. The idea of lepton-hadron symmetry of Gell-Mann (see the end of Chapter 8) was implemented by the Japanese physicist Hara, who introduced a fourth quark next to the up, down and strange quarks proposed by Gell-Mann. That new quark is now called the charmed quark. Glashow, together with Iliopoulos and Maiani, spelled out the interaction of these four quarks with the vector bosons (without a Higgs though). Among others that work explained the hitherto mysterious absence of certain decays of the K-mesons. One speaks of the GIM mechanism.

Also reactions of neutrinos without production of muons or electrons (neutral currents) were discussed. In 1973 neutral current type events were seen in the French bubble chamber Gargamelle, convincing many physicists of the correctness of that part of the Standard Model.

In 1968 I convinced myself of the importance of gauge theories, and made substantial inroads in this complicated mathematical subject. At some point 't Hooft became my PhD student and he then did his work that completed the mathematical understanding of those theories. He delivered a splendid piece of work, and at the time I was very happy with that and proudly introduced him to the physics community at the 1971 Amsterdam conference. Being one of the organizers I was left at liberty to arrange a session of that conference.

The solution is to introduce another vector boson, this time one without electric charge. We assume that it couples to the charged vector bosons and the electrons and a new diagram of the third type is then possible, see below. The coupling of this new neutral particle to the charged vector bosons and the electrons is taken such that the high energy behaviour of the new diagram cancels the bad behaviour of the first diagram above.

The vertex must behave like E, and given that the intermediate vector boson is coupled on both ends to real particles we have indeed the required behaviour, E^2 (E for each of the charged W's, $\sqrt{E}$ for each of the electrons, E for the three vector boson vertex and $1/E^2$ for the intermediate vector boson). Choosing the right sign and magnitude for the coupling constants in the various vertices one may achieve cancellation, and the sum of the two diagrams behaves as a constant at large energies.

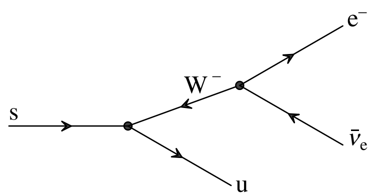


The price to pay is the totally ad hoc introduction of a new particle, a neutral vector boson. But here starts the triumph of gauge theories: a neutral vector boson with the required couplings has indeed been observed. It is commonly called the Z^0. Its mass, 91.187 GeV, is slightly higher than that of the charged W's.

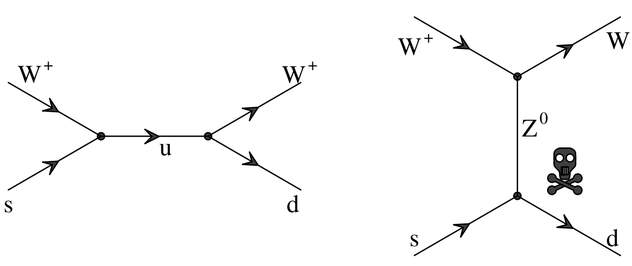
9.9 Charmed Quarks

Another interaction observed is the decay of the Λ. This neutral particle is very much like the neutron, except that it is heavier

(1116 MeV versus 940 MeV for the neutron). The Λ particle has one *d*, one *u* and one *s* (strange) quark, from which you may guess that the *s* quark is about 200 MeV heavier than the *d* quark. The Λ decays in various ways, there being more energy available, but one of its decay modes is quite analogous to neutron decay, namely decay into a proton, electron and antineutrino. That decay can then be understood as a decay of the *s* quark into an *u* quark and a W^-, with, as in the case of neutron decay (or rather *d* quark decay) a subsequent rapid decay of this W^- into an electron and an antineutrino. See figure below.



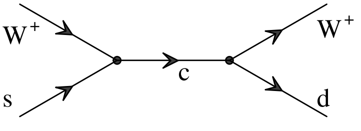
It is found that the coupling constant associated with the s–u–W vertex is smaller by a factor of about $\frac{1}{4}$ as compared to the coupling constant of the d–u–W vertex. This factor has been interpreted as the tangent of some angle, now called the Cabibbo angle θ_c; in that view, now generally accepted, the *d* quark decay has a factor $\cos\theta_c$ in addition to the coupling constant g, while *s* decay has a corresponding factor $\sin\theta_c$. It is interesting to see what must be done in order to achieve the proper high-energy behaviour for the scattering of a *W* from an *s* quark through the process as shown in the first diagram in the figure below. The



quark-W vertices of both Λ and neutron decay are involved here. Given that the quarks are spin $\frac{1}{2}$ particles, just like the electron, we guess just as before a bad high-energy behaviour for this process.

Again a compensating diagram must be found, and naturally the first thing that comes to mind is to try the same trick as found earlier, namely to introduce a diagram involving a Z^0, as shown in the figure above. This however fails. The reason is that there occurs here a new vertex, the $s-Z^0-d$ vertex, which, as far as experiment is concerned, does not exist. In the language of particle theorists this was called "the absence of strangeness changing neutral currents" (the s quark has strangeness, the d has not). How to repair this?

