2 A First Look at Quantum Fields

In this section we look in more detail at some of the key features of quantum fields and their interactions. We illustrate these properties with the simplest quantum force, electromagnetism. Or, to give it its fancy name, *quantum electrodynamics*.

2.1 Matter Fields and Force Fields

We'll meet a bewildering number of names in these lectures, each of them classifying particles according to various properties. But one classification is more important than all others: every type of particle falls into one of two classes called

- Bosons
- Fermions

The distinction between these two kinds of particles lies in the quantum world. Fermions have the property that no two particles can occupy the same quantum state. Roughly speaking, this means that you can't put two fermions on top of each other. This property is known as the *Pauli exclusion principle*. In contrast, there is no such restriction on bosons. You can pile up as many of them as you like, one on top of the other.

(A mathematical aside: if you've done a little quantum mechanics then it's very easy to describe the difference between bosons and fermions. Two identical particles are described by a wavefunction $\psi(\mathbf{x}_1, \mathbf{x}_2)$ which tells you the probability amplitude to find the two particles at positions $\mathbf{x}_1$ and $\mathbf{x}_2$. If the particles are bosons then, when you swap their positions, the wavefunction remains unchanged: $\psi(\mathbf{x}_2, \mathbf{x}_1) = \psi(\mathbf{x}_1, \mathbf{x}_2)$. In contrast, if the particles are fermions then the wavefunction picks up a minus sign when you swap them: $\psi(\mathbf{x}_2, \mathbf{x}_1) = -\psi(\mathbf{x}_1, \mathbf{x}_2)$. This means, in particular, that if you try to bring the two particles together at some point $\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}$, then $\psi(\mathbf{x}, \mathbf{x}) = 0$ for fermions, so there is vanishing probability that the two particles sit on top of each other.)

All the matter particles in the universe – electrons, quarks and neutrinos – are fermions. All the force carrying particles are bosons. In fact, this is more or less a definition of what we mean by a "matter" particle vs a "force" particle. The matter particles obey the Pauli exclusion principle; the force particles do not.

The distinction between bosons and fermions has a couple of familiar consequences. Electrons are fermions and therefore obey the Pauli exclusion principle. This is ultimately responsible for the structure of the periodic table. The electrons can't all sit close to the nucleus, but fill up successive atomic shells, with the electrons in the outer shell — known as valence electrons — largely responsible for the chemical properties of the element.

Photons are bosons, and this means that the Pauli exclusion principle does not apply. A laser is an example of a system in which many photons sit in the same quantum state.

2.1.1 Spin

Particles are endowed with a number of other properties. The most familiar of these is the mass of the particle, but it is not the only one.

Particles also have an inherent angular momentum that we call *spin*. It's not a bad analogy to think of elementary particles as spinning about some axis, much like the Earth spins. But spin is a quantum mechanical property and if you push the spinning-Earth analogy too far then it breaks down. For example, if you ask questions like "how fast is the surface of the particle moving" then you'll get nonsensical answers. Furthermore, there's no way to spin up a particle like a basketball; the magnitude of the spin is something that is fixed and unchanging. You can, however, change the orientation of the axis along which the particle spins.

Like many phenomena in the atomic world, spin is quantised. That means that the spin can't take arbitrary values, but comes in discrete amounts. These are[3]

$$s = 0,\ \frac{1}{2}\hbar,\ \hbar,\ \frac{3}{2}\hbar,\ 2\hbar, \ldots \tag{2.1}$$

In natural units, we just say that a particle has spin 0, or spin $\frac{1}{2}$, and so on. Each particle in nature has a spin with a value taken from this list.

Particles that have a half-integer spin come with a rather strange property. If you rotate them by 360° then they don't quite come back the same as they were before! Instead, their quantum wavefunction comes back to minus itself. This means that you have to rotate the particle by 720° before it comes back to the same state. This is one of the more surprising facts about elementary particles and is a clear departure from our every day experience with classical objects.

[3] I've been a little bit sloppy here. Strictly speaking, the total spin of the particle is $\sqrt{s(s+\hbar)}$ with s taking one of the values listed in (2.1).

There is a deep theorem, originally framed by Pauli, which states that the spin determines whether a particle is a boson or fermion:

The Spin-Statistics Theorem: Particles with integer spin are bosons. Particles with half-integer spin are fermions.

This theorem follows when you combine the laws of quantum mechanics with the rules of special relativity. (The word "statistics" in the name of the theorem is not particularly helpful. Its origin lies in the fact that you get different answers when you count the number of possible states in which bosons or fermions can sit, and this counting is referred to as the "statistics" of the particle. We won't need this interpretation in these lectures. You can learn more in the lectures on Statistical Physics.)

The spins of all the known elementary particles in Nature are:

- Spin 0: The Higgs Boson.
- Spin $\frac{1}{2}$: All matter particles, i.e. the electron, muon and tau, together with the six types of quarks and three neutrinos.
- Spin 1: The photon, gluon and W and Z bosons. In other words, the particles associated to electromagnetism and the weak and strong nuclear forces.
- Spin 2: The graviton.

The remaining properties of particles mostly specify their interactions under the various forces. A familiar example is the electric charge, which determines the strength of a particle's interaction with electromagnetism. We'll describe the electric charge of all particles in Section 2.2, and the interactions with other forces in subsequent sections.

Finally, all of the properties described above, including the fermionic/bosonic nature of the particle, are really properties of the underlying field, which are subsequently inherited by the particle.

2.1.2 The Dirac Equation

All fields with spin $\frac{1}{2}$ — which, as we've just seen, means all fields associated to matter particles — are described by the *Dirac equation*.

We won't explain the mathematics behind the Dirac equation in these lectures, but it's so beautiful that it would be a shame not to show it to you. I've put it in a picture frame to highlight that it's here for decoration as much as anything else.

$$\boxed{(i\gamma^\mu \partial_\mu - m)\,\psi = 0}$$

Here ψ is the quantum field; it depends on space and time. It also has four components, so it's similar to a vector but differs in a subtle way. It is know as a *spinor*. For what it's worth, the parameter m is the mass of the particle, while ∂_μ denotes derivatives and γ^μ are a bunch of 4×4 matrices. If you want to understand what the Dirac equation really means, you can find details in the lectures on Quantum Field Theory.

Dirac originally wrote his equation to describe the electron. But, rather wonderfully, it turns out that this same equation describes muons, taus, quarks and neutrinos. This is part of the rigid structure of quantum field theory. Any particle with spin $\frac{1}{2}$ must be described by the Dirac equation: there is no other choice. It is the unique equation consistent with the principles of quantum mechanics and special relativity.

The Dirac equation encodes all the properties of particles with spin $\frac{1}{2}$. Given such a particle, once you fix the orientation of the spin there are two possible states in which the particle can sit. Roughly speaking, it can spin clockwise or it can spin anti-clockwise. We call these two states "spin up" and "spin down".

The Pauli exclusion principle states that no two fermions can sit in the same quantum state. But the quantum state is determined by both the position and the spin of the electron. This means that an electron with spin down can be in the same place as an electron with spin up, because their spins differ. If you've done some basic chemistry, this should be familiar: both the hydrogen and helium atoms have electrons sitting in the orbit that sits closest to the nucleus. The two electrons in helium necessarily have different spins to satisfy the exclusion principle. But, by the time you get to the third element in the periodic table, lithium, there is no longer room for an additional electron in the closest orbit and the third electron is forced to sit in the next one out.

2.1.3 Anti-Matter

The real pay-off from the Dirac equation comes when you solve it. The most general solution has an interesting property: there is a part which describes the original particles, like the electron. But there is a second part that describes particles with the same mass but with the opposite electric charge. The electron has negative charge, so these other particles must have positive charge. These positively charged electrons are called *positrons*: they are examples of *anti-matter*.

If a particle and anti-particle collide, both are annihilated. Typically, the energy is released in high energy photons. We denote the electron as e^- and the positron as e^+. Their annihilation usually results in the emission of two photons. (The emission of a single photon is not consistent with the conservation of both energy and momentum. For example, in the centre of mass frame, conservation of momentum would mean that the emitted photon would have nowhere to go.) The annihilation process is described by the reaction

$$e^- + e^+ \to \gamma + \gamma$$

The end result of all this is that the Dirac equation actually describes four different types of single particle states: a particle with either spin up or spin down, and an anti-particle with either spin up or spin down. The fact that there are four such states is related to the fact that the field ψ is a vector-like object with four components.

Dirac wrote down his equation in 1928. After a few years of confusion, Dirac himself suggested that these novel solutions should be interpreted as anti-matter. In 1931, he wrote

> "A hole, if there were one, would be a new kind of particle, unknown to experimental physics, having the same mass and opposite charge to an electron."

