3 The Strong Force

In the middle of the 20^{th} century, the number of seemingly "fundamental" particles exploded. The electron, proton and neutron had long been known. These were joined, in 1947, by a new particle called the *pion*, whose role was to keep the proton and neutron bound inside the nucleus.

The discovery of the pion was welcomed: it had long been expected, and its role in the universe was understood. The discovery, the same year, of another particle called the kaon was more confusing. As was the subsequent discovery of the rho meson, the eta, the eta prime, the delta, the lambda and the xi. Before long, physicists were running out of Greek and Roman letters to name these particles. A glance through the current particle data book includes particles with enticing names like $Z_c(3900)$.

By the end of the 1960s, there were hundreds of seemingly elementary particles and the situation looked hopelessly complicated. It was clear that these particles could not all be fundamental, but it was difficult to see any simple underlying explanation. In despair, Freeman Dyson declared

> "I am acutely aware of the fact that the marriage between mathematics and physics , which was so enormously fruitful in past centuries, has recently ended in divorce."

Yet Dyson accepted defeat too soon. The answer was discovered in the early 1970s. In part, the answer lay in the existence of constituent particles called *quarks*. But, equally as important, were the peculiar and unprecedented properties of the force that binds these quarks together. This is the *strong nuclear force*.

3.1 Yang-Mills Theory

Both the strong and weak nuclear forces share a common property with electromagnetism: the force is carried by a field of spin 1. In the case of electromagnetism, this field is described by the Maxwell equations. For the two nuclear forces, it is described by a generalisation of the Maxwell equations known as the *Yang-Mills equations*.

It's not so easy to write down a consistent generalisation of the Maxwell equations. In fact, it turns out that there's more or less a unique way to do it. This is based on a mathematical framework known as group theory. Here we'll give a baby version of this.

Yang-Mills theory is, like its predecessor, based on electric fields $\mathbf{E}(\mathbf{x}, t)$ and magnetic fields $\mathbf{B}(\mathbf{x}, t)$. Each of these is again a 3-dimensional vector,

$$\mathbf{E} = (E_x, E_y, E_z) \quad \text{and} \quad \mathbf{B} = (B_x, B_y, B_z)$$

This is the essence of what it means for a field to be spin 1. The novelty in Yang-Mills theory is that each component of these vectors is now a matrix at each point in space and time, rather than just a number. There are different versions of Yang-Mills theory for different kinds of matrices. There is a Yang-Mills theory based on 2×2 matrices, and one based on 3×3 matrices and so on for each integer N. However, once you decided on the size of the matrix, everything else is fixed.

To write the classical equations of motion, it's best to again bundle the "electric" and "magnetic" fields into a 4×4 matrix $F_{\mu\nu}$, each component of which is now itself an $N \times N$ matrix: i.e a matrix of matrices. The Yang-Mills equations of motion are one of the key equations of physics, and fully deserving of their place in a frame

$$\boxed{\mathcal{D}_\mu F^{\mu\nu} = J^\nu}$$

Here $\mathcal{D}_\mu$ is something like a partial derivative with respect to space and time, but one that also includes some commutator of matrices. (It's known as a covariant derivative.) On the right-hand side sits J^ν, the analog of the electric current which, as we will see shortly, arises from quarks. The Yang-Mills equations are very similar to the Maxwell equations. Indeed, if you choose 1×1 matrices, which are just numbers, then the Yang-Mills equations reduce to the Maxwell equations.

To specify any force described by Yang-Mills theory, we just need to say how big the matrices are. Nature is kind to us: she has chosen to make use only of the simplest matrices.

- Electromagnetism: 1×1 matrices
- Weak Nuclear Force: 2×2 matrices
- Strong Nuclear Force: 3×3 matrices

Isn't that nice!

Before we go on, I should confess that the above discussion was a little imprecise in places. A correct statement is that there exists a version of Yang-Mills theory for every Lie group. Everywhere that I said “size of matrix”, you should replace this with “choice of Lie group” where a “Lie group” is a fancy mathematical object. Matrices provide some simple examples of some Lie groups, but there are also others. Happily, all the groups that we need in particle physics can be reduced to matrices. A more grown-up version of the above list then characterises each force by a group. For what it’s worth, they are:[9]

- Electromagnetism: $U(1)$
- Weak Nuclear Force: $SU(2)$
- Strong Nuclear Force: $SU(3)$

For the rest of this section, we will focus on $SU(3)$ Yang-Mills, relevant for the strong force. The theory of the strong force, interacting with quarks, is known as *Quantum Chromodynamics*, or QCD for short. The 3×3 matrix-valued electric and magnetic fields are sometimes called *chromoelectric* and *chromomagnetic* fields.

3.1.1 Gluons and Asymptotic Freedom

If you solve the Maxwell equations, you find waves propagating at the speed of light. These are light waves. As we have seen, in the quantum theory, these waves are comprised of massless spin 1 particles called photons. The massless nature of the photon is the reason light waves travel at the speed of light.

Similarly, if you solve the classical Yang-Mills equations, you again find waves travelling at the speed of light. In analogy with electromagnetism, we might expect that, in the quantum theory, there are massless particles associated to these waves. But no such massless particles are seen in the world. What’s going on?

[9] A little matrix knowledge can be a confusing thing at this stage. The strong nuclear force is associated to the group $SU(3)$, which consists of 3×3 complex, unitary matrices of determinant 1. This means that the matrix U must obey $U^\dagger U = \mathbb{1}$ and $\det U = 1$ to be in $SU(3)$. But this isn’t the kind of 3×3 matrix that make up the components of the chromoelectric and chromomagnetic fields. Instead, these are Hermitian matrices, namely 3×3 matrices which obey $E_x = E_x^\dagger$. There is, however, a relationship between these kinds of matrices: the exponential of a Hermitian matrix, like e^{iE_x}, is in the group $SU(3)$. Mathematically, this is the difference between a Lie group and a Lie algebra.

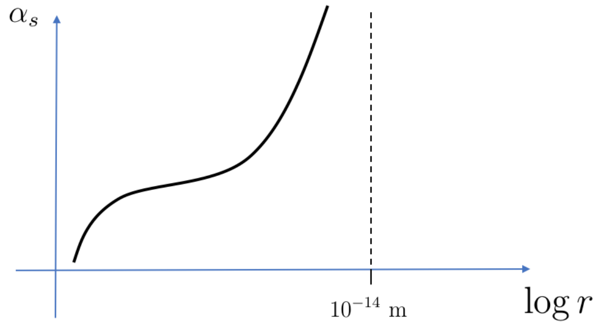
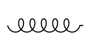


Figure 18. The renormalisation of coupling for the strong force.

The answer to this question lies in a subtle property of Yang-Mills theory, which means that the quantum theory looks very different from its classical counterpart. There are spin 1 particles associated to the $SU(3)$ Yang-Mills theory and these are called *gluons*. But, rather surprisingly, gluons turn out to be massive rather than massless.

To understand this, we first need to appreciate a key difference between Yang-Mills theory and Maxwell theory. The Maxwell equations are linear. One consequence of this is that light wave pass right through each other. In contrast, the Yang-Mills equations are non-linear. (This is not obvious in the framed equation on the previous page. It is hidden in the meaning of the covariant derivative $\mathcal{D}_\mu$.) This means that two classical waves in Yang-Mills will typically scatter off each other in some complicated fashion.

When we turn to the quantum theory, the non-linearity translates to an interaction vertex in the Feynman diagrams for gluons. We depict gluons using curly lines like this  . There are, it turns out, two interaction vertices: one where three gluons interact, and another where four gluons interact:

Just as in QED, there is a dimensionless coupling constant that characterises the strength of this interaction. For QED, this was the fine structure constant α. For the strong force, the coupling is denoted α_s. And, importantly, just like for QED, the value of this coupling depends on the distance scale (or, equivalently, energy scale) at

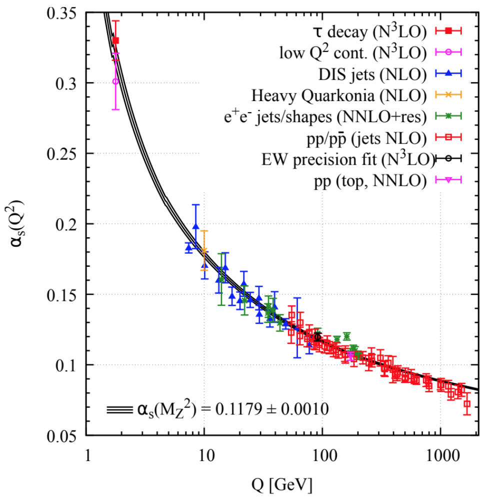


Figure 19. The running of coupling for the strong force, now plotted against energy scale. This figure is taken from the review of QCD by the Particle Data Group.

which it's measured. This is the story of renormalisation that we met previously in Section 2.3. But here is where there's a crucial difference between electromagnetism and the strong force: as we go to larger distances, the strong force gets *stronger*, not weaker. A sketch of the coupling is shown in Figure 18. (This should be contrasted with the running of the fine structure constant, as shown in Figure 10.) The experimental data for the running of the strong coupling is shown in Figure 19, now plotted against energy scale E, which is inversely related to length by $E = 1/r$.

How can we understand this intuitively? For electromagnetism, there was a simple physical picture in which the electric charge gets screened by particle-anti-particle pairs, and so appears smaller as we go to longer distances. For the strong force, the gluons themselves are doing the screening. Except they *anti-screen*, meaning that they cause the force to get stronger the further out we go!

In fact there is an intuitive way to understand this, although it's rather subtle. A clue can be found lurking back in the theory of electromagnetism. Recall that the Maxwell equations contain two parameters: $1/\epsilon_0$ characterises the strength of the electric force, while $1/\mu_0$ characterises the strength of the magnetic force. But these are not independent. They are related by

$$\epsilon_0 \mu_0 = \frac{1}{c^2} \tag{3.1}$$

with c the speed of light. This means that if the strength of the electric force gets weaker, then the magnetic force necessarily gets stronger, and vice versa.

There is a similar story for Yang-Mills. But now the gluons are doing the screening, and couple to both the chromoelectric and the chromomagnetic fields. It turns out that, because they are spin 1 particles, they screen the chromomagnetic fields more strongly than they screen the chromoelectric fields. In other words, the chromomagnetic part of the Yang-Mills interaction gets weaker as we go to larger distances. The relation (3.1) then tells us that the chromoelectric part of the interaction necessarily gets stronger. The upshot is that gluons anti-screen. If you want the gory details, the calculation can be found in Section 2.4 of the lectures on Gauge Theory.

So what is the strength of the strong force? At the energy scale $E \approx 100$ GeV, corresponding to a distance scale of $r \sim 10^{-17}$ m, we have

$$\alpha_s \approx 0.1 \quad \text{when} \quad E \approx 100 \text{ GeV}$$

Even at these fairly high energies, the strength of the force is an order of magnitude larger than QED. If we go to higher energies, or shorter distances, α_s decreases. In fact, as we go to arbitrarily high energies, the strength of the strong force vanishes, $\alpha_s \to 0$. This phenomenon is known as *asymptotic freedom*: it means that at high energies, or short distance scales, the strong force essentially disappears!

However, outside of particle colliders, everything that we observe in the world takes place at distance scales significantly larger than 10^{-17} m, and the strong force only gets stronger as we go to larger distances. But here there is another surprise. According to naive calculations, by the time you get to around $r \sim 10^{-15}$ m, or an energy scale of $E \sim 100$ MeV, the coupling constant appears to get infinitely large!

Above, I used the phrase "naive" calculations, because I have in mind the kind of perturbative Feynman diagram calculations that we described in the previous section. But, as we stressed, these diagrams only make sense when the coupling is small. As soon as the coupling is around $\alpha_s \approx 1$, the very complicated Feynman diagrams, involving lots of loops, are just as important as the simple Feynman diagrams and we have no control over the calculation. But this is exactly what happens in QCD! At very short distances, we're fine and we can do calculations. But at long distances, the theory becomes very challenging. The separation between "easy" and "hard" turns out to be around $r \sim 10^{-14}$ m to 10^{-15} m, but is usually expressed in terms of an energy scale known as the *strong coupling scale*, or *Lambda-QCD*,

$$\Lambda_{\text{QCD}} \approx 200 \text{ MeV}$$

This is a characteristic energy scale of QCD. At energies $E \gg \Lambda_{\text{QCD}}$, the strong force is not particularly strong and we can trust Feynman diagrams. But by the time we get to energies Λ_{QCD}, the strong force lives up to its name. Most phenomena that are due to the strong force have an energy somewhere in the ballpark of Λ_{QCD}

Before we describe some of these phenomena, it's worth pausing to mention that something unusual has happened here. The strength of the strong force, like QED, is characterised by a dimensionless coupling α_s. But the phenomena of renormalisation means that this coupling depends on scale, and the upshot of this is that we ultimately exchange a dimensionless number, α_s, for a dimensionful scale Λ_{QCD}.

