An Introduction to QED and QCD

Jeff Forshaw

Department of Physics and Astronomy
University of Manchester
Manchester M13 9PL

Lectures presented at the School for Young High Energy Physicists,
Rutherford Appleton Laboratory, September 1997.

Contents

1 Introduction

The traditional aim of this course is to teach you how to calculate amplitudes, cross-sections and decay rates, particularly for quantum electrodynamics, QED, but in principle also for quantum chromodynamics, QCD. By the end of the course you should be able to go from a Feynman diagram, such as the one for $e^+e^- \to \mu^+\mu^-$ in Figure 1.1(a), to a number for the cross section.

We will restrict ourselves to calculations at *tree level* but, at the end of the course, we will also look qualitatively at higher order *loop* effects which, amongst other things, are responsible for the running of the QCD coupling. This running means that the QCD coupling appears weaker when measured at higher energy scales and is the reason why we can sometimes do perturbative QCD calculations. As you might guess, the sort of diagrams which are important here have closed loops of particle lines in them: in Figure 1.1(b) is one example contributing to the running of the QCD coupling (the curly lines denote gluons).

In order to do our calculations we will need a certain amount of technology. In particular, we will need to describe particles with spin, especially the spin-1/2 leptons and quarks. We will therefore spend some time looking at the Dirac equation. After this we'll work out how to go from quantum mechanical probability amplitudes to cross sections and decay rates. Then, with these tools in hand, we will look at some examples of tree level QED processes. Here you will get hands-on experience of calculating transition amplitudes and getting from them to cross sections. We then move on to QCD. This will entail a brief introduction to renormalisation in both QED and QCD. We will introduce the idea of the running coupling and look at asymptotic freedom in QCD.

In reference [1] you will find a list of textbooks which may be useful.

1.1 Units and Conventions

I will use natural units, $c = 1$, $\hbar = 1$, so mass, energy, inverse length and inverse time all have the same dimensions.

$$\begin{array}{lll} \text{4-vector} & a^\mu \quad \mu = 0,1,2,3 & a^\mu = (a^0, \mathbf{a}\,) \\ \text{scalar product} & a{\cdot}b = a^0 b^0 - \mathbf{a}{\cdot}\mathbf{b} = g_{\mu\nu} a^\mu b^\nu & \end{array} \tag{1.1}$$

From the scalar product you see that the metric is

$$g = \mathrm{diag}(1, -1, -1, -1), \qquad g^{\mu\lambda} g_{\lambda\nu} = \delta^\mu_\nu = \begin{cases} 1 & \text{if } \mu = \nu \\ 0 & \text{if } \mu \neq \nu \end{cases} \tag{1.2}$$

For $c = 1$, $g^{\mu\nu}$ and $g_{\mu\nu}$ are numerically the same.

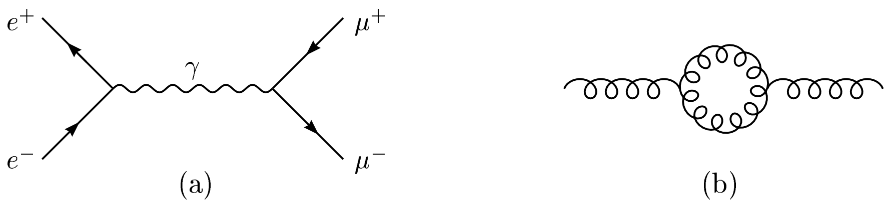


Figure 1.1 Examples of Feynman diagrams contributing to (a) $e^+e^- \to \mu^+\mu^-$ and (b) the running of the strong coupling constant.

From the above, you would think it natural to write the space components of a 4-vector as a^i for $i = 1, 2, 3$. However, for 3-vectors I usually write the components as a_i. This ought not to cause confusion, since for 3-vectors the metric is just the unit matrix. Generally speaking, a little care must be taken in getting the sign right for the '$1, 2, 3$' components of a 4-vector.

Note that ∂_μ is a covector,

$$\partial_\mu = \frac{\partial}{\partial x^\mu}, \quad \partial_\mu x^\nu = \delta^\nu_\mu, \tag{1.3}$$

so $\nabla^i = -\partial^i$ and $\partial^\mu = (\partial^0, -\nabla)$.

My convention for the totally antisymmetric Levi-Civita tensor is

$$\epsilon^{\mu\nu\lambda\sigma} = \begin{cases} +1 & \text{if } \{\mu, \nu, \lambda, \sigma\} \text{ an even permutation of } \{0, 1, 2, 3\} \\ -1 & \text{if an odd permutation} \\ 0 & \text{otherwise} \end{cases} \tag{1.4}$$

Note that $\epsilon^{\mu\nu\lambda\sigma} = -\epsilon_{\mu\nu\lambda\sigma}$, and $\epsilon^{\mu\nu\lambda\sigma} p_\mu q_\nu r_\lambda s_\sigma$ changes sign under a parity transformation (which is obvious because it contains an odd number of spatial components).

1.2 Relativistic Wave Equations

The starting point for this course is the Schrödinger equation which can be written quite generally as

$$H\psi = i\frac{\partial \psi}{\partial t} \tag{1.5}$$

where H is the Hamiltonian (i.e. the energy operator). In this equation ψ is the wave-function describing the single particle probability amplitude. I shall usually reserve the Greek symbol ψ for spin 1/2 fermions and ϕ for spin 0 bosons. So for pions and the like I shall write

$$H\phi = i\frac{\partial \phi}{\partial t}. \tag{1.6}$$

In this course we want to extend the non-relativistic quantum mechanics, familiar to undergraduates, into the relativistic domain. For example, in non-relativistic quantum mechanics you are used to writing

$$H = T + V \tag{1.7}$$

where T is the kinetic energy and V is the potential energy. A particle of mass m and momentum $\mathbf{p}$ has non-relativistic kinetic energy,

$$T = \frac{\mathbf{P}^2}{2m} \tag{1.8}$$

where capital $\mathbf{P}$ is the operator corresponding to momentum $\mathbf{p}$. For a slow moving particle $v \ll c$ (e.g. an electron in a Hydrogen atom) this is adequate, but for relativistic systems ($v \sim c$) the Hamiltonian above breaks down. For a free relativistic particle the total energy E is given by the Einstein equation

$$E^2 = \mathbf{p}^2 + m^2. \tag{1.9}$$

Thus the square of the relativistic Hamiltonian H^2 is simply given by promoting the momentum to operator status:

$$H^2 = \mathbf{P}^2 + m^2. \tag{1.10}$$

So far so good, but now the question arises of how to implement the Schrödinger equation, which is expressed in terms of H rather than H^2. Naively the relativistic Schrödinger equation looks like

$$\sqrt{\mathbf{P}^2 + m^2}\psi(t) = i\frac{\partial \psi(t)}{\partial t} \tag{1.11}$$

but this is difficult to interpret because of the square root. There are two ways forward:

(1) Work with H^2. By iterating the Schrödinger equation we have

$$H^2\phi(t) = -\frac{\partial^2 \phi(t)}{\partial t^2}. \tag{1.12}$$

This is known as the Klein-Gordon (KG) equation. In this case the wavefunction describes spinless bosons.

(2) Invent a new Hamiltonian H_D which is linear in momentum, and whose square is equal to H^2 given above, $H_D^2 = \mathbf{P}^2 + m^2$. In this case we have

$$H_D\psi(t) = i\frac{\partial \psi(t)}{\partial t} \tag{1.13}$$

which is known as the Dirac equation, with H_D being the Dirac Hamiltonian. In this case the wavefunction describes spin 1/2 fermions, as we shall see.

1.3 Wavefunctions and Fields

You may be wondering why I am talking about wavefunctions while in your *field theory* course Dave Dunbar is telling you about fields? Some of you may even be wondering what is the difference between a wavefunction and a field.

The bottom line is that *single particle* wavefunctions work just fine if you want to describe systems where the particle number is conserved. Problems come when you want to allow for relativistic effects. In particular antiparticles, and hence the possibility that particles can annihilate or be pair produced. In fact, these new concepts can be accommodated within the wavefunction approach, but in a way which is not really very satisfactory. We'll take a look at the problems encountered in trying to cling to the wavefunction way of thinking shortly.

Rather than patch things up it's much more appealing simply to ditch the usual interpretation of ψ as a wavefunction and to identify it as a field. This field is then subjected to the usual laws of quantum mechanics. This means elevating the field and its canonical momentum to the status of operators which are then deemed to satisfy the usual commutation relations. This is what Dave has been doing in his course. In the limit that particle number is conserved, the theory is equivalent to the single particle wavefunction approach.

It is important to emphasise that a field is very different from a wavefunction. Think first of a classical field. I find it easier to picture a field by first dividing space into infinitesimally small boxes. In each box is a fictitious particle, and the amplitude of the particle's displacement from equilibrium is the value of the field as the point where the small box is. If the field satisfies some wave equation then you can think of the motions of the individual particles as being influenced by what's going on in the neighbourhood (e.g. consider a vibrating membrane): the particles can be thought of as little harmonic oscillators all coupled together. The field is defined in the limit of vanishingly small boxes, and so describes a system with an infinite number of degrees of freedom.

The Maxwell field of electromagnetism is a famous classical field. We can go ahead an demand that it also be consistent with the laws of quantum mechanics. This means we must quantise the motion of the infinite number of little harmonic oscillators. To do this, we simply impose the commutation relation $[x_i, p_i] = i$, where x_i is the displacement of the i^{th} oscillator and p_i is its momentum. Note that x_i is simply the value of the field at the point where the i^{th} box is located. After doing this we have a theory which is consistent with both relativity and quantum mechanics. Photons emerge as the quanta of the electromagnetic field, and so the theory naturally describes systems with any number of particles.