This bold proposal, was confirmed experimentally just one year later, a development that we will describe in more detail in Section B.) The prediction of anti-matter remains one of the great triumphs of theoretical physics. We now know that all the matter particles in Nature have corresponding anti-particles. In all cases, the conserved charges of the anti-particles are equal and opposite to those of the particles.

The Fallacy of the Dirac Sea

Although Dirac's genius led him to predict anti-particles, the argument that got him there was somewhat flawed.

Dirac's mistake was to misinterpret the meaning of ψ in his equation! He originally wrote down the Dirac equation as a relativistic generalisation of the Schrödinger equation, with ψ viewed as the wavefunction of a single particle. We now know that this is not the right interpretation: ψ should be viewed as a quantum field, whose excitations describe many particles, rather than the wavefunction for a single particle.

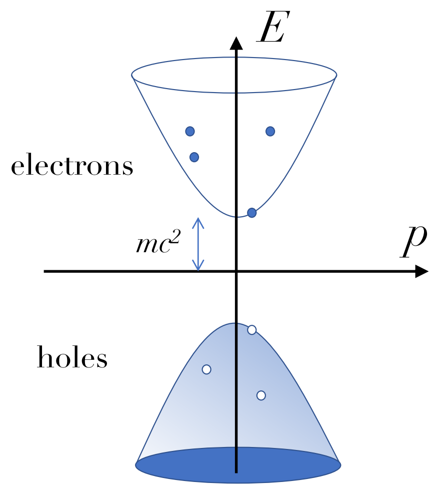


Figure 5. The filled Dirac sea, shown in blue, with holes interpreted as anti-matter.

We can explain what's going on here in a little more detail. The famous Einstein equation $E = mc^2$ tells us the energy of a particle of mass m when the particle is at rest. If the particle is moving with momentum p, then the correct formula is

$$E = \sqrt{p^2c^2 + m^2c^4} \tag{2.2}$$

where, for once, we've left the factors of c in the equation rather than setting $c = 1$. If you solve the Dirac equation, with ψ viewed as a wavefunction, then you find the two sets of solutions, but with energy

$$E = \pm\sqrt{p^2c^2 + m^2c^4}$$

The positive energy solutions are identified as, say, electrons. But what to do with the negative energy solutions? Note that as the particle moves faster, so p increases, and the negative energy solutions become more and more negative. This is problematic. If systems can lower their energy, they usually do. But clearly we don't observe any particles getting faster and faster.

Dirac found a clever, but not altogether convincing, trick to escape from this conclusion. He suggested that the negative energy states were already filled by electrons. Because electrons are fermions, the Pauli exclusion principle means that no other electron is allowed to sit in these states, blocking the possibility for electrons to lose energy by tumbling to ever-lower states. This situation is shown in Figure 5.

In this picture, the vacuum of the universe consists of an infinite number of electrons. This is called the *Dirac sea*. One might worry about why we don't feel the infinite electric charge, but since this situation represents the ground state of the universe, Dirac argued that we reset the clock and say that this is what we mean by neutral. Any charge is now measured relative to this ground state.

Dirac's picture suggests that something novel may happen. We could excite an electron out of the Dirac sea, and into the positive energy states. This is shown in Figure 5. To do this, we need to inject a minimum of $E = 2mc^2$ energy into the system, to bridge the gap between the lower and upper bands in the figure. But if we can somehow achieve this, then we have created an electron out of the vacuum. But we have also created a *hole*, an absence of an electron, in the sea of negative energy states. In a zen-like manoeuvre, we now attribute properties to this absence. Like the excited electron, it can freely move around. Because there's an absence of charge, relative to the vacuum it will appear to have positive charge. Finally, if the electron and the hole come into contact, the electron can drop back down into the negative energy state. It will appear as if the electron and hole have annihilated, releasing at least energy $E = 2mc^2$ in the process.

Dirac's picture of anti-particles is ingenious. But, ultimately, it's not the right way to think about things. If you view the object ψ not as a single-particle wavefunction, but rather as a quantum field, then the energy of both particles and anti-particles turns out to be positive. There is no need to invoke an infinity of electrons, disappearing to the bottom of the sea. Instead, there are no negative energy states: simply particles and anti-particles.

Moreover, it turns out that bosons also have anti-particles. But now there is no counterpart to the Dirac sea argument because bosons don't obey the Pauli exclusion principle. Meanwhile, bosonic anti-particles arise just as straightforwardly in quantum field theory as fermionic anti-particles.

Although Dirac's clever argument is not the right one for fundamental physics, it does turn out to have its uses elsewhere because it's a good description of what happens in solid materials. Any solid is made of atoms, and some number of electrons typically disassociate themselves from the nuclei and wander around which, in the quantum world, means that they fill up the lowest energy levels provided by the surrounding solid. In this context, this is called the *Fermi sea* but it conceptually identical to Dirac's sea. When an electron is excited out of this sea, it leaves behind a hole. This hole – which is the absence of an electron – behaves in many ways like a particle with

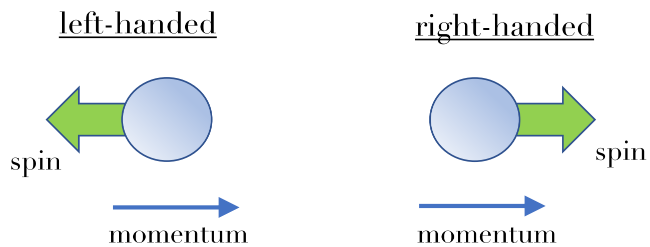


Figure 6. The handedness of a massless particle is determined by the relative direction of its spin and momentum.

positive electric charge. Indeed, there are some materials in which electricity appears to be conducted by positively charged particles. These aren't protons! They are holes.

There is a lesson here which is repeated over and over again in the history of science: a good idea tends to find a place in the world, even if it's not where it was originally intended.

2.1.4 Massless Particles

For massless particles, the story of spin needs to be slightly retold. Before we jump into the details, it's natural to ask: why bother? What spin $\frac{1}{2}$ particles in Nature are massless?

The answer to this question is shocking: all of them! One of the most striking features of the Standard Model is that, at the fundamental level, all the matter particles are massless. In fact, more than that, it turns out that it's not possible to incorporate masses into the theory without first doing some damage to some aspects of the weak force. This damage is achieved by the Higgs boson which, ultimately, is why the fundamental particles appear to have mass. We will describe all of this in Section 4. But, in preparation for that, it will be useful to explain here what becomes of spin when particles are massless.

Solving the Dirac equation, one finds that as a particle gets faster, its spin necessarily becomes oriented along the direction of motion. For massless particles, which travel at the speed of light, there are two options: either the spin points in the same direction as the particle is travelling, or it points in the opposite direction. When the spin points in the same direction, the particle is said to be *right-handed*. When it points in the opposite direction, it is said to be *left-handed*. This is shown in Figure 6.

This distinction is quantified using something called *helicity*. If the particle moves with momentum $\mathbf{p}$ and the spin points in the direction $\mathbf{s}$ then the helicity is defined to be

$$h = \frac{\mathbf{s} \cdot \mathbf{p}}{|\mathbf{p}|}$$

Right-handed particles have helicity $+\frac{1}{2}$; left-handed particles have helicity $-\frac{1}{2}$.

Such a distinction doesn't make sense for massive particles. One can simply overtake the particle and look back to see it moving in the opposite direction, but with the spin remaining the same, so its helicity appears to be flipped. However, you can never overtake a massless particle because it travels at the speed of light, and this means that everyone agrees on the helicity of a massless particle.

In fact, one can go further. It turns out that, for massless particles, it is possible to have "half a Dirac fermion". This is a particle where only, say, the left-handed helicity exists. There is no particle at all with right-handed helicity. The anti-particle would then exhibit the opposite behaviour, only existing in the right-handed state, never left-handed. A particle with these properties is known as a *Weyl fermion*. This idea will play a key role when we discuss the weak force.

I should stress that such Weyl fermions, with fixed helicity, are only possible for massless particles. The particles that we observe, such as electrons, do ultimately have a mass and they achieve this by gluing together two Weyl fermions to form a complete Dirac fermion, with both kinds of spin. We'll learn more about how this happens in Section 4.

2.2 Quantum Electrodynamics

The Dirac equation described in the previous section tells us that matter particles necessarily come with anti-particles. But for these particles to subsequently do something, they must interact. Those interactions happen through forces.

The simplest force in particle physics is electromagnetism. In large part, it is simplest because we have some classical intuition for this force: it is the same force understood many centuries ago by Coulomb, Ampére, Faraday and Maxwell, albeit dressed by some quantum bells and whistles.

The force is mediated by two fields, the electric field $\mathbf{E}(\mathbf{x}, t)$ and the magnetic field $\mathbf{B}(\mathbf{x}, t)$. Each of these is a *vector field*, meaning that at every point in space $\mathbf{x}$ and for every time t, the field is specified by both a magnitude and a direction.