3.1.2 The Mass Gap

When we try to study the strong force on energy scales $E < \Lambda_{\text{QCD}}$, corresponding to distance scales, $r > 10^{-14}$ m, we have a problem. The strong coupling means that the fields are wildly fluctuating on these scales, and our favourite method of Feynman diagrams is no longer useful. This also means that the classical equations of motion are no guide at all for what the quantum theory might look like.

The gluon is the first casualty of strong coupling. As we explained at the beginning of this section, the classical Yang-Mills equations suggest that the gluon should be massless. But the strong coupling effects change this. Instead, the gluon — which is a ripple of the Yang-Mills fields —has mass, given by

$$m_{\text{gluon}} \approx \Lambda_{\text{QCD}}$$

This is sometimes referred to as the *Yang-Mills mass gap*. The "gap" here is one between the ground state and the first excited state. For theories of massless particles, there is no gap because we can have particles of arbitrarily low energy. But for massive particles, the minimum amount of energy needed is $E = mc^2$.

To say that the Yang-Mills mass gap is difficult to prove would be something of an understatement. Demonstrating the mass gap is generally regarded as one of the major open problems in theoretical physics. Indeed, a million dollar Clay mathematics prize awaits anyone who succeeds. Although we do not have any rigorous (or even semi-rigorous) derivations of the mass gap, there is no doubt that it is a property of Yang-Mills. Our best theoretical evidence comes from computer simulations which, in this context, are called *lattice simulations*, reflecting the fact that spacetime is approximated by a grid, or lattice, or points. These simulations show unambiguously that the gluon is massive. You can read more about the lattice, and other approaches to Yang-Mills, in the lectures on Gauge Theory.

Finally, and most importantly, the existence of a mass gap is consistent with experiment, where no massless gluon is seen. Moreover, while the strong force is strong, it is also short-ranged. The characteristic energy scale $\Lambda_{\rm QCD}$ corresponds to a distance scale which in natural units (remember $\hbar = c = 1$) is

$$R_{\rm QCD} = \frac{1}{\Lambda_{\rm QCD}} \approx 5 \times 10^{-15}\ {\rm m}$$

To understand how this scale affects the world around us, we first need to throw in the final key ingredient: quarks.

3.2 Quarks

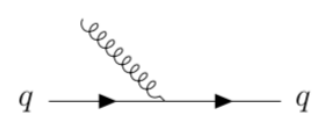
Usually in physics, we can get by without memorising lots of random names. The strong force is the exception, leaving us looking more like botanists than physicists. First, there are the many hundreds of names of different particles. But more important are names of groups of particles, each classifying a different property.

To kick things off, the fermions in the Standard Model are divided into two different types

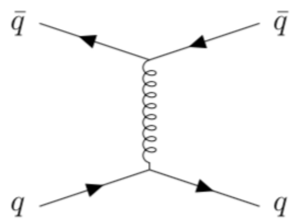
- Quarks: These are particles that feel the strong force.
- Leptons: These are particles that don't.

The leptons are the electron, muon, tau and three species of neutrino. (The name comes from the Greek $\lambda\epsilon\pi\tau\acute{o}\zeta$ meaning small.) Leptons don't interact with the $SU(3)$ Yang-Mills field, and we will ignore them for the rest of this section. In contrast, the six quarks — up, down, strange, charm, bottom and top — do feel the strong force.

In the language of Feynman diagrams, we denote the quarks as a solid line, with an arrow distinguishing quark from anti-quark. This is the same kind of line that we previously used to denote leptons, so we add a label q to show that it's a quark. The interaction between quarks and gluons is then described by the interaction vertex



When evaluating these Feynman diagrams, each vertex contributes a factor of the strong coupling, α_s. We can then use the Feynman diagrams to compute, say, the force between a quark and anti-quark. This comes from the following diagram:



If the quark and anti-quark are separated by a distance $r \ll R_{\rm QCD}$, then evaluating this diagram results in an attractive force that is very similar to the Coulomb force (2.4),

$$F(r) \sim \frac{\alpha_s}{r^2} \quad \text{when } r \ll R_{QCD} \tag{3.2}$$

Here α_s itself also depends on r, albeit logarithmically. We already saw a sketch of this dependence in Figure 18.

However, there's a catch: if the distance between the quarks is too big — bigger than R_{QCD} — then the language of Feynman diagrams stops working. As we increase the separation between quarks to distances greater than $R_{\rm QCD}$, the Coulomb-like expression (3.2) stops being the right one, and it instead changes to

$$F(r) \sim \text{constant} \quad \text{when } r \gg R_{QCD} \tag{3.3}$$

A constant force may not seem like much. But it gets exhausting. This is better seen if we look at the associated energy needed to separate a quark and anti-quark by distance R. For short distances, the energy takes the same form as in electrostatics,

$$V(r) \sim -\frac{\alpha_s}{r} + \text{constant} \quad \text{when } r \ll R_{QCD}$$

But when the quarks experience a constant force, the energy grows linearly

$$V(r) \sim \Lambda_{\rm QCD}^2\, r \quad \text{when } r \gg R_{QCD}$$

Clearly if you want to separate the quark-anti-quark pair by a long distance, then it costs an increasing amount of energy. In particular, it costs an infinite amount of energy to separate them an infinite distance. But taking, say, the anti-quark a long way away is tantamount to leaving the quark on its own. In other words, a solitary quark requires infinite energy! Quarks do not want to be alone: they only occur in bound states with other quarks or anti-quarks. This phenomenon is called *confinement*.

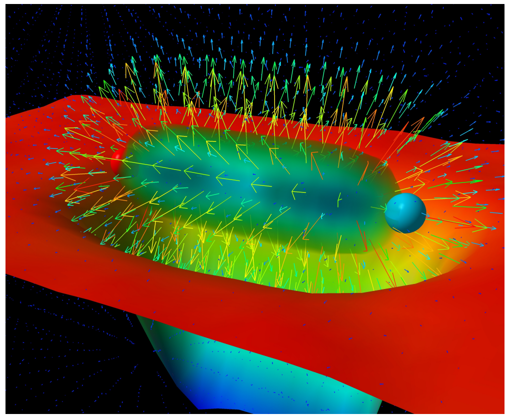
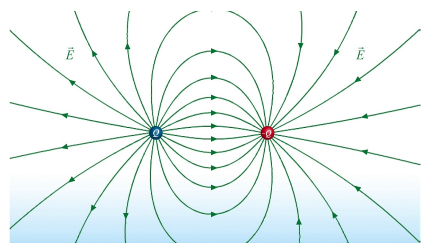


Figure 20. The chromoelectric flux tube between a quark and anti-quark in a meson state, from the QCD simulation of Derek Leinweber

Confinement, like the Yang-Mills mass gap, has so-far resisted a rigorous mathematical derivation. But we again have very clear evidence from numerical simulations, together with a handful of less-than-rigorous mathematical arguments, that confinement occurs. Moreover, this also gives us some intuition for what's going on.



First, let's recall what happens in electrostatics. If we separate a positive and negative electric charge by some distance r, then an electric field is set up between the two. The form of this electric field is shown on the right, and ultimately is responsible for the $F \sim 1/r^2$ Coulomb force law that the charges experience.

In Yang-Mills theory, the chomoelectric field takes a similar form if the quark and anti-quark are separated by a distance $r \ll R_{QCD}$. But as you increase the separation of the quark and anti-quark beyond the critical distance R_{QCD}, the form of the field changes. Instead of the field lines spreading out, the mass of the gluon forces them to bunch together into string-like configurations called *flux tubes*. This can be clearly seen in the computer simulation of QCD shown in Figure 20. It's as if the quark and anti-quark are joined by a piece of string. If you want to separate them further, you have to stretch the flux tube and this costs an energy $V(r) \sim r$ proportional to its length. This is responsible for the confinement of quarks.

3.2.1 Colour

The property that determines how particles experience the electromagnetic force is electric charge. The analogous property for the strong force is called *colour*[10]. Needless to say, this has nothing to do with the colour that we see. It is merely a label given by physicists who grew up without the classical education needed to name things in Greek or Latin.

While electric charge is just a number, colour charge is a little more involved: it is best thought of as a 3-dimensional vector $\boldsymbol{\omega}$ of fixed length. Here it's 3-dimensional because the Yang-Mills fields for the strong force are 3×3 matrices. This vector doesn't point in the three dimensions of space, but instead is something more abstract. If it points in different directions, we think of the quark as carrying a different colour. The exact translation between the vector and colour is pretty arbitrary, but we could take

$$\text{red:}\ \ \boldsymbol{\omega} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad\qquad \text{green:}\ \ \boldsymbol{\omega} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad\qquad \text{blue:}\ \ \boldsymbol{\omega} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

But the vector doesn't have to point exactly in one of these directions. For example, we could consider a vector $\boldsymbol{\omega} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ which should be thought of as a combination (or, in quantum language, a superposition) of red and green. Fortunately, we don't extend the colour analogy so far as to call this "muddy brown".

3.2.2 A First Look at Mesons and Baryons

Each quark carries a colour-vector $\boldsymbol{\omega}$, while the leptons do not. Confinement means that we do not see individual quarks, but they only appear in bound states so that the total colour charge vanishes. Any such composite particle, held together by the strong force, is referred to as a *hadron*. There are two ways in which hadrons can form,

- Mesons: These contain a quark and anti-quark. If the quark has colour vector $\boldsymbol{\omega}_1$ and the anti-quark colour vector $\boldsymbol{\omega}_2$, then they can combine as $\boldsymbol{\omega}_2^\dagger \cdot \boldsymbol{\omega}_1$ to form a colour-neutral state. Schematically, this looks like

  $$\text{meson} = \bar{r}r + \bar{g}g + \bar{b}b$$

  which should be read as "(anti-red)red + (anti-green)green + (anti-blue)blue". The flux tube for such a meson is shown in Figure 20.

[10] Scientists in the US work in units $u = 1$.

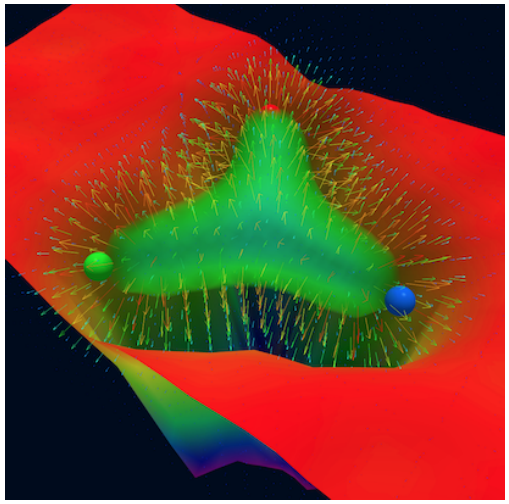


Figure 21. The flux tube between three quarks in a baryon state, from the QCD simulation of Derek Leinweber

- Baryons: These contain three quarks. If the three quarks have colour vectors $\boldsymbol{\omega}_1$, $\boldsymbol{\omega}_2$ and $\boldsymbol{\omega}_3$ then they combine as the triple product $\boldsymbol{\omega}_1 \cdot (\boldsymbol{\omega}_2 \times \boldsymbol{\omega}_3)$ to form a colour neutral state. Schematically, this looks like

$$\text{baryon} = rbg$$

  The fact that the baryon contains 3 quarks, rather than any other number, can be traced to the 3×3 matrices that describe the strong force. The flux tube for a baryon is shown in Figure 21.

To understand the collection of hadrons that emerges after confinement, we first need to look at the masses of quarks. In particular, we should compare the masses to the characteristic scale of the strong interactions, $\Lambda_{QCD} \approx 200$ MeV.

