Elementary Particle Physics Lecture Notes 2013-14

Bobby Samir Acharya

March 4, 2014

Resources

Books:

Halzen-Martin: Quarks and Leptons: An Introductory Course in Modern Elementary Particle Physics

and

Kane: Modern Elementary Particle Physics

and

Thompson: Modern Particle Physics

Will mainly use Halzen-Martin, but Kane and Thompson are also very useful texts.

The known properties of the Standard Model particles and all known particles are listed in the Particle Data Group book: http://pdg.lbl.gov/2012/listings/contents_listings.html

This is an invaluable resource for particle physicists!

Particle Physics. Weeks 1 and 2

Brief Review of Week 1: (refer to the slides for a reminder. Chapter 1 of Halzen-Martin and/or Thompson is also a good summary of some topics we discussed.)

In Week 1 we covered:

Some evidence for the classification of matter into particles: the periodic table, the discovery of the electron and the discovery of the nucleus.

We connected particle physics to the Universe at large: We noted the fact that the $\sim 10^{11}$ stars in the galaxy look alike to first approximation. Similarly, the $\sim 10^{11}$ galaxies are also alike to first approximation, implying that there are roughly $10^{11} \times 10^{11} = 10^{22}$ stars, all approximately like our sun, in the observable Universe. Since our sun's mass is approximately 10^{30} kg and we know it is composed of atoms (e.g. H and He and others) and that an atom has a mass of roughly 10^{-27} kg, we learn that there are approximately $10^{11} \times 10^{11} \times 10^{30} \times 10^{27} = 10^{79}$ protons and neutrons in the Universe!. All of these are described by particle physics.

We introduced the Standard Model of Particle Physics for the first time: we described the three families (or generations or flavours) of quarks and leptons. We discussed their properties such as electric charge and mass. We pointed out the huge range of masses from the lightest charged particle (the electron) to the heaviest one (the top quark). These span six orders of magnitude. If we consider that the neutral fermion masses (the neutrinos) have to be less than about 1 eV/c^2 (though we will use natural units with $c = 1$), then the masses span at least 12 orders of magnitude (an order of magnitude is essentially adding an additional digit, so the number 130 is an order of magnitude larger than the number 16; similarly 2456 is two orders of magnitude smaller than 355573.)

We introduced the strong, weak and electromagnetic interactions; these are the three forces which govern the behaviour of matter at distance scales short enough (equivalently energy scales high enough) that gravity is sub-dominant. We described the forces as being via the exchange of gauge bosons (the photon, gluons, $W^{\pm}, Z$) and introduced the basic Feynman diagrams (containing two fermions and one boson) which gives the basic structure of the Standard Model. We described which particles participate in the three forces. All electrically chared particles participate in electromagnetism, so that means all the fermions except neutrinos. The leptons do not participate in strong interactions, only the quarks and gluons do this. All the particles participate in the weak interactions involving the W and Z bosons. (Note: later we will note that fermions can be "left-handed" or "right-handed" so that only left handed particles (and right handed anti-particles) participate in the interactions with W-bosons.). By putting together these basic Feynman diagrams we can create more complicated ones and these are the processes described by the Standard Model.

Each interaction (strong, weak, electromagnetic) is thus associated with a simple Feynman diagram; at the vertex of the diagram is assigned a "charge". When the boson is the photon, this is the electromagnetic interaction and the "charge" is actually the electric charge, e. For the strong and weak interactions it is a different charge and is denoted g_3 for the strong interaction and g_2 for the weak interaction with W-bosons. For the weak interactions involving the Z-boson we will later see that the strength is given by a particular combination of e and g_2.

We mentioned that the charges actually are not constants, rather they depend upon the energy scale of the process that a particular Feynman diagram is describing. Consider the following two diagrams taken from Halzen-Martin. An electron is not "just an electron". It is continually radiating photons which convert into e^+e^- pairs. An electron is thus surrounded by a "virtual cloud" of electron/positron pairs (see next diagram). The positrons in the cloud will tend to be closer to the original electron as they have opposite charge. Hence, the cloud is polarised. This affects what we mean by charge in a distance/energy dependent way.

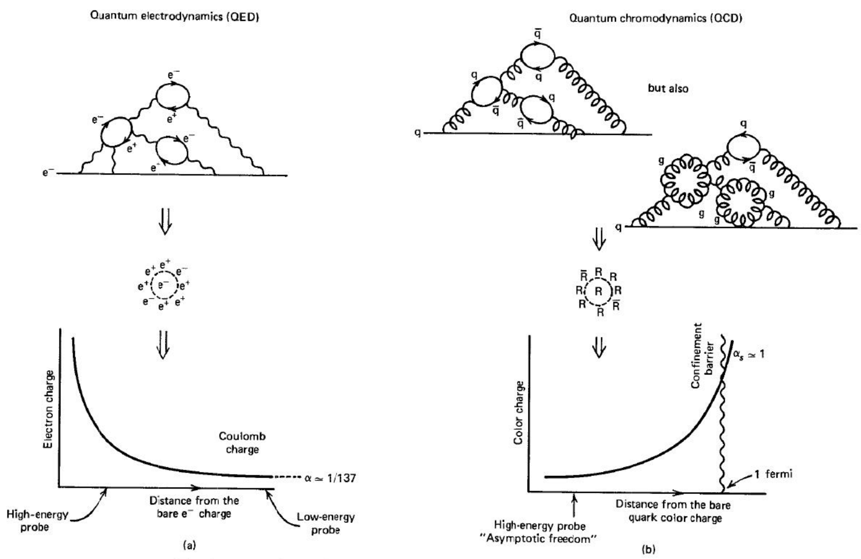


Fig. 1.5 Screening of the (a) electric and (b) color charge in quantum field theory.

We can measure the electron charge by taking a test charge and measuring the force it experiences. Clearly, the result depends upon how close the test charge is to the electron cloud, as the next figure illustrates. The positrons effectively screen the electron from the test charge: a photon which is exchanged between the positive test charge and a cloud positron gives a "repulsive" contribution to the overall force (which is attractive); the cloud electrons also contribute to this effect, but because they are farther out there is a smaller chance that the photons will interact with them so the effect is smaller. As the test charge is brought closer, the effective charge increases as the screening effect becomes less and less relevant. This results in a distance (equivalently energy) dependent charge, which is larger at shorter distances or higher energies.

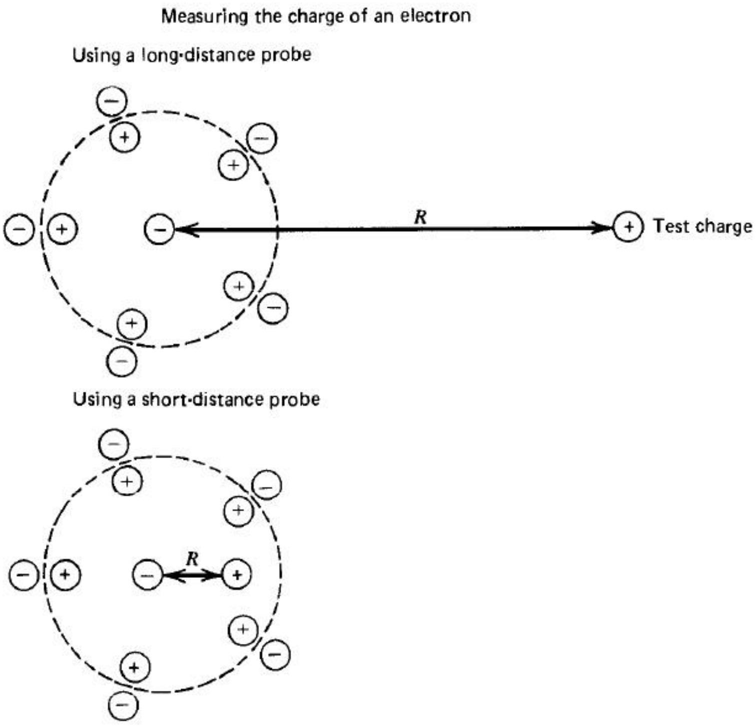


Fig. 1.6 Measuring the charge of an electron.

In the case of the strong interactions (described by quantum chromodynamics or QCD) the story is similar but with a different result. This is illustrated in the other part of figure 1.5. In the strong interactions, gluons can interact with themselves (cf electromagnetism where photons cannot self-interact). This leads to additional contributions to the effective charge which in fact leads to an energy dependance of the charge which is opposite to that in quantum electrodynamics (QED). In QCD the strength of the interaction i.e. the charge increases with distance or decreases with energy – so that asymptotically at arbitrarily high energies the charge goes to zero. This is known as asymptotic freedom. The understanding of this phenomenon led to a Nobel prize in physics for Gross, Politzer and Wilczek in 2004. As a consequence, at low energies it is not possible to directly observe quarks or gluons: the strong interaction is so strong that it binds them into bound states known collectively as "hadrons". One of these, the proton, is long lived. The rest of them, though, do not have long lifetimes: they quickly decay due to the electromagnetic and weak interactions that their constituents also participate in.

The Large Hadron Collider

We gave an overview of the Large Hadron Collider. Set in a roughly circular, 27 km diameter tunnel, about 100 m below ground, the LHC is the world's largest machine. It is also the world's most energetic and powerful particle collider. Powerful superconducting magnets and electric fields are used to accelerate charged particles (usually protons, but also lead ions) around the LHC ring in both directions. The two counter rotating particle beams are made to collide at four interaction points around the ring. Surrounding each interaction point are different particle detectors, designed to measure the results of proton-proton (and other) collisions. The LHC is designed to collide protons at a centre-of-mass energy $\sqrt{s}$ of 14 TeV = 14 $\times 10^{12}$ eV. It first started operation (more than 20 years after its conception) in 2010. It has successfully run in 2011 at $\sqrt{s}$=7 TeV and in 2012 at $\sqrt{s}$=8 TeV. For most of 2013 and 2014 it will be undergoing upgrades in order to ramp up to the design energy.

High energy is one key property that is necessary for a successful LHC. The collider also has a very high luminosity. The higher the luminosity the greater the number of collisions between the particle beams will take place. This is proportional to the number of protons in each of the colliding beams and inversely proportional to the effective area over which the beams are made to collide. Hence, luminosity is measured in units of inverse area cm^{-2} or the instantaneous luminosity is given in $cm^{-2}s^{-1}$. The design luminosity of the LHC is about $10^{34}cm^{-2}s^{-1}$ (or about $10^{41}cm^{-2}$ per year) and last year it was running pretty close to that. High energy is required because we are hoping to produce new particles with a large mass in the collisions and this can't be done if there isn't enough energy. High luminosity is required because the production of new particles is a rare process (otherwise we may have seen them already!) so the larger the number of collisions the greater the probability of producing a new particle. For instance, a Higgs boson is produced about once for every billion collisions at the LHC.

Protons are not elementary particles, but rather "large", complicated bound states of quarks and gluons tightly bound together by the strong interaction. One aim is to get the quarks or gluons from one proton beam to collide "hard" with those of the other proton beam and produce interesting new particles. Colliding protons together is a bit like colliding cans of beans together at high energies to both create new flavours of beans in the process as well as understand which beans are in the cans. As a result, proton-proton collisions are "quite messy" with hundreds of particles being produced in a given collision. The detectors which measure the results of these collisions must therefore be quite sophisticated, comprehensive, versatile instruments capable of recording pretty much anything that could conceivably be spat out of a proton-proton collision.

One of the main aims of the LHC projects was to prove or disprove the existence of the Higgs boson. The theory of the Higgs boson goes back to 1964 when Peter Higgs (who studied at King's !) proposed a mechanism to give masses to gauge bosons. This mechanism was incorporated into what eventually became the Standard Model – and it is the interactions of the fermions and the bosons with the Higgs boson which is responsible for the masses that those particles have. The proof (or disproof) of the existence of the Higgs boson is thus paramount to our understanding of the nature of mass and goes a long way towards answering e.g. the question "why is the mass of the electron about 0.5 MeV/c^2 ?" and towards asking about the nature of mass itself. Particle physicists have been looking for this particle for decades. Remarkably, after analysing data collected in 2011 and 2012, the CMS (Compact Muon Solenoid) and ATLAS (A Toroidal LHC Apparatus) experiments at the CERN LHC have discovered a new particle with properties remarkably similar to that of the Higgs boson. Much work is now underway to measure the detailed properties of this newly discovered particle to verify if it really is the Higgs boson or something else.

Natural Units and Dimensional Analysis

Natural units are a system of units which actually make thinking about physics easier. This is because in these units energy and mass have the same dimensions and these are inverse to the dimensions of lengths and times. In fact, in natural units any physical quantity will have dimensions of mass to some power; equivalently it will have the dimension of length to the inverse power.

These units are defined by "setting Planck's constant $\hbar$ and the speed of light c to one. " We put the previous statement in speech marks because the speed of light and $\hbar$ are not equal to one in a general unit system; however they are fundamental constants of nature. Hence you can view the speed of light as a relationship between metres and seconds. In fact, in SI units metres are defined by "the length of the path travelled by light in vacuum during a time interval of 1/299,792,458 of a second ". Thus c can be viewed as a conversion factor between metres and seconds.

In natural units,

$$1 \text{ s} = 299792458.. \text{ m} \tag{1}$$

which we will often approximate to

$$1 \text{ s} = 3 \times 10^8 \text{ m} \tag{2}$$

Similarly, Plancks constant (over 2π) which has the value

$$\hbar = 1.05457172534 \times 10^{-34} \ Js \tag{3}$$

can be thought of as a conversion factor between Joules and inverse seconds. This leads to the approximate relation

$$1 \text{ J} \sim 10^{34}\text{s}^{-1} \tag{4}$$

Since both $\hbar$ and c involve seconds, combining them allows us to convert between energy $[E]$, 1/time $[T]^{-1}$ and 1/distance $[L]^{-1}$.

One Joule can be expressed in electron volts as

$$1 \ \text{eV} = 1.602176.. \times 10^{-19} \ \text{J} \tag{5}$$

so, $\hbar = 1$ implies

$$\hbar = 1 \sim 10^{-34}\text{J s} \sim 3 \times 10^{-26}\text{J m} \sim 3 \times 10^{-26}\frac{10^{19}}{1.6..}\text{eV m} \sim 2 \times 10^{-7}\text{eV m} \tag{6}$$

$$1 \ \text{m} \sim \frac{1}{2 \times 10^{-7} \ \text{eV}} \tag{7}$$

which means that a distance of one metre is equivalent to an energy scale of 2×10^{-7} eV. Larger distances correspond to smaller energies and vice-versa.

In natural units, the mass of the electron is about 0.5 MeV. This is the inverse of a distance of $\frac{1}{0.5\text{MeV}} \sim 4 \times 10^{-13}$ m This is the Compton wavelength of the electron. The mass of proton is about 0.94 GeV, corresponding to a length scale of order 2×10^{-16} m. This is the characteristic size of a nucleus. The masses of the W-bosons, the Z-boson, the top quark and (if confirmed at the LHC) the Higgs boson are of order 100 GeV – corresponding to a distance of around 2×10^{-18} m. The LHC collisions in 2012 took place at energies of 8 TeV = 8 thousand GeV – a distance scale of almost 10^{-20} m. This makes the LHC the world's most powerful microscope.

Lorentz Invariance, Special Relativity and 4-vector notation

Einstein's postulates of special relativity:

1. The speed of light is a constant, measured to have the same value by all observers

2. The laws of physics are the same for all observers moving with constant relative velocity to one another (i.e. in uniform relative motion).

If the relative motion between observers is along the x direction, then

$$\begin{aligned} t' &= \gamma(t - vx) \\ x' &= \gamma(x - vt) \\ y' &= y \\ z' &= z \end{aligned} \tag{8}$$

where

$$\gamma = \frac{1}{\sqrt{1 - v^2}} \tag{9}$$

The boost along the x-direction with velocity v mixes the space and time coordinates x and t. A key fact about special relativity is that the quantity

$$L^2 \equiv t^2 - x^2 - y^2 - z^2 \tag{10}$$

is invariant under Lorentz transformations. In fact,

$$\begin{aligned} t'^2 - x'^2 - y'^2 - z'^2 &= \gamma^2(t^2 - 2vx + v^2x^2) - \gamma^2(x^2 - 2vt + v^2t^2) - y^2 \quad (11) \\ &= \frac{1}{1 - v^2}(t^2 - x^2)(1 - v^2) \\ &= t^2 - x^2 - y^2 - z^2 = L^2 \end{aligned}$$

The left and right hand side of this equation can be regarded as a measure of spacetime length and are unchanged by Lorentz transformations (note: the overall sign of the length is not important, we could have used $-t^2 + x^2 + ...$).

We can write this in a more compact form.

Assemble the four space and time coordinates into a vector with four components, a 4-vector and call this x_μ.

$$x^\mu \equiv \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \equiv \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

The invariant length can be thought of as a sort of dot product between this vector and itself, but defined with some minus signs. In fact

$$L^2 = \begin{pmatrix} t & x & y & z \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}$$

i.e.

$$L^2 = (x^\mu)^T \eta x^\mu \equiv x^\mu \eta_{\mu\nu} x^\nu \tag{12}$$

where x_μ is viewed as a vector multipling the matrix $\eta_{\mu\nu}$; η is known as the Minkowski "metric":

$$\eta_{\mu\nu} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Note: η is denoted by g in Halzen-Martin.

We often simplify life and omit the η by writing

$$L^2 = x_\mu x^\mu \tag{13}$$

where, now, notice that the μ is a subscript on the first x. x_μ is simply $(x^\nu)^T \eta_{\mu\nu}$ and now the product between x^μ and x_μ is the usual dot product.

The space coordinates $(x, y, z) \equiv x_i$ are the components of a 3-vector. Another familiar 3-vector is the 3-momentum, $p_i \equiv (p_x, p_y, p_z)$ which are the components of an objects momentum in the three spatial directions. In the same way that the x_i combine with another quantity t to form a 4-vector x^μ, the 3-vector p_i combines with another quantity to give a 4-vector p^μ – the 4-momentum. In this case, it is the energy of the object which is p_0:

$$p^\mu \equiv \begin{pmatrix} E \\ p_x \\ p_y \\ p_z \end{pmatrix} \equiv \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix}$$

The quantity,

$$\boxed{M^2 = p^\mu p_\mu = E^2 - p_x^2 - p_y^2 - p_z^2 = E^2 - |p_i|^2} \tag{14}$$

is also Lorentz invariant, like L^2. Hence it will be the same in all relatively uniform frames. Consider now a (frame in which we have a) particle a rest. Being at rest, its momentum is zero, $p_i = 0$. Hence

$$p^\mu \equiv \begin{pmatrix} E \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

its energy is equal to its mass. Hence, in this frame,

$$M^2 = m^2 \tag{15}$$

but, since M^2 is Lorentz invariant, $M^2 = m^2$ in all frames. Hence M is known as the invariant mass. Therefore,

$$\boxed{E^2 = |p_i|^2 + m^2} \tag{16}$$

notice that, in the small velocity limit, $p_i \to 0$ and $E \approx m$. In non-natural units this becomes $E \approx mc^2$, a famous equation but one which is only approximately true.

The energy-momentum 4-vectors of a system are conserved in time: e.g. consider and initial state with two particles A and B with 4-vectors

p_μ^A and p_μ^B which interact, producing a final state with particles C and D. Energy-mometum conservation asserts that:

$$P_A^\mu + P_B^\mu = P_C^\mu + P_D^\mu \tag{17}$$

where the lhs and rhs are both additions of two 4-vectors.

In the case of a Higgs boson decaying into, say, two photons the initial state is one particle (the Higgs boson) whereas the final state is two particles. If we label the Higgs boson as A and the two photons as B and C, at an LHC experiment we will measure the energy and momenta of each of the two photons. But conservation of energy and momentum implies that:

$$P_A^\mu = P_B^\mu + P_C^\mu \tag{18}$$

Therefore, we can reconstruct the energy and momentum of the Higgs by measuring the energy and momentum of the particles it decayed into. The mass of the Higgs boson m_h is the invariant mass of P_A^μ, therefore we have:

$$m_h^2 = P_A^\mu P_{A\mu} = (P_B^\mu + P_C^\mu)(P_{B\mu} + P_{C\mu}) \tag{19}$$

So we can measure the mass of the Higgs boson at the LHC even though the Higgs decayed very quickly.