The equations that describe the dynamics of the electric and magnetic fields are known as the *Maxwell equations*. We won't need them in these lectures but, for completeness, here they are:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad , \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$
$$\nabla \cdot \mathbf{B} = 0 \quad , \quad \nabla \times \mathbf{B} = \mu_0 \left(\mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \tag{2.3}$$

The fields $\mathbf{E}$ and $\mathbf{B}$ react to the presence electric charge density ρ and electric currents $\mathbf{J}$, while ϵ_0 and μ_0 are two constants that characterise the strength of the electric and magnetic forces in a way that we will describe more below.

The equations, as written above, hide the full beauty of the Maxwell equations. A better formulation encodes both the electric and magnetic fields in a 4×4 anti-symmetric matrix called the field strength, which takes the form $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Only then do the Maxwell equations reveal their true simplicity, in a way that deserves hanging in a frame,

$$\boxed{\partial_\mu F^{\mu\nu} = J^\nu}$$

You can learn more about the Maxwell equations and their classical solutions in the lectures on Electromagnetism. Famously, among the solutions to these equations are electromagnetic waves, including visible light. When you look at these solutions through the lens of quantum mechanics, you find that they decompose into particles, known as *photons*.

A photon can come in two different states that we call polarisation. These are entirely analogous to the "spin up" and "spin down" states of the electron. (The fact that both spin 1/2 and spin 1 particles have two internal states is something of a coincidence. For example, it's only true in three spatial dimensions; the counting is different in other dimensions.)

The theory describing the electromagnetic field interacting with the electron field is known as *quantum electrodynamics*, or *QED* for short. It is the theory describing light interacting with matter, and ultimately underpins large swathes of science, including condensed matter physics and chemistry. Happily, it is also the simplest component of the Standard Model.

The Strength of the Interaction

Take two particles carrying electric charge Q_1 and Q_2 and hold them some distance r apart. The relevant solution to the Maxwell equations tells us that the particles experience a force given by

$$F = \frac{Q_1 Q_2}{4\pi\epsilon_0 r^2} \tag{2.4}$$

This is called the *Coulomb force*. The force is repulsive if the two particles carry charge of the same sign; it is attractive if they carry charges of opposite sign.

This formula shows us that the constant ϵ_0 characterises the strength of the Coulomb force. If the value of ϵ_0 was smaller, then Coulomb force would be more powerful. If you look up ϵ_0, you'll find some unhelpful number quoted with unit of Farads per metre. A more useful measure of the strength of the electric force comes from the dimensionless quantity known as the *fine structure constant*,

$$\alpha = \frac{e^2}{4\pi\epsilon_0 \hbar c} \tag{2.5}$$

where e is the electric charge of the electron. It turns out that the value of the fine structure constant is roughly

$$\alpha \approx \frac{1}{137}$$

This is the cleanest way to characterise the strength of the electric force. In particular, in natural units with $\hbar = c = 1$, two electrons held a distance r apart experience a force given by

$$F = \frac{\alpha}{r^2}$$

Maxwell's equations also contain a second constant, μ_0, which characterises the strength of the magnetic interaction. It turns out that this is not independent from ϵ_0. One of the great discoveries of Maxwell is that the two constants are related by

$$\epsilon_0 \mu_0 = \frac{1}{c^2}$$

with c the speed of light. As a side remark, note that if the strength of the electric force $1/\epsilon_0$ were weaker, then the strength of the magnetic force $1/\mu_0$ would necessarily be stronger.

2.2.1 Feynman Diagrams

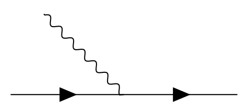
There are some simple cartoons that allow us to figure out what processes are allowed in quantum electrodynamics (and, indeed, in the other forces). These cartoons are called Feynman diagrams.

We will take time to run horizontally, from left to right[4]. We then draw electrons as solid lines with a forward pointing arrow, like this —►— .

Positrons are depicted as solid lines with a backwards pointing arrow, like this —◄—. We'll see the utility of the backwards-arrow notation below. It suggests that it may be possible to think of anti-particles as particles that move backwards in time. There is mathematical sense in which this statement is correct, but it shouldn't be taken too literally.

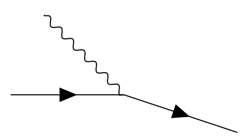
Finally, photons are depicted as wavy lines like this ∿∿∿ .

There is just a single interaction between the electron and photon, from which all other processes can be built. This can be viewed as an electron absorbing a photon, and scattering off in a different direction. It looks like this



The point where the photon hits the electron is referred to as a *vertex*.

Conservation of momentum means that the electron necessarily moves off in a different direction after absorbing the photon. So you might have thought that it would be better to draw the Feynman diagram like this



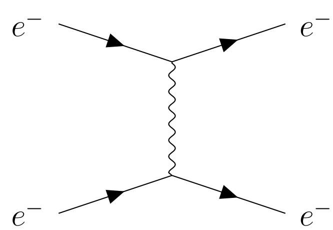
Indeed, sometimes we'll draw diagrams like this. However, the Feynman diagrams should not be read too literally: the paths aren't the actual paths of particles in spacetime. They should be viewed in a more topological fashion, like the London underground map. We'll say more below about what Feynman diagrams are, and what they aren't.

[4]This is the convention used in the Quantum Field Theory lectures, but it's not universal. Some authors prefer time to flow upwards.

Figure 7. This is not a good analogy for virtual particles.

Now the game is as follows: you can describe any process by stitching together the Feynman diagram building blocks above. You can orient the different legs of the diagrams in any way you wish. You just have to make sure that the arrows on the solid lines follow each other. Any process that you can draw can happen, *provided that* it is allowed on grounds of energy and momentum conservation.

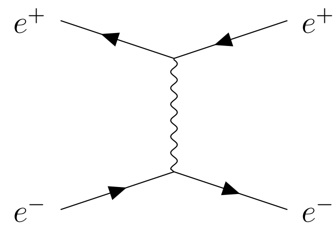
Let's look at some examples. Here is a Feynman diagram describing one electron scattering off another



I've added the name of the particle to the external legs, a practice that will prove useful as we progress. Note that the electrons don't just bounce off each other; there is no direct contact between them. Instead, the electrons scatter by exchanging a photon. Particles that appear only in internal legs of Feynman diagrams, like the photons above, are referred to as *virtual particles*. This is a lesson that we'll see repeated later: all forces can be understood by the exchange of virtual bosonic particles.

In some ways, Feynman diagrams are a little too evocative, and we should be careful not to interpret the diagram above too literally. For example, you shouldn't think of one electron as recoiling as it emits a virtual photon, which is then absorbed by the second, resulting in a repulsive force from Newton's third law, like two people in boats throwing a ball back and forth between themselves. This will then leave you puzzled about how such particle exchange can possibly lead to an attractive force.

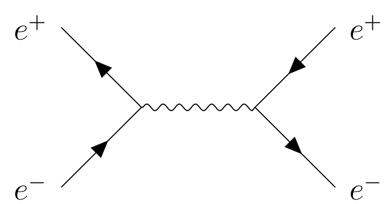
Yet in quantum field theory, there is no problem with virtual particles describing an attractive force. Indeed, the Feynman diagram for the scattering of an electron off a positron is almost identical to the one above:



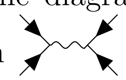
Translating this diagram into mathematics gives the attractive Coulomb force, but this isn't easily captured by the people-in-the-boat analogy. (If you get to the point where you start thinking about people in boats throwing boomerangs backwards and forwards then you might realise that the analogy has clearly been stretched too far.)

In fact, what's really going on here is that both of the scattering diagrams above are a reformulation of the familiar result from classical physics in which one electron experiences a force due to the electric field of another. If you translate the diagram above into mathematics, you will find that it is simply a rewriting of the Coulomb force law (2.4). Viewed this way, the "virtual particles" are merely a handy device to capture the behaviour of the underlying field. If we were to think in terms of fields, then we have no need to discuss virtual particles. Moreover, there are situations — like for the strong force — where the concept of virtual particles is not useful, while the fields remain.

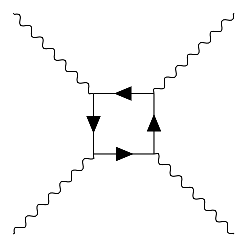
For the scattering of an electron off a positron, there is a second, qualitatively different diagram that also contributes.



This has the interpretation of the electron-positron pair annihilating into a virtual photon, which then turns back into a pair of particles. It turns out that this diagram doesn't contribute to the Coulomb force (2.4), which holds only in the non-relativistic limit where velocities of all particles are low, but does change the scattering behaviour at higher energies. For our purposes, however, it is useful simply to illustrate the utility of the forwards/backwards arrow notation for particles and anti-particles: they capture the conservation of electric charge.