Maxwell's wave equations are not the only equations we can write down which are consistent with relativity. We can also write down the Klein-Gordon and Dirac equations too (there are others, but these are the most relevant ones for particle physics). Quantising the corresponding fields leads to quanta which have spin-0 and spin-1/2 respectively (the Maxwell field leads to spin-1 quanta). Letting the spin-1/2 quanta carry electric charge means that they can interact with the spin-1 quanta and the formalism has no problem dealing with varying numbers of quanta. Of course, I'm skipping the details so as to give you an overview. Dave Dunbar's course aims to provide a fairly comprehensive introduction to this whole area.

A final word of caution. Notation can be a little confusing. People often use the same symbol for both the wavefunction and the field. You should always be aware of the difference and be able to spot from the context which is meant.

1.4 The Klein-Gordon Equation

Let's now take a more detailed look at the KG equation (1.12). In position space we write the momentum operator as

$$\mathbf{p} \to -i\nabla, \tag{1.14}$$

so that the KG equation becomes

$$\left(\Box + m^2\right)\phi(x) = 0 \tag{1.15}$$

where we have introduced the box notation,

$$\Box = \partial_\mu \partial^\mu = \partial^2/\partial t^2 - \nabla^2 \tag{1.16}$$

and x is the 4-vector $(t, \mathbf{x})$.

The operator $\Box$ is Lorentz invariant, so the Klein-Gordon equation is relativistically covariant (that is, transforms into an equation of the same form) if ϕ is a scalar function. That is to say, under a Lorentz transformation $(t, \mathbf{x}) \to (t', \mathbf{x}')$,

$$\phi(t, \mathbf{x}) \to \phi'(t', \mathbf{x}') = \phi(t, \mathbf{x})$$

so ϕ is invariant. In particular ϕ is then invariant under spatial rotations so it represents a spin-zero particle (more on spin when we come to the Dirac equation); there being no preferred direction which could carry information on a spin orientation.

The Klein-Gordon equation has plane wave solutions:

$$\phi(x) = Ne^{-i(Et-\mathbf{p}\cdot\mathbf{x})} \tag{1.17}$$

where N is a normalisation constant and $E = \pm\sqrt{\mathbf{p}^2 + m^2}$. Thus, there are both positive and negative energy solutions. The negative energy solutions pose a severe problem if you try to interpret ϕ as a wavefunction (as indeed we are trying to do). The spectrum is no longer bounded from below, and you can extract arbitrarily large amounts of energy from the system by driving it into ever more negative energy states. Any external perturbation capable of pushing a particle across the energy gap of $2m$ between the positive and negative energy continuum of states can uncover this difficulty. Furthermore, we cannot just throw away these solutions as unphysical since we need them in order to define a complete set of states. Note that if one interprets ϕ as a quantum field there is no problem, i.e. the positive and negative energy modes are just associated with operators which create or destroy particles.

A second problem with the wavefunction interpretation arises when trying to find a probability density. Since ϕ is Lorentz invariant, $|\phi|^2$ doesn't transform like a density. To search for a candidate we derive a continuity equation, rather as you did for the Schrödinger equation in the pre-school problems. Defining ρ and $\mathbf{J}$ by

$$\begin{aligned} \rho &\equiv i\left(\phi^*\frac{\partial\phi}{\partial t} - \phi\frac{\partial\phi^*}{\partial t}\right) \\ \mathbf{J} &\equiv -i\left(\phi^*\nabla\phi - \phi\nabla\phi^*\right) \end{aligned} \tag{1.18}$$

you obtain (see problem) a covariant conservation equation

$$\partial_\mu J^\mu = 0 \tag{1.19}$$

where J is the 4-vector $(\rho, \mathbf{J})$. It is natural to interpret ρ as a probability density and $\mathbf{J}$ as a probability current. However, for a plane wave solution (1.17), $\rho = 2|N|^2E$, so ρ is not positive definite since we've already found E can be negative.

▷ **Exercise 1.1**
Derive the continuity equation (1.19). Start with the Klein-Gordon equation multiplied by ϕ^* and subtract the complex conjugate of the K-G equation multiplied by ϕ.

Thus, ρ may well be considered as the density of a conserved quantity (such as electric charge), but we cannot use it for a probability density. To Dirac, this and the existence of negative energy solutions seemed so overwhelming that he was led to introduce another equation, first order in time derivatives but still Lorentz covariant, hoping that the similarity to Schrödinger's equation would allow a probability interpretation. Dirac's original hopes were unfounded because his new equation turned out to admit negative energy solutions too! Even so, he did find the equation for spin-1/2 particles and predicted the existence of anti-particles.

Before turning to discuss what Dirac did. Let's put things in context. We have found that the Klein-Gordon equation, a candidate for describing the quantum mechanics of

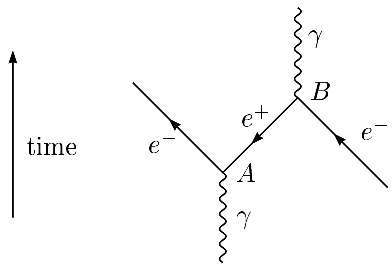


Figure 1.2 Feynman interpretation of a process in which a negative energy electron is absorbed. Time increases moving upwards.

spinless particles admits unacceptable negative energy states when ϕ is interpreted as the single particle wavefunction. We could solve all our problems here and now, and restore our faith in the Klein-Gordon equation by simply re-interpreting ϕ as a quantum field. However we won't do that. There is another way forward (this is the way followed in the textbook of Halzen & Martin) due primarily to Feynman. Causality forces us to ensure that positive energy states propagate forwards in time. Feynman spotted that the negative energy states cause us problems only so long as we think of them as real physical states propagating forwards in time. If we force these negative energy states only to propagate backwards in time then we find a theory which is consistent with the requirements of causality and which has none of the aforementioned problems.

According to Feynman, we should interpret the emission (absorption) of a negative energy particle with momentum p^μ as the absorption (emission) of a positive energy antiparticle with momentum $-p^\mu$. So, in Figure 1.2, for example, an electron–positron pair is created at point A. The positron propagates to point B where it is annihilated by another electron. Another way of describing this picture is to say that the incoming electron interacts to produce an outgoing photon and a negative energy electron. This electron travels back in time where it scatters off the incoming photon to produce the outgoing electron. To someone observing in real time, the negative energy state moving backwards in time looks to all intents and purposes like a positively charged electron with positive energy moving forwards in time.

2 The Dirac Equation

Dirac wanted an equation first order in time derivatives and Lorentz covariant, so it had to be first order in spatial derivatives too. His starting point was to assume a Hamiltonian of the form,

$$H_D = \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3 + \beta m \tag{2.1}$$

where P_i are the three components of the momentum operator $\mathbf{P}$, and α_i and β are some unknown quantities, which as will be seen below cannot simply be commuting numbers. When the requirement that the $H_D^2 = \mathbf{P}^2 + m^2$ is imposed, this implies that α_i and β must be interpreted as 4×4 matrices, as we shall discuss. The first step is to write the momentum operators explicitly in terms of their differential operators, using equation (1.14), then the Dirac equation (1.13) becomes, using the Dirac Hamiltonian in equation (2.1),

$$i\,\frac{\partial\psi}{\partial t} = (-i\,\boldsymbol{\alpha}\cdot\nabla + \beta m)\psi \tag{2.2}$$

which is the position space Dirac equation. Remember that in field theory, the Dirac equation is the equation of motion for the field operator describing spin 1/2 fermions. In order for this equation to be Lorentz covariant, it will turn out that ψ cannot be a scalar under Lorentz transformations. In fact this will be precisely how the equation turns out to describe spin 1/2 particles. We will return to this below.

If ψ is to describe a free particle it is natural that it should satisfy the Klein-Gordon equation so that it has the correct energy-momentum relation. This requirement imposes relationships among the $\boldsymbol{\alpha}$ and β. To see these, apply the operator on each side of equation (2.2) twice, i.e. iterate the equation,

$$-\frac{\partial^2\psi}{\partial t^2} = [-\alpha^i\alpha^j\nabla^i\nabla^j - i\,(\beta\alpha^i + \alpha^i\beta)m\nabla^i + \beta^2 m^2]\psi$$

with an implicit sum over i and j from 1 to 3. The Klein-Gordon equation by comparison is

$$-\frac{\partial^2\psi}{\partial t^2} = [-\nabla^i\nabla^i + m^2]\psi \tag{2.3}$$

If we do not assume that the α^i and β commute then the KG will clearly be satisfied if

$$\begin{aligned} \alpha_i\alpha_j + \alpha_j\alpha_i &= 2\delta_{ij} \\ \beta\alpha_i + \alpha_i\beta &= 0 \\ \beta^2 &= 1 \end{aligned} \tag{2.4}$$

for $i, j = 1, 2, 3$. It is clear that the α_i and β cannot be ordinary numbers, but it is possible to give them a realisation as matrices. In this case, ψ must be a multi-component *spinor* on which these matrices act.

▷ **Exercise 2.1**

Prove that any matrices $\boldsymbol{\alpha}$ and β satisfying equation (2.4) are traceless with eigenvalues ± 1. Hence argue that they must be even dimensional.