Three of the quarks have masses smaller than Λ_{QCD}.

$$\begin{aligned} m_{\text{down}} &= 5 \text{ MeV} \\ m_{\text{up}} &= 2 \text{ MeV} \\ m_{\text{strange}} &= 95 \text{ MeV} \end{aligned}$$

The up and down quark have masses significantly smaller than Λ_{QCD}, while the strange quark is only slightly smaller. Recall from our discussion in 1.2 that the Compton wavelength, $\lambda = \hbar/mc$, can be thought of as the size of a particle. Any particle with

$m < \Lambda_{QCD}$ necessarily has its size $\lambda > R_{QCD}$. That means that there's no way to bring two of these quarks closer than R_{QCD}, so these light quarks will only experience the confining force (3.3).

In contrast, the three heavier quarks have masses

$$\begin{aligned} m_{\text{charm}} &= 1.3 \text{ GeV} \\ m_{\text{bottom}} &= 4.2 \text{ GeV} \\ m_{\text{top}} &= 170 \text{ GeV} \end{aligned}$$

These are all much heavier than Λ_{QCD}. These three quarks all have Compton wavelength $\lambda \ll R_{QCD}$, so it makes sense for them to come close enough to experience the Coulomb-like force (3.2). This means that we might expect the spectrum of hadrons containing charm, bottom and top quarks to be a little different from those containing only the lighter quarks. Indeed, this turns out to be the case.

3.3 Baryons

We'll kick thing off with baryons, containing three quarks. To start, suppose that we have only the up and down quark to work with. There are various ways that we can combine these quarks. First, recall that the each quark has spin $\frac{1}{2}$. When combining quarks, we need to figure out what to do with their spins.

3.3.1 Protons and Neutrons

Suppose that we have two spins in one direction, and the third spin in the opposite direction. This will result in a baryon of spin $\frac{1}{2} + \frac{1}{2} - \frac{1}{2} = \frac{1}{2}$. There are two choices, which result in the two most familiar baryons: the proton (p) and the neutron (n). Their quark content and masses are

$$\begin{aligned} n\ (ddu) &\quad m_n \approx 939.57 \text{ MeV} \\ p\ (uud) &\quad m_p \approx 938.28 \text{ MeV} \end{aligned}$$

These are the two lightest spin $\frac{1}{2}$ baryons. Recall that the down quark has charge $-1/3$ and the up quark charge $+2/3$, so the proton has charge $+1$ while the neutron has no electric charge.

Already, there is something of a surprise here. The up and down quarks each have mass of a few MeV. Yet the proton and neutron each have mass of around 1000 MeV. How is this possible given that the proton and neutron are supposed to contain three quarks each?

The answer to this is simple: when we say that the proton and neutron each contain three quarks, we are hiding a more painful truth. The proton and neutron, and indeed all other hadrons, are in reality enormously complicated objects. The most accurate description is in terms of complicated, strongly interacting fields. In a particle language, we could describe them as containing many hundreds of quarks, anti-quarks and gluons, all interacting in a complicated fashion. The statement that the proton and neutron are composed of three quarks is really shorthand for the fact that they contain three more quarks than anti-quarks. These three additional quarks are sometimes called *valence quarks*, to distinguish them from the surrounding sea of quark-anti-quark pairs.

To put this in perspective, we could ask the following hypothetical question: suppose that the quarks were actually massless. What would the mass of the proton be? The surprising answer is that the mass of the proton would be more or less unchanged, still weighing in at a little less than 940 MeV! The same is true of the neutron. The masses of the proton and neutron care almost nothing about the mass of the three valence quarks: instead they are entirely dominated by the strong coupling scale $\Lambda_{\rm QCD}$ and their mass is few times $\Lambda_{\rm QCD}$.

You may have heard it said that the Higgs is responsible for all the mass in the universe. This is a fairly blatant lie. Later, in Section 4, we will learn that all elementary particles, including the quarks, do indeed get their mass from the Higgs boson. But the overwhelming majority mass in atoms is contained in the protons and neutrons that make up the nucleus, and this mass has nothing to do with the Higgs boson. It is entirely due to the urgent thrashing of strongly interacting quantum fields.

The irrelevance of the quark masses also has a more subtle implication. The masses of the proton and neutron are almost equal, despite the fact that the down quark is twice as heavy as the up. This reflects the fact that, at least as far as the strong interaction is concerned, the two particles behave almost identically. If we do an experiment with protons that is mediated by the strong force, then the same experiment performed with neutrons will yield exactly the same answer. This almost-symmetry of nature is referred to as *isospin*. Note, however, that it only holds for the strong force. The proton and neutron do not behave the same under the electromagnetic force, since the proton is charged while the neutron is not. Neither, it turns out, do they behave the same under the weak force.

There is, however, one place where the masses of the quarks are important. The fact that the down quark is heavier than the up quark does contribute a tiny amount to the mass and is the reason that the neutron is very slightly heavier than the proton. This is important because it means that beta decay proceeds by a neutron decaying into a proton, rather than the other way around. We'll learn more about this in Section 4 where we discuss the weak force.

3.3.2 Delta Baryons

We could ask: why can't we have a baryon with, say, three up quarks? The answer to this lies in the Pauli exclusion principle. The full explanation is a little subtle, but the upshot is that we can have three up quarks in a baryon, but only if all their spins point in the *same* direction. (At first glance this might seem the wrong way around since, in chemistry, the Pauli exclusion principle dictates that electrons in the same orbital state have opposite spins. But quarks have that additional colour degree of freedom, and there is an anti-symmetry there which, in turn, requires a symmetric alignment of spins.)

When all the spins point in the same direction, the baryon itself has spin $\frac{1}{2}+\frac{1}{2}+\frac{1}{2}=\frac{3}{2}$ baryon. Now there are four choices, all of which are known as Delta (Δ) baryons. These particles have more or less equal mass, but different charges

$$\left.\begin{array}{ll} \Delta^{++} & (uuu) \\ \Delta^{+} & (uud) \\ \Delta^{0} & (udd) \\ \Delta^{-} & (ddd) \end{array}\right\} \quad m \approx 1232 \text{ MeV}$$

Here the superscripts, ++, +, 0 and - specify the electric charge of the particle.

We don't see Δ baryons floating around in the world. They have a lifetime of around 10^{-24} seconds, after which they decay, typically into a proton or neutron together with a meson called a pion. For example,

$$\Delta^{++} \to p + \pi^{+} \quad \text{and} \quad \Delta^{-} \to n + \pi^{-}$$

The lifetime of 10^{-24} seconds is much shorter than we can measure and particles with such short lifetimes are called, quite reasonably, *unstable*. Even when moving close to the speed of light Δ baryons don't travel far enough to register directly in particle detectors. Instead, they reveal themselves in more indirect means as so-called *resonances* in certain experiments. We'll describe this further in Interlude C.1.

The lifetime of the Δ baryons is actually the characteristic timescale of the strong force: $T_{QCD} = R_{QCD}/c \approx 10^{-24}$ seconds. If anything happens due to the strong force, it usually happens on roughly this timescale.

3.3.3 Strangeness

Let's now consider baryons that we can construct from the first three quarks: up, down and strange.

We'll first consider baryons with spin $\frac{1}{2}$. In addition to the proton and neutron, we now have four further baryons that contain a single strange quark, called sigma (Σ) baryons

$$\begin{aligned} \Sigma^-(dds) \quad & m \approx 1197 \text{ MeV} \\ \Sigma^0\,(dus) \quad & m \approx 1193 \text{ MeV} \\ \Sigma^+(uus) \quad & m \approx 1189 \text{ MeV} \end{aligned}$$

and the lambda (Λ) baryon

$$\Lambda^0\,(dus) \quad m \approx 1116 \text{ MeV}$$

Again, the superscript labels the electric charge of the baryon. You may have noticed that the quark content of the Σ^0 and Λ^0 are the same. The difference lies in the details of the wavefunctions for the up and down quarks. (Technically, some different minus signs mean that all Σ baryons have isospin 1, while the Λ has isospin 0.)

There are also two types of baryons that contain two strange quarks, called cascade, or xi (Ξ) baryons

$$\begin{aligned} \Xi^-(dss) \quad & m \approx 1322 \text{ MeV} \\ \Xi^0\,(uss) \quad & m \approx 1315 \text{ MeV} \end{aligned}$$

None of these baryons are familiar from our everyday experience. This is because they again decay, typically to protons and pions. However, here there is a surprise: although these new baryons have a mass in the same ballpark as the Δ's, they live for significantly longer. In particular, the $\Sigma^\pm$, the Λ^0 and the $\Xi^{0,-}$ all live for a whopping 10^{-10} seconds.

Now, 10^{-10} seconds may not sound like much. Indeed, it's difficult to imagine having a rich and fulfilling life in this time. But it's an aeon compared to the 10^{-24} seconds that the Δ baryons live. This is a puzzle: why do these new baryons have a such a comparatively long life, even though their masses are comparable to the Δ?

A partial answer to this is to invoke a new conservation law. We know that electric charge is conserved in all interactions. It turns out that there is another quantity – *strangeness* – which is conserved. Or, at the very least, almost conserved. Strangeness is simply a count of the number of strange quarks.

Here "almost conserved" means conserved by the strong interaction. The strong interactions cannot change the number of strange quarks and so particles like $\Sigma^{\pm}$, Λ^0 and $\Xi^{0,-}$ do not decay straight away. The decay only proceeds through the weak interaction, and this takes significantly longer. We'll describe how these decays occur in Section 4. In contrast, particles like the Δ baryons decay directly through the strong interaction, and this happens much faster.

(As an aside: there's always one complication. It turns out that, among the collection of strange baryons, there is one which is unstable: the $\Sigma^0 \to \Lambda^0 + \gamma$ with a lifetime of around 10^{-20} seconds. But this is allowed by the strong force because the number of strange quarks is unchanged. The Λ^0, as we've seen, then waits another 10^{-10} seconds before it too decays.)

As a general rule of thumb, hadrons can decay through one of the three forces: strong (like the Δ's), electromagnetic (like Σ^0) or weak (like $\Sigma^{\pm}$, Λ^0 and Ξ.). The lifetimes of these particles reflect the decay process:

- Strong decay: $\sim 10^{-22}$ to 10^{-24} seconds.
- Electromagnetic decay: $\sim 10^{-16}$ to 10^{-21} seconds.
- Weak decay: $\sim 10^{-7}$ to 10^{-13} seconds.

Where you sit within each range depends on other factors, such as the relative masses of the parent and daughter particles. Particles that live for up to 10^{-10} seconds are referred to (I think, somewhat tongue in cheek) as *stable*. In contrast, any particle that lasts 10^{-20} seconds or shorter is, like the Δ baryon, referred to as *unstable* or a *resonance*.

It should be clear from the discussion, however, that there's nothing very qualitatively different between a stable particle like the Λ and a resonance like the Δ. Both will decay in less than the blink of an eye. But a lifetime of 10^{-10} seconds mean that, with good technology, you can take a photograph of the particle's track in a cloud chamber or bubble chamber. You can see many such photographs in Interludes B and C. When a particle leaves such a vivid trace, it's hard to deny its existence. In contrast, we're never going to take a photograph of something that lasts 10^{-20} seconds. But that doesn't mean that it's any less real! It just leaves its signature in more subtle ways.

3.3.4 The Eightfold Way

There is a clear pattern to the masses of the spin $\frac{1}{2}$ baryons. To highlight this, we can place the 8 baryons in the following shape:

$$\begin{array}{ccccc} & n\ (ddu) & & p\ (uud) & \\ \Sigma^-\ (dds) & & \Sigma^0, \Lambda^0\ (dus) & & \Sigma^+\ (uus) \\ & \Xi^-\ (dss) & & \Xi^0\ (uss) & \end{array}$$

The particles are arranged in rows of increasing strangeness and, correspondingly, in increasing mass. The particles in the top row have $m \approx 940$ MeV; those in the second row $m \approx 1190$ Mev; and those in the final row $m \approx 1320$ MeV. Meanwhile, particles in the same $\backslash$ diagonal have equal charge.

We see that adding a strange quark increases the mass of a baryon by roughly 140 ± 10 MeV. This can be largely, but not entirely, accounted for by the mass of the strange quark, $m_{\text{strange}} \approx 95$ MeV.