The total electric charge of an electron-positron pair is zero, which allows them to annihilate into a (virtual) photon. In contrast, such a process isn't possible for the scattering of two electrons, because their charge is non-zero. In the diagrammatic language, we see this because the corresponding Feynman diagram  doesn't have the arrows matching up, and so is illegal.

There are also interesting processes that we can construct with photons on the external legs. For example, here is a diagram that corresponds to one photon scattering off another



Famously, light doesn't scatter off itself in the classical world. This is important, for it allows us to see! But it's no longer true in the quantum world. The diagram above can be viewed as a light scattering off a particle-anti-particle pair which briefly appear as a vacuum fluctuation. The probability for such a process is small, which is why we don't notice this process every day. But, although small, it is non-zero, and light-by-light scattering has been observed in particle colliders.

2.2.2 What is a Feynman Diagram Really?

All quantum processes have an element of randomness. Particle physics is no different. If you collide two particles together at high energies, there are many possibilities for what may emerge. Quantum field theory allows us to assign probabilities — or, more precisely, quantum amplitudes — to all of these possibilities.

However, there's a hitch. Quantum field theory is hard, and the expressions for these probabilities are ridiculously complicated. In many situations, we have no idea how to compute them. However, for QED we can make progress based on the observation that the interaction strength, as captured by the fine structure constant $\alpha \approx 1/137$, is small. This means that we can expand the complicated probabilities in a perturbative expansion, rather like Taylor expanding a function.

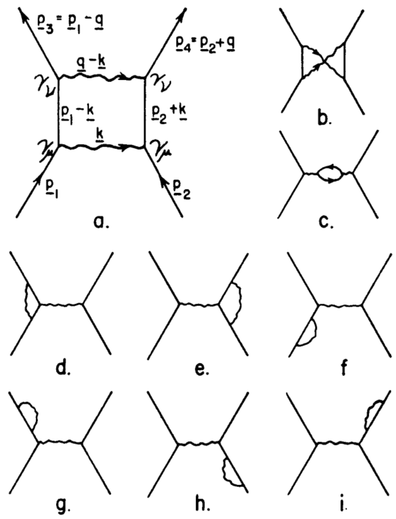
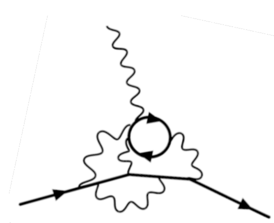


Figure 8. A handful of Feynman diagrams, taken from Feynman's original paper.

The Feynman diagrams are a pictorial way of capturing this perturbative expansion. Suppose that you want to compute the probability for some process to happen, for example the electron-positron scattering described above. The process itself is defined by the external legs of the diagram — these are what tell you, for example, that you start with two particles and end up with two particles. Given this data, you should now write down all possible Feynman diagrams. The diagrams that we drew above are the simplest diagrams, but there are an infinite number of diagrams contributing to any process with an increasingly complicated structure of internal lines. For example, the original vertex can be dressed with all sorts of other lines, to give things that look like this:



Some examples of Feynman diagrams for $e^- + e^+$ scattering are shown in Figure 8

So far, this procedure doesn't sound very helpful. We have to write down an infinite number of ever more elaborate diagrams to describe any process. Moreover, there are rules which translate each diagram into a mathematical expression, usually involving some very complicated and challenging integrals. What saves us is the fact that the more complicated a diagram, the less important it is in some process. To compute the importance of any diagram, you need simply to count the number of vertices. While the exact contribution of any diagram may be very difficult to compute, the happy news is that it is proportional to

$$\alpha^{\#\text{ of vertices}}$$

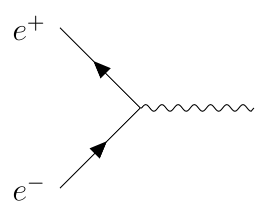
So, for example, a diagram with a single vertex will mean that the probability is something in the ballpark of 1/137. Meanwhile, the the elaborately dressed vertex shown above contributes something proportional to α^9. With $\alpha \approx 1/137$, this kind of diagram barely changes the answer, and so can be safely neglected. The light-by-light scattering diagram that we showed earlier comes in at α^4, explaining why we don't observe this phenomena in every day experience.

There is one important lesson to take from this: the utility of Feynman diagrams is intimately connected to the weakness of the electromagnetic interaction. In situations where the interactions between fields are strong, Feynman diagrams are not the right way to think about the physics.

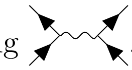
If you want to learn more about how to unmask Feynman diagrams, and turn them back into the underlying equations, then you can find details in lectures on Quantum Field Theory.

Some Examples

To illustrate these ideas, we can compute the relative probabilities that an electron and positron will annihilate to a bunch of photons. First, we need to address a subtlety. It's possible to draw a diagram representing an annihilation to a single photon:



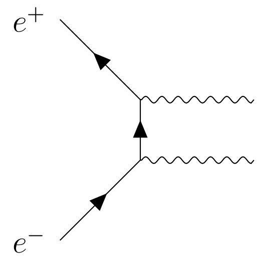
This would appear to be proportional to α. However, if you calculate this diagram, you'll find that it's vanishing. This is because although it is consistent with charge conservation, it's not consistent with the conservation of energy and momentum. To see this, consider the frame in which the electron and positron have equal and opposite momenta. The outgoing photon must then have vanishing momentum, but non-vanishing energy, and this isn't possible.

It's worth pointing out that this diagram can occur as a sub-diagram in other processes. For example, we already used it in electron-positron scattering . In this case, the intermediate photon is virtual and it turns out that there's no requirement for virtual particles to obey the usual energy-momentum relations. You can hand-wave this away by saying that virtual particles can borrow energy for some period of time by virtue of (the slightly dodgy version of) the Heisenberg uncertainty relation: $\Delta E \Delta t \sim \hbar$.

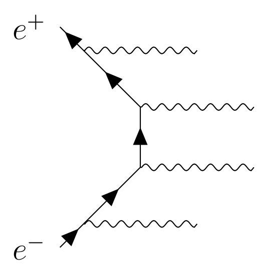
Back to the annihilation of an electron and positron, the simplest process results in two photons:

$$e^+ + e^- \to \gamma + \gamma$$

This is described by the following Feynman diagram:



From our discussion above, we know that this is proportional to α^2. But other processes are possible. For example, the pair could annihilate to any number $n > 1$ of photons. For example, the diagram for annihilation to four photons is



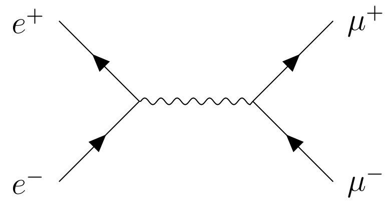
This probability for such a diagram is proportional to α^4. This means that the probability of getting four photons our of a collision is suppressed by a factor of α^2 relative to the probability of getting two photons.

2.2.3 New Particles From Old

Electrons are not the only particles that experience the force of electromagnetism. Any particle that carries electric charge also interacts with the photon. In particular this means that all fermions, except for the neutrinos, feel the force of electromagnetism.

We can easily extend our Feynman diagrams to include these extra particles. By convention, we depict any fermion with a forward arrow and any anti-fermion with a backward arrow, but now we must label these lines with the particle name to show what particle we're talking about. Every particle with electric charge will have an interaction vertex of the form. When evaluating Feynman diagrams, these vertices contribute a factor of $Q^2\alpha$ to the probabilities, where Q is the charge of the particle, in units where the electron has $Q = -1$.

This brings new opportunities. For example, we can collide an electron and positron, but now produce new particles such as a muon-anti-muon pair as shown in the diagram below.



We still have to worry about energy and momentum conservation when evaluating this diagram. If, in the centre of mass frame, the energy of the incoming electron-positron pair is less than $2mc^2$, the rest mass of the muon-anti-muon pair, then the muons cannot be produced. However, when the incoming energy exceeds this threshold, then we can start to produce new particles from old ones. The same kind of process allows us to produce any charged particles from the collision of electrons and positrons. This, of course, is the basis for particle colliders.

2.3 Renormalisation

At the fundamental level our world is built not from particles, but from fields. Moreover, as we stressed in the introduction, these fields froth and foam in the uncertain quantum world. This gives rise to an important phenomenon known as *renormalisation.*

Let's consider a single electron. It gives rise to an electric field which, like the force, varies as an inverse square law,

$$\mathbf{E} = \frac{e}{4\pi\epsilon_0 r^2}\hat{\mathbf{r}}$$

where $\hat{\mathbf{r}}$ is the unit vector pointing radially outwards. Clearly the electric field gets bigger and bigger as we approach the position $r = 0$ of the particle. But what is happening to the electron field near this point? It turns out, that both electric and electron fields start thrashing wildly as we get near to $r = 0$.