In two dimensions a natural set of matrices for the $\boldsymbol{\alpha}$ would be the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.5}$$

However, there is no other independent 2×2 matrix with the right properties for β, so the smallest dimension for which the Dirac matrices can be realised is four. One choice is the *Dirac representation*:

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \qquad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.6}$$

Note that each entry above denotes a two-by-two block and that the 1 denotes the 2×2 identity matrix.

There is a theorem due to Pauli which states that all sets of matrices obeying the relations in (2.4) are equivalent. Since the Hermitian conjugates $\boldsymbol{\alpha}^\dagger$ and $\beta^\dagger$ clearly obey the relations, you can, by a change of basis if necessary, assume that $\boldsymbol{\alpha}$ and β are Hermitian. All the common choices of basis have this property. Furthermore, we would like α_i and β to be Hermitian so that the Dirac Hamiltonian (2.17) is Hermitian.

▷ **Exercise 2.2**
Derive the continuity equation $\partial_\mu J^\mu = 0$ for the Dirac equation with

$$\rho = J^0 = \psi^\dagger \psi, \qquad \mathbf{J} = \psi^\dagger \boldsymbol{\alpha} \psi. \tag{2.7}$$

We will see in section 2.6 that $(\rho, \mathbf{J})$ transforms, as it must, as a four-vector.

2.1 Free Particle Solutions I: Interpretation

We look for plane wave solutions of the form

$$\psi = \begin{pmatrix} \chi(\mathbf{p}) \\ \phi(\mathbf{p}) \end{pmatrix} e^{-i(Et - \mathbf{p}\cdot\mathbf{x})} \tag{2.8}$$

where $\phi(\mathbf{p})$ and $\chi(\mathbf{p})$ are two-component spinors which depend on momentum $\mathbf{p}$ but are independent of $\mathbf{x}$. Using the Dirac representation of the matrices, and inserting the trial solution into the Dirac equation gives the pair of simultaneous equations

$$E \begin{pmatrix} \chi \\ \phi \end{pmatrix} = \begin{pmatrix} m & \boldsymbol{\sigma}\cdot\mathbf{p} \\ \boldsymbol{\sigma}\cdot\mathbf{p} & -m \end{pmatrix} \begin{pmatrix} \chi \\ \phi \end{pmatrix}. \tag{2.9}$$

There are two simple cases for which equation (2.9) can readily be solved, namely

(1) $\mathbf{p} = 0$, $m \neq 0$ which might represent an electron in its rest frame.

(2) $m = 0$, $\mathbf{p} \neq 0$ which might represent a massless neutrino.

For case (1), an electron in its rest frame, the equations (2.9) decouple and become simply,

$$E\chi = m\chi, \qquad E\phi = -m\phi. \tag{2.10}$$

So, in this case, we see that χ corresponds to solutions with $E = m$, while ϕ corresponds to solutions with $E = -m$. In light of our earlier discussions, we no longer need to recoil in horror at the appearance of these negative energy states.

The negative energy solutions persist for an electron with $\mathbf{p} \neq 0$ for which the solutions to equation (2.9) are readily seen to be

$$\phi = \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi, \qquad \chi = \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E-m}\phi. \tag{2.11}$$

Thus the general positive energy solutions with $E = +|\sqrt{\mathbf{p}^2 + m^2}|$ are

$$\psi(x) = \begin{pmatrix} \chi \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi \end{pmatrix} e^{-i(Et-\mathbf{p}\cdot\mathbf{x})}, \tag{2.12}$$

while the general negative energy solutions with $E = -|\sqrt{\mathbf{p}^2 + m^2}|$ are

$$\psi(x) = \begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E-m}\phi \\ \phi \end{pmatrix} e^{-i(Et-\mathbf{p}\cdot\mathbf{x})}, \tag{2.13}$$

for arbitrary constant ϕ and χ. Clearly when $\mathbf{p} = 0$ these solutions reduce to the positive and negative energy solutions discussed previously. As an aside, it is interesting to see how Dirac coped with the negative energy states.

Dirac interpreted the negative energy solutions by postulating the existence of a "sea" of negative energy states. The vacuum or ground state has all the negative energy states full. An additional electron must now occupy a positive energy state since the Pauli exclusion principle forbids it from falling into one of the filled negative energy states. On promoting one of these negative energy states to a positive energy one, by supplying energy, an electron-hole pair is created, i.e. a positive energy electron and a hole in the negative energy sea. The hole is seen in nature as a positive energy positron. This was a radical new idea, and brought pair creation and antiparticles into physics. Positrons were discovered in cosmic rays by Carl Anderson in 1932.

The problem with Dirac's hole theory is that it doesn't work for bosons. Such particles have no exclusion principle to stop them falling into the negative energy states, releasing their energy.

Recall that Feynman tells us to keep both types of free particle solution. One is to be used for particles and the other for the accompanying antiparticles. Let's return to our spinor solutions and introduce some basis spinors. Take the positive energy solution of equation (2.12) and define

$$\sqrt{E+m} \begin{pmatrix} \chi_r \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi_r \end{pmatrix} e^{-ip\cdot x} \equiv u_r(p)e^{-ip\cdot x}. \tag{2.14}$$

For the negative energy solution of equation (2.13), change the sign of the energy, $E \to -E$, and the three-momentum, $\mathbf{p} \to -\mathbf{p}$, to obtain,

$$\sqrt{E+m} \begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi_r \\ \chi_r \end{pmatrix} e^{ip\cdot x} \equiv v_r(p)e^{ip\cdot x}. \tag{2.15}$$

In these two solutions E is now (and for the rest of the course) always positive and given by $E = (\mathbf{p}^2 + m^2)^{1/2}$. The subscript r takes the values $1, 2$, with

$$\chi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \chi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.16}$$

For the simple case $\mathbf{p} = 0$ we may interpret χ_1 as the spin-up state and χ_2 as the spin-down state. Thus for $\mathbf{p} = 0$ the 4-component wavefunction has a very simple interpretation: the first two components describe electrons with spin-up and spin-down, while the second two components describe positrons with spin-up and spin-down. Thus we understand on physical grounds why the wavefunction had to have four components. The general case $\mathbf{p} \neq 0$ is slightly more involved and is considered in the next section.

The u-spinor solutions will correspond to particles and the v-spinor solutions to antiparticles. The role of the two χ's will become clear in the following section, where it will be shown that the two choices of r are spin labels. Note that each spinor solution depends on the three-momentum $\mathbf{p}$, so it is implicit that $p^0 = E$.

2.2 Free Particle Solutions II: Spin

Now it's time to justify the statements we have been making that the Dirac equation describes spin-1/2 particles. The Dirac Hamiltonian in momentum space is given in equation (2.1) as

$$H_D = \boldsymbol{\alpha}\cdot\mathbf{P} + \beta m \tag{2.17}$$

and the orbital angular momentum operator is

$$\mathbf{L} = \mathbf{R} \times \mathbf{P}.$$

Normally you have to worry about operator ordering ambiguities when going from classical objects to quantum mechanical ones. For the components of $\mathbf{L}$ there is no ambiguity.

Evaluating the commutator of $\mathbf{L}$ with H_D,

$$\begin{aligned} [\mathbf{L}, H_D] &= [\mathbf{R} \times \mathbf{P}, \boldsymbol{\alpha}\cdot\mathbf{P}] \\ &= [\mathbf{R}, \boldsymbol{\alpha}\cdot\mathbf{P}] \times \mathbf{P} \\ &= i\boldsymbol{\alpha} \times \mathbf{P}, \end{aligned} \tag{2.18}$$

we see that the orbital angular momentum is not conserved (otherwise the commutator would be zero). We'd like to find a *total* angular momentum $\mathbf{J}$ which *is* conserved, by adding an additional operator $\mathbf{S}$ to $\mathbf{L}$,

$$\mathbf{J} = \mathbf{L} + \mathbf{S}, \quad [\mathbf{J}, H_D] = 0. \tag{2.19}$$

To this end, consider the three matrices,

$$\boldsymbol{\Sigma} \equiv \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} = -i\alpha_1\alpha_2\alpha_3\boldsymbol{\alpha}, \tag{2.20}$$

where the first equivalence is merely a definition of $\boldsymbol{\Sigma}$ and the last equality can readily be verified. The $\boldsymbol{\Sigma}/2$ have the correct commutation relations to represent angular momentum, since the Pauli matrices do, and their commutators with $\boldsymbol{\alpha}$ and β are,

$$[\boldsymbol{\Sigma}, \beta] = 0, \qquad [\Sigma_i, \alpha_j] = 2i\epsilon_{ijk}\alpha_k. \tag{2.21}$$

▷ **Exercise 2.3**

Verify the commutation relations in equation (2.21).

From the relations in (2.21) we find that

$$[\boldsymbol{\Sigma}, H_D] = -2i\boldsymbol{\alpha} \times \mathbf{P}.$$

Comparing this with the commutator of $\mathbf{L}$ with H_D in equation (2.18), you readily see that

$$\left[\mathbf{L} + \tfrac{1}{2}\boldsymbol{\Sigma}, H_D\right] = 0,$$

and we can identify

$$\mathbf{S} = \frac{1}{2}\boldsymbol{\Sigma}$$

as the additional quantity which, when added to $\mathbf{L}$ in equation (2.19), yields a conserved total angular momentum $\mathbf{J}$. We interpret $\mathbf{S}$ as an angular momentum *intrinsic* to the particle. Now

$$\mathbf{S}^2 = \frac{1}{4}\begin{pmatrix} \boldsymbol{\sigma}\cdot\boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\boldsymbol{\sigma} \end{pmatrix} = \frac{3}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and recalling that the eigenvalue of $\mathbf{J}^2$ for spin j is $j(j+1)$, we conclude that $\mathbf{S}$ represents spin-1/2 and the solutions of the Dirac equation have spin-1/2 as promised.