From Schrodingers equation to Relativistic Quantum Theory

Non-relativistic Quantum Mechanics

Non-relativistic quantum mechanics is governed by Schrodingers equation. This equation is simply an expression in waves/fields of the non-relativistic energy-momentum relation:

$$\boxed{E = \frac{p^2}{2m}} \tag{20}$$

If we substitute for E and p the differential operators:

$$E \to i\frac{\partial}{\partial t} \quad \text{and} \quad p_i \to -i\frac{\partial}{\partial x_i} \tag{21}$$

we get the free Schrodinger equation:

$$\boxed{i\frac{\partial \Psi}{\partial t} + \frac{1}{2m}\frac{\partial^2 \Psi}{\partial x_i^2} = 0} \tag{22}$$

where we are acting with the equation on a function of space and time Ψ, called the wavefunction. Then

$$\rho = |\Psi|^2 \tag{23}$$

is the probability density and

$$\rho d^3x \tag{24}$$

the probability of finding the particle in the infinitessimal volume element d^3x.

For applications in which beams of particles interact/collide we will need to calculate the "flux density" of particles $\mathbf{j}$ passing out of some volume V. The conservation of probability ρ asserts that the rate of decrease of the number of particles is equal to the flux coming out of a surface S surrounding the volume V:

$$-\frac{\partial}{\partial t}\int_V \rho dV = \int_S \mathbf{n}.\mathbf{j} dS = \int_V \nabla.\mathbf{j} dV \tag{25}$$

$\mathbf{n}$ is just a unit normal vector to S and the last inequality is Stokes theorem. We therefore have a continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla.\mathbf{j} = 0 \tag{26}$$

Next we can get an expression for $\frac{\partial \rho}{\partial t}$ by taking Schrodingers equation and multiplying on the left by $-i\Psi^*$ and substracting this from the complex-conjugate of Schrodingers equation multiplied by $-i\Psi$. This gives:

$$\frac{\partial \rho}{\partial t} = \frac{i}{2m}(\Psi^* \nabla^2 \Psi - \Psi \nabla^2 \Psi^*) \tag{27}$$

This implies that the probability flux density $\mathbf{j}$ is

$$\mathbf{j} = -\frac{i}{2m}(\Psi^* \nabla \Psi - \Psi \nabla \Psi^*) \tag{28}$$

A simple solution of the free Schrodinger equation is:

$$\Psi = N e^{i(p^j x_j - Et)} = N e^{ip^\mu x_\mu} \tag{29}$$

which describes a free particle with momentum $p_i \equiv \mathbf{p}$ and energy E (notice that the exponent is Lorentz invariant). This has

$$\rho = |N|^2 \quad \text{and} \quad \mathbf{j} = \frac{\mathbf{p}}{m}|N|^2 \tag{30}$$

Relativistic Quantum Mechanics: the Klein-Gordon Equation

We can perform the same exercise with the relativistic E, p relation:

$$E^2 = \mathbf{p}^2 + m^2 \tag{31}$$

This gives the Klein-Gordon equation,

$$\boxed{-\frac{\partial^2 \phi}{\partial t^2} + \nabla^2 \phi = m^2 \phi} \tag{32}$$

ϕ is the wavefunction (or field) for a relativistic particle with mass m, energy E and momentum $\mathbf{p}$. Again, we can multiply the equation by

$-i\phi^*$ and substract this from the complex conjugate equation multiplied by $-i\phi$ to obtain ρ and $\mathbf{j}$:

$$\rho = i(\phi^* \frac{\partial \phi}{\partial t} - \phi \frac{\partial \phi^*}{\partial t}) \tag{33}$$

and

$$\mathbf{j} = -i(\phi^* \nabla \phi - \phi \nabla \phi^*) \tag{34}$$

A free particle solution of the KG equation is again given by

$$\phi_{free} = N e^{i(p^j x_j - Et)} = N e^{i p^\mu x_\mu} \tag{35}$$

hence

$$\rho = 2E|N|^2 \quad \text{and} \quad \mathbf{j} = 2\mathbf{p}|N|^2 \tag{36}$$

The KG equation is Lorentz invariant, since it can be written as

$$\boxed{(\partial^\nu \partial_\nu + m^2)\phi = 0} \tag{37}$$

and this suggests that ρ and $\mathbf{j}$ should combine into the four components of a 4-vector:

$$j^\mu \equiv \begin{pmatrix} \rho \\ j^1 \\ j^2 \\ j^3 \end{pmatrix} = -i(\phi^* \partial^\mu \phi - \phi \partial^\mu \phi^*) = 2p^\mu |N|^2$$

The continuity equation can be written in the Lorentz invariant form

$$\boxed{\partial^\mu j_\mu = 0} \tag{38}$$

Substituting the free particle solution into the KG equation, we see that we can have solutions for both signs of the square root:

$$\boxed{E = \pm\sqrt{\mathbf{p}^2 + m^2}} \tag{39}$$

We need to interpret the negative energy solutions. Furthermore, the negative energy solutions have negative probability density:

$$\boxed{E < 0 \ , \ \ \rho < 0} \tag{40}$$

Pauli and Weisskopf had the idea of multiplying j^μ by the charge, e:

$$j^\mu \longrightarrow -ie(\phi^* \partial^\mu \phi - \phi \partial^\mu \phi^*) \tag{41}$$

In this case "negative ρ" can be associated with particles with the opposite charge.

The final interpretation of the negative energy solutions was given by Feynman and Stuckelberg who asserted: The negative energy particle solutions are actually positive energy anti-particles. Moreover, if the latter propagate forwards in time, the former are propagating backwards in time. To see this, consider an electron of energy E, momentum $\mathbf{p}$ and charge $-e$. It has

$$j^\mu(e^-) = -2e|N|^2(E, \mathbf{p}) \tag{42}$$

A positron, with the same E and $\mathbf{p}$ has

$$\begin{aligned} j^\mu(e^+) &= +2e|N|^2(E, \mathbf{p}) \\ &= -2e|N|^2(-E, -\mathbf{p}) \end{aligned} \tag{43}$$

which is the same as that for an electron with energy $-E$ and momentum $-\mathbf{p}$.

This basically implies that the single particle wavefunction formalism describes both particles and antiparticles simultaneously. See Halzen-Martin 3.5 for some more details.

From Amplitudes to Cross-sections and Lifetimes.

Using basic quantum mechanics and the relativistic wave equation we have seen that partice physics processes can be organised into Feynman diagrams in which the probability per unit time per unit volume is related to the square of the amplitude.

For charged particles scattering by exchanging photons, the amplitude is given by

$$T_{fi} = -i \int d^4x j_1^\mu(x) \frac{-1}{q^2} j_{2\mu}(x) \tag{44}$$

Here, j_1^μ is the current for one of the particles and $j_{2\mu}$ the current for the other. $\frac{-1}{q^2}$ is the propagator of the photon which is exchanged between them.

We derived the currents for both non-relativistic and relativistic particles earlier. In the relativistic case, $j_{2\mu} = -eN_BN_D(p_D + p_B)^\mu e^{i(p_D - p_B).x}$. Putting this, and the analagous expression for the other current into the expression for T_{fi} gives

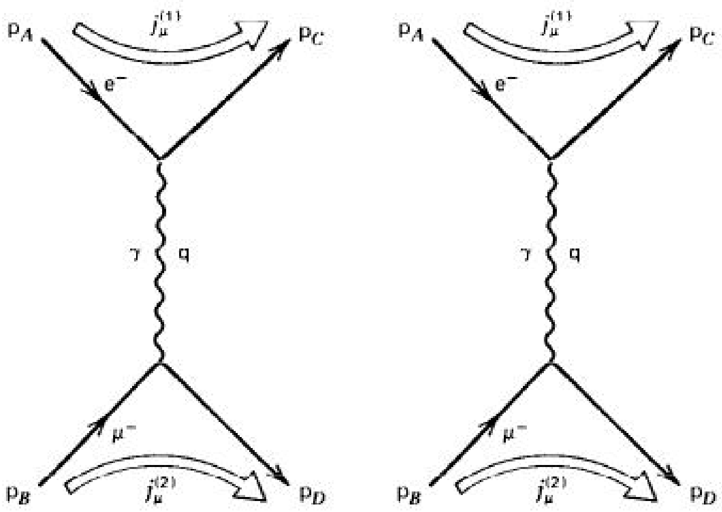
$$T_{fi} = -iN_AN_BN_CN_D(2\pi)^4\delta^4(p_D + p_C - p_A - p_B)\mathcal{M} \tag{45}$$

where

$$-i\mathcal{M} = (ie(p_A + p_C)^\mu)(-i\frac{g_{\mu\nu}}{q^2})(ie(p_B + p_D)^\nu) \tag{46}$$

$\mathcal{M}$ is simply the Feynman diagram translated into a mathematical expression, known as the invariant amplitude. $\mathcal{M}$ is intrinsically physical. The remaining factors in T_{fi} take care of a) the energy and momentum conservation of the process and b) the normalisation of the particle wave functions.

In particle physics experiments we measure quantities like the lifetime of a decaying particle or the rate of production of particles from a collider experiment. We need to be able to calculate quantities like these, starting with the invariant amplitude $\mathcal{M}$ for such a process.



1 Particle Decays

Consider a particle which decays into a number of other particles. e.g. a W-boson can decay into a positron and a neutrino, a neutron can decay as $n \to pe\bar{\nu}_e$, a muon decays as $\mu \to e\nu_\mu\bar{\nu}_e$. These processes can be measured and we would like to understand what the Standard Model predicts for them.

In general, if the initial particle has some fixed energy and momentum, there are many possibilities for what the final state energies and momenta could be for the particles produced in the decay. The initial particle has a specified energy and momentum (obeying $E^2 = p^2 + m^2!$), but, given this, there are many possible final state particle energies and momenta. If we integrate over all of these possible final state energies and momenta, taking into account energy and momentum conservation (which is enforced by the delta-function), we can obtain (for example) the lifetime of the particle.

Independent of the final state energies and momenta, there might be many different combinations of particles that a given particle can decay into. For instance, a W-boson can decay into $(e^+, \nu_e); (\mu^+, \nu_\mu); \ldots\ldots.$

Each possible combination of final state particles (which differ by their "charges") contributes to the total lifetime. The branching ratio (or branching fraction) for a particular final state combination of particles is the (inverse of) the total lifetime of the particle divided by the average lifetime for decaying into the given combination.

Recall (Fermi's Golden rule) that gave us the probability per unit time for a transition from one state to another. If the initial state is a particle and the final state the particles it decays into, this gives us the probability per unit time for the decay to occur.

$$\Gamma = 2\pi |\mathcal{M}|^2 \rho(E) \tag{47}$$

$\rho(E)$ is the density of states. More about this in a moment. Γ is known as the decay width of the particle. The lifetime of the particle is:

$$\tau = \frac{1}{\Gamma} \tag{48}$$

In natural units, $\hbar$ =1 $= c$ and, hence, energies and masses can be measured in units of eV and distances and times in units of eV^{-1}. Thus a lifetime has dimensions of $[\mathrm{E}]^{-1} = [\mathrm{M}]^{-1}$ and Γ has dimensions of mass i.e. $[\Gamma] = [\mathrm{M}]$.

1.1 Particle Number Density

Consider quantum mechanics in a finite region of space - a box with sides of equal length L.

We require that the single particle, free particle wave function, $e^{-ip.x}$ has well defined boundary conditions at the boundaries of the box i.e. when, x, y or $z =$ L. For this we can impose periodic boundary conditions so that the wave function is periodic at the boundaries:

$$\Psi(x, y, z) = \Psi(x + L, y + L, z + L) \tag{49}$$

This implies that

$$(p_x, p_y, p_z) = \frac{2\pi}{L}(n_x, n_y, n_z) \tag{50}$$

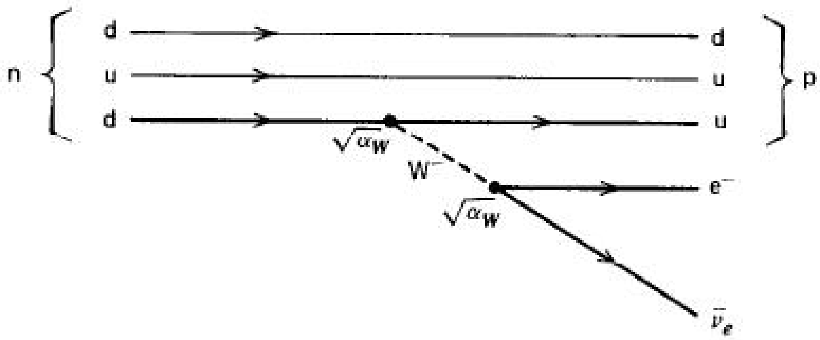
where the n_i are INTEGERS. The momentum of a single particle in a box is QUANTISED.

Hence, the number of states between p_x and $p_x + dp_x$ is $\frac{L}{2\pi} dp_x$.

Therefore, the number density of states for a single particle is:

$$dn = \frac{L^3}{(2\pi)^3} d^3p = \frac{V}{(2\pi)^3} d^3p \qquad (51)$$

1.2 The Lifetime of the Neutron



Though stable inside most atoms, a free neutron is unstable. It decays into a proton, an electron and an anti-neutrino:

$$n \longrightarrow p e^- \bar{\nu}_e \qquad (52)$$

This is β-decay and also occurs within certain radioactive elements. The lifetime of the neutron at rest is about 900 seconds = 900 $\times$ 3$\times$ 10^8 m $\approx$ 900 $\times$ 3$\times$ 10^8 $\times 10^6$eV^{-1} $\approx$ 2.7 $\times$ 10^{17}eV^{-1} = 2.7 $\times 10^{26}$GeV^{-1}.

We will estimate the lifetime ignoring the masses of the electron and neutrino. The mass difference between a neutron and proton is around 1 MeV. The kinematics of this process is as follows. The energy of the neutron must equal that of the proton plus electron plus neutrino:

$$E_n = E_p + E_e + E_\nu \qquad (53)$$

If the initial neutron is at rest, $E_n = m_n$ and, since we ignore the masses of the e and ν we have that $E_e = |p_e| \equiv p$ and $E_\nu = |p_\nu| \equiv q$.

The momentum of the proton is equal and opposite to the sum of the momenta of the electron and neutrino, so is completely fixed. Hence, when we integrate over the final state particles, we only need to integrate over the electron and neutrino momenta. We will define $E_n - E_p$ to be E_0, an energy of order 1 MeV.

$$\Gamma = 2\pi|\mathcal{M}|^2\rho = 2\pi\int|\mathcal{M}|^2\frac{d^3p}{(2\pi)^3}\frac{d^3q}{(2\pi)^3} \tag{54}$$

where the $\int$ means we will integrate over p and q.

A quick reminder about spherical polar coordinates is in order!

1.3 Spherical polar coordinates on the side

Consider the three space coordinates $x_i = (x, y, z)$. We can re-write these in terms of two angles (θ, ϕ) and a radius r:

$$\begin{aligned} x &= rsin\theta cos\phi & (55)\\ y &= rsin\theta sin\phi & (56)\\ z &= rcos\theta & (57) \end{aligned}$$

where $r > 0$, $0 \leq \theta \leq \pi$, $0 \leq \phi \leq 2\pi$. Now, instead of specifying x, y and z to specify a particular point, we can specify r, θ and ϕ instead. This is a coordinate transformation. Note, that we can do this for the components of any three-vector, e.g. momentum.

Notice that:

$$\begin{aligned} r &= (x^2 + y^2 + z^2)^{1/2} & (58)\\ \phi &= tan^{-1}(y/x) & (59)\\ \theta &= cos^{-1}(z/r) & (60) \end{aligned}$$

So, in particular, $r = |x_i|$, the length of the vector.

Important to note is how the integration measure gets transformed:

$$dxdydz \longrightarrow r^2 sin\theta dr d\phi d\theta \tag{61}$$

We can integrate over the two angles:

$$\int r^2 dr \int d\phi \int sin\theta d\theta = 4\pi r^2 \int dr \tag{62}$$

So, the integral over space becomes equivalent to an integral over r:

$$\int d^3x = 4\pi \int r^2 dr \tag{63}$$

Similarly, instead of the 3-vector x_i we could consider the 3-momentum p_i:

$$\int d^3p = 4\pi \int p^2 dp \tag{64}$$

end of scholium on spherical polar coordinates. back to the density of final states.

Now, back to neutron decay!

Assuming the angular dependence of the matrix element is trivial, we can replace d^3p with $4\pi p^2 dp$, and, similarly for q. Moreover, $q = E_0 - E_e = E_0 - p$ (since we ignore the electron mass), so there is no actual integral over q and we only have to integrate over p. Therefore, we have

$$\Gamma \approx \int_0^{E_0} |\mathcal{M}|^2 \frac{(E_0 - p)^2}{2\pi^3} p^2 dp \tag{65}$$

The integral over p is easy and the end result is:

$$\Gamma \approx \frac{|\mathcal{M}|^2 E_0^5}{60\pi^3} \tag{66}$$

Since Γ has dimensions of mass, dimensional analysis tells us that for this process the dimension of $|\mathcal{M}|^2$ is mass^{-4}.

Let us therefore relabel $|\mathcal{M}|^2$ as $\frac{1}{M_*^4}$, the goal being to estimate what this mass M_* is. Historically, M_* is related to Fermi's constant as $G_F \sim M_*^{-2}$.

So, we now have:

$$\Gamma \approx \frac{E_0^5}{M_*^4 60\pi^3} \tag{67}$$

Since E_0 is of order an MeV, we can write

$$\Gamma \approx \frac{\text{MeV}}{60\pi^3}\frac{\text{MeV}^4}{M_*^4} \tag{68}$$

The second factor is dimensionless. Since $60\pi^3 \approx 1860 \approx 2000$ we have

$$\Gamma \approx 5\times 10^{-4}\text{MeV}\frac{\text{MeV}^4}{M_*^4} = 5\times 10^{2}\text{eV}\frac{\text{MeV}^4}{M_*^4} \tag{69}$$

To estimate M_*, we use the fact that the lifetime of the neutron is about 900s. This gives

$$M_* \approx 100\text{GeV} \tag{70}$$

This is a remarkable result. It suggests that there is something interesting happening at a scale which is 10^5 times larger than the energy transferred in this process and 100 times larger than the mass scale of the proton and neutron. This is essentially the mass scale of the W-boson (which is 80 GeV)!!! We have discovered the mass scale of the Standard Model of Particle Physics!.

Consider now, the muon. It decays according to the following Feynman diagram.

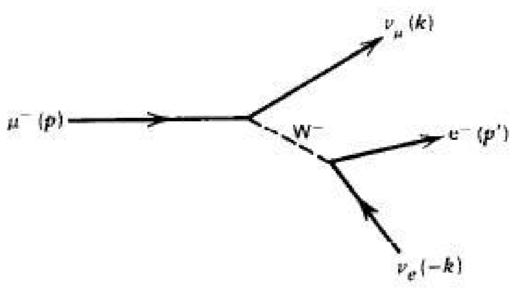


Fig. 12.5 Muon decay.

i.e. $\mu^- \to e^-\nu_\mu\bar{\nu}_e$

Why does the muon decay? The answer is that there is nothing from preventing this process to occur, so it will. The electron cannot decay, because there is no combination of lighter particles with the correct charges for it to decay to. Hence, its mass and charge prevent it from decaying. The muon, on the other hand has the same charge as an electron, but weighs 200 times more.

What we see from the diagram is that, even though a W-boson weighs 80 GeV, 800 times the muon mass, quantum mechanics allows this to occur: the energy-time uncertainty relation,

$$\Delta E \Delta t \geq \frac{1}{2} \tag{71}$$

allows a W-boson to be created for a very short amount of time. Note that the W-boson which "propagates" in the neutron and muon decay diagrams is not on-shell i.e. it does not obey $E^2 = p^2 + m_W^2$. It is the uncertainty relation that allows off-shell W's to propagate.

The diagram is completely analagous to the neutron decay diagram. In analogy with that one, we would estimate

$$\Gamma(\mu^- \to e^- \nu_\mu \bar{\nu}_e) \approx \frac{m_\mu^5}{M_*^4 60\pi^3} \tag{72}$$

$M_* \sim M_W$. Hence we see that the presence of the W-boson in the intermediate state contributes a factor of $1/M_W^2$ to the matrix element. This is the W-boson propagator for a process in which the energies are much smaller than M_W. The actual propagator is proportional to $\frac{1}{q^2 - M_W^2}$ c.f. the photon propagator.