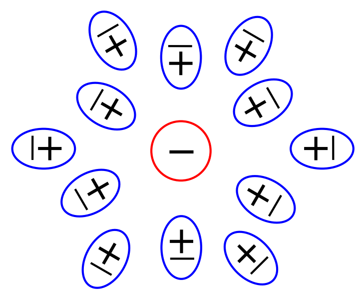
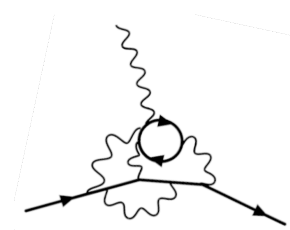


Figure 9. The renormalisation of electric charge.

While it's challenging to talk about quantum fields, we can build some intuition by reverting to the language of particles. As we get closer to the electron, the electric field gets stronger and, as a result, the energy density stored in the field gets larger and larger. At some point — a distance of around 10^{-12} m — the energy density is so large that an electron-positron pair can be produced from the vacuum.

There is a general rule in quantum field theory that anything that can happen does, in fact, happen. This means a single electron is surrounded by a swarm of particle-anti-particle pairs, continually popping in and out of the vacuum. As we get closer still, muons, taus and even quarks will also appear in the mix. We learn that any simple picture you may have of a single particle giving rise to the electric field is really very far from the truth: it is impossible to enforce any kind of social distancing in the quantum world.

This story should really be viewed in terms of quantum fields. When we talk of a swarm of particle-anti-particle pairs, it is really a metaphor for the quantum field being excited in a tangled and elaborate fashion. Just as the vacuum is something complicated in quantum field theory, so too is the notion of a single particle. In the language of Feynman diagrams, these particle-anti-particle pairs are captured by the diagrams that include loops of particles, like the one shown on the right.



The excited quantum field has an important consequence for the strength of the electromagnetic interaction. Again, we can understand this in the language of particles. The swarm of particle-anti-particle pairs will not be oriented randomly around the electron. Instead, the positrons, which carry positive charge, will be attracted to the

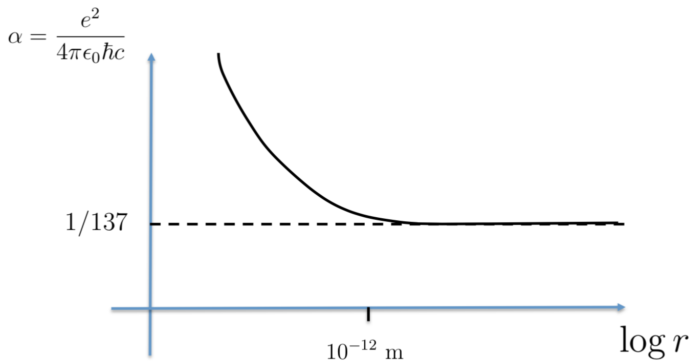


Figure 10. The renormalisation of fine structure constant.

electron in the centre, while the electrons will be repelled, as shown in Figure 9. The net effect is that as you get closer to the electron, you find more and more negative charges outside you. Which must mean that the original charge must have been bigger than we see. This is an effect known as *screening*. This swarm of particle-anti-particle pairs is actually hiding the true charge of the electron in the centre.

In fact, an excellent analogy of this phenomenon arises in metals. Take a positive charge and place it in a metal. The mobile electrons will enthusiastically cluster around, screening the positive charge so that it can't be detected at large distances. This is very similar to what happens with the electron in the vacuum and is one of many situations in which ideas in particle physics are mirrored in condensed matter physics.

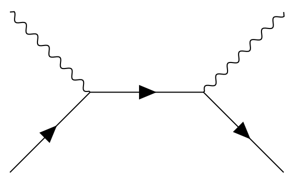
Because the effective charge of the electron gets bigger at shorter distances, so too does the interaction strength as captured by the fine structure constant (2.5). In fact, we learn that the fine structure constant is very badly named, since it's not in fact constant at all. At distances larger than $r \gtrsim 10^{-12}$ m, it plateaus to the usually quoted value of $\alpha \approx 1/137$. But as you go to smaller scales, the strength of the electromagnetic interaction increases logarithmically. For example, the strength of the interaction has been well measured at the scale of the weak force, which is roughly $r \approx 10^{-17}$ m, where it is found to be $\alpha \approx 1/127$. A sketch of the variation of the fine structure constant — often referred to as *running* — is shown in Figure 10.

The lesson of renormalisation as described above is a general one. It turns out that none of the dimensionless physical constants of nature are, in fact, constant. All of them depend on the distance scale you're looking at.

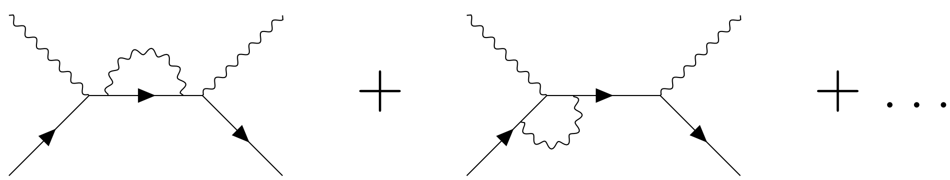
2.3.1 The Long, Confusing History of Renormalisation

While the description of renormalisation described above is fairly straightforward, the mathematics underlying it is not. For this reason, our forefathers had to travel a long and tortuous road to make sense of quantum field theory in general, and the issue of renormalisation in particular.

The story starts in the late 1920s, soon after the original development of quantum mechanics. The quantum pioneers — Heisenberg, Dirac, Pauli and others — tried to apply their ideas to the interaction of light and matter. They tried to ask very simple questions, like the probability for a photon to scatter off an electron. While they didn't have the diagrammatic tools later introduced by Feynman, they did understand that we could approach the problem in a perturbative expansion, starting with a process which we now draw like this:



They found that this calculation gave pretty good agreement with the experiments. But then they tried to do better, and compute the leading corrections. In diagrammatic language, this means evaluating diagrams like this



However, here they ran into a problem. Each of these subsequent diagrams was proportional to α^4, as expected. But the proportionality constant was infinity. That made it very hard to argue that the contributions from these diagrams was smaller than the first.

The quantum heroes worked on this problem for well over a decade, but made little progress. Looking back, many of their ideas were simply too crazy. Having forged one revolution they were, like Che Guevara, all keen for the next. Bohr wanted to get rid of energy conservation. Heisenberg wanted to make spacetime non-commutative. Pauli wanted to invade Bolivia. Yet the answer they were seeking did not, ultimately, require an overhaul of the foundations of physics. It needed a different approach.

The war intervened, and as life returned to normal a new generation of physicists took up the problem, people like Tomonaga in Japan, and Schwinger, Feynman and Dyson in the US. They were helped in no small way by a new experimental result, discovered by Willis Lamb. The eponymous Lamb shift is a tiny, but detectable change to the energy levels of hydrogen due to the problematic one-loop Feynman diagrams. Their solution was a slow-burn revolution, one that took many decades to play out as the power of quantum field theory became clear. However, at the time it didn't feel like a revolution. It felt like a con. Their solution was this:

$$\infty - \infty = \text{finite}$$

In other words, they found a mathematical procedure that allowed you to subtract one infinity from another, leaving an unambiguous finite answer. They called this process *renormalisation*. The results were nothing short of spectacular.

The poster boy for renormalisation is a quantity known as the magnetic moment of the electron. If you place an electron in a magnetic field, then the spin of the particle will precess, as shown in figure. The speed at which the spin precesses is characterised by a dimensionless number g known as the electron's magnetic moment.

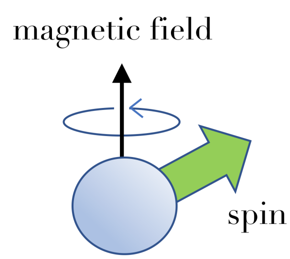


Figure 11. The precession of a spin

In the grand scheme of things, this number is not particularly important. However, it has played a key role in the development of quantum field theory because it is a quantity that we can determine with some accuracy, both experimentally and theoretically. After many decades of painstaking work, the experimental result for the electron magnetic moment is

$$g_{\text{expt}} = 2.0023193043617 \pm 3$$

Meanwhile, after many decades of extraordinarily challenging calculations, evaluating increasingly complex Feynman diagrams up to corrections of order α^5, the theoretical result is

$$g_{\text{theory}} = 2.00231930436\ldots$$

The agreement is awe inspiring. In most areas of science you jump up and down with joy if you get the first digit right. In economics you don't even need that. Yet here there is agreement between theory and experiment to 12 significant figures[5]!

Despite this runaway success, there was something a little disquieting about renormalisation. The idea that you can take one infinity away from another, to leave something finite was not mathematically legitimate. This was Feynman's take:

> "The shell game that we play to find [the answer] is technically called renormalization. But no matter how clever the word, it is what I would call a dippy process! Having to resort to such hocus-pocus has prevented us from proving the theory of quantum electrodynamics is mathematically self- consistent. .... I suspect that renormalization is not mathematically legitimate."