We worked in the Dirac representation of the matrices for convenience, but the result is necessarily independent of the representation.

Now consider the u-spinor solutions $u_r(p)$ of equation (2.14). Choose $\mathbf{p} = (0, 0, p_z)$ and write

$$u_\uparrow = u_1(p) = \begin{pmatrix} \sqrt{E+m} \\ 0 \\ \sqrt{E-m} \\ 0 \end{pmatrix}, \qquad u_\downarrow = u_2(p) = \begin{pmatrix} 0 \\ \sqrt{E+m} \\ 0 \\ -\sqrt{E-m} \end{pmatrix}. \tag{2.22}$$

It is easy to see that,

$$S_z u_\uparrow = \frac{1}{2}u_\uparrow, \qquad S_z u_\downarrow = -\frac{1}{2}u_\downarrow.$$

So, these two spinors represent spin up and spin down along the z-axis respectively. For the v-spinors, with the same choice for $\mathbf{p}$, write,

$$v_\downarrow = v_1(p) = \begin{pmatrix} \sqrt{E-m} \\ 0 \\ \sqrt{E+m} \\ 0 \end{pmatrix}, \qquad v_\uparrow = v_2(p) = \begin{pmatrix} 0 \\ -\sqrt{E-m} \\ 0 \\ \sqrt{E+m} \end{pmatrix}, \tag{2.23}$$

where now,

$$S_z v_\downarrow = \frac{1}{2}v_\downarrow, \qquad S_z v_\uparrow = -\frac{1}{2}v_\uparrow.$$

This apparently perverse choice of up and down for the v's is actually quite sensible when one realises that a negative energy electron carrying spin $+1/2$ backwards in time looks just like a positive energy positron carrying spin $-1/2$ forwards in time.

2.3 Normalisation and Gamma Matrices

We have included a normalisation factor $\sqrt{E+m}$ in our spinors. With this factor,

$$u_r(p)^\dagger u_s(p) = v_r(p)^\dagger v_s(p) = 2E\delta^{rs}. \tag{2.24}$$

This corresponds to the standard relativistic normalisation of $2E$ particles per unit volume (we shall justify this a bit later on). It also means that $u^\dagger u$ transforms like the time component of a 4-vector under Lorentz transformations as we will see in section 2.6.

▷ **Exercise 2.4**
Check the normalisation condition for the spinors in equation (2.24).

There is a much more compact way of writing the Dirac equation, which requires that we get to grips with some more notation. Define the *gamma matrices*,

$$\gamma^0 = \beta, \qquad \boldsymbol{\gamma} = \beta\boldsymbol{\alpha}. \tag{2.25}$$

In the Dirac representation,

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \boldsymbol{\gamma} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}. \tag{2.26}$$

In terms of these, the relations between the $\boldsymbol{\alpha}$ and β in equation (2.4) can be written compactly as,

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \tag{2.27}$$

Combinations like $a_\mu\gamma^\mu$ occur frequently and are conventionally written as,

$$\not{a} = a_\mu\gamma^\mu = a^\mu\gamma_\mu,$$

pronounced "a slash." Note that γ^μ is not, despite appearances, a 4-vector. It just denotes a set of four matrices. However, the notation is deliberately suggestive, for when combined with Dirac fields you can construct quantities which transform like vectors and other Lorentz tensors (see the next section).

Let's close this section by observing that using the gamma matrices the Dirac equation (2.2) becomes

$$(i\not{\partial} - m)\psi = 0, \tag{2.28}$$

or in momentum space,

$$(\not{p} - m)\psi = 0. \tag{2.29}$$

The spinors u and v satisfy

$$\begin{aligned} (\not{p} - m)u_r(p) &= 0, \\ (\not{p} + m)v_r(p) &= 0. \end{aligned} \tag{2.30}$$

▷ **Exercise 2.5**
Derive the momentum space equations satisfied by $u_r(p)$ and $v_r(p)$.

2.4 Lorentz Covariance

We want the Dirac equation (2.28) to preserve its form under Lorentz transformations (LT's). Let $\Lambda^\mu{}_\nu$ represent a LT:

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu \tag{2.31}$$

A familiar example of a LT is a boost along the z-axis, for which

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix},$$

with as usual $\beta = v$ (in units of c) and $\gamma = (1-\beta^2)^{-1/2}$. LT's can be thought of as generalised rotations.

The requirement is

$$(i\gamma^\mu\partial_\mu - m)\psi(x) = 0 \quad \longrightarrow \quad (i\gamma^\mu\partial'_\mu - m)\psi'(x') = 0,$$

where $\partial_\mu = \Lambda^\sigma{}_\mu \partial'_\sigma$. This last equality follows because

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \frac{\partial x'^\sigma}{\partial x^\mu}\frac{\partial}{\partial x'^\sigma} = \Lambda^\sigma{}_\mu \frac{\partial}{\partial x'^\sigma}$$

and equation (2.31) has been used in the last step.

We know that 4-vectors get their components mixed up by LT's, so we expect that the components of ψ might get mixed up too:

$$\psi(x) \to \psi'(x') = S(\Lambda)\psi(x) = S(\Lambda)\psi(\Lambda^{-1}x') \tag{2.32}$$

where $S(\Lambda)$ is a 4×4 matrix acting on the spinor index of ψ. Note that the argument $\Lambda^{-1}x'$ is just a fancy way of writing x, so each component of $\psi(x)$ is transformed into a linear combination of components of $\psi(x)$.

It is helpful to recall that, for a vector field, the corresponding transformation is

$$A^\mu(x) \to A'^\mu(x')$$

where $x' = \Lambda x$. This makes sense physically if one thinks of space rotations of a vector field. For example the wind arrows on a weather map of England are an example of a vector field: at each point on the map there is associated an arrow. Consider the wind direction at a particular point on the map, say Abingdon. If the map of England is rotated, then one would expect on physical grounds that the wind vector at Abingdon always point in the same physical direction and have the same length. In order to achieve this, both the vector itself must rotate, and the point to which it is attached (Abingdon) must be correctly identified after the rotation. Thus the vector at the point x' (corresponding to Abingdon in the rotated frame) is equal to the vector at the point x (corresponding to Abingdon in the unrotated frame), but rotated so as to keep the physical sense of the vector the same in the rotated frame (so that the wind always blows towards Oxford, say, in the two frames). Thus having correctly identified the same point in the two frames all we need to do is rotate the vector:

$$A'^\mu(x') = \Lambda^\mu_\nu A^\nu(x).$$

A similar thing also happens in the case of the 4-component spinor field above, except that we do not (yet) know how the components of the wavefunction themselves must transform, i.e. we do not know S.

We now need to figure out what S is. To determine S we rewrite the Dirac equation in terms of the primed variables (just a mathematical substitution):

$$(i\gamma^\mu \Lambda^\sigma{}_\mu \partial'_\sigma - m)\psi(\Lambda^{-1}x') = 0. \tag{2.33}$$

The key step is to realise that matrices, $\Gamma^\sigma \equiv \gamma^\mu \Lambda^\sigma{}_\mu$ satisfy the same anticommutation relations as the γ^μ's in equation (2.27), i.e.

$$\{\Gamma^\mu, \Gamma^\nu\} = 2g^{\mu\nu}. \tag{2.34}$$

▷ **Exercise 2.6**
Check relation (2.34).

This is an important result, since it means that the Γ matrices constitute an acceptable representation of the gamma matrices. Now any two equivalent representations of the gamma matrices are related by a transformation:

$$\Gamma^\mu = M^{-1}(\Lambda)\gamma^\mu M(\Lambda). \tag{2.35}$$

This allows us to rewrite equation (2.33) as

$$(i\gamma^\mu \partial'_\mu - m)M(\Lambda)\psi(\Lambda^{-1}x') = 0.$$

So we see that if $S(\Lambda) = M(\Lambda)$ then it follows that

$$(i\gamma^\mu \partial'_\mu - m)\psi'(x') = 0, \tag{2.36}$$

and the Dirac equation does indeed preserve its form in the primed frame.

We still haven't solved for S explicitly (we have just shown that there is a solution). To find S we need to solve equation (2.35), which may be written as,

$$\gamma^\sigma \Lambda^\mu{}_\sigma = S^{-1}(\Lambda)\gamma^\mu S(\Lambda). \tag{2.37}$$

For an infinitesimal LT, it can be shown that,

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu - \epsilon_{ij}(g^{i\mu}\delta^j{}_\nu - g^{j\mu}\delta^i{}_\nu) \tag{2.38}$$

where ϵ_{ij} is an infinitesimal parameter and the pair (i,j) label the six types of transformation, i.e. 3 boosts and 3 rotations.

For example a boost along the z-axis corresponds to $i=0$, $j=3$, since in this case,

$$\begin{aligned}\Lambda^\mu{}_\nu &= \delta^\mu{}_\nu - \epsilon_{03}(g^{0\mu}\delta^3{}_\nu - g^{3\mu}\delta^0{}_\nu)\\ &= \begin{pmatrix} 1 & 0 & 0 & -\epsilon \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\epsilon & 0 & 0 & 1 \end{pmatrix},\end{aligned}$$

which is the usual matrix for an infinitesimal Lorentz boost, with $\beta = \epsilon_{03}$ and $\gamma = 1$.