So, if we hadn't done the calculation above we would have seen that the decay width is proportional to $\frac{1}{M_W^4}$ and hence, dimensional analysis would tell us that there is a m_μ^5 present. The $60\pi^3$ would require knowing that the final state "phase-space" gives an additional suppression.

The actual properly calculated muon decay width gives the result:

$$\Gamma(\mu^- \to e^- \nu_\mu \bar{\nu}_e) \approx \frac{m_\mu^5}{M_*^4 192\pi^3} \equiv G_F^2 \frac{m_\mu^5}{192\pi^3} \tag{73}$$

where G_F is now defined to be equal to M_*^{-2}.

Use this to calculate the muon lifetime and compare it to the value in the PDG (Particle Data book).

Calculate the lifetime of the τ-lepton.

1.4 Lorentz Invariant Phase Space Factor

Recall that the probability density, $\rho = 2E|N|^2$. The integral of ρ in a volume V produces a factor of V, so we could choose the constant N such that this cancels, i.e.

$$N = \frac{1}{\sqrt{V}} \tag{74}$$

and

$$\int_V d^3x\rho(E) = 2E \tag{75}$$

This normalisation is thus one in which there are $2E$ particles in a volume V. Hence, with this normalisation, the number density becomes:

$$dn = \frac{V}{2\pi^3}\frac{d^3p}{2E} \tag{76}$$

note: if we had not chosen a value for N, we would still have the same E dependence, but there would be a $\frac{1}{N^2}$ present as well. The final result doesn't actually depend on N

The last expression for dn is Lorentz invariant! To see this, consider a Lorentz boost in the z direction with velocity v:

$$E' = \gamma(E - vp_z) \tag{77}$$

$$p'_z = \gamma(p_z - vE) \tag{78}$$

Then,

$$dp'_z = \gamma dp_z - \gamma v dE \tag{79}$$

so

$$\frac{dp'_z}{dp_z} = \gamma(1 - v)\frac{dE}{dp_z} \tag{80}$$

Now

$$\frac{dE}{dp_z} = \frac{d}{dp_z}(p_x^2 + p_y^2 + p_z^2 + m^2)^{1/2} = p_z(p_x^2 + p_y^2 + p_z^2 + m^2)^{-1/2} = \frac{p_z}{E} \quad (81)$$

Therefore:

$$\begin{aligned} \frac{dp'_z}{dp_z} &= \gamma(1 - v\frac{p_z}{E}) && (82) \\ &= \gamma\frac{(E - vp_z)}{E} && (83) \\ &= E'/E && (84) \end{aligned}$$

Therefore

$$dp'_z/E' = dp_z/E \quad (85)$$

and

$$\frac{d^3p'}{E'} = \frac{d^3p}{E} \quad (86)$$

We now write the general formula for the differential decay width (its differential because it is before we integrate over final state particles) in a relativistic, Lorentz invariant system.

Consider a particle A which decays into n particles. The transition probability (or differential decay width) is given by:

$$d\Gamma = \frac{1}{2E_A}|\mathcal{M}|^2 \frac{d^3p_1}{(2\pi)^3 2E_1} .... \frac{d^3p_n}{(2\pi)^3 2E_n}(2\pi)^4\delta^4(p_A - p_1 - p_2 - ... - p_n) \quad (87)$$

$2E_A$ is the number of decaying particles per unit volume, $\mathcal{M}$ the invariant amplitude for the process. Note that there are no factors of the arbitrarily chosen normalisation volume V. Since $d\Gamma$ is measurable, the result should not depend on V (one can demonstrate this, but we will not do this in class).

2 Cross-sections

In a particle collider experiment, we collide beams of particles with some given flux (also called luminosity and then we measure the particles that come out of the collisions. The total number of events is obviously proportional to the flux since e.g. if we increase the number of protons in the LHC beams we will increase the total number of collision events that we get. Thus, the luminosity is something that we control as experimenters in a laboratory. It is not an intrinsically physical entity.

However, the proportionality "constant" in the relation

$$N_{events} \propto \text{Luminosity} \equiv \mathcal{L} \quad (88)$$

is an intrinsically physical quantity. This is called the cross-section and is usually labeled by σ.

$$N_{events} = \mathcal{L}\sigma \quad (89)$$

The left hand side is dimensionless, a whole number if we count the number of events (or collisions) after a given time interval. Or we could consider the number of events per second (or any other unit of time). The luminosity is thus a flux of particles per unit time. This has dimensions of $[\mathcal{L}] = [L]^{-2}[T]^{-1} = [M]^3$, where the last equality is because we use natural units. The dimensions of luminosity are like this because it is essentially the number of particles going through a given area i.e. number of particles per unit area per second.

So, by dimensional analysis, the cross-section σ has dimensions of AREA. In fact, that is where its name comes from. It is, effectively, the area over which the interaction takes place.

2.1 Dimensional analysis calculations of σ

Lets get a feel for σ by estimating it for various processes. Our estimates are based on Fermi's Golden Rule plus dimensional analysis in natural units. We will also compare our results to the "properly calculated" results as well as the actual experimental observations.

1. $\sigma(e^+e^- \rightarrow \mu^+\mu-)$ at high energies.

A Feynman diagram at leading order for this is the same as the interaction between muons and electrons diagram above, but turned on its side. We would like to estimate the cross-section for this at high centre of mass energies $\sqrt{s}$ much greater than m_μ.

By now, we know that the answer for this will by proportional to the "square of the Feynman diagram" for this process. This tells us that:

$$\sigma \propto e^4 \sim \alpha^2 \tag{90}$$

because we have a factor of the charge at each vertex. α is dimensionless, a number with no units. But σ has dimension $[M]^{-2}$. The only other scale in this problem is the center of mass energy $\sqrt{s}$. Hence we expect that

$$\sigma \sim \frac{\alpha^2}{s} \tag{91}$$

The actual "properly calculated" leading order result is

$$\sigma = \frac{4\pi}{3}\frac{\alpha^2}{s} \tag{92}$$

so, we were off by a factor of four or so. Not bad!

The cross-section for this process has been measured at various energies, as shown in the figure. The smooth line on this graph shows the theoretical prediction from the above equation and the different 'markers' are actual, experimentally measured values. The agreement between the theory and experiment is very good. Notice the vertical lines which eminate from the different experimental points.

Since no experiment is perfect, every reported measurement is uncertain by a certain amount depending on either the limitations of the experimental apparatus and/or the size of the data sample available. The latter, statistical uncertainty, is reduced when a larger data sample becomes available. The vertical lines on the graph are the total uncertainty on the measurement made, so the actual cross-section is somewhere in between the "band" represented by the line.

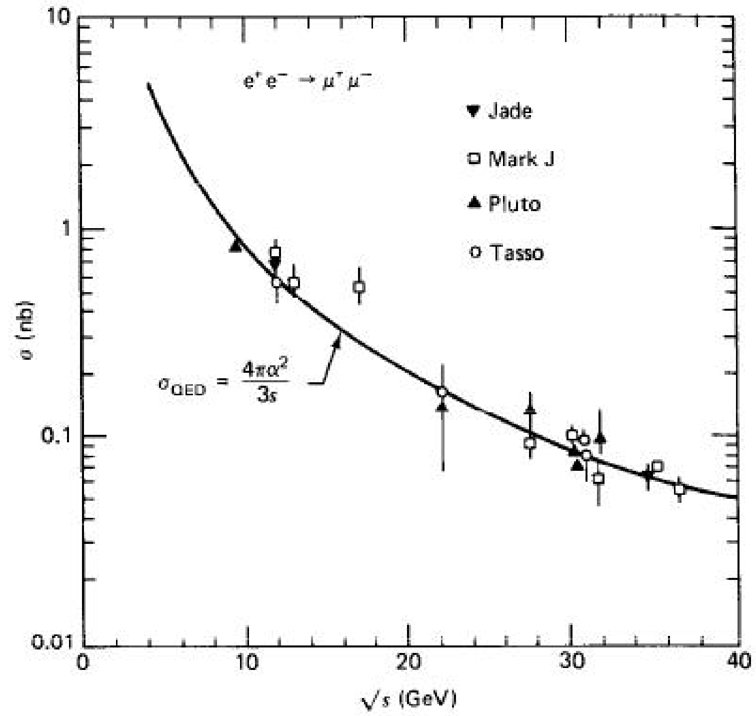


Fig. 6.6 The total cross section for $e^-e^+ \to \mu^-\mu^+$ measured at PETRA versus the center-of-mass energy.

Putting actual numbers to the cross-section we get:

$$\sigma(e^+e^- \to \mu^+\mu^-) \sim \frac{4 \times 10^{-32}}{s/\mathrm{GeV}^2}\mathrm{cm}^2 \tag{93}$$

So in order to produce a few muon pairs in electron-positron scattering at a centre-of-mass energy of one GeV, you need to have a luminosity of order $10^{32}\mathrm{cm}^{-2}$. At higher energies, since the cross-section decreases quadratically with energy, you need much higher luminosities to produce the same number of muons.

2. $\sigma(\nu N \to X)$

Here we want to consider neutrinos interacting with nucleons in matter. Neutrinos are not electrically charged so they don't couple to photons. Nor do they participate in the strong nuclear interactions (like quarks and gluons do). But they do undergo weak interactions, eg via exchanging W-bosons and Z-bosons.

2.2 Centre-of-Mass Frame and Laboratory Frame

The centre of mass frame for a collision of two particles is one in which the two particles have equal and opposite momentum and equal energies. In this frame the Lorentz 4-vectors for the two particles are

$$
\begin{aligned}
p_1^\mu &= (1/2E_{cm}, p^i) && (94)\\
p_2^\mu &= (1/2E_{cm}, -p^i) && (95)
\end{aligned}
$$

We can calculate the Lorentz invariant quantity s:

$$s = (p_1 + p_2)_\mu (p_1 + p_2)^\mu = E_{cm}^2 \tag{96}$$

The laboratory frame is one in which one of the particles is at rest and the other is moving.

$$
\begin{aligned}
p_1^\mu &= (M, 0) && (97)\\
p_2^\mu &= (E_{lab}, p_{lab}^i) && (98)
\end{aligned}
$$

In the lab frame

$$s = (E_{lab} + M)^2 - E_{lab}^2 + m^2 \approx 2E_{lab}M \tag{99}$$

where we assumed E to be much larger than M or m the two particle masses.

Okay, back to neutrino cross-section

Since neutrinos interact via the weak interaction, the $|\mathcal{M}|^2$ will be proportional to G_F^2 hence we will have that

$$\sigma(\nu N) \sim G_F^2 \tag{100}$$

But σ must have mass dimension minus two. At high energies, the only other scale in the problem is s which has mass dimension two.

Therefore, we expect that

$$\sigma(\nu N) \sim G_F^2 s \tag{101}$$

In most experimental situations with neutrinos we are normally in the lab frame, scattering a beam of neutrinos off a fixed target. e.g. the nucleons could be a "block" of matter and the neutrino beam is "fired" into it. Therefore $s \sim 2E_\nu m_N$

Using the fact that $m_N \sim 1$ GeV and that $G_F \sim 10^{-5}$ GeV^{-2} we get

$$\sigma(\nu N) \sim \text{afew} \times 10^{-38} \frac{E_\nu}{\text{GeV}} \text{cm}^2 \tag{102}$$

which is again in agreement with the "proper calculation" to within a factor of 10.

Notice that this is a much smaller cross-section than the previous one we estimated for a fixed centre-of-mass.

How far can neutrinos propagate through matter?

Imagine a neutrino which has been emitted by the sun (or any other star in the galaxy) and arrives at the Earth. We can use the result above to estimate how far a neutrino can propagate in the Earth before it actually interacts with a proton or neutron in the Earth.

Obviously the reaction rate is proportional to both the cross-section for the reaction per nucleon (as estimated above) and the density of nucleons i.e. the density of the Earth, ρ. The greater the reaction rate, the shorter the distance a neutrino can propagate before interacting. Thus, we have, the average propagation distance L before an interaction takes place is:

$$L \propto \frac{1}{\rho\sigma} \tag{103}$$

where σ is calculated above for neutrinos interacting with nucleons and ρ is the mass per unit volume of the matter through which the neutrino propagates.

Now we use dimensional analysis. The length L has dimensions of $[L] = [M]^{-1}$. So, the RHS has to have the same dimensions. This will help us to fix the proportionality constant in the above. ρ has dimensions of $[M]^4$ and $[\sigma] = [M]^{-2}$. Thus, if

$$L = \frac{C}{\rho\sigma} \tag{104}$$

the dimension of C is $[C] = [M]$. Therefore we are looking for a quantity which plays an important role in the interaction between a neutrino and a nucleon with the dimensions of mass. The obvious candidate is the nucleon mass, $m_N \sim 1$ GeV.

We therefore find:

$$L = \frac{m_N}{\rho\sigma} \tag{105}$$

Notice that $\frac{\rho}{m_N}$ is essentially the number of atoms per unit volume in the Earth. Let us call this N. Hence we see that

$$L = \frac{1}{N\sigma} \tag{106}$$

If we take $\rho \sim 10^3 kg/m^3 \sim 10^{30}$ GeV$/m^3$. Since $m_N \sim$GeV we have that $N \sim 10^{30}/m^3$. We have calculated σ above. For a neutrino with energy of order 1 GeV

$$\sigma \sim 10^{-38} cm^2 = 10^{-42} m^2 \tag{107}$$

Hence, neutrinos with energies of order a GeV propagate roughly $10^{12} m$ through water before interacting! This is a billion kilometres. Most of the neutrinos from the sun have energies which are one hundred or more times less than a GeV and, hence they propagate much further.

3. $\sigma(pp \to X)$

The next cross-section we estimate is the cross-section for two hadrons (e.g. two protons) to interact. This is relevant for hadron colliders such as the CERN LHC. This cross-section is different to those above because hadrons are not point particles. Rather, they are bound states of quarks, anti-quarks and gluons, bound together by the strong nuclear force. The strong nuclear force is the $SU(3)$ part of the Standard Model. The remarkable thing about the strong nuclear force is that all hadrons have masses which are of order a GeV. In fact, most of the particles described in the PDG are hadrons (either mesons or baryons). If you look at their masses, they are all within one order of magnitude of the proton mass. This reflects the fact that the strong nuclear force is characterised by a scale $\Lambda \sim$ GeV. This is known as the QCD scale since the underlying theoretical description of the strong nuclear interaction is called Quantum Chromodynamics.

Λ is essentially the binding energy of the quarks, anti-quarks and gluons inside any hadron. Since the u d and s quarks have masses which are much smaller than Λ, the masses of hadrons made of these quarks are mostly binding energy. Therefore your mass, and the masses of all the stars in the Universe is binding energy of the strong nuclear force. The b and c quarks have masses of order Λ itself, so c and b hadrons have masses which are not just binding energy. The t quark, which is the most massive known elementary particle ($m_t \sim 173$GeV $\pm$ 1 GeV) actually decays before it has time to "hadronise" and form a hadron. This is because $\tau_t = \frac{1}{\Gamma_t} < \frac{1}{\Lambda}$.

Exercise: Estimate the decay length of a b-hadron. Use the formula for muon decay. Compare it to some of the b-hadron lifetimes in the PDG.

We want to calculate the cross-section for scattering two hadrons which interact via the strong nuclear interaction. We have just seen that everything about the strong nuclear force is characterised by a single scale Λ. Hence, we expect that the effective cross-section for strong nuclear interactions is also determined by Λ. Hence,

$$\sigma(pp \to X) \sim \frac{1}{\Lambda^2} \tag{108}$$

It is Λ^{-2} on dimensional grounds. This has the dimensions of a cross-section. Since GeV$^{-1} \sim 10^{-15}$m,

$$\sigma(pp \to X) \sim 10^{-30} m^2 = 10^{-26} cm^2 \tag{109}$$

Now, the in 2012, the LHC was running with an instantaneous luminosity of about $\mathcal{L} = 10^{33} cm^{-2} s^{-1}$ at a centre of mass energy of 8 TeV.

Hence, with our rough estimate, we expect $\mathcal{L}\sigma \sim 10^7$ events every second! Actually, our estimate is around a factor of 10 smaller than the actual answer so we are producing even more collisions than that.

The greater the number of LHC collisions, the greater the probability of creating a "rare" event such as the production of a Higgs boson. The cross-section for producing a Higgs boson with a mass of around 126 GeV at the LHC is about $10^{-35}cm^2$. This means that we have to "sift through" around a billion events for every Higgs boson produced. The search for the Higgs is thus very much like looking for a needle in a haystack.

Exercise: The Higgs boson mass is approximately 126GeV. How many Higgs bosons were produced in the 2012 run of the LHC? For this you need to find out how much data was recorded i.e. the total integrated luminosity.

3 Symmetries

3.1 Symmetries Commute with the Hamiltonian

Consider the Schrodinger equation

$$i\frac{d\Psi}{dt} = H\Psi \tag{110}$$

Suppose there is an Hermitian operator, K, with expectation value $\langle K \rangle$.

Q: when is $\langle K \rangle$ conserved?

By this we mean, when is $\langle K \rangle$ a constant of motion i.e.

$$\frac{d}{dt}\langle K \rangle = 0 \tag{111}$$

This implies

$$0 = \frac{d}{dt}\langle K \rangle = \frac{d}{dt}\int \Psi^* K \Psi d^3x \tag{112}$$

Hence

$$\int \frac{d\Psi^*}{dt} K\Psi d^3x + \int \Psi^* K \frac{d\Psi}{dt} d^3x = 0 \qquad (113)$$

Since

$$-i\frac{d\Psi^*}{dt} = (H\Psi)^* = \Psi^* H \qquad (114)$$

we have that

$$-\int \Psi^* HK\Psi d^3x + \int \Psi^* KH\Psi d^3x = 0 \qquad (115)$$

Therefore:

$$KH - HK = 0 \equiv [K, H] \qquad (116)$$

i.e. the operator K commutes with the Hamiltonian.
This implies that eigenstates of K are also eigenstates of H.

$$H\Psi = E\Psi \qquad (117)$$
$$K\Psi = k\Psi \qquad (118)$$

This implies that the states transformed into each other by K have the same energy:

$$H(K\Psi) = E(k\Psi) \qquad (119)$$

ie Ψ and $(K\Psi)$ are degenerate in energy.

3.2 Lagrangians and Equations of Motion

In classical mechanics one considers generalised coordinates $q_i(t)$ of a particle. Then the Lagrangian

$$L = T - V \qquad (120)$$

which is the difference between Kinetic and Potential energy leads to the Euler-Lagrange equations of motion

$$\frac{d}{dt}\left(\frac{dL}{d\dot{q}_i}\right) - \frac{dL}{dq_i} = 0 \tag{121}$$

We can use this formalism to obtain the relativistic wave equations such as the Klein-Gordon equation, the Maxwell equations and the Dirac equation.

Instead of considering L to be a function of discrete coordinates q_i, we consider Lagrangians which are functions of the fields which are continuous functions of both x_i and t i.e. of x_μ.

For example, for the Klein-Gordon equation L is a function of $\phi(x_\mu)$ as well as the derivatives $\frac{\partial\phi}{\partial x_\mu} \equiv \partial_\mu\phi$:

$$L(q_i, \dot{q}_i, t) \to L(\phi, \partial_\mu\phi, x_\mu) \tag{122}$$

L is obtained from a Lagrangian density $\mathcal{L}$ integrated over space

$$L = \int d^3x \mathcal{L}(\phi, \partial_\mu\phi) \tag{123}$$

Integrating over time gives the action, usually called S:

$$S = \int dt L = \int d^4x \mathcal{L} \tag{124}$$

By varying S wrt ϕ and $\partial_\mu\phi$ and $\partial_\mu\phi^*$ we obtain the Euler-Lagrange equations (this is derived at the end of the notes in the section on Noethers theorem):

$$\partial_\mu\left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\right) - \frac{\delta\mathcal{L}}{\delta\phi} = 0 \tag{125}$$

The Lagrangian density for the KG equation is

$$\mathcal{L} = \partial_\mu\phi^*\partial^\mu\phi - m^2\phi^*\phi \tag{126}$$

Substituting this into the Euler-Lagrange equations gives

$$\partial_\mu\partial^\mu\phi + m^2\phi = 0 \tag{127}$$

Note:

$$\left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)}\right) = \partial^\mu\phi^* \qquad (128)$$

The Lagrangian density for Maxwells equations in vacuum is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \qquad (129)$$

Here we consider $\mathcal{L}$ as a function(al) of fields A_μ and derivatives $\partial_\mu A_\nu$. ie since A_μ has four components, we treat each component as a separate field.