The physical meaning of renormalisation took several more decades to uncover. And it came from an unusual place: the attempt to understand boiling water! In particular, when water is close to the critical point, it turns out that the physics can be understood using very similar Feynman diagram techniques to those employed in particle physics. But this time there are no infinities. That's because when you get to the place in the calculation where infinities might rear their head, you need to remember that water isn't infinitely divisible, but is made of atoms. When you take this into account, the infinities in the diagrams simply aren't there.

But the same should also be true of quantum field theories in particle physics. There is no reason to think that our theories are valid to arbitrarily high energies, or arbitrarily short distance scales. A modern perspective on renormalisation absorbs this lesson. Just as Newtonian mechanics comes with a health warning, stating that you shouldn't trust it in extreme situations — when speeds become too large, masses become too heavy, or particles become too small — so too does quantum field theory. No quantum field theory should be trusted to arbitrarily small distance scales, or arbitrarily high energies. That would be hubris. Instead, we should admit that there is an energy scale beyond which our theory no longer applies. This energy is called the *cut-off*. For

[5]The same calculations for the magnetic moment of the muon give $g_{\text{theory}} = 2.00233183602$ and $g_{\text{expt}} = 2.00233184122$. Although the agreement is impressive, it fails at the 10^{th} significant figure. This is one of the very few discrepancies between theory and experiment. If this disagreement is borne out, it may be due to the effects of new particles, beyond those of the Standard Model. If nothing else, this discrepancy should serve to show just how astonishingly good the Standard Model really is: there are surely no other areas of science where people have sleepless nights over a failure in the 10^{th} digit.

the Standard Model, the cut-off is somewhere above 1 TeV. Simply injecting a little humility into the proceedings, and admitting that we don't yet understand everything, is sufficient to remove the infinities.

But now there is a new problem. We must make sure that no physical answer depends on our choice of the cut-off. After all, the cut-off is an expression of our ignorance, the limit of our current knowledge. It would be very unsatisfactory if something physical like, say, the electron mass depended on what we don't know about physics at high energies.

It took many scientists many years to figure out how to include a level of our ignorance in the theory, without anything depending on our ignorance. The different parts of the story were finally pieced together in the early 1970s by Kenneth G. Wilson. His work is surely the most influential piece of theoretical physics in the latter part of the 20^{th} century, and now has applications from particle physics to biological physics to gravitational physics. Wilson's insight was that nothing you can physically measure depends on the choice of cut-off providing that physical quantities are not constant: they must change depending on the scale at which you explore the world. This is way that we introduced renormalisation in the previous section. Moreover, the complicated and dubious calculations which seemingly gave $\infty - \infty =$ finite can be reinterpreted in a more palatable way as telling us how quantities change with scale.

This new approach is sometimes referred to as the *Wilsonian renormalisation group*, to distinguish it from the older approach of Feynman and others. You can read more about this in the lectures on Statistical Field Theory.

The upshot of this is that if you do quantum field theory carefully, there are no infinites. But neither are there (dimensionless) constants of Nature. Instead, one key lesson of quantum field theory is that the universe in which we live is organised by scale. If you want to write down a theory of our world, then you need to explicitly state the scale at which the theory holds. Change the scale, and you must change the theory. Or, at the very least, you must change the parameters of the theory.

There is a final twist to this. The Feynman quote on the previous page is from 1985, fifteen years after Wilson did his crucial work and three years after Wilson was awarded the Nobel prize. I don't know why Feynman still held that opinion at that time. It is conceivable that he was unaware of the importance of Wilson's work. After all, the detailed calculations are not wildly different from those Feynman himself did decades earlier, and perhaps he did not appreciate the all-important change of emphasis. Or perhaps he simply didn't like the truth getting in the way of a good story.

B Interlude: Looking to the Sky

Radioactivity was a gift to physicists. Beta decay provided the first insights into the weak force, a topic we will return to in Section 4, while beams of alpha particles — which can reach energies up to 5 MeV — were used as a microscope to peek into the nucleus for the first time. As we learned in Interlude A, both the proton and the neutron were discovered by bombarding other elements with alpha particles.

But 5 MeV can only get you so far. Looking inwards requires higher energies. The smaller the distance scale you want to explore, the higher the energy you need. Ultimately, progress was made by constructing particle accelerators, but physicists first made use of another of Nature's gifts: cosmic rays.

The Rise of the Balloon

Our world is constantly bombarded by charged particles from the cosmos. These particles – which are collectively known as *cosmic rays* – are mostly protons or helium nuclei, with the occasional electron and heavier nuclei thrown in for good measure. They travel enormous distances before reaching us, originating far outside our galaxy in supernova explosions or in the accretion discs which surround supermassive black holes in the centre of other galaxies.

When cosmic rays hit the upper atmosphere, they create a shower of new particles, many of which survive the journey down to Earth where they can be detected. Theodore Wulf was the first to realise that there was something interesting to be explored. In 1910 he built a simple electrometer to detect ionised particles in the atmosphere. At the time, it was thought that this ionising radiation was emitted by the Earth. Wulf had the simple but brilliant idea to test this by climbing the Eiffel tower to perform his experiment. He found that the amount of radiation did indeed drop, but nowhere near as quickly as one would expect. Something was afoot.

The challenge was accepted two years later by Victor Hess. Needing to get to greater heights, he took up ballooning. At a height of 1 km, he found that the radiation was more or less the same as on Earth. At a height of 5 km, he found the radiation was nine times greater. To test various hypotheses, he flew in the day and he flew at night. He even flew during a solar eclipse. He concluded that either there was some unknown substance hiding in the upper atmosphere, or the radiation had an extraterrestrial source. He called it *ultraradiation*.

The name didn't stick. In Caltech, Robert Millikan (of oil drop fame) turned his attention to this new phenomenon. He had a better name: cosmic rays. More importantly, he also had resources and a team of brilliant experimenters who could explore their implications.

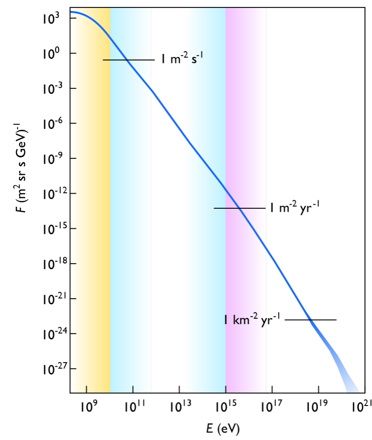


Figure 12. The distribution of cosmic ray energies. Image from Wikipedia.

With hindsight, it's very clear why cosmic ray showers provided such an opportunity for particle physics. Radioactivity offers alpha particles with energies up to 5 MeV. Cosmic rays have no such limitations. A plot of the energy vs the flux of cosmic rays is shown to the right. As you can see, cosmic rays with energies of 1 GeV are common place, but energies extend up to 10^{11} GeV, way beyond what we can create in colliders. To find interesting physics, we just need to have our detectors in the right place at the right time.

B.1 The Positron

In May 1931, after struggling for some years with the meaning of the negative energy solutions of his equation, Dirac finally made the bold leap and predicted the existence of anti-electrons. In September 1932, Carl Anderson announced the discovery of a new particle, with the same mass as the electron, but opposite charge. He later named this particle the positron. Rather surprisingly, the discovery of anti-matter owed essentially nothing to its earlier theoretical prediction.

Anderson's interest was in cosmic rays. Unlike Hess, however, he had no intention of getting in a balloon. With a good detector, he needed to climb no higher than his third storey office to study the showers from cosmic rays.

Anderson's detector was the cloud chamber. We already briefly met a preliminary version of the cloud chamber in Section A.2 when describing J.J. Thomson's discovery of the electron. In the intervening years, it was perfected and was usually referred to as the *Wilson cloud chamber* after its inventor, C.T.R. Wilson. When a charged particle passes through the chamber, it leaves behind a path of ionised gas particles, around which droplets subsequently condense. The result is that an elementary particle leaves behind a misty trail, visible to the naked eye, like the contrails left by a plane in sky.

Working at Caltech under the guidance of Millikan, Anderson built a cloud chamber sitting within a magnetic field of 25,000 Gauss. The purpose of the magnetic field was

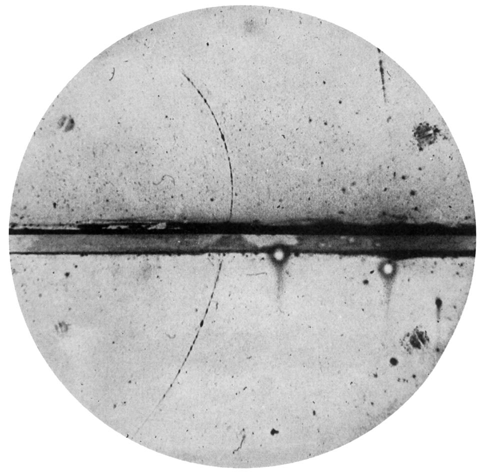


Figure 13. Anderson's first published picture of a positron. You can see the layer of lead running through the middle. The positron entered the cloud chamber at the bottom left, was slowed by the lead, with the trail visibly more curved after it exited. Somewhat unusually, this positron arose from a cosmic ray collision below the detector.

to bend the trajectory of a charged particle, allowing one to get a handle on the ratio e/m of charge to mass. Anderson found trajectories bending in both directions. Those that bent in one direction were clearly negatively charged electrons, coming down from the sky. But what about those that bent in the other direction? They were too light to be protons. However, they could have been electrons coming up from the ground. It seemed unlikely because, as Milikan pointed out: "Everyone knows that cosmic ray particles go down. They don't go up except in very rare circumstances".