The combinations $(i,j) = (0,1)$ and $(0,2)$ corresponds to boosts along the x and y axes respectively. The remaining, combinations $(2,3)$, $(3,1)$ and $(1,2)$, correspond to infinitesimal anti-clockwise rotations through an angle ϵ_{ij} about the x, y and z axes respectively. It's a nice exercise to check this out.

We are at liberty to write

$$S(\Lambda) = 1 + i\epsilon_{ij}s^{ij} \tag{2.39}$$

and our task is to determine the set of matrices s^{ij}. To do this, substitute the expression for S into equation (2.35) (and remember that $S^{-1}(\Lambda) = 1 - i\epsilon_{ij}s^{ij}$). It is then not too hard to show that the solution is

$$s^{ij} = \frac{i}{4}[\gamma^i, \gamma^j] \equiv \frac{1}{2}\sigma^{ij}. \tag{2.40}$$

Here, I have taken the opportunity to define the matrix σ^{ij}. Thus S is given explicitly in terms of gamma matrices for a general LT.

▷ **Exercise 2.7**
Verify that equation (2.35) relating Γ and γ is satisfied by s^{ij} defined through equations (2.39) and (2.40). The result

$$[A,[B,C]] = \{\{A,B\},C\} - \{\{A,C\},B\}$$

might prove useful.

We are aiming to find quantities which are Lorentz invariant, or transform as vectors or tensors under LT's. To this end, we'll find it useful to introduce the Pauli and Dirac adjoints. The Pauli adjoint $\overline{\psi}$ of a *spinor* ψ is defined by

$$\overline{\psi} \equiv \psi^\dagger \gamma^0 = \psi^\dagger \beta. \tag{2.41}$$

The Dirac adjoint of a *matrix* A is defined by

$$(\overline{\psi} A \phi)^* = \overline{\phi}\, \overline{A} \psi. \tag{2.42}$$

For Hermitian γ^0 it is easy to show that

$$\overline{A} = \gamma^0 A^\dagger \gamma^0. \tag{2.43}$$

Some properties of the Pauli and Dirac adjoints are

$$\begin{aligned} \overline{(\lambda A + \mu B)} &= \lambda^* \overline{A} + \mu^* \overline{B}, \\ \overline{AB} &= \overline{B}\,\overline{A}, \\ \overline{A\psi} &= \overline{\psi}\,\overline{A}. \end{aligned}$$

With these definitions, $\overline{\psi}$ transforms as follows under LT's:

$$\overline{\psi} \to \overline{\psi}' = \overline{\psi} S^{-1}(\Lambda) \tag{2.44}$$

▷ **Exercise 2.8**

(1) Verify that $\gamma^{\mu\dagger} = \gamma^0 \gamma^\mu \gamma^0$. This says that $\overline{\gamma}^\mu = \gamma^\mu$.

(2) Using (2.39) and (2.40) verify that $\gamma^0 S^\dagger(\Lambda) \gamma^0 = S^{-1}(\Lambda)$, i.e. $\overline{S} = S^{-1}$. So S is not unitary in general, although it *is* unitary for rotations (when i and j are spatial indices). This is because the rotations are in the unitary $O(3)$ subgroup of the nonunitary Lorentz group. Here you show the result for an infinitesimal LT, but it is true for finite LT's.

(3) Show that $\overline{\psi}$ satisfies the equation

$$\overline{\psi}\,(-i\overleftarrow{\not\partial} - m) = 0$$

where the arrow over $\not\partial$ implies the derivative acts on $\overline{\psi}$.

(4) Hence prove that $\overline{\psi}$ transforms as in equation (2.44).

Note that result (2) of the problem above can be rewritten as $\overline{S}(\Lambda) = S^{-1}(\Lambda)$, and equation (2.35) for the similarity transformation of γ^μ to Γ^μ takes the form,

$$\overline{S} \gamma^\mu S = \Lambda^\mu{}_\nu \gamma^\nu. \tag{2.45}$$

Combining the transformation properties of ψ and $\overline{\psi}$ in equations (2.32) and (2.44) we see that the bilinear $\overline{\psi}\psi$ is Lorentz invariant. In section 2.6 we'll consider the transformation properties of general bilinears.

Let me close this section by recasting the spinor normalisation equations (2.24) in terms of Dirac inner products. The conditions become

$$\begin{aligned} \overline{u}_r(p)u_s(p) &= 2m\delta^{rs} \\ \overline{u}_r(p)v_s(p) &= \overline{v}_r(p)u_s(p) = 0 \\ \overline{v}_r(p)v_s(p) &= -2m\delta^{rs} \end{aligned} \tag{2.46}$$

▷ **Exercise 2.9**
Verify the normalisation properties in the above equations (2.46).

2.5 Parity

In the next section we are going to construct quantities bilinear in ψ and $\overline{\psi}$, and classify them according to their transformation properties under LT's. We normally use LT's which are in the connected Lorentz Group, $SO(3,1)$, meaning they can be obtained by a continuous deformation of the identity transformation. Indeed in the last section we considered LT's very close to the identity in equation (2.38). However, the full Lorentz group consists not only of the $SO(3,1)$ transformations but also includes the discrete operations of parity (space inversion), P, and time reversal, T:

$$\Lambda_P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \Lambda_T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

LT's satisfy $\Lambda^T g \Lambda = g$ (see the preschool problems), so taking determinants shows that $\det\Lambda = \pm 1$. LT's in $SO(3,1)$ have determinant 1, since the identity does, but the P and T operations have determinant -1.

Let's now find the action of parity on the Dirac wavefunction and determine the wavefunction ψ_P in the parity-reversed system. According to the discussion of the previous section, and using the result of equation (2.45), we need to find a matrix S satisfying

$$\overline{S}\gamma^0 S = \gamma^0, \qquad \overline{S}\gamma^i S = -\gamma^i.$$

It's not hard to see that $S = \overline{S} = \gamma^0$ is an acceptable solution, from which it follows that the wavefunction ψ_P is

$$\psi_P(t, \mathbf{x}) = \gamma^0 \psi(t, -\mathbf{x}). \tag{2.47}$$

In fact you could multiply γ^0 by a phase and still have an acceptable definition for the parity transformation.

In the non-relativistic limit, the wavefunction ψ approaches an eigenstate of parity. Since

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

the u-spinors and v-spinors at rest have opposite eigenvalues, corresponding to particle and antiparticle having opposite *intrinsic* parities.

2.6 Bilinear Covariants

Now, as promised, we will construct and classify the bilinears. You might like to ponder why we are so interested in bilinears, i.e. objects which carry no spinor indices and which involve only 2 spinor fields (recall that the starting point of any field theory is writing down a Lagrangian density). To begin, note that by forming products of the gamma matrices it is possible to construct 16 linearly independent 4×4 matrices. Any constant 4×4 matrix can then be decomposed into a sum over these basis matrices. In equation (2.40) we have defined

$$\sigma^{\mu\nu} \equiv \frac{i}{2}[\gamma^\mu, \gamma^\nu],$$

and now it is convenient to define

$$\gamma_5 \equiv \gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{2.48}$$

where the last equality is valid in the Dirac representation. This new matrix satisfies

$$\gamma_5^\dagger = \gamma_5, \qquad \{\gamma_5, \gamma^\mu\} = 0.$$

Now, the set of 16 matrices

$$\{1, \gamma_5, \gamma^\mu, \gamma^\mu\gamma_5, \sigma^{\mu\nu}\}$$

form a basis for gamma matrix products.

To see that there are 16 matrices: there is 1 unit matrix, 1 γ_5 matrix, 4 γ^μ matrices and 4 $\gamma^\mu\gamma_5$ matrices, and 6 $\sigma^{\mu\nu}$ matrices (see equation (2.40) for the definition of $\sigma^{\mu\nu}$). Note that $\sigma^{\mu\nu}$ is anti-symmetric to avoid double counting with the unit matrix. A little thought is needed to convince yourself that any constant 4×4 matrix can be obtained from linear combinations of these 16 matrices.

Using the transformations of ψ and $\overline{\psi}$ from equations (2.32) and (2.44), together with the similarity transformation of γ^μ in equation (2.45), the 16 fermion bilinears and their transformation properties can be written as follows:

$$\begin{array}{rcll}
\overline{\psi}\psi & \rightarrow & \overline{\psi}\psi & \text{S scalar} \\
\overline{\psi}\gamma_5\psi & \rightarrow & \det(\Lambda)\,\overline{\psi}\gamma_5\psi & \text{P pseudoscalar} \\
\overline{\psi}\gamma^\mu\psi & \rightarrow & \Lambda^\mu{}_\nu\overline{\psi}\gamma^\nu\psi & \text{V vector} \\
\overline{\psi}\gamma^\mu\gamma_5\psi & \rightarrow & \det(\Lambda)\,\Lambda^\mu{}_\nu\overline{\psi}\gamma^\nu\gamma_5\psi & \text{A axial vector} \\
\overline{\psi}\sigma^{\mu\nu}\psi & \rightarrow & \Lambda^\mu{}_\lambda\Lambda^\nu{}_\sigma\overline{\psi}\sigma^{\lambda\sigma}\psi & \text{T tensor}
\end{array} \tag{2.49}$$

▷ **Exercise 2.10**

Verify the transformation properties of the bilinears in equation (2.49).