Recall that there is an approximate symmetry, relating up and down quarks, called isospin. If we squint, we could enhance this to a larger, and even more approximate, symmetry relating up, down and strange quarks. This subsumes the previous isospin symmetry, but isn't quite as a good because the strange quark is significantly heavier than the other two. Indeed, it only holds if we're willing to pretend that 1320 MeV $\approx$ 940 MeV. Nonetheless, it is true that, at least as far as the strong force is concerned, all 8 baryons in the table have more or less the same properties. For example, a given meson will scatter off all 8 in pretty much the same way. The symmetry, first proposed by Gell Mann and, independently, Ne'eman in 1961 is called $SU(3)$ flavor symmetry or, more poetically, the *eightfold way*.

The $SU(3)$ flavor symmetry relating the three lightest quarks is not to be confused with the $SU(3)$ colour symmetry that underlies the strong force. It's just that the number 3 appears a bunch of times in the Standard Model. (Don't read anything mystic into this. It's just a small number.) With decades of hindsight, the eightfold way appears more accidental than fundamental. Nonetheless, it was important historically as a useful organising principle in cataloging the many hundreds of hadrons that were discovered.

The same almost-symmetry is sitting within the baryons of spin $\frac{3}{2}$. These can be placed in the following pattern, again organised by increasing strangeness

$$
\begin{array}{ccccccc}
\Delta^- \ (ddd) & & \Delta^0 \ (ddu) & & \Delta^+ \ (duu) & & \Delta^{++} \ (uuu) \\
& \Sigma^{\star -} \ (dds) & & \Sigma^{\star 0} \ (dus) & & \Sigma^{\star +} \ (uus) & \\
& & \Xi^{\star -} \ (dss) & & \Xi^{\star 0} \ (uss) & & \\
& & & \Omega^- \ (sss) & & &
\end{array}
$$

Once again, the masses in each row are roughly constant: $m \approx 1232$ MeV in the first row; $m \approx 1385$ MeV in the second; $m \approx 1533$ MeV in the third; and with the Ω^- weighing in at $m \approx 1672$ MeV. In each case, the increase is again roughly 140 to 150 Mev and is due to the extra strange quark.

Notice that the middle 7 baryons in this table have the same quark content as the spin $\frac{1}{2}$, strangeness 1, baryons that we met previously. These should be thought of excitations of these previous baryons, with the spin of one of the constituent quarks flipped, changing the overall spin from $\frac{1}{2}$ to $\frac{3}{2}$.

The real novelties in the table above are the three outliers, in which all quarks are the same. As we mentioned above, the Pauli exclusion principle prohibits the existence of spin $\frac{1}{2}$ baryons with three identical quarks, so they appear here for the first time. Two are particularly important: the Δ^{++} was the first particle to be found without charge ± 1 (or 0) and helped enormously in piecing together the story of the underlying quarks. The Ω^- baryon, meanwhile, holds a special place in the history of science because Gell-Mann used the simple quark model described above to predict its mass and properties before it was discovered experimentally. He therefore followed Mendeleev and Dirac in predicting the existence of a "fundamental" particle of nature (where, as should by now be clear, the meaning of the word "fundamental" is time-dependent).

There are eight baryons with spin $\frac{1}{2}$, and ten baryons with spin $\frac{3}{2}$, with the former lending its name to the "eightfold way" used to describe the whole enterprise. But, underlying this set-up are just 3 quarks — up, down and strange — and, as we mentioned above, an approximate $SU(3)$ symmetry that rotates them. Why do we start with the number 3, and end up with 8 and 10 baryons respectively? In fact, there is a good reason for this, although it's difficult to explain without going into the mathematics of group theory. It turns out that any group has a bunch of different ways in which it can express itself, known as *representations*. And these representations come only with very specific numbers. Although it's not obvious, the numbers 8 and 10 are naturally associated to the group $SU(3)$.

Finally, I mention that, in principle, it should be possible to calculate the masses of the proton, neutron and all other baryons directly from our knowledge of QCD dynamics. While this is somewhat beyond what we can do with pen and paper, we can simulate QCD on a computer, and get pretty accurate predictions for the masses of baryons that we've seen above, certainly good to the 5% level. The same is true for the mesons that we will meet in the next section. There is now no doubt that the complexity seen in the hadron spectrum can be entirely explained by the dynamics of QCD.

More Baryons

The lists above do not exhaust the baryons that have been discovered. There are further baryons containing charm and bottom quarks. For example, in addition to the Σ^+, comprised of uus, there is also a Σ_c^+ comprised of uuc and Σ_b^+ comprised of uub, and similar stories for many of the other baryons. It is, however, difficult to argue for any approximate symmetry among these baryons, since the mass of the heavy quarks is much greater than those of the lighter quarks and greater than Λ_{QCD}.

There are also excited states of all these baryons, in which the quarks orbit each other, not dissimilar to the way in which the electrons orbit the proton in the excited states of the hydrogen atom.

There are not, however, baryons containing top quarks. The top quark is so heavy that such baryons are predicted to decay in around 10^{-25} seconds, even faster than the characteristic timescale $T_{QCD} \approx 10^{-23}$ seconds of the strong force. This means that such "top baryons" decay before they even form. Needless to say, none have been observed.

3.4 Mesons

We now turn to mesons, bound states of a quark and an anti-quark. Many hundreds have been discovered. Here we describe some of the most important.

3.4.1 Pions

We will again start by assuming that we've only got up and down quarks to play with. Once again, we can put the quark spins in the same direction, or in opposite directions.

We start by putting the spin of the quark and anti-quark in opposite directions. This results in mesons with spin 0. There are three such mesons that we can build from the up and down quarks, known as *pions*. Their masses and quark content are given by

$$\begin{array}{lll} \pi^+ & (\bar{d}u) & m \approx 139 \text{ MeV} \\ \pi^0 & \frac{1}{\sqrt{2}}(\bar{u}u - \bar{d}d) & m \approx 135 \text{ MeV} \\ \pi^- & (\bar{u}d) & m \approx 139 \text{ MeV} \end{array}$$

The π^- is the anti-particle of π^+ and the two have exactly the same mass. The neutral pion, π^0 is a combination of up and down quarks as shown. It has no electric charge, but a very similar mass to the $\pi^\pm$. The similar masses reflect the isospin symmetry which says that, as far as the strong force is concerned, the up and down quarks have the same properties.

Despite their similar masses, the neutral and charged pions have rather different lifetimes. The neutral pion decays through the electromagnetic force to two photons

$$\pi^0 \rightarrow \gamma + \gamma$$

It has a lifetime of around 10^{-17} seconds. In contrast, the charged pions π^+ and π^- decay through the weak force. We'll see in Section 4 that they typically decay to a muon and a neutrino

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \quad \text{and} \quad \pi^- \rightarrow \mu^- + \bar{\nu}_\mu$$

They live for 10^{-8} seconds, an eternity in the subatomic world and much longer than any of the baryons except the proton and neutron.

There is one, very important characteristic that distinguishes mesons from baryons. Mesons, made of a quark and anti-quark, have integer spin and are therefore bosons. Baryons, made of three quarks, have half integer spin and are therefore fermions.

Back at the beginning of Section 2, we explained that fermions are "matter particles" while bosons are "force particles". It should come as no surprise to learn that baryons, like protons and neutrons, are matter particles. After all, you're made up of them. But it may be less familiar to hear that mesons, like the pion, are force particles. What force do they mediate?

The answer to this is quite lovely: the pions give rise to an attractive force between the baryons. In particular, they give rise to an attractive force between any collection of protons and neutrons. It is this force that binds the protons and neutrons together inside the nucleus.

The existence of a scalar particle, mediating the interaction between protons and neutrons, was predicted by Yukawa in 1935, more than a decade before the pion was discovered. Yukawa observed that a massive scalar particle would give rise to a Coulomb-type force, but with an exponential suppression due to the mass. The potential energy between any two particles takes the form

$$V(r) \sim -\frac{e^{-mr}}{r} \tag{3.4}$$

This is called the *Yukawa potential*. When $m = 0$, it agrees with the more familiar Coulomb potential. For very small distances, $r \ll 1/m$, it is more or less the same as the Coulomb potential. But the force drops off very quickly at distances $r \gg 1/m$. Yukawa had the simple insight that the force that binds the nucleus together should exert itself over distances comparable to the size of the nucleus. From this, he predicted the mass of the pion to be around 200 times the mass of the electron, or about 100 MeV. As we see, he was not far off.

The force that binds the nucleus together is usually simply referred to as the strong nuclear force. But it would be better to give it a different name — say "mesonic force", or "Yukawa force" — to highlight the fact that it is really a residual, secondary effect. At the fundamental level the strong force is mediated by gluons and binds quarks together. But it binds them together in two ways: one to create baryonic matter particles, and another to create mesonic force particles. The upshot is that there are two layers to the strong force: we start with one force and a set of matter particles — gluons interacting with quarks — and end up with a very different force and a new set of matter particles — the mesonic force interacting with protons and neutrons. In this sense, both the particles in the nucleus, and the force that holds them together, are *emergent* phenomena, arising from something more fundamental underneath.

We might, then wonder: do similar transformations await us as we go to yet smaller scales? Could there be some other, very different degrees of freedom on the smallest scales, from which the Standard Model emerges. The answer, of course, is: we don't know.

3.4.2 The Eightfold Way Again

Let's now throw the strange quark into the mix. In addition to the three pions, there are five further spin 0 mesons we can build. (Actually six, but one of these, the η' has slightly different properties so we'll postpone its discussion for now.) These are a collection of particles called *kaons*

$$\begin{array}{ll} K^+ \ (\bar{s}u) & m \approx 494 \text{ MeV} \\ K^0 \ (\bar{s}d) & m \approx 498 \text{ MeV} \\ \bar{K}^0 \ (\bar{d}s) & m \approx 498 \text{ MeV} \\ K^- \ (\bar{u}s) & m \approx 494 \text{ MeV} \end{array}$$

The charged kaons are, like the charged pions, relatively long lived: their lifetime is around 10^{-8} seconds. They too decay via the weak force.

The lifetime of the neutral kaons is a somewhat more complicated story: rather curiously they appear to have two different lifetimes, either 10^{-7} seconds or 10^{-10} seconds, depending on how you count! That's kind of weird. Moreover, rather unexpectedly, it turns out to be an indirect hint of one of the deepest properties of the Standard Model: the fundamental laws of physics are not the same if you run them forwards and backwards in time! We will postpone discussion of this topic to Section 4.3.4.

Finally, there is one meson that is a combination of up, down and strange quarks called, uninspiringly, the eta (η) meson. Its mass and quark content are

$$\eta : \quad \tfrac{1}{\sqrt{6}}(\bar{u}u + \bar{d}d - 2\bar{s}s) \qquad m \approx 548 \text{ MeV}$$

This is similar in spirit to the π^0 meson: for each type of quark there is also an anti-quark sitting within the meson. This allows it to decay quickly, in 10^{-19} seconds, to two photons.

Together with the pions, these 8 mesons sit in a pretty pattern governed by the eight-fold way symmetry relating the three fundamental quarks. We again construct rows of increasing strangeness, where a strange anti-quark $\bar{s}$ counts as negative strangeness:

$$\begin{array}{ccccc} & K^0 \ (\bar{s}d) & & K^+ \ (\bar{s}u) & \\ \\ \pi^- \ (\bar{u}d) & & \pi^0, \eta & & \pi^+ \ (\bar{d}u) \\ \\ & K^- \ (\bar{u}s) & & \bar{K}^0 \ (\bar{d}s) & \end{array}$$

We haven't written the quark content of the π^0 and η only in an attempt to keep table looking vaguely aesthetic.

Viewed purely in terms of masses, the eightfold way looks less convincing for the mesons than for the baryons. An additional strange quark or anti-quark now costs roughly 350 MeV, and the kaons and η are two to three times heavier than the pions. Despite this, it turns out that the masses of all these mesons are well understood theoretically, with the eightfold way an important part of the derivation.

Although I won't recount the full story here, there is part of it that is worth highlighting. Recall that the masses of the proton and neutron are set almost entirely by Λ_{QCD}, rather than the masses of the up and down quark. Indeed, as we mentioned previously, if the up and down quarks were massless then the mass of the proton and neutron would remain pretty much unchanged. At first glance, it looks like the same might be true of the mesons above. After all, the mass of the pions is much closer to $\Lambda_{QCD} \approx 200$ MeV than to $m_{\rm up}$ and $m_{\rm down}$, which are a few MeV. However, it turns out that this guess is completely wrong! If the mass of the up and down quarks vanish, then the mass of the pions would also vanish! Similarly, if the mass of the up, down and strange quarks all vanished, then all 8 mesons above would have vanishing mass. This may seem like a theoretical curiosity since, in the real world, the masses of the quarks are distinctly not zero. Nonetheless, it turns our that this simple observation is enough to govern many of the properties of these mesons. You can read more about this beautiful story in the chapter on *chiral symmetry breaking* in the lectures on Gauge Theory.