To better understand what was going on, Anderson placed a thin, horizontal layer of lead in the the cloud chamber. This wasn't thick enough to stop the particles completely, but it did cause them to lose a significant amount of energy as they passed through. This meant that the particle would be travelling more slowly after it passed through the lead, and so its trail would bend in a tighter curve. In this way, Anderson was able to determine the direction of the particle and, hence, its charge. He found, in the words of his original paper, an "easily deflectable positive".

Anderson's results were soon confirmed by others, notably Patrick Blackett and Giuseppe Occhialini in Cambridge. Indeed, the week before Anderson dropped his bombshell, Blackett and Occhialini published a paper in Nature entitled "*Photography of Penetrating Corpuscular Radiation*" in which they boasted about their new toy, a cloud chamber in a high magnetic field, rigged up to work only when an accompanying

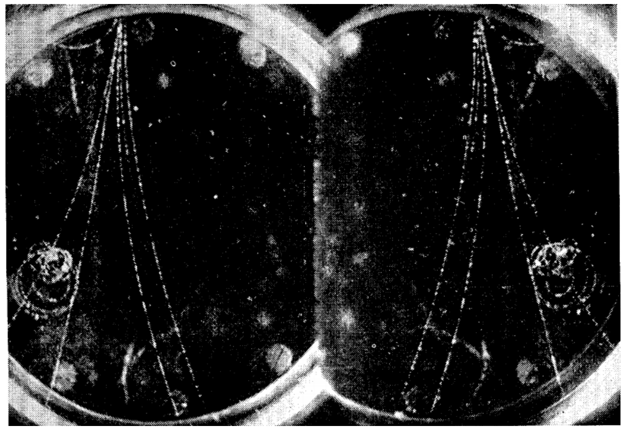


Figure 14. This cloud chamber track, showing electrons and positrons veering in opposite directions, was taken by Anderson and his student Seth Neddermeyer in 1936. They were still avoiding balloons, but climbed to Pike's Peak in the Rocky mountains, 4300 m above sea level, where, as they stated in the abstract, "the proportion of such tracks is considerably greater than at Pasadena."

Geiger counter triggered. They illustrated their paper with a few uninspiring photographs to show that the machine worked. It was only after they heard of Anderson's result that they realised they focussed on the wrong images: among their photographic plates were positrons in "great abundance". The pain of this oversight must have been felt even more acutely given that they were colleagues of Dirac. As a (very!) small compensation, they did quickly find something that Anderson had missed: the creation of electron-positron pairs within the chamber.

Later, when asked if they were aware of Dirac's theory when performing their experiment, Blackett replied that he ...

> "... could not recall but that it did not matter anyway because nobody took Dirac's theory seriously."

This seemed to be the prevailing attitude, at least among experimenters. Neither Anderson's original discovery paper, nor a longer paper published in 1933 in which he introduces the name *positron*, mentions the theory of Dirac[6]. Anderson later recalled:

[6]As a slightly bitchy aside, it appears that theory wasn't Anderson's strong point. His longer paper on the positron ends with the "the proton will then in all probability be represented as a complex particle consisting of a neutron and positron".

> "The Dirac work was not an important ingredient in deciding which way the experiments should be carried out or what should be done experimentally."

Rutherford, begrudging as always towards theorists, took things one grumpy step further:

> "It seems to me that in some way it is regrettable that we had a theory of the positive electron before the beginning of the experiments. Blackett did everything possible not to be influenced by theory, but the way of anticipating results must inevitably be influenced to some extent by theory. I would have liked it better if the theory had arrived after the experimental fact had been established."

B.2 Expecting a Meson

At this point our story of discovery gets somewhat out of sync with the main thread of the lecture notes. To understand what happened next, we must first make some comments on the strong nuclear force. We will describe the theory underlying this force in much more detail in Section 3.

The first hint that a new force was needed to explain the structure of the nucleus came from (who else?) Rutherford. In 1917, he started a series of experiments, following the set-up of Geiger and Marsden but replacing their sheets of metal with hydrogen.

The charge on a hydrogen nucleus is almost 80 times smaller than that of the gold used in the original Geiger-Marsden experiment. This means that the Coulomb repulsion is significantly smaller and the α particle can get much closer to the nucleus. Rutherford noted that the scattering of α-particles no longer agreed with his formula (A.2) that had worked so successfully in the past. Nor did it agree with a more detailed study by Darwin that assumes a Coulomb repulsion, but allows for scattering of the nucleus as well as the α-particle.

We now know that this is because the α-particle and hydrogen nucleus get close enough to experience the strong force. However, the world wasn't quite ready for a new force and Rutherford originally suggested that the effect could be explained by some deformation of the α-particle.

As time went on, this interpretation became increasingly untenable. At a meeting at the Royal Society in 1929, Rutherford stated clearly

> "The hydrogen and helium nucleus appears to be surrounded by a field of force of unknown origin"

But what are the properties of this force field?

The key insight was made in 1934 by the Japanese physicist Hideki Yukawa. He realised that if there were a new spin 0 particle[7] of mass m then it would give rise to a potential energy between particles which varies with the separation r as

$$V(r) \sim \frac{e^{-r/R}}{r}$$

This is now known as the *Yukawa potential*. It has the property that it quickly goes to zero for $r \gg R$, where the range R of the potential is inversely related to the mass as

$$R = \frac{\hbar}{mc}$$

This is the same relationship between energy and length that we met earlier in when describing the Compton wavelength (1.2). The data available at the time suggested that the strong force had a range of about $R \approx 2 \times 10^{-15}$ m. This meant that if you wanted to explain the strong force in terms of some field, you should be looking for a new particle approximately 200 times heavier than the electron, or

$$m \approx 100 \text{ MeV}$$

The idea languished until 1937, when just such a particle was found.

B.3 The Muon and the Pion

The discovery of the muon didn't happen overnight. There was no smoking gun event that people could point to and shout "Eureka". Instead it was more of a slow burn as, from 1934 to 1937, an increasing number of cosmic ray tracks had absorption properties that didn't seem to fit theoretical expectations.

The prime driver in these discoveries was, once again, Carl Anderson, now working with his recently-graduated student, Seth Neddermeyer. As they understood more about cosmic rays, they found tracks that didn't lose energy as quickly as theorists predicted. But it wasn't clear whether the theorists should be trusted, or whether the data contained something more interesting.

[7] This is draped in a little bit of hindsight. Yukawa's original paper suggests a massive spin 1 particle, but only looked at the contribution from the first component, analogous to focussing on the just the electrostatic potential and ignoring magnetic fields. Later, in 1937 with Sakata, he developed the theory with a massive spin 0 field.

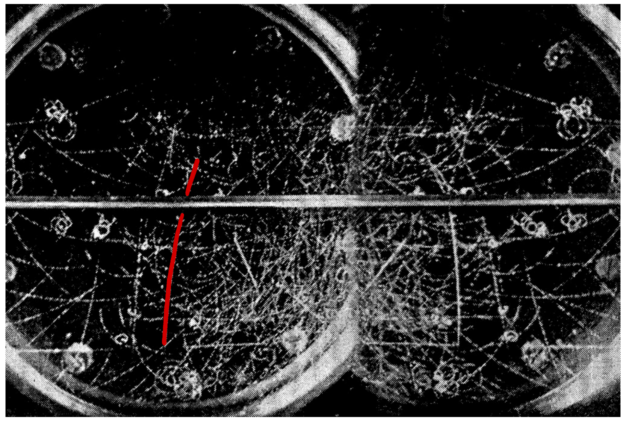


Figure 15. A cloud chamber picture from the same 1936 paper as Figure 14. The track daubed in red by my ipad pen was difficult to interpret as either an electron or a proton.

Already in 1936, Anderson and Neddermeyer published results which didn't conform to expectations. They pointed out that the red track shown in Figure 15 is too ionizing to be identified as an electron, but travels further than expected from a proton. In 1937, they finally bit the bullet, concluding that "there exist particles of unit charge but with a mass ... larger than that of a normal free electron and much smaller than that of a proton."