Observe that $\overline{\psi}\gamma^\mu\psi = (\rho, \mathbf{J})$ is just the current we found earlier in equation (2.7).

2.7 Charge Conjugation

There is one more discrete invariance of the Dirac equation in addition to parity. It is charge conjugation, which takes you from particle to antiparticle and vice versa. For scalar fields the symmetry is just complex conjugation, but in order for the charge conjugate Dirac field to remain a solution of the Dirac equation, you have to mix its components as well:

$$\psi \to \psi_C = C\overline{\psi}^T.$$

Here $\overline{\psi}^T = \gamma^{0T}\psi^*$ and C is a matrix satisfying the condition

$$C\gamma_\mu^T C^{-1} = -\gamma_\mu.$$

In the Dirac representation,

$$C = i\gamma^2\gamma^0 = \begin{pmatrix} 0 & -i\sigma^2 \\ -i\sigma^2 & 0 \end{pmatrix}.$$

I refer you to textbooks such as [1] for details.

When Dirac wrote down his equation everybody thought parity and charge conjugation were exact symmetries of nature, so invariance under these transformations was essential. Now we know that neither of them, nor the combination CP, are respected by the standard electroweak model.

2.8 Neutrinos

In the particle data book you will find only upper limits for the masses of the three neutrinos, and in the standard model they are massless. Let's look therefore at solutions of the Dirac equation with $m = 0$. From equation (2.9) we have in this case

$$E\phi = \boldsymbol{\sigma}\cdot\mathbf{p}\,\chi, \qquad E\chi = \boldsymbol{\sigma}\cdot\mathbf{p}\,\phi. \tag{2.50}$$

These equations can easily be decoupled by taking the linear combinations and defining in a suggestive way the two component spinors ν_L and ν_R,

$$\nu_R \equiv \chi + \phi, \quad \nu_L \equiv \chi - \phi \tag{2.51}$$

which leads to

$$E\nu_R = \boldsymbol{\sigma}\cdot\mathbf{p}\,\nu_R, \qquad E\nu_L = -\boldsymbol{\sigma}\cdot\mathbf{p}\,\nu_L. \tag{2.52}$$

Since $E = |\mathbf{p}|$ for massless particles, these equations may be written

$$\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{|\mathbf{p}|}\nu_L = -\nu_L, \qquad \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{|\mathbf{p}|}\nu_R = \nu_R \tag{2.53}$$

Now $\frac{1}{2}\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{|\mathbf{p}|}$ is known as the *helicity* operator (i.e. it is the spin operator projected in the direction of motion of the momentum of the particle) we see that the ν_L corresponds to solutions with negative helicity, while ν_R corresponds to solutions with positive helicity. In other words ν_L describes a left-handed neutrino while ν_R describes a right-handed neutrino, and each type of neutrino is described by a two-component spinor.

The two-component spinors describing neutrinos transform very simply under LT's,

$$\nu_L \to e^{\frac{i}{2}\boldsymbol{\sigma}\cdot(\boldsymbol{\theta}-i\boldsymbol{\phi})}\nu_L \tag{2.54}$$

$$\nu_R \to e^{\frac{i}{2}\boldsymbol{\sigma}.(\boldsymbol{\theta}+i\boldsymbol{\phi})}\nu_R \tag{2.55}$$

where $\boldsymbol{\theta} = \mathbf{n}\theta$ corresponds to space rotations through an angle θ about the unit $\mathbf{n}$ axis, and $\boldsymbol{\phi} = \mathbf{v}\phi$ corresponds to Lorentz boosts along the unit vector $\mathbf{v}$ with a speed $v = \tanh\phi$. Note that these transformations are consistent with the fact that it is not possible to boost past a massless particle (i.e. its helicity cannot be reversed).

However, under parity transformations they become transformed into each other:

$$\nu_L \leftrightarrow \nu_R. \tag{2.56}$$

So a theory which involves only ν_L without ν_R (such as the standard model) manifestly violates parity.

Although massless neutrinos can be described very simply using two component spinors as above, they may also be incorporated into the four-component formalism as follows. From equation (2.2) we have, in momentum space,

$$|\mathbf{p}|\psi = \boldsymbol{\alpha}\cdot\mathbf{p}\,\psi.$$

For such a solution,

$$\gamma_5\psi = \gamma_5\frac{\boldsymbol{\alpha}\cdot\mathbf{p}}{|\mathbf{p}|}\psi = 2\frac{\mathbf{S}\cdot\mathbf{p}}{|\mathbf{p}|}\psi,$$

using the spin operator $\mathbf{S} = \frac{1}{2}\boldsymbol{\Sigma} = \frac{1}{2}\gamma_5\boldsymbol{\alpha}$, with $\boldsymbol{\Sigma}$ defined in equation (2.20). But $\mathbf{S}\cdot\mathbf{p}/|\mathbf{p}|$ is the projection of spin onto the direction of motion, i.e. the helicity, and is equal to $\pm 1/2$. Thus $(1+\gamma_5)/2$ projects out the neutrino with helicity $1/2$ (right handed) and $(1-\gamma_5)/2$ projects out the neutrino with helicity $-1/2$ (left handed):

$$\frac{(1+\gamma_5)}{2}\psi \equiv \psi_R, \quad \frac{(1-\gamma_5)}{2}\psi \equiv \psi_L, \tag{2.57}$$

define the four-component spinors ψ_R and ψ_L.

To date, only left handed neutrinos have been observed, and only left handed neutrinos appear in the standard model. Since

$$\gamma^0\frac{1}{2}(1-\gamma_5)\psi = \frac{1}{2}(1+\gamma_5)\gamma^0\psi,$$

any theory involving only left handed neutrinos necessarily violates parity (as we saw before in the two-component formalism).

Finally note that in the Dirac representation which we have been using,

$$\gamma^5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{2.58}$$

and the relation between the two-component and four-component formalisms is via the change of variables in equation (2.51). However there exists a representation in which this change of variables is done automatically and the (massless) Dirac equation falls apart into the two two-component equations discussed above. In this chiral representation,

$$\gamma^5 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{2.59}$$

and hence,

$$\frac{(1+\gamma_5)}{2}\psi = \begin{pmatrix} 0 \\ \nu_R \end{pmatrix}, \quad \frac{(1-\gamma_5)}{2}\psi = \begin{pmatrix} \nu_L \\ 0 \end{pmatrix}. \tag{2.60}$$

We have identified ν_R and ν_L as the two-component spinors discussed previously. These results are also applicable to the electron in the approximation that its mass is neglected, by the simple transcription $\nu_R \to e_R$, $\nu_L \to e_L$. In fact in the standard model the electrons start out massless, so these results will be of use to Tim Morris in his course.

> The standard model (and the minimal supersymmetric standard model) contains only left handed massless neutrinos, and neutrino mass terms are forbidden by gauge symmetry, at least given the limited number of fields present in the standard model. If extra fields (e.g. right handed neutrinos) are added then neutrino masses become possible. If neutrino oscillations are confirmed as the solution to the solar neutrino problem, or are discovered in laboratory experiments, then such a modification would become a necessity.

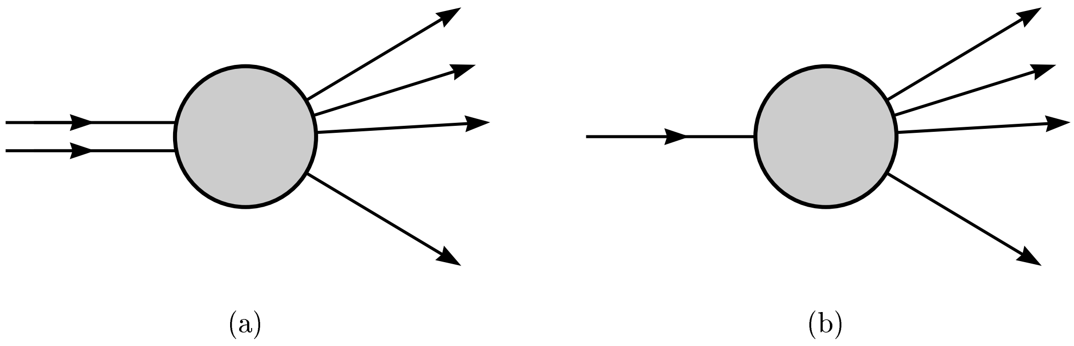


Figure 3.1 Scattering (a) and decay (b) processes.

3 Cross Sections and Decay Rates

Dave Dunbar has already discussed how to compute scattering amplitudes in quantum field theory (see Section 4 of his notes), i.e. how to compute the matrix element

$$i\mathcal{M}_{fi}(2\pi)^4\delta^4(P_f - P_i) \stackrel{T\to\infty}{=} \langle f\ ; +T/2|\,\hat{U}(-T/2, T/2)|\, i\ ; -T/2\rangle, \qquad (3.1)$$

where $\hat{U}(-T/2, T/2)$ is the operator which determines how the initial state i makes the transition to the final state f, i.e.

$$\hat{U}(-T/2, T/2) = \mathrm{T}\exp\left[-i\int_{-T/2}^{T/2} \hat{H}_I\ dt\right]. \qquad (3.2)$$

As Dave has illustrated, it's quite a lengthy procedure to derive particular scattering amplitudes from first principles (i.e. expanding in a perturbation series, using Wicks theorem etc.). Fortunately, there is a quicker way, which involves using Feynman rules. Later on we'll actually calculate some scattering amplitudes for QED, but only after Dave has illustrated how to get at the Feynman rules.