In the presence of a current j_μ there is an additional interaction term

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - j^\mu A_\mu \qquad (130)$$

3.3 Noether's Theorem

Consider a small transformation in a field which leaves the Lagrangian invariant

$$\Psi \to \Psi + i\alpha\Psi \qquad (131)$$

$$0 = \delta\mathcal{L} = \frac{\delta\mathcal{L}}{\delta\Psi}\delta\Psi + \left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\Psi)}\right)\delta(\partial_\mu\Psi) + c.c. \qquad (132)$$

so

$$\begin{aligned} 0 &= i\alpha\Psi\frac{\delta\mathcal{L}}{\delta\Psi} + i\alpha\left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\Psi)}\right)(\partial_\mu\Psi) + .. \qquad (133) \\ &= i\alpha\left[\frac{\delta\mathcal{L}}{\delta\Psi} - \partial_\mu\left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\Psi)}\right)\right]\Psi + i\alpha\partial_\mu\left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\Psi)}\Psi\right) + ... \end{aligned}$$

where, to get to the last line from the previous one we use that:

$$\partial_\mu\left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\Psi)}\Psi\right) = \left(\partial_\mu\frac{\delta\mathcal{L}}{\delta(\partial_\mu\Psi)}\right)\Psi + \left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\Psi)}\right)\partial_\mu\Psi \qquad (134)$$

Going back to the previous expression, the equation before the one above, there are several key points:

1. The last term is proportional to a total derivative. Hence, it only contributes to the action at the boundary of space-time ie at infinity. Requiring this term to vanish at infinity implies that: the action is extremised ($\delta S = 0$) exactly when the Euler-Lagrange equations are satisfied (the term in square brackets) . We have thus derived the Euler-Lagrange equations.

2. In the case that we require that $\Psi \rightarrow \Psi + i\alpha\Psi$ is a symmetry of the action, then, because the terms in square brackets vanish due to the Euler-Lagrange equations, the total derivative term must vanish everywhere not just at infinity. Hence,

3. When a variation of the fields is a SYMMETRY of the action, there exists a conserved quantity:

4. This conserved quantity is a Lorentz 4-vector, $\left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\Psi)}\Psi\right)$ which obeys the equation

$$\partial_\mu \left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\Psi)}\Psi\right) = 0 \tag{135}$$

5. $\left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\Psi)}\Psi\right)$ is identified with the current j^μ.

Any constant times j^μ is also conserved and we put in the charge to identify it with the current we discussed earlier in the course.

$$j^\mu = \frac{ie}{2}\left(\frac{\delta\mathcal{L}}{\delta(\partial_\mu\Psi)}\Psi\right) \tag{136}$$

Exercise: Verify that this gives the same expression as we had for the current in the Klein-Gordon case.

<u>The Lagrangian for Scalar QED ie a charged KG field</u>

Recall (chapter 4 of book) that in order to consider the motion of a particle of charge $-e$ in an electromagnetic field generated by a vector potential A^μ we replace the derivative ∂_μ by

$$\partial_\mu \rightarrow \partial_\mu - ieA_\mu \tag{137}$$

We call the rhs of this expression a covariant derivative. This is usually denoted by D_μ:

$$D_\mu \equiv \partial_\mu - ieA_\mu \tag{138}$$

Recall that, by making the above replacement in the Klein-Gordon equation we obtained

$$\partial_\mu \partial^\mu \phi + m^2 \phi = -V\phi \tag{139}$$

where

$$V = -ie(\partial_\mu A^\mu - A^\mu \partial_\mu) - e^2 A_\mu A^\mu \tag{140}$$

i.e. we get a potential V for the field ϕ. Since the modified equation of motion was obtained by replacing ∂_μ with D_μ, the Lagrangian density for a charged scalar is

$$\begin{aligned} \mathcal{L} &= D_\mu \phi (D^\mu \phi)^* - m^2 \phi \phi^* && (141) \\ &= \partial_\mu \phi \partial^\mu \phi^* - ieA_\mu \phi \partial^\mu \phi^* + ie\partial_\mu \phi A^\mu \phi^* - e^2 A_\mu A^\mu \phi \phi^* - m^2 \phi \phi^* \end{aligned}$$

Notice that the second and third terms in the last expression combine to give

$$\mathcal{L} = \partial_\mu \phi \partial^\mu \phi^* - j^\mu A_\mu - e^2 A_\mu A^\mu \phi \phi^* - m^2 \phi \phi^* \tag{142}$$

ie we have written them in terms of the conserved current. Thus, combining this Lagrangian with that of Maxwell's theory we have the full Lagrangian for scalar QED:

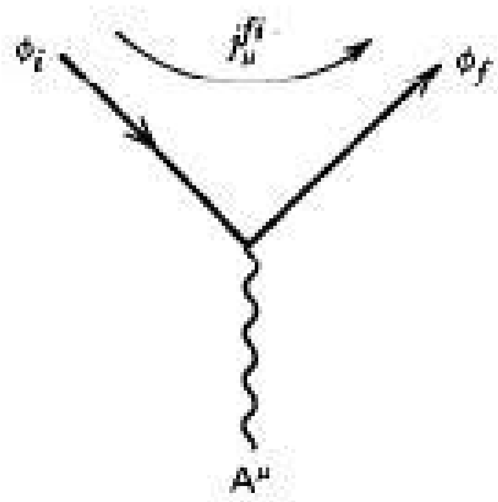
$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + D_\mu \phi (D^\mu \phi)^* - m^2 \phi \phi^* \tag{143}$$

Feynman Diagrams and the Lagrangian

When we studied scalar QED before introducing it in Lagrangian form, we saw that $j^\mu A_\mu$ appears in the invariant amplitude $\mathcal{M}$ for a process.

The fact that $j^\mu A_\mu$ appears in $\mathcal{L}$ suggests that we can just "read off" the vertices allowed in Feynman diagrams from $\mathcal{L}$. This is a general rule for any Lagrangian! We just read off the Feynman rules from $\mathcal{L}$.

In this example, the three point vertex between the photon and two charged particles is represented in $\mathcal{L}$ by the presence of the $j^\mu A_\mu$ term.



Symmetries of Scalar QED

The Lagrangian for scalar QED has various symmetries.

Lorentz Invariance. Since all the Lorentz indices are contracted ($\mathcal{L}$ is a scalar), the Lagrangian is invariant under Lorentz transformations.

Internal Symmetry. In addition to this "spacetime symmetry" it is invariant under an internal symmetry i.e. one which does not act on the coordinates, but just on the fields. This is intrinsic to electromagnetism and the other forces as we will see.

Gauge Symmetry

We are going to consider a transformation of ϕ by a unitary, 1-by-1 matrix, a $U(1)$ transformation.

Any such matrix U is of the form $U(\alpha) = e^{i\alpha}$. α can take any continuous value between zero and 2π.

Clearly, under

$$\phi \to U\phi \tag{144}$$

we have that

$$D_\mu \phi \to U D_\mu \phi \tag{145}$$

The $D_\mu \phi (D^\mu \phi)^*$ term is clearly invariant under this transfomation since it is of the form $D_\mu \phi$ times its complex conjugate, and $UU^* = 1$. Similarly $\phi\phi^*$ is invariant, so the mass term is invariant. Therefore $\mathcal{L}$ is invariant under this transformation of ϕ.

Noethers theorem proves that there is a conserved quantity when a Lagrangian is invariant under a symmetry transformation. In fact, when α is small i.e. when $U \approx 1 + i\alpha$ we see that the conserved current j^μ is precisely that which we derived before.

Now, we would like to consider the case that α is different from point to point in spacetime. i.e. we make $\alpha = \alpha(x_\nu)$ – a function of the coordinates. Clearly this will change the above conclusions because we will get terms proportional to derivatives of α.

$$\begin{aligned} D_\mu \phi &\to U\partial_\mu \phi + iU\partial_\mu \alpha \phi - ieUA_\mu \phi & (146) \\ &= UD_\mu \phi + iU\partial_\mu \alpha \phi & (147) \end{aligned}$$

Thus, because of the term proportional to the derivative of α the Lagrangian is no longer invariant.

However, the unwanted term in the transformation of $D_\mu \phi$ can be removed if A_μ also transforms:

$$A_\mu \to A_\mu + \frac{1}{e}\partial_\mu \alpha \tag{148}$$

which can be verified by replacing this transformed A_μ in the "unwanted" term.

Therefore, we have that

$$D_\mu \phi \to e^{i\alpha(x)} D_\mu \phi \tag{149}$$

and the Klein-Gordon terms in the Lagrangian are invariant under this gauge transformation (this is the name for transformations whose parameters are functions of the coordinates.

What about the Maxwell term in the Lagrangian?

Let us consider the electromagnetic field strength, $F_{\mu\nu}$.

Since

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{150}$$

the field strength transforms into

$$\begin{aligned} F_{\mu\nu} &\to \partial_\mu A_\nu - \frac{1}{e}\partial_\mu \partial_\nu \alpha - \partial_\nu A_\mu + \frac{1}{e}\partial_\nu \partial_\mu \alpha & (151) \\ &= \partial_\mu A_\nu - \partial_\nu A_\mu = F_{\mu\nu} & (152) \end{aligned}$$

Therefore Maxwell's Lagrangian also gauge invariant!

Gauge Symmetry as a Principle

If we use gauge symmetry as a principle then it has far reaching consequences.

1. The covariant derivative must be introduced otherwise the kinetic energy term would not be gauge invariant.

2. This requires the introduction of a vector field A_μ which couples to the matter current. A_μ is usually called the gauge field.

3. If we consider the kinetic energy of the gauge field, then gauge invariance requires it to be of the form $F_{\mu\nu}F^{\mu\nu}$ (or more generally a function thereof). Thus, Maxwell's equations follow from gauge symmetry plus Lorentz symmetry

4. The photon is massless

This last point is crucial. It provides an explanation for why the photon essentially behaves as a massless particle. (Experimentally of course one cannot prove that the photon is exactly massless. Rather, one obtains an upper limit on its mass. The current upper limit is about 10^{-18} eV.)

To see why the photon is massless, we ask: what would a mass term look like?. Well, in analogy with the mass term of the KG equation it would be of the form:

$$\partial_\nu \partial^\nu A_\mu + m^2 A_\mu = 0 \tag{153}$$

This mass term would arise from a term in the Lagrangian of the form

$$\Delta\mathcal{L} \sim -m^2 A_\mu A^\mu \tag{154}$$

Such a term is clearly not invariant under the gauge transformation of A_μ, which is

$$A_\mu \to A_\mu + \frac{1}{e}\partial_\mu \alpha \tag{155}$$

In fact, combining all of these points, the most general gauge and Lorentz invariant Lagrangian which is quadratic in the fields and their derivatives is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + D_\mu\phi(D^\mu\phi)^* - m^2\phi\phi^* \tag{156}$$

Thus: the symmetries determine the Lagrangian and, hence, the physics

This is a key point in particle physics. The Lagrangian for the Standard Model is essentially determined by its symmetries. In other words, symmetries determine the physics of all elementary particles!

Beyond $U(1)$ Gauge Invariance

We would now like to consider generalising $U(1)$ gauge theory (i.e. QED) to $U(N)$ gauge theory.

That is to say that U – the transformation matrix – will become a Unitary $N \times N$ matrix U^i_j where i, j run from 1 to N each.

An $N \times N$ matrix acts naturally on N component vectors v_i. Hence we should introduce N complex scalar fields ϕ_i on which these act.

Under

$$\phi_i \to U_i^j \phi_j \tag{157}$$

we would like to impose a condition that a suitable covariant derivative transforms acting on ϕ_i transforms in the same way

$$(D_\mu \phi)_i \to U_i^j (D_\mu \phi)_j \tag{158}$$

This would be the $N \times N$ generalisation of the $U(1)$ case.

But what is this covariant derivative?

If we try to introduce a gauge field, in general it is a matrix of gauge fields i.e. we can have up to $N \times N$ gauge fields:

$$(D_\mu \phi)_i = \partial_\mu \phi_i - ig(G_\mu)_i^j \phi_j \tag{159}$$

that is, for each of the values of i and j, $(G_\mu)^i_j$ is a different gauge field.

Unitary matrices, exponentials, group generators and all that

In order to understand a little better the structure we would like to look at some of the simple symmetry groups like $SU(2)$ and $SU(3)$. First, though we begin with $U(1)$:

$$\exp i\theta = 1 + i\theta - \frac{\theta^2}{2} - i\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + ... = \cos\theta + i\sin\theta \tag{160}$$

Now consider the rotation matrix

$$R(\theta) \equiv \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{161}$$

We want to Taylor expand this matrix

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 - \frac{\theta^2}{2} + \frac{\theta^4}{4!} + ... & \theta - \frac{\theta^3}{3!} + ... \\ -\theta + \frac{\theta^3}{3!} + ... & 1 - \frac{\theta^2}{2} + \frac{\theta^4}{4!} + ... \end{pmatrix} \tag{162}$$

Just as $e^{i\theta}$ is the exponential of a 1-by-1 matrix, the rotation matrix above is the exponential of a two-by-two matrix:

$$R(\theta) = \exp(i\theta T) = \mathbf{1} + i\theta T - \frac{\theta^2}{2}T^2 - i\frac{\theta^3}{3!}T^3 + ... \tag{163}$$

where

$$T = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{164}$$

and T^2 is the matrix product of T with itself.

So, since any rotation can be written as $\exp i\theta T$ we say that T generates the rotations.

Let us now consider some other examples of this, because we will need matrices like T to define the covariant derivative properly.

Let U be a N-by-N unitary matrix ie

$$U^\dagger . U = \mathbf{1} \tag{165}$$

We will now assume that

$$U = \exp iM \tag{166}$$

What properties does M have?
Since

$$U^\dagger = \exp(-iM^\dagger) \tag{167}$$

unitarity implies that

$$M = M^\dagger \tag{168}$$

So, M is Hermitian.
If, additionally, we require that $\det(U) = 1$ i.e. that U is special unitary, then one can show that the trace of M is zero

$$\det U = 1 \leftrightarrow \mathrm{Tr} M = 0 \tag{169}$$

$SU(2)$

For the case $N = 2$, one can show that M is a linear combination of the Pauli matrices:

$$M = \alpha^a \sigma_a \tag{170}$$

where

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{171}$$

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{172}$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{173}$$

That is to say that the Pauli matrices generate $SU(2)$ transformation matrices!
An important fact about the Pauli matrices is that they obey an algebra:

$$[\sigma_a, \sigma_b] = 2i\epsilon_{ab}{}^c \sigma_c \tag{174}$$

A note on ϵ_{abc}

ϵ_{abc} is the same as $\epsilon_{ab}{}^{c}$ – we put the third index up to remind us that we are summing over c. ϵ_{abc} is totally antisymmetric i.e. if we interchange neighbouring indices then we get a minus sign:

$$\epsilon_{abc} = -\epsilon_{bac} = \epsilon_{bca} = -\epsilon_{cba} \;\; etc \tag{175}$$

This implies that all three indices of ϵ_{abc} must take different values in order to get a non-zero result otherwise it would not be totally antisymmetric.

Therefore ϵ_{abc} is non-zero if and only if (abc) is a permutation of (123). Finally,

$$\epsilon_{123} = 1 \tag{176}$$

For all of the other five permutations of (123) the value of ϵ_{abc} can be obtained from antisymmetry. Thus, for example $\epsilon_{132} = -1$.

Back to $SU(N)$.

In general, for an $SU(N)$ matrix

$$U = \exp iM \tag{177}$$

where M is traceless and Hermitian, there are $N^2 - 1$ generators T_a such that

$$M = \alpha^a T_a \tag{178}$$

and the T_a 's obey an algebra

$$[T_a, T_b] = i f_{ab}{}^{c} T_c \tag{179}$$

where the $f_{ab}{}^{c}$ are constants called the structure constants. For $SU(N)$ the algebra defined by the above equation is called the Lie Algebra of $SU(N)$. If you choose a basis for this algebra, you can explicitly calculate the structure constants in that basis. They are totally antisymmetric, like ϵ_{abc}. For $SU(3)$ there is a basis for the eight 3-by-3 T_a matrices which is used a lot in particle physics called the Gell-Mann basis. You can find these eight matrices in the book.

Finally

$$D_\mu = \partial_\mu + igT_a G^a_\mu \tag{180}$$

where the second term is an N-by-N matrix since T_a is a matrix. There are $N^2 - 1$ gauge fields G^a_μ.

The Standard Model is a gauge theory with $SU(3)$ gauge symmetry, $SU(2)$ gauge symmetry and $U(1)$ gauge symmetry. There are $8+3+1 = 12$ gauge bosons. These are the eight gluons, the two W-bosons (W^+ and W^-), the neutral Z boson and the photon.

The full covariant derivative is thus

$$D_\mu = \partial_\mu - i\frac{Y}{2}g_1 B_\mu - ig_2\frac{\sigma_j}{2}W^j_\mu - ig_3\frac{\lambda_a}{2}G^a_\mu \tag{181}$$

B_μ is the gauge boson of the $U(1)$. The photon is a linear combination of B_μ and W^3_μ. The Z-boson is the opposite linear combination.

Y is called the hypercharge. The charge under $SU(2)$ is called isospin I. The proton has isospin 1/2 and the neutron $-1/2$.

The proton and neutron transform as a doublet under the $SU(2)$ of the Standard Model:

$$\begin{pmatrix} p \\ n \end{pmatrix} \to \exp(iM)\begin{pmatrix} p \\ n \end{pmatrix} \tag{182}$$

Electric charge is a linear combination of Y and I.

$$Q = I + Y/2 \tag{183}$$

Both p and n have to have the same Y which is one.

Similarly, in the Standard Model u and d quarks transform under $SU(2)$

$$\begin{pmatrix} u \\ d \end{pmatrix} \to \exp(iM)\begin{pmatrix} u \\ d \end{pmatrix} \tag{184}$$

Note: this is not strictly speaking correct: fermions can be left-handed or right-handed as we will see when we study the Dirac equation. This is related to the fact they are not scalars, but fermions i.e. it is related to their spin under rotations. The correct statement is that the left-handed up and down quarks transform under $SU(2)$ as above. In fact, right handed fermions do NOT transform atall under the $SU(2)$ i.e. for them I=0. Thus, when the covariant derivative acts on right-handed fermions, the term proportional to the B_μ boson is zero.

Hence left-handed u-quarks have $I = 1/2, Y = 1/6$ whereas right-handed u-quarks have $I = 0, Y = 4/3$. Both of these charges give $Q = 2/3$. Similarly, e_L^- the left-handed electron has $I = -1/2, Y = -1$ whereas e_R^- has $Y = -2$. Left-handed electron-neutrinos are $SU(2)$ partners of left-handed electrons, hence they have $I = 1/2, Y = -1$.

No right-handed neutrinos have been directly observed to exist (yet). The reason is the following. If they did exist, they would not transform under $SU(2)$ like all of the other right-handed neutrinos, hence have both $I = 0$ and $Y = 0$. Therefore, such neutrinos would not couple to B_μ or the $SU(2)$ gauge fields W_μ^i. Since neutrinos do not participate in the strong interactions, the $SU(3)$ part of the covariant derivative would also not act on the right-handed neutrino field. Therefore, the right-handed neutrino does not feel any force directly: its equation of motion does not include a gauge field. Such neutrinos are also called sterile neutrinos, for reasons which should hopefully be clear. Since they do not couple to gauge bosons it is very difficult to produce or detect them in a laboratory.

Beyond U(1) Gauge Invariance

We would now like to consider generalising $U(1)$ gauge theory (i.e. QED) to $U(N)$ gauge theory.

That is to say that U – the transformation matrix – will become a Unitary $N \times N$ matrix U_j^i where i, j run from 1 to N each.