There is one final spin 0 meson that is a bit of a loner. It is even given a rubbish name. The eta-prime (η') meson has quark content and mass given by

$$\eta' \quad \frac{1}{\sqrt{3}}(\bar{u}u + \bar{d}d + \bar{s}s) \qquad m \approx 958 \text{ MeV}$$

The eta-prime is significantly heavier than the other 8 mesons, despite having a very similar quark content to the eta meson. There is again a beautiful story behind this associated to the so-called *axial anomaly*, one of the more subtle and deep aspects of quantum field theory. This too is described in the lectures on Gauge Theory.

More Mesons

So far we have only discussed spin 0 mesons, in which the spins of the constituent quark and anti-quark point in opposite directions. We could also arrange for these spins to be aligned. In this case, we end up with mesons of spin 1.

For example, there is a collection of three spin 1 mesons containing only the up and down quarks. These are called rho mesons, and can be viewed as excitations of the three pions. They have masses ~ 770 MeV and decay quickly to pairs of pions.

There are many many further mesons, including excitations of those already mentioned, and mesons that involve the charm and bottom quark. Once again, the top quark decays too quickly and does not form mesons.

Two sets of these mesons deserve a special mention. The first is *charmonium*, a bound state of charm and anti-charm quark. It also goes by the dual name J-psi (J/ψ),

$$J/\psi\ (\bar{c}c) \qquad m \approx 3.1 \text{ GeV}$$

Its lifetime is around 10^{-21} seconds. The discovery of this particle in 1974 was the first glimpse of the charm quark and will be described in Section C.3.

There are a collection of lighter mesons that contain just a single charm quark. These are called (somewhat peculiarly) *D-mesons*. The lightest are:

$$\begin{aligned} D^0\ (c\bar{u}) \qquad & m \approx 1865 \text{ MeV} \\ D^+\ (c\bar{d}) \qquad & m \approx 1869 \text{ MeV} \end{aligned}$$

These are remarkably long lived particles, with the D^+ living 10^{-12} seconds, and the D^0 about half this time. The long lifetime is because these particles decay only through a somewhat subtle property of the weak force. We will learn more about this in Section 4.3.

Similarly, the bottom quark was first discovered in *bottomonium*, also known as the upsilon (Υ)

$$\Upsilon\ (\bar{b}b) \qquad m \approx 9.5 \text{ GeV}$$

This has a lifetime of 10^{-20} seconds. Once again, it is neither the lightest nor the longest lived meson containing a b-quark. The lightest *B-mesons* are

$$B^+\ (u\bar{b}) \text{ and } B^0\ (d\bar{b}) \qquad m \approx 5280 \text{ MeV}$$

Despite being significantly heavier, they actually live (very) slightly longer than the D-mesons, with a lifetime of around 1.5×10^{-12} seconds. Again, this is down to intricacies of the weak force.

C Interlude: The Rise of the Machine

From the 1920s onwards, it became clear that the alpha particles emitted in radioactivity, with their 5 MeV of energy, would not suffice to understand the nucleus. As we saw in Interlude B, in the short term most discoveries would come courtesy of cosmic rays. But, long term, new accelerator technology was needed.

The cathode ray tube is, in many ways, the first particle accelerator, albeit one built before the constituents were even known to be particles. The key idea is a simple one: if you drop a large voltage over some distance then charged particles will pick up speed. You can then use the resulting beam to smash into other things, as Röntgen did in his discovery of X-rays.

A variant of this idea was taken forward by John Cockcroft and Ernest Walton. Working in Cambridge in 1932, under the ever watchful eye of Rutherford, they built a voltage multiplier capable of accelerating protons to 700 keV. They then used this to great effect, inducing the first artificial transmutation of a nucleus,

$$p + {}^7\mathrm{Li} \;\longrightarrow\; 2\alpha$$

In more popular terminology, they succeeded in splitting the atom. These days, the "high-tension laboratory", where Cockcroft and Walton performed their later experiments, has been converted into the "Cockcroft lecture theatre", where we teach Newtonian mechanics and special relativity to first year undergraduates.

Both the cathode ray tube and the Cockcroft-Walton accelerator are linear accelerators: the charged particles move in a straight line. The next great breakthrough – due to Ernest Lawrence – was a simple one: make the particles bend.

C.1 The Cyclotron

Lawrence was inspired by a simple fact in classical mechanics. Take a charged particle of mass m and charge q, restricted to move in the (x, y)-plane, with a magnetic field B in the z-direction. If you give the particle an initial kick of speed v, then it moves in a circle and comes back to its starting position in time

$$T = \frac{2\pi m}{qB} \tag{C.1}$$

The lovely fact is that this time doesn't depend on the speed v! If you set the particle off with a bigger velocity, then it will travel in a bigger circle, but always come back in the same time.

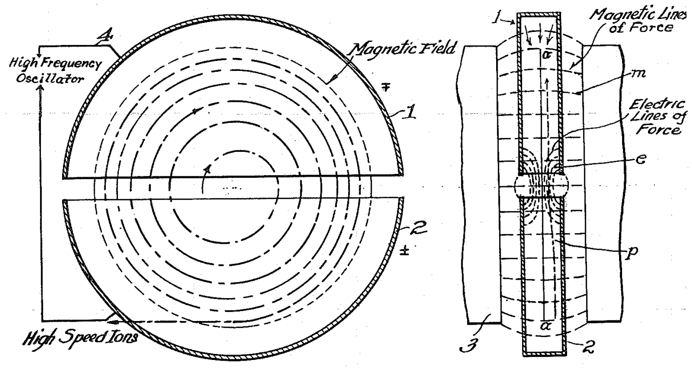


Figure 22. The principle behind the cyclotron, from Ernest Lawrence's 1934 patent application.

This is the key principle behind the *cyclotron*. The charged particles move perpendicular to a fixed magnetic field. They are trapped in two boxes, called *dees*, so named because they are shaped like the letter D. The two dees are placed back to back, with a small gap between them like so: ᗡD. A voltage is placed across the gap, ensuring that the particles are accelerated every time they cross from one dee to the other. The rub, of course, is that first they cross the gap in one direction, then in the other. This means that if you want them to be accelerated each time, rather than decelerated, then you have to flip the polarity of the voltage in the time it takes them to travel a semi-circle. The good news, as we've seen in (C.1), is that this time doesn't depend on how fast the particles are going. This means that you can tune the AC voltage to a frequency, resonant with the particles and the particles will speed up, travelling in ever wider circles until they spit out the end where they can be used for whatever purpose is needed.

A prototype cyclotron, known as the 9-inch, was first constructed in Berkeley in 1930. By 1932, Lawrence and his student, Stanley Livingston, had succeeded in building an 11-inch machine that could reach energies of 1 MeV and were able to quickly reproduce the Cockcroft-Walton results splitting the atom. For the rest of the 1930s, Lawrence exhibited a single-minded focus on reaching higher and higher energies. By 1936 he reached 8 MeV, by 1939 20 MeV. Indeed, such was his desire to reach higher energies, it seemed to barely occur to Lawrence that he should, perhaps, occasionally pause to do some science with his machines.

Furthermore, there is a limit to how far the key equation (C.1) can be pushed. As protons reach energies of around 25 MeV, the effects of special relativity have to be taken into account and their mass starts to increase with velocity. This knocks the timing, since the faster protons now take slightly longer to make their orbit. To compensate, one has to tune the magnetic field to keep the timing in sync. This kind of machine, still with particles spiralling outwards but with a varying magnetic field, goes by the catchy name of a *synchrocylclotron*, or SC for short.

Soon after the war, flush with money from the military, Lawrence completed his first SC, a 184-inch monster machine with a magnet that weighed 10,000 tons and accelerated alpha particles to 380 MeV, later increased to 720 MeV. The question was: what to do with it?

The Neutral Pion

The neutral pion π^0 holds two claims to fame. It was the first particle to be predicted on grounds of symmetry. And it was the first particle to be discovered in a collider on Earth, rather than in a cosmic ray shower.

First the experimental situation. Recall from Interlude B.3 that the charged pion was discovered in cosmic ray showers by Powell and his team in 1947. In fact, rather embarrassingly, by that point Lawrence's SC had been producing pions for over a year. But, such was the focus on reaching higher energies, no one had put any thought into building detectors. It was only when Cécil Lattes, one of Powell's collaborators, moved to Berkeley in 1948 that they placed photographic emulsion plates in the accelerator and found pions in great numbers.

The neutral pion is a different matter altogether. Since it carries no electric charge, it leaves no tracks and neither cloud chambers nor photographic emulsions can be used to find it. Instead, it can only be seen by more indirect means through its decay to two photons

$$\pi^0 \longrightarrow \gamma + \gamma$$

The smoking gun is the detection of two simultaneous photons, whose energy, momentum and angular momentum correlations can be traced to the decay of a single particle. The neutral pion was discovered in this manner in 1950 by Steinberger, Panofski and Steller, using the Berkeley SC.

Next, the theory. In Section 3 we explained the *eightfold way*, the idea that there is an approximate symmetry between the up, down and strange quarks which leads to patterns in the masses of baryons and mesons. We also briefly mentioned a precursor to this idea, first pointed out in the early 1930's by Heisenberg. This is the idea that, at least as far as the strong force is concerned, the neutron and proton behave in almost identical fashion. They have similar mass. Moreover, the binding force between nn, pp and np is more or less the same. Even with the very limited experimental data of the early 1930s, Heisenberg intuited that this was important. At the time the symmetry between the proton and neutron was called *isotopic spin*[11], these days shortened to just *isospin*.

However, the idea of isospin runs into trouble when you appreciate the obvious: the proton and neutron have different electric charges. If the strong force is mediated by charged pions $\pi^{\pm}$ alone, as was thought by the late 1930s (recall, people still thought that the muon was responsible at this point!) then charge conservation meant that the interactions experienced by the proton and neutron would necessarily be different. The way out was suggested by the theorist Nicholas Kemmer in 1938: you get to keep isospin symmetry, but only if there exists a third, neutral pion π^0.

Resonances

Neutral particles are not the only ones that are hard to see. Any hadron or meson that decays through the strong force, rather than the weak force, will have a lifetime of around 10^{-24} seconds. That's not a huge amount of time, even by the standards of particle physics. Even allowing for relativistic time dilation, these particles are not going to travel far enough to snap a photograph of them. Instead, we need more indirect methods to detect them.

The method of choice is to observe the effect of these new particles on the old. To explain this I first need to introduce a new concept: that of *cross-section*.

When two elementary particles come close to each other, there is some probability that they will interact. This interaction may result in them scattering off each other, or transforming – even if only briefly – into some other particle. Roughly speaking, the *cross-section* is the probability that they interact in some way, rather than pass through each other.

We can also speak less roughly. The cross-section, as the name suggests, is actually an area. As such it can't quite be a probability (which must be dimensionless) but it's closely related. The idea is that the cross-section is the area – or size – that the particles present to each other as they approach. As an analogy, if you're throwing balls in an attempt to hit some target, it's the cross-sectional area of the target that

[11] This is, of course, a daft name. A much better name would be *isobaric spin*, since isobars are elements with the same atomic mass but different combinations of protons and neutrons.

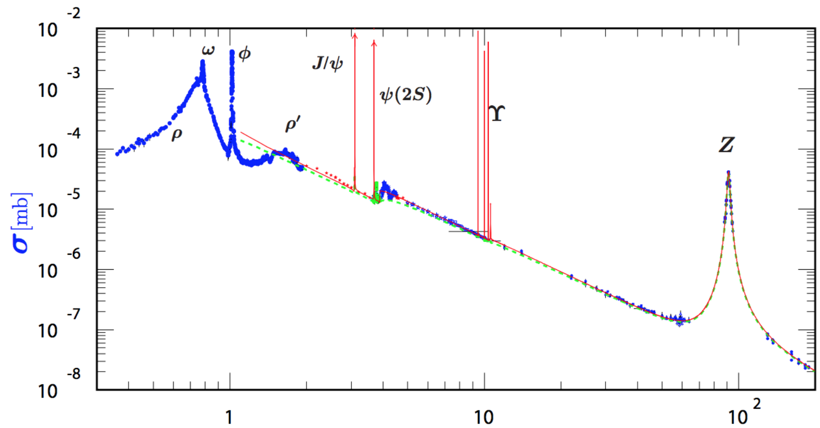


Figure 23. A collection of data from many experiments studying $e^+e^- \longrightarrow$ hadrons. The horizontal axis is the centre of mass energy, measured in GeV. A number of different mesons can be clearly seen at low energy, with the Z-boson at high energy. This plot was taken from the particle data group plots of cross-sections and related quantities.

will determine your success. A barn door has a bigger cross-section than a beer can. This, in turn, means that games where you try to hit a barn door are somewhat less entertaining than those where you try to hit a beer can.