For now let us assume that we've done the work and have computed $\mathcal{M}_{fi}$. Our task in this section is to convert this into a scattering cross section (relevant if there is more than 1 particle in the initial state) or a decay rate (relevant if there is just 1 particle in the initial state), see Figure 3.1.

The probability for the transition to occur is the square of the matrix element, i.e.

$$\text{Probability } = [i\mathcal{M}_{fi}(2\pi)^4\delta^4(P_f - P_i)]^2. \qquad (3.3)$$

Attempting to take the squared modulus of the amplitude produces a meaningless square of a delta function. This is a technical problem because our amplitude is expressed between plane wave states. These states are states of definite momentum and so extend throughout all of space-time. In a real experiment the incoming and outgoing states are localised (e.g. they might leave tracks in a detector). To deal with this properly you can construct normalised wavepacket states which do become well separated in the far past and the far future. We won't deal with wavepackets, instead we'll put our system in a box of volume $V = L^3$. We also imagine that the interaction is restricted to act only over a time of order T. The final answers come out independent of V and T, reproducing the ones we would get if we worked with localised wavepackets. We are in good company here: Nobel Laureate Steven Weinberg says in his recent book, when discussing cross sections and decay rates, "... (as far as I know) no interesting open problems in physics hinge on getting the fine points right regarding these matters."

In infinite spacetime with plane wave states the transition amplitude from i to f is given by (3.1). However in our box of finite size L and for our finite time T the amplitude is given by equation (3.1) but with the Dirac delta functions replaced by well behaved functions:

$$(2\pi)^4\delta^4(P_f - P_i) \to I(E_f - E_i, T)I^3(\mathbf{P}_f - \mathbf{P}_i, L) \tag{3.4}$$

where for example,

$$I(E_f - E_i, T) = \frac{1}{(\frac{E_f - E_i}{2})} \sin\left(\frac{(E_f - E_i)T}{2}\right). \tag{3.5}$$

This function has the property that, as $T \to \infty$,

$$I(E_f - E_i, T) \to 2\pi\delta(E_f - E_i) \tag{3.6}$$

and also

$$I^2(E_f - E_i, T) \to 2\pi T\delta(E_f - E_i) \tag{3.7}$$

with analogous results for $I(\mathbf{P}_f - \mathbf{P}_i, L)$. Thus in our space-time box we have the approximate result,

$$\left|(2\pi)^4\delta^4(P_f - P_i)\right|^2 \simeq VT\,(2\pi)^4\delta^4(P_f - P_i). \tag{3.8}$$

Now, there is a further subtle point which needs to be addressed. If we have chosen to normalise the fields so as to correspond to $2E$ particles per unit volume (as we did for the spinor fields earlier), then we need to get rid of this factor by dividing the amplitude squared by $2EV$ per particle.

To see this note that the Dirac probability density, $\rho = \psi^\dagger\psi$, when integrated over the volume of the box gives

$$\int_{\text{box}} u^\dagger u = 2EV$$

and we have used equation (2.24).

The transition rate, i.e. the probability per unit time, is thus

$$\frac{1}{T}|\mathcal{M}_{fi}|^2 VT(2\pi)^4\delta^4(P_f - P_i)\prod_{f=1}^{N}\left[\frac{1}{2E_f V}\right]\prod_{\text{in}}\left[\frac{1}{2E_i V}\right]. \tag{3.9}$$

3.0.1 The Number of Final States

For a single particle final state, the number of available states dn in some momentum range $\mathbf{k}$ to $\mathbf{k} + d\mathbf{k}$ is, in the box normalisation,

$$dn = \frac{d^3\mathbf{k}}{(2\pi)^3}V. \tag{3.10}$$

This result is proved by recalling that the allowed momenta in the box have components which can only take on discrete values such as $k_x = 2\pi n_x/L$ where n_x is an integer. Thus $dn = dn_x dn_y dn_z$ and the result follows.

For a two particle final state we have

$$dn = dn_1 dn_2$$

where

$$dn_1 = \frac{d^3\mathbf{k}_1}{(2\pi)^3} V, \quad dn_2 = \frac{d^3\mathbf{k}_2}{(2\pi)^3} V,$$

where dn is the number of final states in some momentum range $\mathbf{k}_1$ to $\mathbf{k}_1 + d\mathbf{k}_1$ for particle 1 and $\mathbf{k}_2$ to $\mathbf{k}_2 + d\mathbf{k}_2$ for particle 2. There is an obvious generalisation to an N particle final state,

$$dn = \prod_{f=1}^{N} \frac{d^3\mathbf{k}_f\, V}{(2\pi)^3}. \tag{3.11}$$

3.0.2 Lorentz Invariant Phase Space (LIPS)

The transition rate for transitions into a particular element of final state phase space is thus given by, using equations (3.11) and (3.9),

$$dW = |\mathcal{M}_{fi}|^2 (2\pi)^4 \delta^4(P_f - P_i) V \prod_{f=1}^{N} \left[\frac{1}{2E_f V}\right] \prod_{\text{in}} \left[\frac{1}{2E_i V}\right] \prod_{f=1}^{N} \frac{d^3\mathbf{k}_f\, V}{(2\pi)^3}. \tag{3.12}$$

This can be re-written as

$$dW = |\mathcal{M}_{fi}|^2 V \prod_{\text{in}} \left[\frac{1}{2E_i V}\right] \times (LIPS)\ , \tag{3.13}$$

where the LIPS is,

$$LIPS = (2\pi)^4 \delta^4(P_f - P_i) \prod_{f=1}^{N} \frac{d^3\mathbf{k}_f}{(2\pi)^3\, 2E_f}. \tag{3.14}$$

Observe that everything in the transition rate is Lorentz invariant save for the initial energy factor and the factors of V (using $d^3k\,/2E = d^4k\,\delta^4(k^2 - m^2)\theta(k^0)$, which is manifestly Lorentz invariant, where $E = (\mathbf{k}^2 + m^2)^{1/2}$). For a one particle initial state the factor of V cancels, and we can breath a sigh of relief (after all we would not expect physical quantities to depend on the size of our artificial box). For a two initial particle scattering situation the factors of V will also cancel in the physical cross section as we will show in the next section.

▷ **Exercise 3.1**
Show that the expression for two-body phase space in the centre of mass frame is given by

$$\frac{d^3\mathbf{k}_1}{(2\pi)^3\, 2E_1} \frac{d^3\mathbf{k}_2}{(2\pi)^3\, 2E_2} (2\pi)^4 \delta^4(P - k_1 - k_2) = \frac{1}{32\pi^2 s} \lambda^{1/2}(s, m_1^2, m_2^2) d\Omega^*, \tag{3.15}$$

where $s = P^2$ is the centre of mass energy squared, $d\Omega^*$ is the solid angle element for the angle of one of the outgoing particles with respect to some fixed direction, and

$$\lambda(a, b, c) = a^2 + b^2 + c^2 - 2ab - 2bc - 2ca. \tag{3.16}$$

3.1 Cross Sections

The total cross section for a static target and a beam of incoming particles is defined as the total transition rate for a single target particle and a unit beam flux. The differential cross section is similarly related to the differential transition rate. We have calculated the differential transition rate with a choice of normalisation corresponding to a single "target" particle in the box, and a "beam" corresponding also to one particle in the box. A beam consisting of one particle per volume V with a velocity v has a flux N_0 given by

$$N_0 = \frac{v}{V}$$

particles per unit area per unit time. Thus the differential cross section is related to the differential transition rate in equation (3.13) by

$$d\sigma = \frac{dW}{N_0} = dW \times \frac{V}{v} \tag{3.17}$$

where as promised the factors of V cancel in the cross section.

Now let us generalise to the case where in the frame where you make the measurements. The "beam" has a velocity v_1 but the "target" particles are also moving with a velocity v_2. In a colliding beam experiment for example v_1 and v_2 will point in opposite directions in the laboratory. In this case the definition of the cross section is retained as above, but now the beam flux of particles N_0 is effectively increased by the fact that the target particles are moving towards it. The effective flux in the laboratory in this case is given by

$$N_0 = \frac{|\vec{v}_1 - \vec{v}_2|}{V}$$

which is just the total number of particles per unit area which run past each other per unit time. I denote the velocities with arrows to remind you that they are vector velocities which must be added using the vector law of velocity addition not the relativistic law. In the general case, then, the differential cross section is given by

$$d\sigma = \frac{dW}{N_0} = \frac{1}{|\vec{v}_1 - \vec{v}_2|}\,\frac{1}{4E_1E_2}|\mathcal{M}_{fi}|^2 \times LIPS \tag{3.18}$$

where we have used equation (3.13) for the transition rate, and the box volume V has again cancelled.[1] The amplitude-squared and phase space factors are manifestly Lorentz invariant. What about the initial velocity and energy factors? Observe that

$$E_1E_2(\vec{v}_1 - \vec{v}_2) = E_2\mathbf{p}_1 - E_1\mathbf{p}_2.$$

In a frame where $\mathbf{p}_1$ and $\mathbf{p}_2$ are collinear,

$$|E_2\mathbf{p}_1 - E_1\mathbf{p}_2|^2 = (p_1{\cdot}p_2)^2 - m_1^2m_2^2,$$

and the last expression is manifestly Lorentz invariant. Hence we can define a Lorentz invariant differential cross section. The total cross section is obtained by integrating over the final state phase space:

$$\sigma = \frac{1}{|\vec{v}_1 - \vec{v}_2|}\,\frac{1}{4E_1E_2}\sum\int_{\text{final states}}|\mathcal{M}_{fi}|^2 \times LIPS. \tag{3.19}$$

[1] Because the result is independent of the dimensions of the box, you can think of making the box as large as you like, i.e. big enough to fit your experiment inside. This means that there is no reason to worry about the box.