An $N \times N$ matrix acts naturally on N component vectors v_i. Hence we should introduce N complex scalar fields ϕ_i on which these act.

Under

$$\phi_i \to U_i^j \phi_j \tag{185}$$

we would like to impose a condition that a suitable covariant derivative transforms acting on ϕ_i transforms in the same way

$$(D_\mu\phi)_i \to U_i^j (D_\mu\phi)_j \tag{186}$$

This would be the $N \times N$ generalisation of the $U(1)$ case.

But what is this covariant derivative?

If we try to introduce a gauge field, in general it is a matrix of gauge fields i.e. we can have up to $N \times N$ gauge fields:

$$(D_\mu\phi)_i = \partial_\mu\phi_i - ig(G_\mu)_i^j\phi_j \tag{187}$$

that is, for each of the values of i and j, $(G_\mu)_j^i$ is a different gauge field.

Unitary matrices, exponentials, group generators and all that

In order to understand a little better the structure we would like to look at some of the simple symmetry groups like $SU(2)$ and $SU(3)$. First, though we begin with $U(1)$:

$$\exp i\theta = 1 + i\theta - \frac{\theta^2}{2} - i\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + ... = \cos\theta + i\sin\theta \tag{188}$$

Now consider the rotation matrix

$$R(\theta) \equiv \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{189}$$

We want to Taylor expand this matrix

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 - \frac{\theta^2}{2} + \frac{\theta^4}{4!} + ... & \theta - \frac{\theta^3}{3!} + ... \\ -\theta + \frac{\theta^3}{3!} + ... & 1 - \frac{\theta^2}{2} + \frac{\theta^4}{4!} + ... \end{pmatrix} \tag{190}$$

Just as $e^{i\theta}$ is the exponential of a 1-by-1 matrix, the rotation matrix above is the exponential of a two-by-two matrix:

$$R(\theta) = \exp(i\theta T) = \mathbf{1} + i\theta T - \frac{\theta^2}{2}T^2 - i\frac{\theta^3}{3!}T^3 + ... \tag{191}$$

where

$$T = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{192}$$

and T^2 is the matrix product of T with itself.

So, since any rotation can be written as $\exp i\theta T$ we say that T generates the rotations.

Let us now consider some other examples of this, because we will need matrices like T to define the covariant derivative properly.

Let U be a N-by-N unitary matrix ie

$$U^\dagger . U = \mathbf{1} \tag{193}$$

We will now assume that

$$U = \exp iM \tag{194}$$

What properties does M have?
Since

$$U^\dagger = \exp(-iM^\dagger) \tag{195}$$

unitarity implies that

$$M = M^\dagger \tag{196}$$

So, M is Hermitian.
If, additionally, we require that $\det(U) = 1$ i.e. that U is special unitary, then one can show that the trace of M is zero

$$\det U = 1 \leftrightarrow \mathrm{Tr} M = 0 \tag{197}$$

SU(2)

For the case $N = 2$, one can show that M is a linear combination of the Pauli matrices:

$$M = \alpha^a \sigma_a \tag{198}$$

where

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{199}$$

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{200}$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{201}$$

That is to say that the Pauli matrices **generate** $SU(2)$ transformation matrices!
An important fact about the Pauli matrices is that they obey an algebra:

$$[\sigma_a, \sigma_b] = 2i\epsilon_{ab}{}^c \sigma_c \tag{202}$$

A note on $\epsilon_{\mathbf{abc}}$

ϵ_{abc} is the same as $\epsilon_{ab}{}^{c}$ – we put the third index up to remind us that we are summing over c. ϵ_{abc} is totally antisymmetric i.e. if we interchange neighbouring indices then we get a minus sign:

$$\epsilon_{abc} = -\epsilon_{bac} = \epsilon_{bca} = -\epsilon_{cba} \;\; etc \tag{203}$$

This implies that all three indices of ϵ_{abc} must take different values in order to get a non-zero result otherwise it would not be totally antisymmetric.

Therefore ϵ_{abc} is non-zero if and only if (abc) is a permutation of (123). Finally,

$$\epsilon_{123} = 1 \tag{204}$$

For all of the other five permutations of (123) the value of ϵ_{abc} can be obtained from antisymmetry. Thus, for example $\epsilon_{132} = -1$.

Back to SU(N).

In general, for an $SU(N)$ matrix

$$U = \exp iM \tag{205}$$

where M is traceless and Hermitian, there are $N^2 - 1$ generators T_a such that

$$M = \alpha^a T_a \tag{206}$$

and the T_a 's obey an algebra

$$[T_a, T_b] = i f_{ab}{}^{c} T_c \tag{207}$$

where the $f_{ab}{}^{c}$ are constants called the structure constants. For $SU(N)$ the algebra defined by the above equation is called the Lie Algebra of $SU(N)$. If you choose a basis for this algebra, you can explicitly calculate the structure constants in that basis. They are totally antisymmetric, like ϵ_{abc}. For $SU(3)$ there is a basis for the eight 3-by-3 T_a matrices which is used a lot in particle physics called the Gell-Mann basis. You can find these eight matrices in the book.

Finally

$$D_\mu = \partial_\mu + igT_aG^a_\mu \tag{208}$$

where the second term is an N-by-N matrix since T_a is a matrix. There are $N^2 - 1$ gauge fields G^a_μ.

The Standard Model is a gauge theory with $SU(3)$ gauge symmetry, $SU(2)$ gauge symmetry and $U(1)$ gauge symmetry. There are $8+3+1 = 12$ gauge bosons. These are the eight gluons, the two W-bosons (W^+ and W^-), the neutral Z boson and the photon.

The full covariant derivative is thus

$$D_\mu = \partial_\mu - i\frac{Y}{2}g_1B_\mu - ig_2\frac{\sigma_j}{2}W^j_\mu - ig_3\frac{\lambda_a}{2}G^a_\mu \tag{209}$$

B_μ is the gauge boson of the $U(1)$. The photon is a linear combination of B_μ and W^3_μ. The Z-boson is the opposite linear combination.

Y is called the hypercharge. The charge under $SU(2)$ is called isospin I. The proton has isospin 1/2 and the neutron $-1/2$.

The proton and neutron transform as a doublet under the $SU(2)$ of the Standard Model:

$$\begin{pmatrix} p \\ n \end{pmatrix} \to \exp(iM) \begin{pmatrix} p \\ n \end{pmatrix} \tag{210}$$

Electric charge is a linear combination of Y and I.

$$Q = I + Y/2 \tag{211}$$

Both p and n have to have the same Y which is one.

Similarly, in the Standard Model u and d quarks transform under $SU(2)$

$$\begin{pmatrix} u \\ d \end{pmatrix} \to \exp(iM) \begin{pmatrix} u \\ d \end{pmatrix} \tag{212}$$

Note: this is not strictly speaking correct: fermions can be left-handed or right-handed as we will see when we study the Dirac equation. This is related to the fact they are not scalars, but fermions i.e. it is related to their spin under rotations. The correct statement is that the left-handed up and down quarks transform under $SU(2)$ as above. In fact, right handed fermions do NOT transform atall under the $SU(2)$ i.e. for them I=0. Thus, when the covariant derivative acts on right-handed fermions, the terms proportional to the W_μ bosons is zero.

Hence left-handed u-quarks have $I = 1/2, Y = 1/6$ whereas right-handed u-quarks have $I = 0, Y = 4/3$. Both of these charges give $Q = 2/3$. Similarly, e_L^- the left-handed electron has $I = -1/2, Y = -1$ whereas e_R^- has $Y = -2$. Left-handed electron-neutrinos are $SU(2)$ partners of left-handed electrons, hence they have $I = 1/2, Y = -1$.

No right-handed neutrinos have been directly observed to exist (yet). The reason is the following. If they did exist, they would not transform under $SU(2)$ like all of the other right-handed neutrinos, hence have both $I = 0$ and $Y = 0$. Therefore, such neutrinos would not couple to B_μ or the $SU(2)$ gauge fields W_μ^i. Since neutrinos do not participate in the strong interactions, the $SU(3)$ part of the covariant derivative would also not act on the right-handed neutrino field. Therefore, the right-handed neutrino does not feel any force directly: its equation of motion does not include a gauge field. Such neutrinos are also called sterile neutrinos, for reasons which should hopefully be clear. Since they do not couple to gauge bosons it is very difficult to produce or detect them in a laboratory.

Spin, Fermions and the Dirac Equation

Spin

In quantum mechanics, particles have an intrinsic spin under spatial rotations of space. This is analagous to the rotation of "classical objects" such as planets or spiniing tops, but the spin of elementary particles is quantised (comes in discrete amounts). In natural units the allowed spins are half-integer or integer. Bosons have integer spins; fermions half-integer spins.

T*he Algebra of Rotations*

We saw in the last lecture that rotations in a plane by an angle θ are generated by a matrix:

$$R(\theta) = \exp(i\theta T) = \mathbf{1} + i\theta T - \frac{\theta^2}{2}T^2 - i\frac{\theta^3}{3!}T^3 + ... \tag{213}$$

where

$$T = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{214}$$

We can use this observation to generate all rotations in three dimensional space. These are given by 3-by-3 matrices. Rotations in the (x, y) plane around the z-axis are of the form

$$R_z(\theta_z) = \begin{pmatrix} \cos\theta_z & \sin\theta_z & 0 \\ -\sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{215}$$

Show that these are generated by

$$T_3 \equiv \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{216}$$

Similarly, a rotation around the x-axis

$$R_x(\theta_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & \sin\theta_x \\ 0 & -\sin\theta_x & \cos\theta_x \end{pmatrix} \tag{217}$$

is generated by

$$T_1 \equiv \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \tag{218}$$

and rotations around the y-axis

$$R_y(\theta_y) = \begin{pmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{pmatrix} \tag{219}$$

are generated by

$$T_2 \equiv \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \tag{220}$$

Thus, any rotation in three dimensional space can be generated by a linear combination of T_1, T_2, T_3. The generators T_a obey an algebra:

$$[T_a, T_b] = i\epsilon_{ab}{}^c T_c \tag{221}$$

We have seen this algebra before. The Pauli matrices satisfy

$$[\sigma_a, \sigma_b] = 2i\epsilon_{ab}{}^c \sigma_c \tag{222}$$

Hence,

$$[\frac{1}{2}\sigma_a, \frac{1}{2}\sigma_b] = i\epsilon_{ab}{}^c \frac{1}{2}\sigma_c \tag{223}$$

Therefore the T_a and $\frac{1}{2}\sigma_a$ obey precisely the same algebra, even though the T's are 3-by-3 matrices and the Pauli matrices are 2-by-2.

We have shown that the algebra of rotations is the same as that of SU(2)

Every 3-by-3 rotation matrix can be specified by choosing $(\theta_x, \theta_y, \theta_z)$. From these we can specify a 2-by-2 $SU(2)$ matrix. This gives a correspondence between rotations and elements of $SU(2)$. For example, $R_x(\theta_x)$ corresponds to the matrix

$$M_x(\theta_x) = \exp\left(\frac{i\theta_x}{2}\sigma_1\right) \tag{224}$$

The 3-by-3 rotation matrices act on 3-component vectors, such as the spatial coordinates or the momentum of a particle. 3-by-3 rotations obey

$$R(\theta_x,\theta_y,\theta_z) = R((\theta_x+2\pi,\theta_y,\theta_z) = R(\theta_x,\theta_y+2\pi,\theta_z) = R(\theta_x,\theta_y,\theta_z+2\pi) \tag{225}$$

So a vector returns to itself after a full 2π rotation.

However,

$$M_x(\theta_x+2\pi) = -M_x(\theta_x) \tag{226}$$

Mathematically, 2-component objects which transform under **spatial** rotations via $SU(2)$ matrices are called spinors. Fermion wave functions are represented by spinors. Fermions are said to have spin "one-half": they come back to themselves after two 2π rotations in space. In fact, in the Pauli basis, the entries of $\frac{1}{2}\sigma_3$ give the spin of the two components of the fermion wave function, which are $\pm\frac{1}{2}$.

Spins are Classified

A set of three N-by-N matrices which obey the algebra:

$$[W_a, W_b] = i\epsilon_{ab}{}^{c}W_c \tag{227}$$

is called a representation of the Lie algebra of $SU(2)$ (or the algebra of rotations) with dimension N. The Pauli matrices are a 2-dimensional representation, the T's a 3-dimensional one.

All possible sets of matrices obeying the algebra can be completely classified. In fact, there are representations for any integer N. The N components of the vector on which the W's act have spins which are $(N/2-1/2, N/2-3/2, N/2-5/2...., -N/2+1/2)$. The highest spin of a representation i.e. $N/2-1/2$ is also used to label the representation and is usually called the "spin of the representation". Any positive integer or half-integer is the highest spin of a set of W's obeying the algebra above.

So, representations with odd N have spins which are all integers. Those with N even, have all the spins strictly taking half-integer values. The odd N representations are said to be bosonic whilst the half-integer spin representations are called fermionic.

Scalar fields transform in the spin zero representation, $N = 1$. Quarks and Leptons in the spin $1/2$ representation, $N = 2$ and gauge fields, being vector fields, transform in the spin 1 representation, $N = 3$

Notice that in the 2-by-2 representation given by the Pauli matrices, one of the generators $(\frac{1}{2}\sigma_3)$ is diagonal. It's eigenvalues are the spins of the representation, in this case $\frac{1}{2}$ and $-\frac{1}{2}$. In the 3-by-3 representation given above, all three generators are off-diagonal, however their eigenvalues are $(1, 0, -1)$ so the spins of the 3-dimensional representation are 1,0 and -1.

Since the rotations are part of the Lorentz symmetry, **spin is conserved** in physical processes.

So, when a W-boson of spin one decays into two particles A and B, the only possibilities are that A and B have spins 0 and 1, 1 and 0, or 1/2 and 1/2. In the Standard Model, only the last possibility is realised because in order to realise the first two, the spin 0 particle would have to be the Higgs (the only scalar field present) and the spin 1 particle would have to have be charged because charge is also conserved. The only charged spin 1 particles in the Standard Model are the W-bosons, which obviously, cannot decay into themselves.

The Dirac Equation

Dirac wanted to write a wave equation consistent with Lorentz invariance that was linear in the time derivative as in the Schrodinger equation. He wanted to do this to remove the negative energy solutions of the Klein-Gordon equation. He succeeded in finding a Lorentz covariant equation linear in $\frac{\partial}{\partial t}$ but the negative energy solutions still remained. We know now (as we discussed) that the negative energy solutions are associated with anti-particles. The Dirac equation describes relativistic fermions. We will derive the Dirac equation for free fermions, then, after exploring it's properties we will write the Dirac equation for fermions interacting with gauge fields. Then we will explore the interactions between fermions and gauge fields and explore some of the consequences.

The starting point of Dirac's argument was the hypothesis that the equation should be of the form:

$$H\Psi = (\alpha_i p_i + \beta m)\Psi \tag{228}$$

where the left hand side is, as usual, linear in $\frac{\partial}{\partial t}$. α_i and β are coefficients to be determined.

The term proportional to mass must be present, since a particle with mass m has energy m when its momentum is zero. The term proportional to the momentum must be present since we are looking for an equation which treats space and time (and hence momentum and energy) on an equal footing. We determine the coefficients by demanding that, for a free particle solution

$$H^2\Psi = (p_i p^i + m^2)\Psi \tag{229}$$

From the first equation we have that

$$\begin{aligned} H^2\Psi &= (\alpha_i p_i + \beta m)(\alpha_j p_j + \beta m)\Psi \\ &= (\alpha_i^2 p_i^2 + (\alpha_i\alpha_j + \alpha_j\alpha_i)p_i p_j + (\alpha_i\beta + \beta\alpha_i)p_i m + \beta^2 m)\Psi \end{aligned} \tag{230}$$

Where the first term on the rhs of the second line are the three terms in the product when $i = j$, the second and third terms are the remaining six possibilities when $i \neq j$.

So, we learn that

- $\alpha_1, \alpha_2, \alpha_3$ and β all anti-commute with each other.
- $\alpha_1^2 = \alpha_2^2 = \alpha_3^2 = \beta^2 = \mathbf{1}$

So, the coefficients cannot be numbers. Dirac realised that if they were matrices then solutions to these conditions indeed exist. With $m \neq 0$, (it can be shown, but we won't do so) the simplest solution has the α_i and β 4-by-4 matrices. Thus, Ψ is now understood to be a vector with four components i.e.

$$\Psi \equiv \Psi_A = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{pmatrix} \tag{231}$$

Ψ is called a Dirac spinor. There are actually different solutions for the matrices, though physical results don't depend on this choice.

One representation for the (α_i, β) is called the Dirac-Pauli representation and is given by

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \quad \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \tag{232}$$

where I is the 2-by-2 identity matrix.

Another useful representation is the Weyl representation:

$$\alpha_i = \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \quad \beta = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \tag{233}$$

Covariant Form of the Dirac Equation

We can multiply the original form of the Dirac equation on the left by β to obtain

$$i\beta\frac{\partial}{\partial t}\Psi = -i\beta\alpha_i\frac{\partial}{\partial x_i}\Psi + m\Psi \tag{234}$$

which can be written as

$$(i\gamma^\mu\partial_\mu - m)\Psi = 0 \tag{235}$$

where we have introduced the so-called Dirac γ-matrices

$$\gamma^\mu \equiv (\beta, \beta\alpha_i) \tag{236}$$

and we have anticipated the fact that the four γ-matrices transform under Lorentz transformations as a Lorentz 4-vector.

In the above form we have suppressed the spinor matrix labels on γ^μ and the spinor labels on Ψ. If we put these back in we have:

$$\sum_{K=1}^{4}\left[\sum_\mu i(\gamma^\mu)_{JK}\partial_\mu - m\delta_{JK}\right]\Psi_K = 0 \tag{237}$$

The original relations on the α_i and β required by relativistic invariance are equivalent to

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu} \tag{238}$$

where $g^{\mu\nu}$ is the Minkowski metric.

Currents

We now proceed as we did with the KG equation to obtain a form of the current.

For this we need the Hermitian conjugate of the Dirac equation:

$$-i\frac{\partial\Psi^\dagger}{\partial t}\gamma^0 - i\frac{\partial\Psi^\dagger}{\partial x_k}(-\gamma^k) - m\Psi^\dagger = 0 \tag{239}$$

This does not look terribly covariant since, whilst γ^0 is Hermitian, the γ^k are anti-Hermitian. However, by defining

$$\bar{\Psi} \equiv \Psi^\dagger\gamma^0 \tag{240}$$

we see that

$$i\partial_\mu\bar{\Psi}\gamma^\mu + m\bar{\Psi} = 0 \tag{241}$$

We now multiply the Dirac equation on the left by $\bar{\Psi}$, the conjugate equation above by Ψ and add the two. The mass terms cancel and we see that

$$\bar{\Psi}\gamma^\mu\partial_\mu\Psi + (\partial_\mu\bar{\Psi})\gamma^\mu\Psi = \partial_\mu(\bar{\Psi}\gamma^\mu\Psi) = 0 \tag{242}$$

so that $\bar{\Psi}\gamma^\mu\Psi$ is conserved.

Putting in the required factor of the charge, we then will have, for a fermion of charge $-e$ the charge current density is

$$j^\mu = -e\bar{\Psi}\gamma^\mu\Psi \tag{243}$$

Solutions of the Dirac Equation

We now turn to solutions of the Dirac equation. It is convenient to introduce the Feynman "slash" notation which is defined as $\not{V} \equiv \gamma^\mu V_\mu$ for any four-vector V_μ.

Thus, the Dirac equation can be written as

$$(i\not{\partial} - m)\Psi = 0 \tag{244}$$

Since (see exercise sheets) the Dirac equation implies the KG equation, we have

$$(\partial_\mu \partial^\mu + m^2)\Psi_A = (\Box + m^2)\Psi_A = 0 \tag{245}$$

and we can write

$$\Psi = u(p_\mu)e^{-ip^\mu x_\mu} \tag{246}$$

where the 4-particle spinor $u(p)$ is independent of x and satisfied the momentum space Dirac equation

$$(\not{p} - m)u = 0 \tag{247}$$

The Hamiltonian H is obtained from Diracs original ansatz

$$Hu = (\alpha_i p^i + \beta m)u = Eu \tag{248}$$

There are actually four independent solutions of this equation. Two have $E > 0$ and two have $E < 0$. To see this, consider fermions at rest $p_i = 0$. Let us use the Dirac-Pauli rep for the γ-matrices. Then

$$Hu = \beta m u = \begin{pmatrix} m & 0 & 0 & 0 \\ 0 & m & 0 & 0 \\ 0 & 0 & -m & 0 \\ 0 & 0 & 0 & -m \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}$$

which obviously has two positive and two negative eigenvalues. The two positive energy solutions are interpreted as describing an $E > 0$ electron, the two $E < 0$ solutions are interpreted as describing an $E > 0$ positron.