Importantly, the cross-section for some interaction depends on the energy of the colliding particles. Typically, in particle physics one finds that the cross-section drops as the energy of the incoming particles increases. But, every now and then, one sees a pronounced bump in the cross-section. This bump is called a *resonance* and is the telltale sign that there's a new, third particle, appearing in the story.

The bumps appear when the energy of the incoming particles is tuned to the mass of some new particle. This allows for a new interaction, in which the two incoming particles briefly morph into the new one which will then, typically, subsequently decay. These decay products may be the original particles, in which case it will just look like they've scattered off each other, or they may be something new entirely.

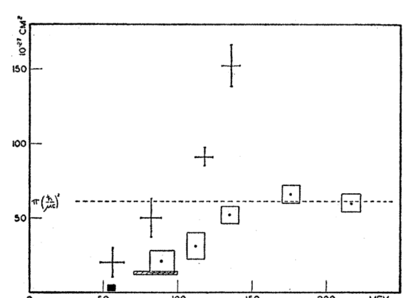
The early data on resonances is not particularly clean. (We'll give an example below.) However, the idea is clearly illustrated in Figure 23 which collects together the results of many decades worth of experiments of e^+e^- collisions, where the end products are hadrons. The horizontal axis depicts the (log of the) energy, measured in GeV. The vertical axis depicts the (log of the) cross-section. You can see the overall downwards trend of the cross-section as the energy increases, but most striking are the various peaks. At lower energies these peaks correspond to meson states that we met in the last section. Way up, at close to 100 GeV, we see the Z-boson that mediates the weak force. We'll discuss this more in the following section.

The shape of the resonance contains a lot of information about the underlying particle. The energy at which the resonance occurs tells us the mass of the particle. Meanwhile, the width of the resonance tells us its lifetime: the bigger the width, the quicker the particle decays. Returning to Figure 23, you can see that the width is barely discernible on the red spikes in the middle. This is because, as we saw in the last section, these are the lightest states containing charm and bottom quarks respectively. The presence of these new quarks limits their decay options, resulting in a much longer lifetime than might naively be expected.

Many of the particles that we'll meet as these lectures progress were detected through their resonance effect on scattering. You can read more about the basics of resonances in quantum mechanical scattering in the lectures on Topics in Quantum Mechanics.

Delta Baryons

The first novel particle detected as a resonance was the collection of four Δ-baryons. Viewed through today's lens of the quark model, they are Δ^{++} (uuu), Δ^{+} (uud), Δ^{0} (udd) and Δ^{-} (ddd). Each has mass 1232 MeV and a lifetime of around 5×10^{-24} seconds.



The year was 1952, the place Chicago. Enrico Fermi and his team had built a synchrocyclotron, based on the principles introduced by Lawrence. They extracted a beam of pions from the machine and directed it at hydrogen gas to watch how the pions scattered off protons. They didn't see anything as distinctive as a bump. They did, however, see a rise in the cross-section. Most strikingly, there was a clear difference in the cross-sections for π^+ (shown as crosses in the data to the right) and π^- (shown as rectangles)[12] . This needed an explanation.

[12]This figure is taken from the paper Total Cross Sections of Positive Pions in Hydrogen" by Anderson, Fermi, Long, and Nagle.

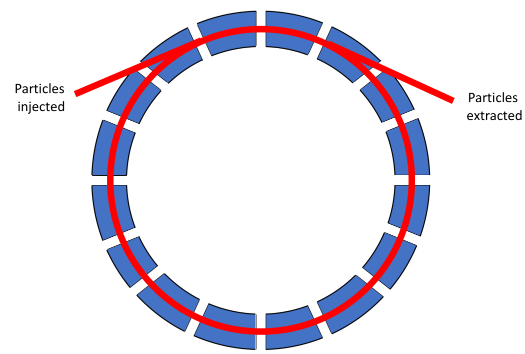
In fact, this effect had been anticipated earlier by a theorist called Keith Brueckner. He looked at the three scattering processes

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

By invoking a set of intermediate states, which we now call the Δ-baryons, he argued that the processes could proceed through the creation of Δ^{++} or Δ^0 and should occur with relative probabilities in the ratio 9:2:1. This was what the Chicago team observed. Further analysis of the angular distribution of the scattering showed that the intermediate states must have spin 3/2. Historically, the discovery of the Δ^{++}, with its striking +2 charge, played an important role in elucidating the underlying structure of quarks.

C.2 The Synchrotron

Cylcotrons, with the particles spiralling outwards, can reach energies no higher than 1 GeV or so. The next phase in accelerator development was the *synchrotron*. Here, particles travel in a circle of fixed radius. Their path is again bent by magnets, and accelerated in the gaps between the magnetics by electric fields. To keep the particle travelling in a circle, the magnetic field must be synchronised with the particle's velocity, so that the magnetic field becomes stronger as the particles pick up speed. A cartoon picture of the synchrotron is shown on the right.



The development of the synchrotron brought many new technical challenges, including the difficult issue of stabilising the beam. Key to the whole endeavour is the concept of *phase stability*, in which an alternating voltage is placed over the gaps acts to ensure that all particles converge on the same speed, with the faster ones slowing slightly and the tardy ones picking up speed. This results in particles sitting in bunches, rather than the continuous beam of the earlier cyclotron.

Another difference from the cyclotron is that you don't get to start the particles from rest. They must be injected into the synchrotron at some other accelerator. This, then, is the fate of accelerators upon retirement: they become injection machines for the next generation.

The early proton synchrotrons had the cool names, before the push for dull acronyms became too strong to ignore. In Brookhaven, the Cosmotron reached energies of 3 GeV; in Berkeley, the Bevatron 6 GeV. These began a long succession of machines, culminating more than half a century later, with the Tevatron at Fermilab reaching 1 TeV and the LHC at CERN reaching 13 TeV.

More Anti-Matter

The 6 GeV reached by the bevatron was not chosen arbitrarily. It was built with a particular science case in mind: the creation of anti-protons. The 6 GeV threshold allows their production through

$$p + p \longrightarrow p + p + p + \bar{p}$$

The first proton on the left-hand side was in the beam; the second proton on the left-hand side was sitting in a fixed copper target. The resulting debris mostly consists of pions, with the occasional anti-proton lying within. The challenge was to find them.

You might have thought that the best way to detect anti-protons would be to watch them annihilate with protons. In fact, it turned out to be significantly simpler to identify them through their mass and charge. To this end, the experimenters first set up a series of magnets, designed to deflect unwanted positively charged particles and focus only negatively charged particles with very specific momentum into the detectors. The first set of detectors consisted of *scintillation counters*, an instrument for measuring the photons emitted by the ionised tracks left by a charged particle. Two scintillation counters were placed 12 m apart. Both anti-protons and pions in the beam had the same momentum, which meant that the heavier protons were slower and so took longer to travel the 12 m between the detectors. A whole 10^{-8} seconds longer. That was the first clue.

Next, the beam entered a pair of *Cerenkov detectors*. These detect an effect known as *Cerenkov radiation* which is emitted when particles travel through a material faster than light can travel through that material. The first Cerenkov detector fired whenever the particle was travelling too fast to be an anti-proton. The second Cerenkov detector had a special design which meant that it fired *only* for a small window of velocities, tuned to that of the anti-proton, and so failed to fire for the faster mesons.

With this elaborate set-up, over the course of two weeks in October 1955, Emilio Segrè and his team of Owen Chamberlain, Clyde Wiegand and Tom Ypsilantis found 60 anti-protons, nestled among 3.5 million pions. They announced the discovery at a press conference on October 19th 1955.

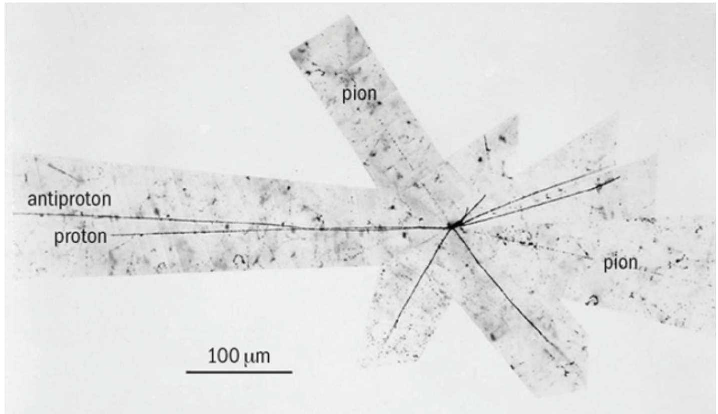


Figure 24. A photographic emulsion showing $p\bar{p}$ annihilation. The original (in vertical orientation) can be found in the 1956 Nuovo Cimento paper "On the Observation of an Antiproton Star in Emulsion Exposed at the Bevatron" by Chamberlain et. al.

This story also has some human twists. Indeed, as machines got bigger, so too did the opportunity for machinations. The clever magnetic lens design used in the experiment was due to an Italian physicist, Oreste Piccioni who, at the time, was a visitor in Berkeley. Although the group lead by Segrè adopted this design, they were reluctant to allow Piccioni to join their team when he returned to Berkeley full time in the summer of 1955. Instead, Piccioni joined a rival team, also searching for the anti-proton, lead by Edward Lofgren.

The two teams took it in turns to use the bevatron, two weeks on, two weeks off. The machine broke in the middle of Segrè's second run and they could take no data. When the machine was finally repaired, nice-guy Lofgren yielded his time to allow Segrè to complete his run. Two weeks later they announced the discovery of the anti-proton. Four years after that, Segrè and Chamberlain collected their Nobel prize. Nice guys, it turns out, rarely fared well in the increasingly ruthless world of experimental particle physics.

However, there was some joy for both Lofgren and Piccioni. One year later, they were among the team who discovered the anti-neutron through the annihilation processes

$$p + \bar{p} \longrightarrow n + \bar{n}$$

It appears, however, that this brought little comfort to Piccioni. Some years later, he sued Segré and Chamberlain for \$125,000 of their Nobel prize money. The case made it to the US supreme court, before being dismissed.

Figure 25. This bubble chamber photograph, taken at the Lawrence Berkeley Laboratory, shows a high energy γ ray colliding with an electron to produce an e^+e^- pair, which spiral in opposite directions. The original electron receives a huge kick and flies off to the right. At some point it emits a gamma ray which then turns into a second e^+e^- pair, this time with higher energy, visible as the V signature to the right of the picture.

Mesons and Baryons in Bubbles

The next leap forward was in detector technology. In the 1950s, Donald Glaser proposed the idea of a *bubble chamber*, a successor to the cloud chamber. The detector is filled with a liquid – ideally liquid hydrogen – that is kept under pressure at constant temperature, slightly below its boiling point. Just before the particles pass through, the pressure is reduced by a small amount, lowering the boiling point so that the liquid is *super-heated*, meaning that it remains in the liquid phase even though the temperature is above boiling. As a charged particle passes through, it gives the liquid the nudge it needs to start boiling, leaving behind a trail of bubbles. Examples of the particles' graceful arcs, as they spiral in an applied magnetic field, can be seen in Figures 25 and 26.

In addition, this was a time of increased automation. The number of events that could be recorded in a bubble chamber was far greater than in previous detectors. Teams of highly skilled, poorly paid women, scouring photographic emulsions for interesting forks and kinks just wasn't going to cut it anymore. Instead, both analysis and data storage required the use of computers. Experiments could be done in one institution, and analysed elsewhere.

Results came quickly. First, a collection of vector (i.e. spin 1) mesons were discovered, starting with the ρ and ω. The η meson was discovered at Johns Hopkins university, using data borrowed from Berkeley. (η translates to the letter H, which is short for "Hopkins" apparently.)