Well, the solution was to postulate yet another particle, one with properties close to that of the u quark, but with suitably chosen couplings to the W's. The figure shows the construction; it is completely analogous to the diagram with the intermediate u quark given above.



The new quark is called a "charmed" quark.

The coupling constant of the $s-c-W$ vertex is assumed to be like the one in neutron decay, i.e. with a factor $\cos\theta_c$, while for the $c-d-W$ vertex it is taken to be $-\sin\theta_c$. Due to this extra minus sign the diagrams almost cancel at high energy, and their sum behaves neatly like a constant.

As if asked for, experimenters discovered this c quark, or rather, since quarks are never seen singly, discovered particles that contained these c quarks. The mass of the c quark is of the order of 1500 MeV, much heavier than the masses of the d quark

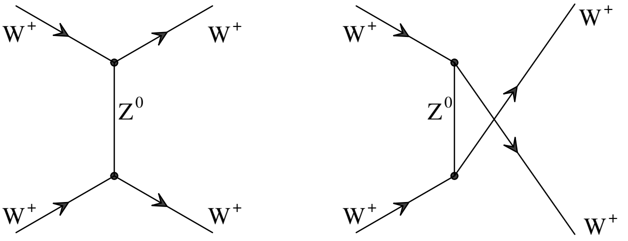
(7.5 MeV), the *u* quark (5 MeV) or the *s* quark (200 MeV). Even so the *c* quark is much lighter than the *b* (bottom, 5000 MeV) and *t* (top, 175,000 MeV) quarks found since then. Incidentally, these masses are not well established, in particular not the lighter ones, because quarks never appear singly, and there are substantial energies related to the binding mechanism peculiar to quark interactions. By necessity the quark masses must be derived from the experimentally accessible particles, quark composites, and that always requires elaborate arguments.

The discovery of the *c* quark was the second major victory for the gauge idea. Its couplings were found to be precisely as given above, involving the appropriate factors $\cos\theta_c$ and $\sin\theta_c$. But the story does not end here.

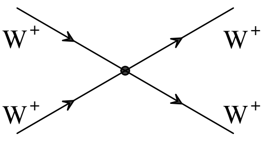
9.10 The Higgs Particle

We now turn to processes involving vector bosons only. Of course, our considerations are here of a purely hypothetical nature, since one cannot observe in the laboratory any of the scattering processes discussed here. Vector bosons have very short lifetimes and hence cannot in practice be used as projectiles or targets. Anyway, we shall now consider vector bosons scattering from each other, in particular W^+–W^+ scattering. There are two diagrams contributing to this process (see figure below), and the behaviour at high energy is really bad. That is because compared to previous cases there are now only spin 1 particles (as compared to the occurrence of spin $\frac{1}{2}$ particles in the previous diagrams), and higher spin makes for worse behaviour. Drastic remedies are needed.

Recall that the W–W–Z^0 vertex has an energy dependence (a factor E). The powercounting gives a behaviour as E^4: a factor E for any of the four external W's, a factor of $1/E^2$ for the intermediate Z^0 (it is coupled to vertices without other virtual particles), and a factor E^2 coming from the two vertices.



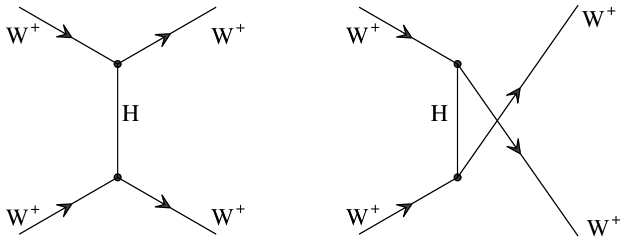
The worst part can be cured by introducing a totally new type of diagram, namely a direct four-W interaction as shown in the figure.



By carefully adjusting the coupling constant associated with this new vertex an almost complete cancellation can be achieved. However, there is still trouble remaining, although much reduced.

The vertex itself is not energy dependent, and the four W's give a factor E^4, so it has at least the required behaviour. The E^4 part can be cancelled. This one new diagram is not enough. A part behaving as E^2 remains. Rather, since its dimensions have to be the same as those of the E^4 part, it is a behaviour of the form M^2E^2, or $M_0^2E^2$ where M and M_0 are the masses of the W and Z^0 bosons, respectively. These masses are the only available parameters with the dimension of an energy. How to compensate the remaining E^2 part?

The solution is to postulate yet another particle, called the Higgs boson H. It is to be a spinless neutral particle, coupling to the W's with a strength so chosen as to produce the required compensation. The figure shows the two possible diagrams.