For non-zero momentum, the equation becomes

$$Hu = \begin{pmatrix} m & \sigma_i p^i \\ \sigma_i p^i & -m \end{pmatrix} \begin{pmatrix} u_A \\ u_B \end{pmatrix} = E \begin{pmatrix} u_A \\ u_B \end{pmatrix} \tag{249}$$

where

$$u \equiv \begin{pmatrix} u_A \\ u_B \end{pmatrix} \tag{250}$$

we therefore have that

$$\sigma_i p^i u_B = (E - m) u_A \tag{251}$$

$$\sigma_i p^i u_A = (E + m) u_B \tag{252}$$

$$\tag{253}$$

To solve the equations we pick the two positive energy solutions u^1 and u^2 to have

$$u_A^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad u_A^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{254}$$

Then the remaining components of u^1 and u^2, namely u_B^1 and u_B^2 are determined as

$$u_B^s = \frac{\sigma_i p^i}{E + m} u_A^s \tag{255}$$

So, u^1 and u^2, the positive energy solutions of the Dirac equation are given by

$$u^s = N \begin{pmatrix} u_A^s \\ \frac{\sigma_i p^i}{E+m} u_A^s \end{pmatrix} \tag{256}$$

Similarly, to obtain the two negative energy solutions u^3 and u^4 we take the lower components (u_B) of the Dirac spinor to be

$$u_B^3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad u_B^4 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{257}$$

This gives the negative energy solutions as

$$u^{s+2} = N \begin{pmatrix} \frac{-\sigma_i p^i}{E-m} u_A^s \\ u_A^s \end{pmatrix} \tag{258}$$

Helicity

Consider the following 4-by-4 matrix:

$$\Lambda \equiv \frac{1}{2}\begin{pmatrix} \sigma_i \hat{p}^i & 0 \\ 0 & \sigma_i \hat{p}^i \end{pmatrix} \tag{259}$$

where $\hat{p}^i = \frac{p^i}{|p|}$ satisfies $\hat{p}^i \hat{p}_i = 1$ i.e. is the unit vector which points in the same direction as p^i. Λ commutes with the Hamiltonian. Therefore, the eigenvalues of Λ are conserved.

$\frac{1}{2}\sigma_i \hat{p}^i$ is clearly the spin projected in the direction of motion. We call this the **helicity** of the state. The possible eigenvalues of $\frac{1}{2}\sigma_i \hat{p}^i$ are just $\lambda = \pm\frac{1}{2}$.

We say that fermions with helicity $-1/2$ are left-handed and those with helicity $1/2$ are right-handed.

Massless Dirac Equation and Helicity

The Dirac spinor has four components, whereas by considerations of spin we expected two. The four are divided into two sets of two: left-handed and right-handed. Thus the Dirac spinor for a fermion describes both the left- and right-handed helicity states. We can see this a bit more clearly by considering the Dirac equation for $m = 0$ i.e. massless fermions. Then β drops out and the α_i have to satisfy

$$\alpha_i \alpha_j + \alpha_j \alpha_i = 2\delta_{ij} \ \ , \ \ \alpha_i^\dagger = \alpha_i \tag{260}$$

These conditions can be solved by 2-by-2 matrices, $\alpha_i = \sigma_i$. Thus the Dirac equation in the massless case can be satisfied by a 2-component spinor, ϕ and we have:

$$H\phi = \sigma_i p^i \phi \tag{261}$$

In this case, $E^2 = p_i p^i$ so, $|E| = |p|$. Hence we have

$$\pm |p|\phi = \sigma_i p^i \phi \tag{262}$$

Consider now the positive energy solution $E = |p|$, ϕ^+. It has

$$\sigma_i \hat{p}^i \phi^+ = \phi^+ \tag{263}$$

Therefore ϕ^+ is right-handed in this case. The second solution ϕ^- has negative energy $E = -|p|$, so in this case

$$\sigma_i \hat{p}^i \phi^- = -\phi^- \quad (264)$$

When the energy is $-E$ and the momentum $-p_i$, we have

$$\sigma_i (-\hat{p}^i) \phi^- = -\phi^- \quad (265)$$

and describes a left-handed anti-fermion. So, if we apply this to a neutrino in situations where $|E| >> m$, ϕ describes a right-handed neutrino ν_R and a left-handed anti-neutrino $\bar{\nu}_L$.

We could have also chosen $\alpha_i = -\sigma_i$. In this case we would have

$$E\chi = -\sigma_i p^i \chi \quad (266)$$

The two solutions $\chi^\pm$ describe a left-handed neutrino ν_L and a right-handed anti-neutrino $\bar{\nu}_R$. Only left-handed neutrinos and right-handed anti-neutrinos have been observed in nature.

We can combine χ and ϕ into a four component Dirac spinor u as:

$$u = \begin{pmatrix} \chi \\ \phi \end{pmatrix} \quad (267)$$

with

$$\alpha_i = \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \quad (268)$$

and, hence our solution for the α_i is part of the Weyl representation of the γ-matrices.

For any Dirac spinor, we can make a matrix P_L which projects onto the left-handed component of u. If we introduce a fifth γ-matrix, γ^5, defined by

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \quad (269)$$

Then

$$\gamma^5\gamma^5 = \mathbf{1}, \quad \gamma^5\gamma^\mu + \gamma^\mu\gamma^5 = 0 \quad (270)$$

In the Weyl representation

$$\gamma^5 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix} \quad (271)$$

Thus,

$$P_L \equiv \frac{1}{2}(\mathbf{1} - \gamma^5) = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \tag{272}$$

and

$$P_L u = P_L \begin{pmatrix} \chi \\ \phi \end{pmatrix} = \begin{pmatrix} \chi \\ 0 \end{pmatrix} \tag{273}$$

Similarly

$$P_R \equiv \frac{1}{2}(\mathbf{1} + \gamma^5) = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \tag{274}$$

$$P_R u = P_R \begin{pmatrix} \chi \\ \phi \end{pmatrix} = \begin{pmatrix} 0 \\ \phi \end{pmatrix} \tag{275}$$

Parity Violation in Weak interactions

In the Standard model, only left-handed fermions (and right-handed anti-fermions) couple to W-bosons. Hence, the form of the current which couples, for instance the electron and neutrino to a W-boson takes the form:

$$J^\mu = \bar{\Psi}_e \gamma^\mu \frac{1}{2}(\mathbf{1} - \gamma^5)\Psi_\nu \tag{276}$$

So that the right-handed component of Ψ_ν is not present. Nor is the left-handed component of $\bar{\Psi}_e$ which, because of the "bar" describes the positron. Hence, a right-handed electron does not interact with neutrinos in the Standard Model. Similarly only left-handed up and down quarks interact with W-bosons. Right-handed ones do not.

Parity is the operation under which we reflect all space coordinates through the origin: $x_i \to -x_i$. Under parity, $p_i \to -p_i$ but the spin operators σ_i are unchanged. Hence parity exchanges left-handed and right-handed fermions.

The fact that right-handed neutrinos have not been observed means that parity is violated in Nature in this subtle way. The parity violating form of the charged current above has been experimentally verified.

For instance, one can consider β-decay of Cobalt nuclei:

$$^{60}Co \rightarrow ^{60} Ni^{*} + e^{-} + \nu_e \tag{277}$$

The nuclear spins of the Cobalt were aligned by a magnetic field and an asymmetry in the direction of the emitted electrons observed. This asymmetry changed sign when the magnetic field was reversed. This was interpreted as a left-handed electron and right-handed anti-neutrino. The fact that there was an asymmetry is a strong indication that no right-handed electrons were produced and hence they do not couple to W-bosons!

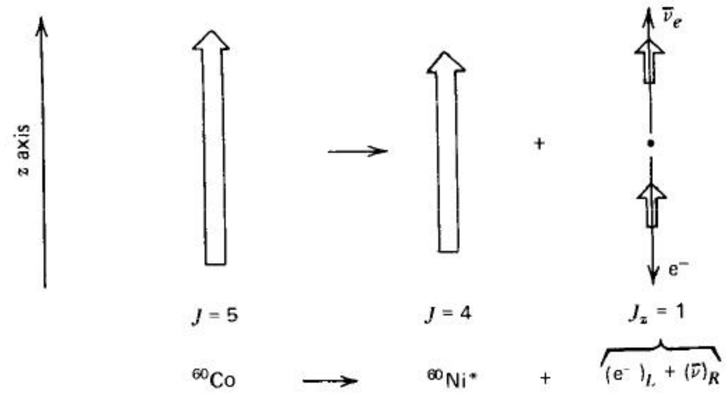


Fig. 12.3 The ^{60}Co experiment: the electron is emitted preferentially opposite the direction of the spin of the ^{60}Co nucleus.

Coupling the Current to the Gauge Fields

The free Dirac equation for a massless fermion can be derived from a Lagrangian

$$\mathcal{L} = i\bar{\Psi}(\partial\!\!\!/ \Psi) \tag{278}$$

In order to obtain the gauge invariant interactions between the fermions in the Standard Model and the gauge fields of the Standard Model, we simply replace the ordinary derivative which appears in $\partial\!\!\!/$ with the covariant derivative.

$$\partial\!\!\!/ \longrightarrow \gamma^{\mu} D_{\mu} = D\!\!\!\!/ \tag{279}$$

where

$$D_\mu = \partial_\mu - i\frac{Y}{2}g_1 B_\mu - ig_2\frac{\sigma_j}{2}W^j_\mu - ig_3\frac{\lambda_a}{2}G^a_\mu \tag{280}$$

Let us ignore g_1 and g_3 for now and calculate the terms in the Lagrangian proportional to g_2.

We need the formula

$$-ig_2\frac{\sigma_j}{2}W^j_\mu = -i\frac{g_2}{2}\begin{pmatrix} W^3_\mu & W^1_\mu - iW^2_\mu \\ W^1_\mu + iW^2_\mu & -W^3_\mu \end{pmatrix} \tag{281}$$

which we re-write as:

$$-i\frac{g_2}{2}\begin{pmatrix} W^3_\mu & W^1_\mu - iW^2_\mu \\ W^1_\mu + iW^2_\mu & -W^3_\mu \end{pmatrix} = -i\frac{g_2}{2}\begin{pmatrix} W^3_\mu & \sqrt{2}W^+_\mu \\ \sqrt{2}W^-_\mu & -W^3_\mu \end{pmatrix} \tag{282}$$

we have introduced $W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu)$ which are the actual positive and negative charge eigenstates.

All the left-handed fermions pair up as doublets under the $SU(2)$ gauge symmetry i.e.

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L, \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L \tag{283}$$

Remember that each of the entries of these doublets are 4-component Dirac spinors. We have suppressed the spinor labels in the last equation.

Similarly the Lagrangian depends on the "bars" of all of these multiplets and hence the couplings to the W-bosons also involves the right-handed anti-particles to all of the above e.g. right handed-positrons also couple to W's, but left-handed positrons do not. Let's write the $SU(2)$ gauge interactions for the muon doublet.

It is

$$\begin{pmatrix} \bar{\nu}_\mu & \bar{\mu}^- \end{pmatrix} \gamma^\nu \times g_2 \begin{pmatrix} W^3_\nu & \sqrt{2}W^+_\nu \\ \sqrt{2}W^-_\nu & -W^3_\nu \end{pmatrix} \frac{1}{2}(\mathbf{1} - \gamma^5) \times \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix} \tag{284}$$

The "i"'s have multiplied to 1. The $\times$'s are there to remind us we have to multiply the matrices together. The $\frac{1}{2}(\mathbf{1} - \gamma^5)$ is there to project onto the left-handed component of the fermion doublet to the right of the matrix of W-bosons. The expression above is equivalent to

$$\begin{aligned} \sqrt{2}\frac{g_2}{2}\bar{\mu}_L^- \gamma^\nu \nu_{\mu L} W_\nu^- &+ \sqrt{2}\frac{g_2}{2}\bar{\nu}_{\mu L}\gamma^\nu \mu_L^- W_\nu^+ \\ -\frac{g_2}{2}\bar{\mu}_L^- \gamma^\nu \mu_L^- W_\nu^3 &+ \frac{g_2}{2}\bar{\nu}_{\mu L}\gamma^\nu \nu_{\mu L} W_\nu^3 \end{aligned} \tag{285}$$

The first line above is the charged current interaction. The second is a neutral current interaction. If we were to change any of the four coefficients, the Lagrangian would no longer be $SU(2)$ gauge invariant. The relative strengths of these interactions $\sqrt{2} : 1$ is thus fixed by symmetry.

For all of the six fermion doublets, there is a similar expression for the $SU(2)$ interaction terms in the Lagrangian.

The interactions we have derived lead to some remarkable consequences:

- The couplings of W^+ and W^- are **universal** for both quarks and leptons. e.g. the interaction between W^+, u and $\bar{d}$ is the *same* as the interaction between W^+, μ and $\bar{\nu}$.
- W^+ can decay into $(e^+, \bar{\nu}_e), (\mu^+, \bar{\nu}_\mu), (\tau^+, \bar{\nu}_\tau), (u, \bar{d}), (c, \bar{s})$
- Similarly for W^-.
- W bosons cannot decay into $(t, \bar{b})$ because $m_t \sim 173\text{GeV}\pm 1\text{GeV}$ and $m_W \sim 80.4\text{GeV}$.
- Since m_W is much larger than $m_e, m_\mu, m_\tau, m_u, m_d, m_c, m_s$ the decay width of the $W^\pm$ doesn't "care" about the fermion masses
- So the partial decay widths

$$\begin{aligned} \Gamma(W^+ \to e^+\bar{\nu}) &= \Gamma(W^+ \to \mu^+\bar{\nu}) & (286) \\ &= \Gamma(W^+ \to \tau^+\bar{\nu}) & (287) \end{aligned}$$

- The decays into quarks are not just two decay channels: $(u\bar{d})$ and $(c, \bar{s})$, but three each, since there are three u quarks, three d quarks, three c quarks and three s quarks. This is because the quarks transform under the $SU(3)$ gauge symmetry (the leptons do not). So the decays into quarks are actually six channels.

- This gives a total of three leptonic and six hadronic decay channels, nine in total.
- If $\Gamma(W^+ \to all)$ is the total W decay width, the Standard Model predicts that

$$\begin{aligned}
\frac{\Gamma(W^+ \to e^+\bar{\nu})}{\Gamma(W^+ \to all)} &= \frac{\Gamma(W^+ \to \mu^+\bar{\nu})}{\Gamma(W^+ \to all)} && (288)\\
&= \frac{\Gamma(W^+ \to \tau^+\bar{\nu})}{\Gamma(W^+ \to all)} && (289)\\
&= 1/9 && (290)
\end{aligned}$$

and

$$\frac{\Gamma(W^+ \to hadrons)}{\Gamma(W^+ \to all)} = 6/9 = 2/3 \qquad (291)$$

- These branching fractions have been measured and agree with the Standard Model predictions to within 1%. Have a look at the PDG.
- This excellent agreement between theory and experiment represents a thorough check of the structure of the Standard Model. The interactions are derived entirely from symmetry principles. If we changed the number of leptons the result would change. If there were four u quarks instead of three (which would be the case if $SU(3)$ were replaced with $SU(4)$) the result would change.
- Note: gauge symmetry, Lorentz invariance and charge conservation allows the possibility of flavour-changing decays that involve different quark families e.g. $W^+ \to c\bar{b}$. In fact these decays also occur, but are suppressed by so-called CKM-mixing (Cabibbo-Kobayashi-Maskawa). We will not discuss this in any detail.

Now, let us also add the terms proportional to g_1, the hypercharge coupling constant.

These are the interactions between fermions and the hypercharge gauge boson B_μ. These have two contributions: one from left-handed fermions and one from the right-handed fermions. For example, in the case of the left-handed electron neutrino doublet, which has $Y = -1$, we have

$$\frac{g_1}{2}\left(\bar{\nu}_e \;\; \bar{e}^-\right)_L \gamma^\mu Y B_\mu \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L \tag{292}$$

where $Y = -1$. The right-handed electron contribution to the neutral current is

$$\frac{g_1}{2}\bar{e}_R\gamma^\mu Y e_R B_\mu \tag{293}$$

where $Y = -2$ for e_R. There is no neutrino contribution in the above, since there is no ν_R.

If we add all the contributions to the neutral current interactions for electrons, we get

$$\bar{e}_L\gamma^\mu e_L\left[-\frac{g_1}{2}B_\mu - \frac{g_2}{2}W^3_\mu\right] - \bar{e}_R\gamma^\mu e_R\left[g_1 B_\mu\right] \tag{294}$$

On the other hand, the basic QED interaction for an electron is of the form

$$\mathcal{L} \sim -eA_\mu\left(\bar{e}_L\gamma^\mu e_L + \bar{e}_R\gamma^\mu e_R\right) \tag{295}$$

where A_μ is the Maxwell gauge field. This **has** to emerge from the Standard Model. In order to see it, consider the neutral current interactions for neutrinos:

$$\bar{\nu}_L^- \gamma^\nu \nu_L^- \left[\frac{-g_1}{2}B_\mu + \frac{g_2}{2}W^3_\nu\right] \tag{296}$$

Neutrinos, being neutral, do not couple to A_μ. This motivates us to define two new linear combinations of B_μ and W^3_μ. One of these will be A_μ, the other combination will be called Z_μ. The correctly normalised combinations are

$$A_\mu = \frac{g_2 B_\mu + g_1 W^3_\mu}{\sqrt{g_2^2 + g_1^2}} \tag{297}$$

$$Z_\mu = \frac{-g_1 B_\mu + g_2 W^3_\mu}{\sqrt{g_2^2 + g_1^2}} \tag{298}$$

With these re-definitions we have two contributions to the electron neutral current interaction

$$A_\mu(\bar{e}_L\gamma^\mu e_L\left[\frac{-g_1g_2}{\sqrt{g_1^2+g_2^2}}\right] \;+\; \bar{e}_R\gamma^\mu e_R\left[\frac{-g_1g_2}{\sqrt{g_1^2+g_2^2}}\right])$$

$$Z_\mu(\bar{e}_L\gamma^\mu e_L\left[\frac{g_1^2-g_2^2}{2\sqrt{g_1^2+g_2^2}}\right] \;+\; \bar{e}_R\gamma^\mu e_R\left[\frac{g_1^2}{\sqrt{g_1^2+g_2^2}}\right])$$

Notice that the left-handed and right-handed components of the electron couple to the photon with precisely the same strength! This is required by the parity invariance of the electromagnetic interaction. We therefore have that

$$\left[\frac{-g_1g_2}{\sqrt{g_1^2+g_2^2}}\right] = -e \tag{299}$$

Notice also that the left-handed and right-handed electron currents couple with **different** strengths to the Z-boson. Thus, these couplings violate parity.

Furthermore, the neutral current coupling to the neutrino does not involve the photon atall:

$$\bar{\nu}_L^-\gamma^\nu\nu_L^-\left[\frac{-g_1}{2}B_\mu+\frac{g_2}{2}W_\nu^3\right]\sim\bar{\nu}_L\gamma^\mu\nu_L\left[\frac{-\sqrt{g_1^2+g_2^2}}{2}Z_\mu\right] \tag{300}$$

Thus, as long as the combination which represents A_μ is massless, and that one can arrange for Z_μ to be sufficiently massive, all phenomena described accurately by QED will be correctly reproduced.

One can (see exercises) similarly calculate the couplings of the quarks to the W's, the Z and the photon.

Mass Terms

From the fact that $\mathcal{L}$ has mass dimension $[M]^4$, we see from the Dirac Lagrangian that fermions have mass dimension $[M]^{3/2}$ i.e. 3/2. A mass term for a fermion with mass m would be of the form

$$\mathcal{L}\sim m\bar{\Psi}\Psi \tag{301}$$

Indeed, a term like this gives exactly the mass term in the Dirac equation.