By 1939, the data seemed to suggest that this new particle had a mass 200 times heavier than the electron. The connection to Yukawa's proposed particle was obvious and a number of names were suggested for this new particle, including "yukon". Physical Review, pedantic as ever, insisted on "mesotron". Physicists ultimately converged on "meson", with the meso- from the Greek "mid".

However, as time went on, less and less about this new meson made sense. Models of nuclear binding worked much better with a particle that was slightly heavier and decayed significantly quicker. More brutally, experiments in 1946 showed that the interaction of the new meson with nuclei was around 10^{12} times weaker than that predicted by Yukawa's theory. That was too large a discrepancy to overlook!

Resolution

Happily the situation was resolved not long after. The discovery was made in 1947 in Bristol, England by a group of scientists led by Cecil Powell. Whenever Powell's collaborators are mentioned, people always include Giuseppe Occhialini (who recall, just missed out on the discovery of the positron) and sometimes César Lattes. But

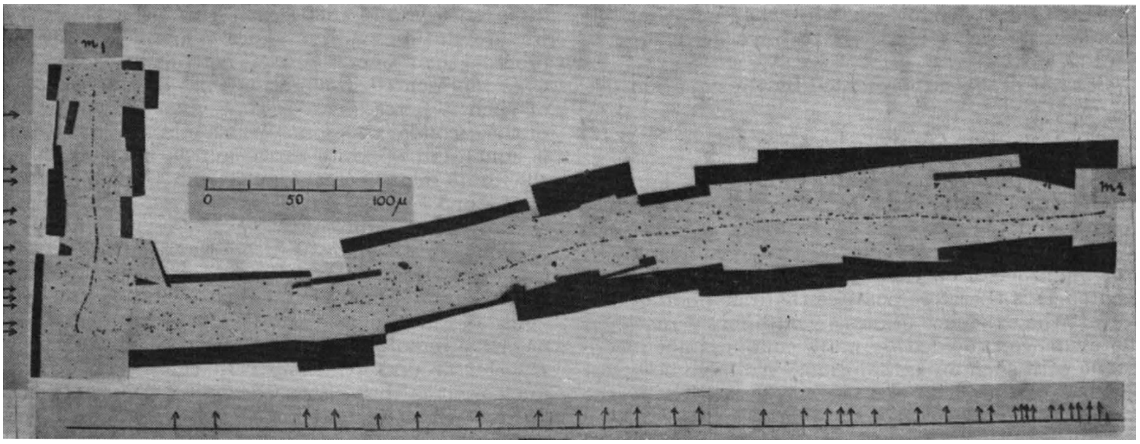


Figure 16. The discovery of the charged pion. It enters in the top left (labelled m_1), slows in the bromide and comes to rest, before decaying into a muon that flies off to the right (labelled m_2) and an anti-neutrino which is invisible in the picture. The caption in the paper starts with the comment "Observation by Mrs I. Roberts".

they rarely mention the fourth author on the paper, a poor graduate student by the name of Hughes Muirhead. And they certainly never mention the people behind the scenes who did the hard work: a team of women who painstakingly studied the images under microscopes to find the interesting events[8].

The discovery made by Powell's team was possible, as always, because they had a new piece of kit. Powell developed a new way of detecting particles by coating a glass plate with a photographic emulsion. When a charged particle passes through, it activates the emulsion, leaving behind a trail of silver bromide. They exposed their photographic plates to cosmic rays at high altitude, in balloons and on mountains, including Jungfraujoch and Kilimanjaro. When developed, the plates revealed a new meson, one both heavier and more short-lived, which decays quickly to the earlier meson. The V-shaped tracks were first found by Marietta Kurz, but these sat towards the edge of the emulsion and so were considered incomplete. A few days later, two clear L-shaped tracks were found by Irene Roberts. This was the long-sought meson.

Now, of course, we had two "mesons", rather than one. It quickly became clear that the particle discovered by Roberts had the properties expected of Yukawa's meson and, as we explain in the next section, can be viewed as the glue that binds together the proton and neutron in the nucleus. This became known as the π-meson, or pion. It is

[8] Actually, this statement was true when I wrote it, but then I decided to edit the Wikipedia page.

not an elementary particle but is, we now know, composed of quarks. Meanwhile, the particle discovered by Anderson and Neddermeyer is something new entirely, completely unrelated to the nuclear force. It became known as the μ-meson although, as time passed, the word meson was dropped. It is now simply the muon.

The muon-pion mix-up, which lasted a decade, is entirely due to a coincidence in their mass. We now know that the mass and lifetime of the two is

$$\begin{aligned} \pi^{\pm}: &\quad M \approx 140 \text{ MeV} \quad \text{and} \quad T \approx 2 \times 10^{-8} \text{ s} \\ \mu^{\pm}: &\quad M \approx 106 \text{ MeV} \quad \text{and} \quad T \approx 2 \times 10^{-6} \text{ s} \end{aligned}$$

The pion decays primarily as $\pi^- \longrightarrow \mu^- + \bar{\nu}_\mu$ and $\pi^+ \longrightarrow \mu^+ + \nu_\mu$. (You'll have to wait until we discuss the weak force in Section 4 to understand how this decay occurs.) Moreover, as we explain in Section 3.4, we have a good understanding of the pion in terms of its constituent quarks. We even understand why it has the mass that it does. In contrast, the muon remains a mystery, a repetition of the electron at a higher scale whose existence is as surprising to us today as in the 1940s.

B.4 The Beginning of the Deluge

Cosmic rays had still more surprises in store for physicists searching for elementary particles. The first came later in 1947, when George Rochester and Clifford Butler, working in Blackett's laboratory in Manchester, made a careful study of around 5000 photographs that they had taken the over the previous year. Among them they discovered two with peculiar features.

The first, shown on the left of Figure 17, contains a forked track, seemingly appearing from nowhere. This is due to a neutral invisible particle which subsequently decays into two charged particles which are either muons or pions. The second, shown on the right of Figure 17, shows a marked kink, strongly suggesting that a charged particle decayed into a different charged particle (again, either a muon or pion), but also an invisible neutral particle which left without leaving a track. These new particles were estimated to have masses between 770 m_e and 1600 m_e. They were dubbed *V-particles* on account of the V-shaped tracks that they left.

It was not long before further V-particles were discovered, some lighter than the proton, some heavier. Indeed, at some point it seemed like there would be no end to these new particles. In collecting his Nobel prize in 1955, Willis Lamb quipped

> "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. "

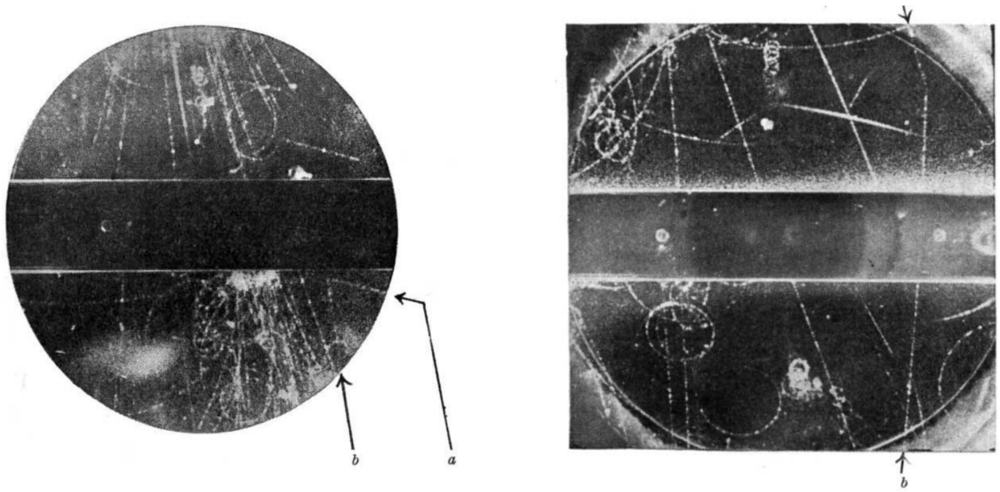


Figure 17. The discovery of kaons. A neutral kaon decays on the left, a charged kaon on the right.

For the next decade or so, particle physics entered a phase of taxonomy. The goal was to classify particles, first in terms of their mass and lifetimes, and then in terms of their decays, looking for patterns that would bring some order to the mess. It took many years before the things fell nicely into place and it was appreciated that all these particles could be understood in terms of yet smaller constituents called quarks. The V-particles shown in Figure 17 are now called *kaons* and were the first particles seen that contain a strange quark. We will tell the story of quarks in the next section.

The discovery of V-particles marks the beginning of a new era in theoretical physics. It also marks the end of an era in experimental physics. By the mid-1950's the energies and fluxes from accelerators were more than competitive with those in cosmic rays, and man-made muons, pions and V-particles were readily available. We will chart the rise of accelerator and detector technology in Interlude C.