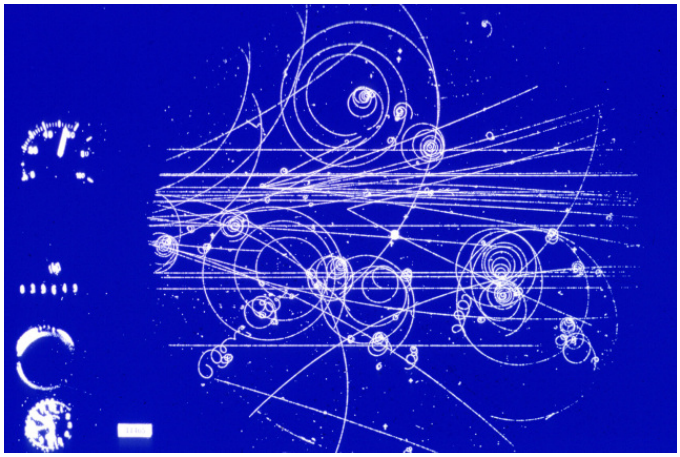


Figure 26. This photo was taken in 1960 at CERN. A stream of π^- mesons enter from the left. One of these pions hits a proton in the liquid hydrogen spraying new particles. Among these is a neutral Λ baryon, which doesn't leave a track but reveals its existence a little to right as it decays into a proton and pion, producing the characteristic V shape seen in the middle of the picture. The results of this decay have high energy and travel in straight line. Other, lower energy charged particles spiral in the magnetic field.

Baryons were also seen in quantity, both long-lived strange baryons and, indirectly, shorter lived resonances. An important breakthrough happened in 1964, when the Ω^- baryon was discovered. The Ω^- contains three strange quarks and its mass, lifetime and decay modes had been predicted earlier using the quark model.

The Ω^- discovery is shown in Figure 27. An incoming kaon collides with a proton in the liquid, yielding

$$K^-\,(\bar{u}s) + p(uud) \;\longrightarrow \Omega^-\,(sss) + K^+\,(u\bar{s}) + K^0(d\bar{s})$$

There is then a succession of further baryon decays, with

$$\begin{aligned}
\Omega^-\,(sss) &\longrightarrow \Xi^0\,(uss) + \pi^-\,(\bar{u}d) \\
\Xi^0\,(uss) &\longrightarrow \Lambda^0\,(uds) + 2\gamma \\
\Lambda^0\,(uds) &\longrightarrow p\,(uud) + \pi^-\,(\bar{u}d)
\end{aligned}$$

Each of these decays changes strangeness by -1, and so happens only through the weak force. Correspondingly, the lifetime is around 10^{-10} seconds for each, long enough for them to leave a trail a bubbles.

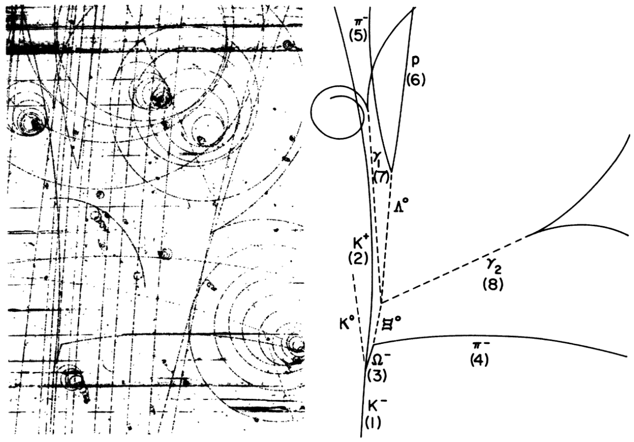


Figure 27. The original bubble chamber picture, and an accompanying line tracing, showing the discovery of the Ω^- baryon. This is taken from the 1964 paper "Observation of a hyperon with strangeness minus three" by Barnes et. al. Here "et al" refers to an additional 30 authors, heralding the large collaborations that were to come.

C.3 Quarks

The cornucopia of hadrons discovered throughout the 50s and 60s gave enough clues to lead to the quark model. The observed masses and lifetimes could be roughly accounted for by postulating three different types of constituent spin $\frac{1}{2}$ particles – up, down and strange – each of which carries three different internal degrees of freedom called colour. But that still left open the question of whether these quarks were real entities, or just mathematical accounting tricks.

Most physicists at the time assumed that quarks were merely a useful fiction. Their concern was the obvious one: if quarks are real, then why don't we observe them in isolation, where their fractional electric charge would stand out like a sore thumb. Gell-Mann ends the 1964 paper in which he first proposed quarks with the sentence

> "A search for stable quarks of charge $-\frac{1}{3}$ or $+\frac{2}{3}$ and/or stable di-quarks of charge $-\frac{2}{3}$ or $+\frac{1}{3}$ at the highest energy accelerators would help to reassure us of the non-existence of real quarks.

Zweig, who independently had the same idea in the same month – January 1964 – referred to quarks as "aces" and was marginally more optimistic. The final sentence of his paper reads

> "There is also the outside chance that the model is a closer approximation to nature than we may think, and that fractionally charged aces abound within us."

The situation was resolved by a series of experiments that, although initially designed to study resonances of the proton, turned out to be perfectly placed to instead explore its inner structure. These experiments, which ran from 1967 to 1973, took place at the Stanford Linear Accelerator, better known as SLAC, a 3 km long machine that accelerated electrons to 20 GeV, effectively by getting them to surf the crest of an electromagnetic wave.

The SLAC experiments didn't see anything as striking as fractional electric charge. Indeed, their data was, at first, murky and complicated. However, over the course of several years it became increasingly apparent that their results could only be explained if the electrons were scattering off some point-like constituent inside the proton and neutron. As we now explain, these experiments involve a process known as ...

Deep Inelastic Scattering

By the mid 1960s, it was apparent that the proton is not a point-like object but has a size of around 10^{-15} m. This was seen, for example, in the *elastic* scattering of electrons off protons. Here, the word "elastic" means that the electron has the same energy after the collision as before. For example, the Geiger-Marsden experiment that uncovered the structure of the atom involved the elastic scattering of alpha particles off the nucleus.

In contrast, in *inelastic* scattering, the electron collides with more destructive force, knocking the proton into a higher excited state, such as a spin $\frac{3}{2}$ baryon, or breaking it apart completely. This process can roughly be characterised as

$$e + p \longrightarrow e + \text{other stuff}$$

where we don't too much care about the other stuff. We care only about what happens to the electron. If the electron comes in at very high energies then it buries deep into the proton where it can probe whatever lies inside.

If, as was originally thought, the proton was structureless, then the cross-section for electron-proton scattering should rapidly decrease with energy. And this is indeed seen in elastic scattering. However, when the electron energy gets high enough to enter the inelastic regime, something different occurs. The first clue that something was afoot was simply that more electrons were scattered at low angles than expected. However, when the data was plotted in a particular way, something more surprising stood out.

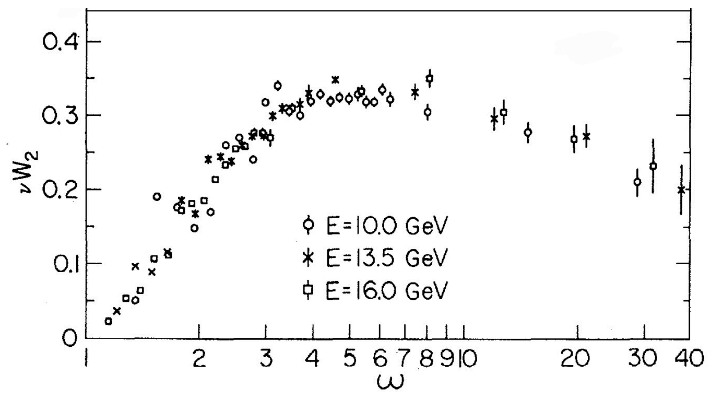


Figure 28. It's neither pretty, nor obvious, but this data provided the first hint of the existence of quarks. The proton structure function is plotted on the vertical axis, while $\omega = 1/x$ is on the horizontal axis. This graph is taken from the 1969 paper "Observed Behavior of Highly Inelastic Electron-Proton Scattering".

To explain this, I need to first describe a little bit of scattering theory. The cross-section for electron-proton scattering depends on an object called the *proton structure function*. And this, in turn, depends on the energy E of the incoming electron, the energy E' of the outgoing electron and the angle θ by which the electron is deflected. In general, any cross-section depends on these variables in two combinations

$$\nu = E - E' \quad \text{and} \quad q^2 = 4EE'(1 - \cos\theta)$$

The surprise was that, at high energies, the proton structure function doesn't depend on both ν and q^2: it depends only on the dimensionless ratio

$$x = \frac{q^2}{2M\nu}$$

where M is the mass of the proton. The initial data, shown in Figure 28, plots the proton structure function for scattering by a small angle ($\theta = 6°$ in this case.) The fact that the data for three different energies all lie on the same curve is showing that the structure function depends only on the combination x. This is known as *Bjorken scaling*, after the theorist Bjorken who first suggested this behaviour (and first suggested that the experimenters plot their data in this unlikely manner.)

The question, of course, is what does scaling mean? The intuitive idea, largely due to Feynman, is that scaling is telling us directly that the electron is scattering off a point-like object inside the proton. Very roughly speaking, this is because the proton structure function doesn't depend on any scale, because x is a dimensionless variable. But the only object with no scale is a point. (Feynman's original argument is characteristically creative and involves going to a frame of reference in which the point-like object in the proton has infinite momentum.)

Feynman referred to the point-like objects inside the proton as *partons*, noticeably avoiding the term "quark" in his original paper. I do not know if this was for the purposes of scientific agnosticism or scientific antagonism. (Gell-Man and Feynman were colleagues, collaborators and, perhaps above all, rivals!) Either way, the omission was appropriate. Simple models in which the proton and neutron are each composed of three quarks do not provide a good fit to the data. This is because, as we explained in the previous section, the proton and neutron are themselves enormously complicated objects. The cartoon picture in which they each contain three quarks is a long way from reality and ignores the morass of gluons and quark-anti-quark pairs that also sit inside. Deep inelastic scattering provides the tool to probe this complexity. Good fits to the data could only be achieved by including partons that are gluons and quark-anti-quark pairs, in addition to the three valence quarks.

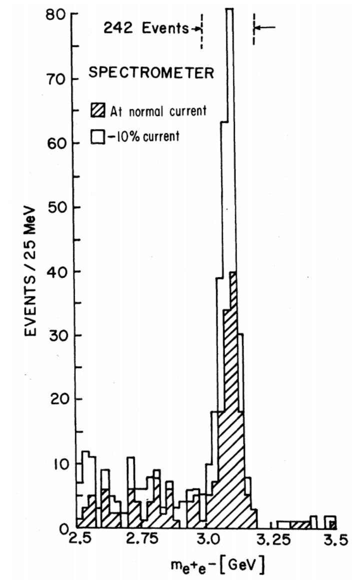
Deep inelastic scattering also provides a method to indirectly test the electric charges carried by the partons. It turns out that the cross-section is proportional to the sum of the squares of the charges of the partons. For a proton, with uud quarks, this sum gives $\left(\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 = 1$, while for a neutron with udd, it is $\left(\frac{2}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 = \frac{2}{3}$. The simplest quark model then predicts that the cross-section for neutrons should be $\frac{2}{3}$ that of protons. After taking into the account the many subtleties described above, this is confirmed by experiment.

Although the parton model provided a good explanation for the experimental results, there was one mystery that remained. The partons that lead to exact scaling behaviour are free particles. Subsequent deviations from scaling suggested that there were some interactions between the partons, but these were necessarily small. Yet how could partons be confined within the proton if the interactions were so small? This issue was resolved by the discovery of asymptotic freedom in Yang-Mills like theories by Gross, Wilczek and, independently, Politzer in 1973.

Charmed

In November 1974, a new quark was found. This was the charm. It came as a surprise to many physicists and went a long way towards cementing belief in the quark model.

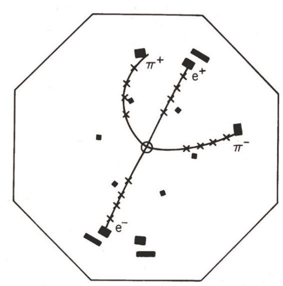
The discovery was made simultaneously on the West coast in SLAC and on the East coast in Brookhaven. Since both teams got to name the particle, it now goes by the double-barrelled J/ψ.