One can show (see exercises) that

$$\bar{\Psi}\Psi = \bar{\Psi}_L\Psi_R + \bar{\Psi}_R\Psi_L \tag{302}$$

so a mass term couples the left-handed and right-handed components of the fermion together.

But in the Standard model, the left and right handed fermions transform differently under $SU(2)$ and $U(1)_Y$. Therefore:
Mass terms in the Standard Model are forbidden by gauge invariance!

This is a deep and significant point.

Check this by looking at how $\bar{\Psi}_L\Psi_R + \bar{\Psi}_R\Psi_L$ transforms under $U(1)_Y$ for the case the $\Psi = e^-$, the Dirac spinor for the electron. It is important to remember that $\Psi = e = \begin{pmatrix} e_L^- \\ e_R^- \end{pmatrix}$ ie the Dirac spinor is the 4-component spinor containing both the left and right handed fermions.

So, in order for a mass term to be generated, the gauge symmetry must be broken. This is consistent with the fact that the W and Z bosons must also be massive – something which is not possible unless the gauge symmetry is actually broken. The gauge symmetry breaking and mass generation is accomplished by introducing the Higgs boson. The Higgs boson is associated with a Higgs field – this is a scalar field which transforms under $SU(2)$ exactly as the left-handed quarks and leptons and has hypercharge $Y = +1$. The remarkable thing is that the Lagrangian including the Higgs is gauge invariant, but the vacuum state of the Standard Model is not!

This is the first academic year in which particle physics courses can state with reasonable confidence that the Higgs boson actually exists! This is thanks to the remarkable success of both the Large Hadron Collider and the ATLAS and CMS experiments which are based at two of the hadron collision points.

Mass and the Higgs Boson

The Higgs boson is described by a scalar field in the Standard Model. It is therefore described by a Klein-Gordon equation with a covariant derivative. We have studied this before. We will discuss a simplified model to begin with, namely a $U(1)$ gauge theory with a Higgs boson field ϕ of charge q. The Lagrangian is:

$$\begin{aligned}
\mathcal{L} &= D_\mu\phi(D^\mu\phi)^* - m^2\phi\phi^* &(303)\\
&= \partial_\mu\phi\partial^\mu\phi^* - iqA_\mu\phi\partial^\mu\phi^* + iq\partial_\mu\phi A^\mu\phi^* - q^2A_\mu A^\mu\phi\phi^* - m^2\phi\phi^*
\end{aligned}$$

Since ϕ has mass dimension $[M]^1$ and $|\phi|^2 = \phi * \phi$ is gauge invariant under $\phi \to e^{iq\theta}\phi$, we can add one more gauge invariant interaction to $\mathcal{L}$:

$$\Delta\mathcal{L} = -\lambda|\phi|^4 \tag{304}$$

which represents a self-interaction of the Higgs with strength given by the dimensionless coupling λ. So, the potential for ϕ is

$$V(\phi) = m^2|\phi|^2 + \lambda|\phi|^4 \tag{305}$$

If ϕ were real, instead of complex, V looks like

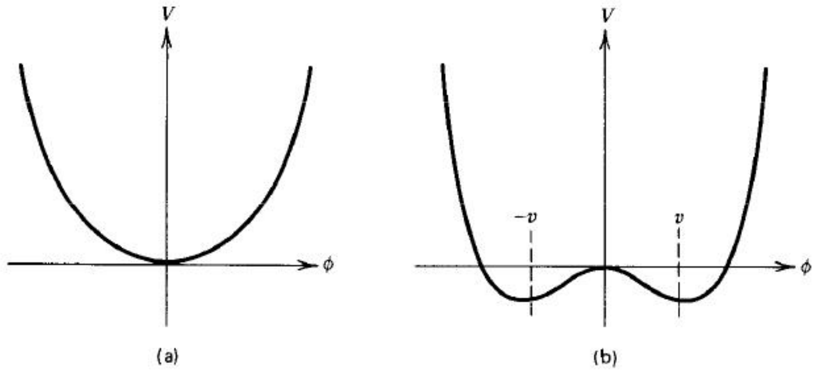


Fig. 14.3 The potential $V(\phi) = \frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4$ for (a) $\mu^2 > 0$ and (b) $\mu^2 < 0$, and $\lambda > 0$.

Note: the book uses $\mu^2/2$ instead of m^2 and $\lambda/4$ instead of λ.

- The shape of the potential takes a very different form depending on whether or not m^2 is +ve or -ve.
- When $m^2 > 0$ there is one minimum at $\phi = 0$

- When $m^2 < 0$ there are two minima, both at **non − zero** values of ϕ.
- $V(\phi)$ is invariant under the symmetry $\phi \to -\phi$.
- When $m^2 > 0$, the minimum is still invariant under the symmetry
- When $m^2 < 0$ any of the two minima break the symmetry!

In gauge theory, however, ϕ is actually complex. The potential, for $m^2 < 0$ looks like

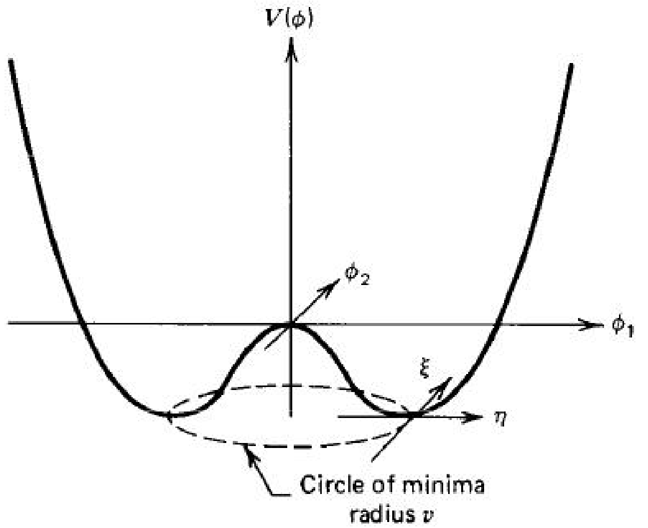


Fig. 14.5 The potential $V(\phi)$ for a complex scalar field for the case $\mu^2 < 0$ and $\lambda > 0$.

- When $m^2 > 0$ there is one minimum at $\phi = 0$
- When $m^2 < 0$ there is a whole circle of minima, all at **non − zero** values of ϕ.
- $V(\phi)$ is invariant under the gauge symmetry $\phi \to e^{iq\theta}\phi$.
- When $m^2 > 0$, the minimum is still invariant under the gauge symmetry
- When $m^2 < 0$ any of the minima **break the gauge symmetry** since a non-zero value of ϕ transforms to a non-zero, different value!

• This is called spontaneous symmetry breaking

The system will minimize the energy. Therefore, if the parameter $m^2 < 0$, the vacuum spontaneously breaks the gauge symmetry. Note that, close to the minimum of the potential, the potential is well approximated by

$$V \sim -4m^2|\phi|^2 \tag{306}$$

which, since $m^2 < 0$ describes a positive mass-squared particle! This is the Higgs boson.

Going back to the original Lagrangian, there is an interesting term:

$$\mathcal{L} = -q^2 A^\mu A_\mu |\phi|^2 \tag{307}$$

Now, since the value of ϕ in the minimum is $< |\phi| >= v = \sqrt{-m^2/2\lambda}$ this leads to

$$\mathcal{L} = -q^2 A^\mu A_\mu v^2 \tag{308}$$

This is a mass term for the gauge field and it arises from a gauge invariant Lagrangian! The mass of the $U(1)$ gauge boson is

$$M = \sqrt{2}qv = q\frac{|m|}{\sqrt{2\lambda}} \tag{309}$$

The factor of $\sqrt{2}$ is present since a mass term for a real boson, such as A_μ is of the form

$$\mathcal{L} = \frac{M^2}{2} A^\mu A_\mu \tag{310}$$

Therefore the Higgs mechanism as described in this simple model can give masses to gauge bosons without the Lagrangian violating gauge symmetry.

In the full Standard Model, the Higgs field is an $SU(2)$ doublet with $Y = 1$

$$\phi = \begin{pmatrix} \phi^1 \\ \phi^2 \end{pmatrix} \tag{311}$$

Therefore it couples to both W-bosons and Z_μ. In the vacuum the gauge symmetry is broken and both $W_\mu^\pm$ and Z_μ get a mass both proportional to v. The Z and W masses are different in this case since both g_1 and g_2 enter into the formula for their masses.

The physical Higgs particle, being neutral, does not couple to the Maxwell field A_μ, which is therefore massless. Hence, the Higgs in the Standard Model induces the symmetry breaking pattern:

$$SU(2) \times U(1)_Y \longrightarrow U(1)_{em} \tag{312}$$

What about fermion masses?

Well, let us go back to our simple $U(1)$ model and assume that we have a left-handed "electron" e_L with charge -1 and a right-handed "electron" e_R of charge -2 (just like the hypercharge in the Standard Model).

Without the Higgs field, the bare mass term

$$m_e\bar{\Psi}\Psi = m_e\bar{\Psi}_L\Psi_R + m_e\bar{\Psi}_R\Psi_L \tag{313}$$

is not gauge invariant.

However, if the charge q of the Higgs field ϕ is equal to one, then

$$\mathcal{L}_{yukawa} = y_e\phi\bar{\Psi}_L\Psi_R + y_e\phi^*\bar{\Psi}_R\Psi_L \tag{314}$$

with a dimensionless coupling constant y_e, is gauge invariant ($0 = 1 + 1 -2 = -1 + 2 -1$)! Since it is gauge invariant and has mass dimension $[M]^4$ we should include it. Such terms are called Yukawa interactions and y_e the Yukawa coupling.

Remarkably, in the vacuum we get

$$\mathcal{L}_{yukawa} = y_e v\bar{\Psi}_L\Psi_R + y_e v\bar{\Psi}_R\Psi_L \tag{315}$$

which is a mass-term for the electron. The electron mass is

$$m_e = y_e v \tag{316}$$

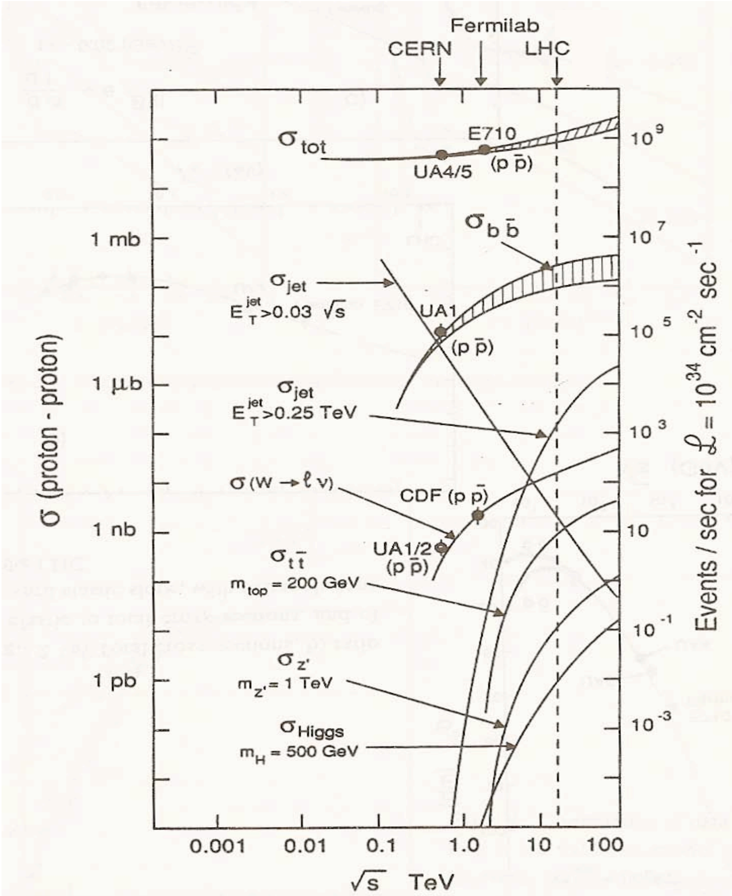
In the Standard Model, v is of order 200GeV in order that the W and Z bosons have the correct masses. Therefore $y_e \sim 0.25 \times 10^{-5}$. Mass terms for all the other fermions arise in the same way by introducing y_μ, y_t, y_d etc. By fixing these parameters appropriately the model correctly describes all the particle masses.

So, according to the Standard Model, all elementary particles get their masses via their interactions with the Higgs boson. The stronger the coupling between a particle and the Higgs, the more massive that particle is. e.g. the top quark Yukawa coupling is of order one, the coupling between the Higgs and the Z-boson is of order 1/2 , but the coupling with the muon is of order 0.5×10^{-3}.

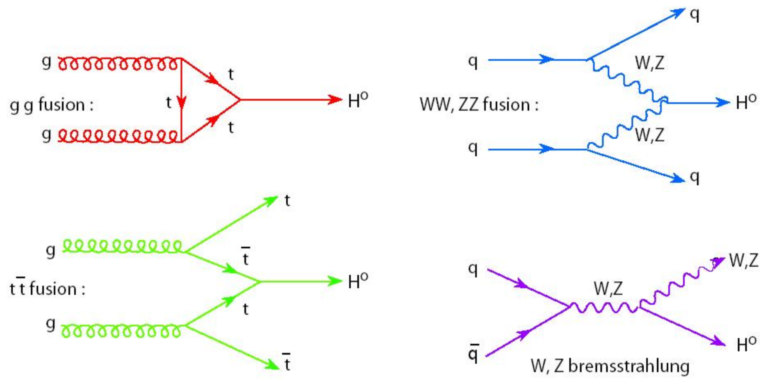
In essence, the Higgs mechanism and symmetry breaking and particle masses all occur in the Standard Model as soon as we propose the existence of an elementary charged scalar particle (as long as we assume a potential with $m^2 < 0$). The fact that the Standard Model fits all particle data thus far so extremely well, encourages (many of) us to believe that the Higgs boson actually exists and will be found at the CERN Large Hadron Collider.

The last sentence was written in October 2011 and I've left it in as a reminder. Now, we are fairly confident that we have discovered the Higgs boson.

Processes and Production Rates at the LHC



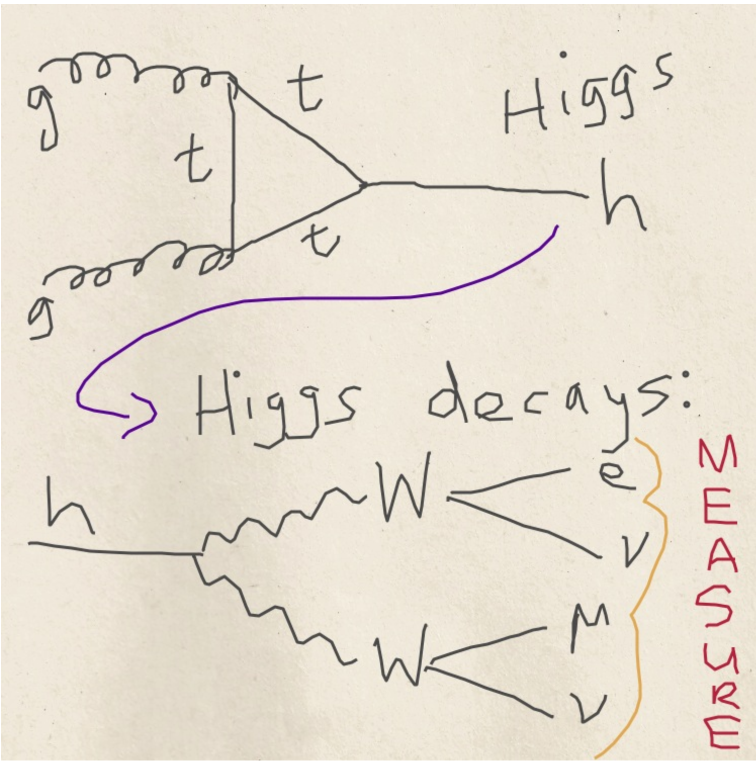
Producing Higgs Bosons with the LHC

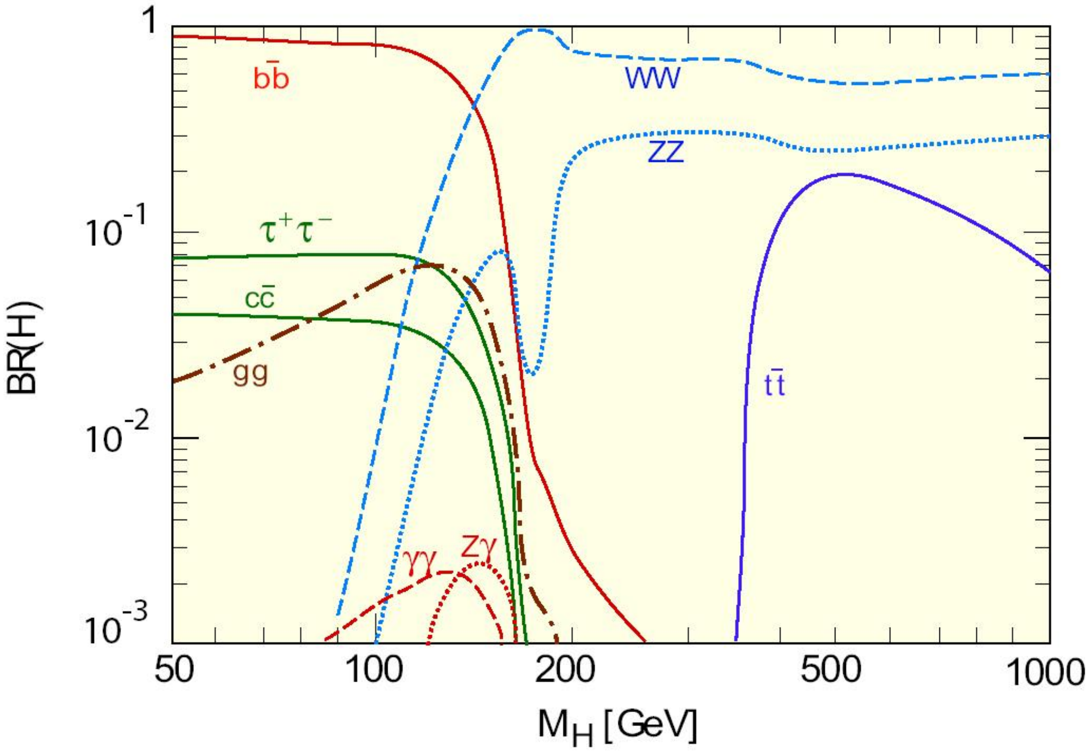


Once produced, the Higgs decays to lighter Standard Model Particles.

Two photons, two tau leptons, two Z bosons,...

We then measure these decay products and "reconstruct" the Higgs!





The above figure shows the branching ratios of the Higgs boson as a function of its mass. These are the probablities for the Higgs to decay into a certain set of particles as a function of its mass. Consisder the region from 100 to 130 GeV. Here the decays are dominated by the $b\bar{b}$ final state. This is because the mass of the Higgs is not large enough to decay into two Z bosons or W bosons or two top quarks. But it can decay into any of the other fermion/anti-fermion pairs. It decays into $b\bar{b}$ most of the time because its coupling to bottom quarks is much larger than that of any of the other fermions (except the top quark, but it can't decay to $t\bar{t}$) because it doesn't have enough mass.

As we increase the mass of the Higgs, beyond 130 GeV, decays into W's become available and start to dominate until 180 GeV when Z's also enter. Finally at twice the top quark mass, decays into $t\bar{t}$ are allowed and dominate.

The actual Higgs mass as measured by the LHC experiments is around 125 GeV $\pm$ 2 GeV or so. This region is interesting because, with enough data many of the final states are available and many have actually been observed. If the Higgs mass had turned out to be 190 GeV, only WW and ZZ would have been observable.

Notice that, even thought the Higgs is electrically neutral, that decays to two photons i.e. $\gamma\gamma$ appear on the graph. These decays actually occur due to higher order (one loop) processes in the Standard Model. This is why the $\gamma\gamma$ channel has a small branching fraction. However, even though it is small, experimentally it is easier to study than $b\bar{b}$ at a hadron collider. This is because the background to $b\bar{b}$ is much larger than that of $\gamma\gamma$. In fact, the calorimeters of ATLAS and CMS were designed to find the Higgs boson in the $\gamma\gamma$ final state. These instruments turned out to perform even better than anticipated and the Higgs was first seen in the $\gamma\gamma$ channel.