It turns out that the Higgs particle must be coupled to the vector bosons with a strength proportional to the mass of the particle that it couples to. This peculiar feature, typical for all the couplings of the Higgs particle, raises many interesting questions.

9.11 General Higgs Couplings

Is this the end of the story? Not quite. There remain many little problems of the nature sketched before, but it would carry us too far to enter into a detailed discussion here. Suffice it to say that the Higgs particle must also be coupled to the neutral vector boson (the Z^0) and to the quarks etc. as well. In short, it must be coupled to **any** particle having a mass. Moreover, the coupling must always be proportional to the mass of the particle to which it is coupled.

To date the Higgs particle has not been observed experimentally. Unfortunately the theory has nothing to say about its mass, except that it should not be too high (less than, say, 1000 GeV), or else its compensating actions set in too late. The present experimental lower limit for the Higgs mass is roughly 100 GeV. The new collider being built at CERN (the LHC, colliding protons each with an energy of 7000 GeV) might give us information on this Higgs particle. It will not be easy to produce Higgs particles, because the proton contains only u and d quarks, and these, because of their low masses, couple only weakly to this Higgs particle. Higher order processes, involving virtual (heavy) W's and Z^0's are needed to produce this particle.

The demonstration that all bad energy behaviour can be made to vanish with only those particles discussed above, and no others is usually referred to as the proof that this theory is renormalizable. On reading the previous discussion one may easily have the impression that there is no end to new hypothetical particles that must be introduced. But no, this is it! The Higgs particle is the last one needed.

It is perhaps necessary to state explicitly to what extent the discussion above reflects the historical development. We have sketched a theory involving many particles, with their interactions so orchestrated and tuned as to have a renormalizable theory. The result is a theory possessing a high degree of symmetry. The historical development was quite the opposite of that suggested by our treatment. The symmetry was discovered and investigated some 20 years before its consequence, a renormalizable theory, was finally understood.

9.12 Speculations

Because this Higgs particle seems so intimately connected to the masses of all elementary particles, it is tempting to think that somehow the Higgs particle is responsible for these masses. Up to now we have no clue as to where masses come from: they are just free parameters fixed by experiment. It requires no great imagination to suppose that the Higgs particle might have something to do with gravitation, and indeed, theoretical models suggest a strong involvement of the Higgs particle in the structure of the Universe, otherwise thought to be shaped by gravitation. Some theorists believe that the Higgs particle does not really exist, but that it somehow mimics a much more complicated reality, involving gravitation in a fundamental way.

These are very exciting and interesting questions and speculations. We are looking forward to LHC experiments, noting that so far theorists have not been able to come up with any credible theory that answers all or some of these questions, including questions concerning the magnitude of the masses, the Cabibbo angle, the existence of all these quarks, the grouping of these particles into families, and so on. There is clearly so much that we do not know! Even so, we have certainly made enormous advances in understanding the structure of the interactions between the elementary particles.

9.13 ρ-Parameter

In addition to W^+–W^+ scattering other processes may be considered, such as for example W^-–Z^0 scattering, or Z^0–Z^0 scattering. All these processes can be made to have decent high energy behaviour, but to cure all of them using only one Higgs particle requires a relation between the charged and neutral vector boson masses, usually expressed in the form $\rho = 1$, with $\rho = M^2/(M_0^2 \cos^2\theta_w)$. Higher order quantum corrections slightly modify this relation. In this equation M and M_0 are the masses of the W (80.3 GeV) and the Z^0 (91.2 GeV) respectively. To explain the angle θ_w appearing here would require a detailed discussion about the interplay of weak and electromagnetic interactions, due to the fact that wherever the Z^0 couples to charged particles on both ends, the photon can take its role. Experimentally one finds $\sin^2\theta \approx 0.2315$, and we conclude this discussion with the observation that ρ comes out to the predicted value so that there is no need to have more than one Higgs particle.

It is interesting to note that the higher order corrections to the equation $\rho = 1$ involve among others, the mass of the top quark in a most peculiar way: the correction becomes bigger as the top quark mass is heavier. Here we have a quantum effect that increases if the intermediate state is energywise further away! Many years before the top quark was actually observed the measured magnitude of the quantum corrections was used to predict the top quark mass. That prediction agrees quite well with the experimental value.

The reason that the radiative correction grows with the top mass is a very typical consequence of a gauge theory structure. The top quark has a function in the scheme, for if it is not there certain diagrams grow in an intolerable way. So, if you try to eliminate the top quark from the theory (by making it very heavy) you are left with an infinity. The figure shows the relevant diagrams, which concern momentary dissociation of the W^+ and Z^0 into a quark-antiquark pair. Such diagrams are called self-energy diagrams. The effect we discuss involves the first diagram minus $\cos^2\theta_w$ times the sum of the second and third diagram (since the first diagram gives a correction to M^2, the other two to M_0^2). The top quark is now essential; without the top quark only the second diagram would be there, and this diagram all by itself gives an infinity.