The J/ψ weighs in at 3.1 GeV, with a lifetime of almost 10^{-20} seconds. This is much longer than expected for a particle that is so heavy, a fact that reveals itself in the very narrow resonance shown on the right.

It didn't take long to understand that this resonance is so narrow because it contains a new quark-anti-quark pair, bound as $\bar{c}c$ meson. It is now also known as *charmonium*. In fact, in many ways the charmed mesons are easier to understand than their lighter cousins. Because the charm quark is so heavy (about 1.3 GeV), it has a very small Compton wavelength. This means that two charm quarks in a meson sit so closely together that they are in the asymptotically free regime of QCD, where the coupling strength is fairly small. This makes it easier to understand their properties.

Within a few weeks of the original J/ψ discovery, a collection of further resonances had been found, all agreeing well with theoretical expectations. These new resonances were found only at SLAC so their name took preference, and they are called ψ'. In a cute twist, the reconstructed event decay of a ψ' takes the eponymous form shown on the right.



While most physicists had not expected the charm quark, there were some who had previously argued for its existence on theoretical grounds. The most compelling was due to Glashow, Iliopoulos and Maiani (or GIM) who, in 1970, analysed a subtle property of the weak force known as *flavour changing neutral currents*. (We'll describe some basic properties of the weak force in Section 4.) An example of this phenomenon is the decay $K^0 \longrightarrow \mu^+ + \mu^-$. This happens very rarely, but it was difficult to understand why: theoretical expectations suggest that a process involving the exchange of an up quark and W-boson should give a much larger decay rate. The GIM mechanism proposed the existence of a charm quark, with charge $+2/3$, and showed that this gives a contribution to the decay almost exactly cancels the contribution from the up quark.

After the discovery of the charm quark, many things fell into place. The key ideas underlying the Standard Model, from asymptotic freedom of the strong force, to symmetry breaking of the weak force (described in Section 4) had all been developed previously. But the discovery of the charm quark prompted a synthesis of these ideas, with all phenomena described by four quarks, two charged leptons, and two neutrinos, coupled through forces associated to $SU(3) \times SU(2) \times U(1)$. The inappropriately modest name *Standard Model* dates from this time. For this reason, physicists who lived through the discovery of charm sometimes refer to this period as the *November Revolution*.

Colliders

The discovery of charm also involved an important experimental leap: colliders.

All the accelerators that we've discussed so far take a beam of particles and smash it into a fixed target. This has the obvious advantage that it's easy to hit a fixed target. But it also has a disadvantage because conservation of momentum means that any particles you create fly off at close to the speed of light. It's not so much the fact they're moving that's the problem, but that much of the beam energy is wasted because it goes into the kinetic energy of the final product. If you succeed in accelerating particles of mass m to an energy E which then hits a fixed target, then the energy available to create new particles is $\sqrt{2mE}$. Clearly it would be better if we could make use of more of that precious beam energy.

These problems go away in a *collider*. This consists of two beams with equal and opposite momentum, which are then brought together at one or more intersection points where the collision takes place.

Needless to say, the technological hurdles in getting this to work are formidable. Not least is the problem of *luminosity*, meaning the number of collision events taking place. The density of particles in each beam is significantly less than that of a fixed target and, correspondingly, the number of collisions is greatly reduced. To compensate for this, each beam should consist of many pulses of particles, with collisions happening repeatedly and often. This, in turn, requires that the beams have long lifetimes. Typical numbers involve beam lifetimes of several hours, allowing them to circulate 10^{10} times or so. These accelerators also go by the name of *storage rings*.

The discovery of charm at Brookhaven used a standard fixed target experiment. A beam of electrons ploughed into a beryllium target, and they found their J particle as the huge peak in the centre of mass energy of e^+e^- pairs.

In contrast, the SLAC team used an e^+e^- collider, called SPEAR, short for the Stanford Positron Electron Asymmetric Rings. These storage rings took electrons and positrons from the linear collider and brought them into head-on collisions. They discovered their ψ particle as a resonance in the cross-section for $e^+ + e^- \longrightarrow$ hadrons.

As an aside, the acronym "asymmetric" in the acronym SPEAR dates from an earlier proposal in which e^+ and e^- were accelerated in different rings. After a budget cut, the design changed to a single ring, but the acronym stayed.

Round Three

Back in 1932, there were 100 days where physicists could revel in a simplistic world containing only electrons, protons and neutrons. Then the discovery of positrons burst their bubble.

In the 1970s, physicists had a little over a year in which they could believe in a nice symmetrical world with two generations of fermions, each containing two quarks, an electron-type particle and a neutrino. Late in 1975, a group working at the SPEAR experiment found something odd and entirely unexpected. Their original paper doesn't beat around the bush, opening with the blunt statement:

> "We have found 64 events of the form
>
> $$e^+ + e^- \longrightarrow e^\pm + \mu^\mp + \geq 2 \text{ undetected particles}$$
>
> for which we have no conventional explanation."

The paper suggests that these events could be due to a new charged, heavy lepton or to a new charged heavy boson. It took some years to realise that the former is the case: the original e^+e^- pair collide to form what we now call the *tau leptons*,

$$e^+ + e^- \longrightarrow \tau^+ + \tau^-$$

The taus subsequently decay as, for example,

$$\tau^- \longrightarrow \nu_\tau + e^- + \bar{\nu}_e \quad \text{and} \quad \tau^+ \longrightarrow \bar{\nu}_\tau + \mu^+ + \nu_\mu$$

giving rise to the observed signature.

The discovery of the τ lepton upset the balance. It would be another 20 years until it was restored, with the bottom quark, top quark and tau neutrino filling out the set. The bottom quark was discovered not long after. In 1977, a fixed target experiment at Fermilab found a strong, narrow resonance at 9.5 GeV, the upsilon Υ with quark content $\bar{b}b$.

Finding the top quark was another matter. A rough guesstimate for its mass can be made by a quick glimpse at the first five quarks:

$$\begin{array}{lll} m_d = 4.7 \text{ MeV} & m_s = 96 \text{ MeV} & m_b = 4.2 \text{ GeV} \\ m_u = 2.2 \text{ MeV} & m_c = 1.3 \text{ GeV} & m_t = \text{?} \end{array}$$

Given this pattern, what would you guess for m_t? Perhaps 40 GeV? In the 1970's, a couple of e^+e^- colliders failed to find the top quark at 30 GeV. In the 1980's, a p^+p^- collider failed to find it at 80 GeV. By this time it was clear that the top quark was so heavy that it would not form a detectable $\bar{t}t$ meson state like J/ψ or Υ. This is because (see Section 4) it could decay directly to a W-boson through

$$t \longrightarrow W^+ + b \quad \text{and} \quad \bar{t} \longrightarrow W^- + \bar{b}$$

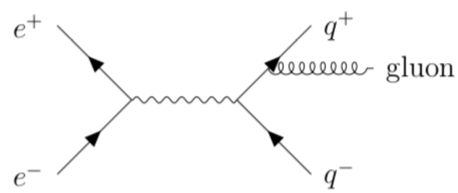
This decay happens in a shorter time scale than that associated to the strong force. This means that, if we want to find the top quark, we need to find clear evidence for the decay products that it leaves behind.

Because the top quarks are created as $\bar{t}t$ pairs, they will decay to $\bar{b}b$ pairs, together with a W^+W^- pair. The W-bosons themselves then decay, sometimes to quarks $W^\pm \to \bar{q}q$ and sometimes to leptons $W^\pm \to l^\pm\nu$ where $l = e, \mu$ or τ. The question is: how to see these decay products?

We know what charged leptons look like, and we know that neutrinos are just going to escape unnoticed. So the real question is: how do we see the quarks? Because these are confined, we can't see quarks directly. However, when quark-anti-quark pairs are formed from collisions, something rather dramatic happens. Each one of the pair flies off in a different direction but, because they hate to be alone, they pull further quark-anti-quark pairs from the vacuum as they go. The end result is that each quark morphs into a collection of hadrons, all moving in roughly the same direction. This is called a *jet* and the phenomenon of quarks turning into a multitude of mesons and baryons is called *hadronisation*.

Jets were first seen in SPEAR e^+e^- collisions in 1975 and proved yet more evidence for the reality of quarks. With a lot of work, it is sometimes possible to go backwards and, from the jet, reconstruct the kind of quark that started it.

As an aside to our main story, in the late 1970s, the first indirect evidence for the existence of the gluon came through a process associated to the following Feynman diagram:



One of the emitted quarks radiates a gluon, but the gluon can no more live on its own than the quarks. The result is 3 jets; two from the quarks, one from the gluon.

Back to the story of the top, the cleanest signature of top production occurs when both W bosons decay into leptons, giving two leptons and two jets

$$t\bar{t} \longrightarrow l^+ + l^- + \nu + \bar{\nu} + \bar{b} + b$$

This has a low background from other processes, but happens only rarely in top decay as well. Another option is that one of the W-bosons decays into leptons and the other into quarks, giving

$$\begin{aligned} t\bar{t} \longrightarrow & \quad l^+ + \nu + \bar{q} + q + \bar{b} + b \\ & \text{or } \; l^- + \bar{\nu} + \bar{q} + q + \bar{b} + b \end{aligned}$$

which has a single lepton, and four jets. This process, which is depicted in Figure 29, happens more often, but also has a higher background. The final decay, with no lepton and six jets gets swamped by the background.

After a number of hints in the early 1990s, the discovery of the top quark was finally announced in 1995. It was found at the Tevatron in Fermilab, a p^+p^- collider that reached energies of 1 TeV. Collisions took place at two similar but complimentary detectors, situated in different places around the ring, and were analysed by two independent rival collaborations known as D∅ (pronounced dee-zero) and CDF. A top-quark event from the CDF collaboration is shown on the next page.

The top took so long to track down because it was much heavier than anyone had anticipated. It finally weighed in at

$$m_t = 170 \text{ GeV}$$

This remains the heaviest fundamental particle that we know. Why is it so much heavy? We have, I think it's fair to say, no idea. Surely it is telling us something important.

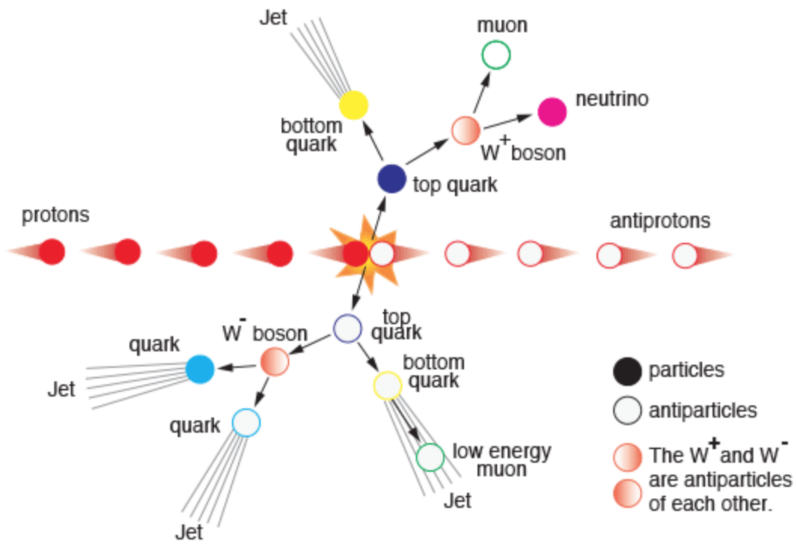
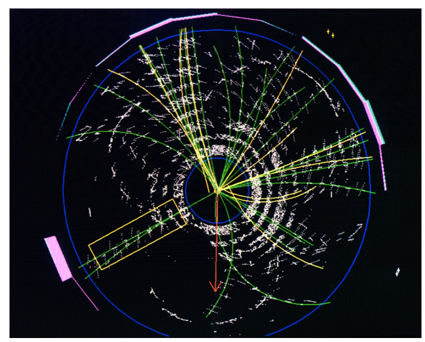


Figure 29. A schematic picture of a top quark event, taken from the Fermilab website.



With the tau, bottom and top quark in place, all that was left was the tau neutrino. This was discovered by the DONUT experiment at Fermilab in 2000. They used the Tevatron to create a beam of neutrinos which included ν_τ. These were then directed at nuclear emulsion targets where they collided with iron nuclei. The tell-tale sign of a ν_τ neutrino is the creation of a τ lepton, which leaves a small 1 mm track in the emulsion. Four such events were seen.