How is the mass of the Higgs measured in the diphoton channel?

First, we select events with two photons and nothing else.

We require that these photons are "very clean" and have energies/momenta of at least 15 GeV or so.

Since we measure the energies and momenta of the photons, for each such event we have a Lorentz 4-vector for each photon: P_1^μ and P_2^μ.

From these we construct the invariant mass squared:

$$M^2 = (P_1^\mu + P_2^\mu)(P_{1\mu} + P_{2\mu}) = (E_1 + E_2)^2 - (p_{1x} + p_{2x})^2 - ... \tag{317}$$

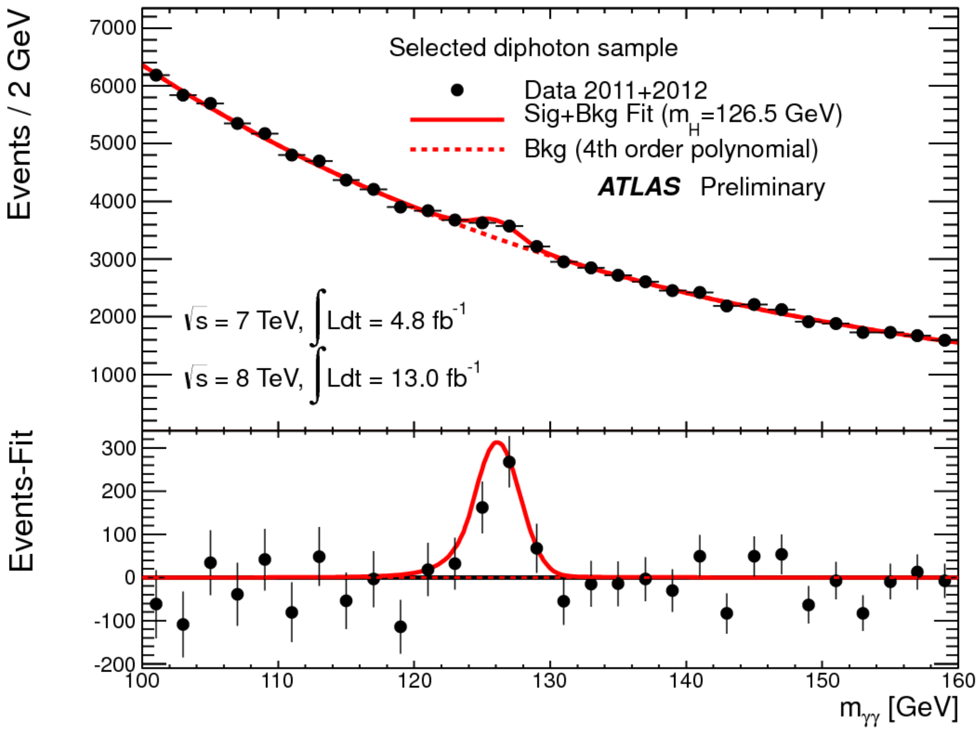
The reason for this is that, if the Higgs decays to two photons then Lorentz invariance implies that

$$P_h^\mu = P_1^\mu + P_2^\mu \tag{318}$$

i.e. that the Higgs' Lorentz vector equals the sum of the Lorentz vectors of the two photons and the Higgs Lorentz vector satisfies

$$P_h^\mu P_{h\mu} = m_h^2 \tag{319}$$

Now, events with diphotons can be produced by many other processes leading to a large background to this search. But, if we understand precisely enough the background and the signal-to-background rate is large enough then we should be able to see the "Higgs peak" above the background. This is an example of what has been seen:



You can see a clear peak in the data around 126 GeV.

SU(5) the Standard Model and Proton Decay

Consider the set of special, unitary, 5-by-5 matrices i.e. $SU(5)$. This set contains both $SU(3)$, $SU(2)$ and $U(1)$.
$SU(3)$ is the subset of of $SU(5)$ matrices of the form:

$$\begin{pmatrix} M_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \tag{320}$$

where M_3 is a 3-by-3 $SU(3)$ matrix, the $\mathbf{1}$ is the 2-by-2 identity matrix and the other entries are all zero.
Similarly, $SU(2)$ is contained in the subset

$$\begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & M_2 \end{pmatrix} \tag{321}$$

Moreover, these $SU(3)$ and $SU(2)$ subsets of $SU(5)$ commute with each other. There is also a $U(1)$ subset which commutes with both $SU(3)$ and $SU(2)$ subsets:

$$\begin{pmatrix} e^{iq\theta} & 0 & 0 & 0 & 0 \\ 0 & e^{iq\theta} & 0 & 0 & 0 \\ 0 & 0 & e^{iq\theta} & 0 & 0 \\ 0 & 0 & 0 & e^{-i\frac{3}{2}q\theta} & 0 \\ 0 & 0 & 0 & 0 & e^{-i\frac{3}{2}q\theta} \end{pmatrix} \tag{322}$$

Note that the determinant is 1 implies that the sum of the five $U(1)$ charges is zero.
The fact that the determinant of an $SU(5)$ matrix is 1, implies that, when $SU(5)$ acts on a 5-vector, the five charges of the 5-components of that vector under any $U(1)$ subset will add to zero.
We have shown that the Standard Model gauge symmetry $SU(3) \times SU(2) \times U(1)_Y$ is a subset of a "simpler" symmetry $SU(5)$. Could this be a useful idea to consider? Could the gauge symmetry of our world be $SU(5)$? This would be remarkable because the three SM couplings g_1, g_2, g_3 would be unified into one "grand unified coupling" g_{GUT}.

Furthermore, it explains an unexplained fact about the Standard Model particles, as we will see.

$SU(5)$ acts naturally on 5-component vectors. So, in order to use $SU(5)$ as a gauge symmetry, we have to arrange five of the Standard Model fermions into the components of this vector. Consider the following 5-vector of fermions

$$\begin{pmatrix} d_R^1 \\ d_R^2 \\ d_R^3 \\ \bar{\nu}_L \\ \bar{e}_L \end{pmatrix} \tag{323}$$

where $\bar{e}_L$ is the right-handed positron with charge $+e$ and $, d_R^2, d_R^3$ are the three "colours" of right-handed quarks. Under the $SU(3)$ subset, the right-handed down quarks transform as they do under the Standard Model $SU(3)$. Similarly the anti-neutrino/positron doublet transform correctly under the $SU(2)$ subset.

If $SU(5)$ is a gauge symmetry, it must contain Maxwell's $U(1)_{em}$. Whatever this $U(1)$ is, the fact that $SU(5)$ has determinant one implies that the electric charge of the right-handed down quarks is minus $1/3$ times the electric charge of the positron! Equivalently, $SU(5)$ explains why the proton and positron have the same electric charge!

$$Q_{el}(p) = Q_{el}(e^+) \tag{324}$$

Similarly, if we identify $U(1)_Y$ with the $U(1)$:

$$\begin{pmatrix} e^{iq\theta} & 0 & 0 & 0 & 0 \\ 0 & e^{iq\theta} & 0 & 0 & 0 \\ 0 & 0 & e^{iq\theta} & 0 & 0 \\ 0 & 0 & 0 & e^{-i\frac{3}{2}q\theta} & 0 \\ 0 & 0 & 0 & 0 & e^{-i\frac{3}{2}q\theta} \end{pmatrix} \tag{325}$$

We learn that the hypercharge of the positron/antineutrino doublet is minus $3/2$ times that of the right-handed down quarks i.e. that

$$Y(d_R) = -\frac{2}{3} Y(e_R^+) \tag{326}$$

What about the other fermions?

Well, we have already seen that $SU(2)$ can be generated by 2-by-2 matrices (Pauli) and by 3-by-3 matrices (rotation generators). So, a fixed symmetry can be realised by matrices of different dimensions. In fact, for $SU(2)$ there are matrix representations of all sizes i.e. k-by-k where k is any integer. Similarly, the other symmetry groups $SU(N)$ for $N > 2$ can also be generated by matrices whose sizes are larger than N-by-N. e.g. $SU(3)$ can be generated by 6-by-6, 8-by-8 and larger matrices. For $SU(5)$ which interests us here, after the 5-by-5 representation, there is also a 10-by-10 one. Though there are lots of larger ones, these are the two smallest sets of generators. The remarkable fact is that these two representations "explain" all of the charges of the Standard Model fermions.

We usually denote the representations by their dimensions. So, for instance the representation of $SU(2)$ by the Pauli matrices is denoted as a **2** of $SU(2)$; the **3** of $SU(3)$ is the representation of SU(3 under which the quarks transform, and so on. So, under the Standard Model gauge symmetry $SU(3) \times SU(2) \times U(1)_Y$ the fermions of a single generation transform as

$$(\mathbf{3}, \mathbf{2})_{\mathbf{1/3}} + (\mathbf{3}^*, \mathbf{1})_{-\mathbf{4/3}} + (\mathbf{3}^*, \mathbf{1})_{\mathbf{2/3}} + (\mathbf{1}, \mathbf{2})_{-\mathbf{1}} + (\mathbf{1}, \mathbf{1})_{\mathbf{2}} \qquad (327)$$

Note that we have written the representation of all the left-handed fermions or anti-fermions since it makes sense to combine all the left-handed spinors together. The first term is the left handed, up-down quark doublet. This is because the left-handed quarks transform in the **3** of $SU(3)$, the u and d combine to transform as a **2** of $SU(2)$ and the subscript 1/3 is their hypercharge. The second term is the anti-up quark (hence the * on the **3**); this is the anti particle to the right handed up quark and is hence left-handed like the first term. The second is the anti-down quark. The third is the left-handed neutrino-electron doublet and the last is the (left-handed) anti-(right-handed) electron i.e. the left handed positron.

When we put the right-handed down quark and the right-handed anti-electron/neutrino doublet into the 5-vector, we learned that, under the $SU(3) \times SU(2) \times U(1)_Y$ subset of $SU(5)$:

$$\mathbf{5} = (\mathbf{3}, \mathbf{1})_{-\mathbf{2/3}} + (\mathbf{1}, \mathbf{2})_{\mathbf{1}} \tag{328}$$

Similarly we can consider the conjugate of this:

$$\mathbf{5}^* = (\mathbf{3}^*, \mathbf{1})_{\mathbf{2/3}} + (\mathbf{1}, \mathbf{2}^*)_{-\mathbf{1}} \tag{329}$$

Finally, the remarkable thing is that, if we take the ten-dimensional representation of $SU(5)$, under the SM subset we have

$$\mathbf{10} = (\mathbf{3}, \mathbf{2})_{\mathbf{1/3}} + (\mathbf{3}^*, \mathbf{1})_{-\mathbf{4/3}} + (\mathbf{1}, \mathbf{1})_{\mathbf{2}} \tag{330}$$

We therefore learn that

$$\mathbf{5}^* + \mathbf{10} = (\mathbf{3}, \mathbf{2})_{\mathbf{1/3}} + (\mathbf{3}^*, \mathbf{1})_{-\mathbf{4/3}} + (\mathbf{3}^*, \mathbf{1})_{\mathbf{2/3}} + (\mathbf{1}, \mathbf{2})_{-\mathbf{1}} + (\mathbf{1}, \mathbf{1})_{\mathbf{2}} \tag{331}$$

so that the two simplest representations of $SU(5)$, when written in terms of the $SU(3) \times SU(2) \times U(1)_Y$ subset reproduce exactly the charges of all the Standard Model fermions!

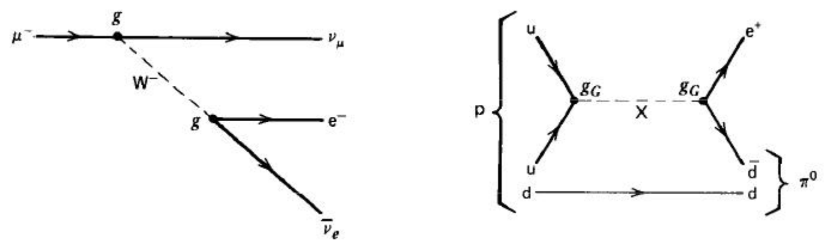
We have thus far discussed how the Standard Model particles fit into $SU(5)$ representations. We have discussed the 12 "directions" inside an $SU(5)$ matrix which correspond to the 8+3+1=12 generators of $SU(3) \times SU(2) \times U(1)_Y$. What about the "other" elements of $SU(5)$ which are *not* in $SU(3) \times SU(2) \times U(1)_Y$? $SU(5)$ is 24-dimensional i.e. has 24 generators and 24 gauge bosons. There are thus 12 gauge bosons predicted by $SU(5)$ which are *beyond* the Standard Model. Can we estimate the masses of these gauge bosons?

In order to do that we have to consider what sorts of interactions are mediated by these gauge bosons. To see those we should consider $SU(5)$ transformations which are NOT in $SU(3) \times SU(2) \times U(1)_Y$. Consider the following example of an $SU(5)$ transformation:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} d_R^1 \\ d_R^2 \\ d_R^3 \\ \bar{\nu}_L \\ \bar{e}_L \end{pmatrix} = \begin{pmatrix} d_R^1 \\ d_R^2 \\ \bar{e}_L \\ \bar{\nu}_L \\ -d_R^3 \end{pmatrix} \tag{332}$$

This interchanges a down quark with a positron! Therefore, the 12 $SU(5)$ gauge bosons which are not identified with Standard Model gauge bosons mediate interactions between quarks and leptons. If a quark can turn into a lepton by emitting an $SU(5)$ gauge boson, then the proton (which is absolutely stable in the Standard Model) would become unstable. Thus, the idea of grand unification of the fundamental forces predicts protons will decay. We can estimate the lifetime of the proton in terms of the $SU(5)$ coupling constant and the masses of the (fictitious) $SU(5)$ gauge bosons. The diagram and the calculation are very similar to the muon decay calculation that we have done before.

Muon Decay ($\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu$) at $Q^2 \ll M_W^2$	Proton Decay ($p \rightarrow \pi^0 e^+$) at $Q^2 \ll M_X^2$



$\frac{G}{\sqrt{2}} = \frac{g^2}{8M_W^2}$ (12.15)	$\frac{G_G}{\sqrt{2}} = \frac{g_G^2}{8M_X^2}$
$\Gamma(\mu \rightarrow e\bar{\nu}_e\nu_\mu) = \cdots G^2 m_\mu^5$ (12.42)	$\Gamma(p \rightarrow \pi e) = \cdots G_G^2 m_p^5$
$= \cdots \frac{m_\mu^5}{M_W^4}$	$= \cdots \frac{m_p^5}{M_X^4}$

The figure shows the analogy with muon decay in calculating the proton lifetime. The diagram on the right shows a process in which $p \to \pi^0 e^+$. The two u-quarks in p annihilate into the $SU(5)$ gauge boson X. X then decays into $e^+, \bar{d}$ via the interaction we described above. At this point, the d-quark combines with the produced $\bar{d}$ to make a neutral hadron, the pion.

The basic matrix element comes from the $uu \to X \to e^+\bar{d}$ part of the diagram and

$$\mathcal{M} \propto \frac{g_{GUT}^2}{M_X^2} \tag{333}$$

where g_{GUT} is the $SU(5)$ gauge coupling and M_X the mass of the gauge boson. Thus the decay width

$$\Gamma(p \to e^+\pi^0) \propto |\mathcal{M}|^2 \sim \frac{g_{GUT}^4}{M_X^4} \tag{334}$$

This is a very low momentum process relative to M_X. Since the proton mass is much larger than both the pion and positron masses, the only other relevant scale in the problem is the proton mass m_p. Since Γ has mass dimension $[M]^1$ it must be that

$$\Gamma(p \to e^+\pi^0) \sim \frac{g_{GUT}^4 m_p^5}{M_X^4} \tag{335}$$

We can thus use this formula, together with the experimental limit on the lifetime of the proton to constrain the combination $M \equiv \frac{M_X}{g_{GUT}}$, though in practice g_{GUT} is not usually that small so effectively we constrain M_X.

Measuring/Setting a limit on the proton lifetime

The current limit on the proton lifetime for the decay mode we are considering is about $\tau_p > 1.6 \times 10^{33}$ years!

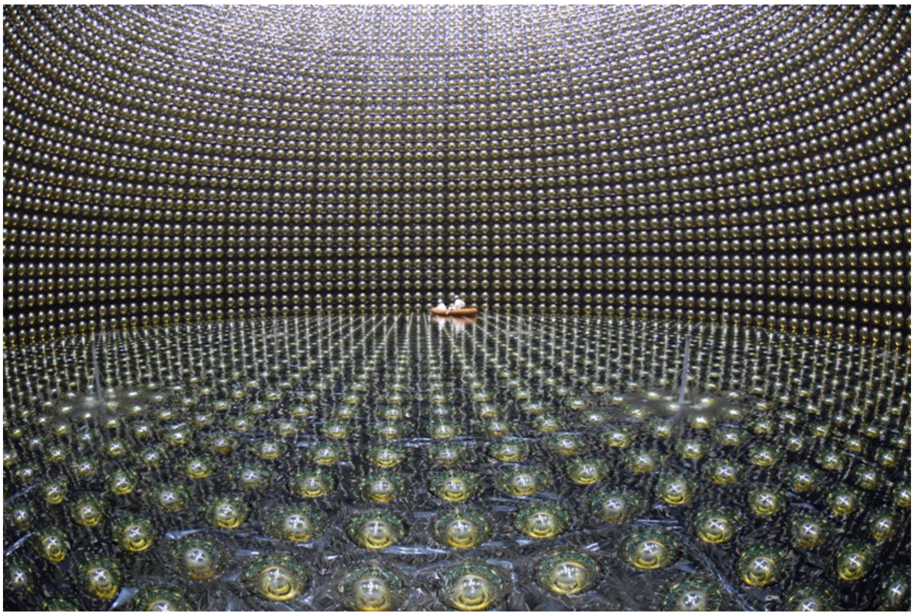
Q: How can we obtain such a strong limit, 23 orders of magnitude longer than the age of the Universe (about 10^{10} years) ?

A: By watching a very large number of protons for as long as possible and seeing that none have decayed yet!

One way to look for proton decay is to surround a tank of water with "photon detectors " ie photo multiplier tubes. The positron will interact with the electrons in the water to produce photons and the π^0 almost always decays to two photons.

Since $1m^3$ of water contains around 10^{30} protons, if no protons are observed to decay after one year we get a limit of about 10^{30} years. So, by putting photo-multiplier tubes around larger amounts of water one can obtain a stronger limit. The SuperKamiokande experiment has 50,000 tons of water and this is where the best limit on the proton lifetime comes from.

SuperK also detects neutrinos: when a neutrino comes in to the water it can sometimes exchange a Z-boson with an electron in the water. If the original neutrino is energetic enough, the electron will recoil with high enough momentum such that it will travel faster than the speed of light in water. When this happens, the electron will emit what is called Cerenkov radiation, which are photons which can be detected in the photomultiplier tubes. Similarly, if the incoming neutrino exchanges a W-boson with an electron or quark in the water, it will convert into an electron, muon (or tau if energetic enough) and this could also emit Cerenkov light if fast enough.



We therefore have that

$$\frac{M^4}{m_p^5} > 1.6 \times 10^{33} yrs = 50.5 \times 10^{39} s = 7.65 \times 10^{64} (GeV)^{-1} \tag{336}$$

Approximating $m_p \sim 1GeV$ we have that

$$M^4 > 7.65 \times 10^{64} GeV \tag{337}$$

or

$$M > 1.7 \times 10^{16} GeV \tag{338}$$

In other words we obtain a limit on the mass scale which is 16 orders of magnitude larger than the proton mass, 14 orders of magnitude larger than the masses of the W and Z bosons! This very large scale of order 10^{16} GeV is known as the GUT scale – the scale of grand unification. IF grand unification is correct, it is at this scale that the additional bosons predicted by grand unification get their masses i.e. the $SU(5)$ symmetry is broken at a very high scale.

Anyway, from the point of view of the course, the idea behind this calculation is to give a) in idea about how theoretical physics think – using the idea of gauge symmetry and pushing it to its final conclusions and b) to show how one can use the experimental data to get an idea of what theoretical parameters have to be and c) to understand the basic idea behind proton decay experiments.