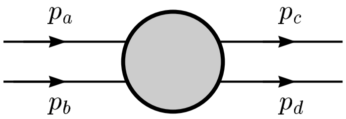


Figure 3.2 $2 \to 2$ scattering.

A slight word of caution is needed in deciding on the limits of integration to get the total cross section. If there are identical particles in the final state then the phase space should be integrated so as not to double count.

3.1.1 Two-body Scattering

An important special case is $2 \to 2$ scattering (see Figure 3.2),

$$a(p_a) + b(p_b) \to c(p_c) + d(p_d).$$

▷ **Exercise 3.2**
Show that in the centre of mass frame the differential cross section is,

$$\frac{d\sigma}{d\Omega^*} = \frac{\lambda^{1/2}(s, m_c^2, m_d^2)}{64\pi^2 s\, \lambda^{1/2}(s, m_a^2, m_b^2)} |\mathcal{M}_{fi}|^2. \tag{3.20}$$

Invariant $2 \to 2$ scattering amplitudes are frequently expressed in terms of the *Mandelstam variables*, defined by

$$\begin{aligned} s &\equiv (p_a + p_b)^2 = (p_c + p_d)^2, \\ t &\equiv (p_a - p_c)^2 = (p_b - p_d)^2, \\ u &\equiv (p_a - p_d)^2 = (p_b - p_c)^2. \end{aligned} \tag{3.21}$$

In fact there are only two independent Lorentz invariant combinations of the available momenta in this case, so there must be some relation between s, t and u.

▷ **Exercise 3.3**
Show that

$$s + t + u = m_a^2 + m_b^2 + m_c^2 + m_d^2.$$

▷ **Exercise 3.4**
Show that, for two body scattering of particles of equal mass m,

$$s \geq 4m^2, \qquad t \leq 0, \qquad u \leq 0.$$

3.2 Decay Rates

With one particle in the initial state the total transition rate is

$$W = \frac{1}{2E} \sum \int_{\text{final states}} |\mathcal{M}_{fi}|^2 \times LIPS.$$

Only the factor $1/2E$ is not manifestly Lorentz invariant. In the rest frame of a particle of mass m we have

$$\Gamma \equiv \frac{1}{2m} \sum \int_{\text{final states}} |\mathcal{M}_{fi}|^2 \times LIPS. \tag{3.22}$$

This is the definition of the "decay rate." In an arbitrary frame we find, $W = (m/E)\Gamma$, which has the expected Lorentz dilation factor. In the master formula (equation (3.13)) this is what the product of $1/2E_i$ factors for the initial particles does.

Actually although the result (3.22) is correct our derivation is not quite right. To get the answer, we needed to consider an initial state at sometime in the distant past, but this state is unstable so it would have decayed before the interaction! I refer to Peskin & Schroeder for a nice discussion of this subtlety.

4 Quantum Electrodynamics

4.1 Quantising the Dirac Field

Dirac Field Theory is defined to be the theory whose field equations correspond to the Dirac equation. We regard the two Dirac fields $\psi(x)$ and $\overline{\psi}(x)$ as being dynamically independent fields and postulate the Dirac Lagrangian density:

$$\mathcal{L} = \overline{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x). \tag{4.1}$$

It's easy to show that the Euler-Lagrange equation

$$\frac{\partial}{\partial x^\mu}\frac{\partial\mathcal{L}}{\partial(\partial_\mu\overline{\psi})} - \frac{\partial\mathcal{L}}{\partial\overline{\psi}} = 0 \tag{4.2}$$

leads to the Dirac equation.

The canonical momentum is

$$\pi(x) = \frac{\partial\mathcal{L}}{\partial\dot{\psi}(x)} = i\psi^\dagger(x) \tag{4.3}$$

and the Hamiltonian density is

$$\mathcal{H} = \pi\dot{\psi} - \mathcal{L} = \psi^\dagger i\frac{\partial\psi}{\partial t}. \tag{4.4}$$

Rather than interpret ψ as a wavefunction (and thereby have to keep in mind notions of negative energy states moving backwards in time), we shall follow Dave Dunbar, and regard ψ as a quantum field. We need to quantise this field. Now, naively we would try to impose the usual equal time commutation relations, i.e.

$$[\psi_i(\mathbf{x},t), \pi_j(\mathbf{y},t)] = i\delta_{ij}\delta^3(\mathbf{x}-\mathbf{y}), \tag{4.5}$$

$$[\psi_i(\mathbf{x},t), \psi_j(\mathbf{y},t)] = 0, \tag{4.6}$$

$$[\pi_i(\mathbf{x},t), \pi_j(\mathbf{y},t)] = 0, \tag{4.7}$$

where i and j label the spinor components of ψ and π. This is a recipe for disaster. In particular, there is no ground state, i.e. excitations of the vacuum can have negative energies. The only way to cure the problem is to impose anti-commutation relations:

$$\{\psi_i(\mathbf{x},t), \pi_j(\mathbf{y},t)\} = i\delta_{ij}\delta^3(\mathbf{x}-\mathbf{y}). \tag{4.8}$$

$$\{\psi_i(\mathbf{x},t), \psi_j(\mathbf{y},t)\} = 0, \tag{4.9}$$

$$\{\pi_i(\mathbf{x},t), \pi_j(\mathbf{y},t)\} = 0. \tag{4.10}$$

There is a very nice discussion in Peskin & Schroeder on this (Chapter 3). In particular, they show how anti-commutation relations really are the only solution.

As in Dave's course, the Heisenberg equations of motion for the field operators have solution

$$\psi(\mathbf{x},t) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{1}{2E}\sum_{\alpha=1,2}[b_\alpha(\mathbf{k})u_\alpha(\mathbf{k})e^{-ik.x} + d^\dagger_\alpha(\mathbf{k})v_\alpha(\mathbf{k})e^{ik.x}] \tag{4.11}$$

$$\overline{\psi}(\mathbf{x},t) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{1}{2E}\sum_{\alpha=1,2}[b^\dagger_\alpha(\mathbf{k})\overline{u}_\alpha(\mathbf{k})e^{ik.x} + d_\alpha(\mathbf{k})\overline{v}_\alpha(\mathbf{k})e^{-ik.x}] \tag{4.12}$$

and the anti-commutation relations imply that

$$\left\{b_\alpha(\mathbf{k}), b^\dagger_{\alpha'}(\mathbf{k}')\right\} = (2\pi)^3\, 2E\, \delta^3(\mathbf{k}-\mathbf{k}')\delta_{\alpha\alpha'}, \tag{4.13}$$

$$\left\{d_\alpha(\mathbf{k}), d^\dagger_{\alpha'}(\mathbf{k}')\right\} = (2\pi)^3\, 2E\, \delta^3(\mathbf{k}-\mathbf{k}')\delta_{\alpha\alpha'}, \tag{4.14}$$

$$\{b_\alpha(\mathbf{k}), b_{\alpha'}(\mathbf{k}')\} = 0, \tag{4.15}$$

$$\left\{b^\dagger_\alpha(\mathbf{k}), b^\dagger_{\alpha'}(\mathbf{k}')\right\} = 0, \tag{4.16}$$

$$\{d_\alpha(\mathbf{k}), d_{\alpha'}(\mathbf{k}')\} = 0, \tag{4.17}$$

$$\left\{d^\dagger_\alpha(\mathbf{k}), d^\dagger_{\alpha'}(\mathbf{k}')\right\} = 0. \tag{4.18}$$

The total Hamiltonian is

$$H = \mathrm{N}\int d^3\mathbf{x}\,\mathcal{H} \tag{4.19}$$

The prefix N denotes normal ordering. This is the way we remove the ambiguity associated with the order of operators. Normal ordering means we are to put all creation operators to the left of all the annihilation operators. If this means moving an anticommuting (fermion) operator through another such operator then we need to remember to pick up a minus sign.

After a bit of algebra we can get

$$H = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{1}{2E}E\sum_{\alpha=1,2}[b^\dagger_\alpha(\mathbf{k})b_\alpha(\mathbf{k}) + d^\dagger_\alpha(\mathbf{k})d_\alpha(\mathbf{k})]. \tag{4.20}$$

If we had tried to impose commutation relations, the $dd^\dagger$ term would have entered with a minus sign in front, which would signal that something has gone wrong. In particular, it would mean that $d^\dagger$ creates particles of negative energy. This is not supposed to happen in the quantised field theory. Note that we could not fix the problem by simply re-labelling $d \leftrightarrow d^\dagger$ since that would not be consistent with the commutations relations imposed on ψ and π.

So, in order to quantise the Dirac field we are necessarily led to the introduction of anti-commutation relations. Remarkably we find that we have automatically taken into account the Pauli inclusion principle! For example,

$$\left\{b^\dagger_\alpha(\mathbf{k}), b^\dagger_{\alpha'}(\mathbf{k}')\right\} = 0$$

implies that it's not possible to create two quanta in the same state, i.e.

$$b^\dagger_\alpha(\mathbf{k})b^\dagger_\alpha(\mathbf{k})|0\rangle = 0.$$

This intimate connection between spin and statistics is a direct consequence of desiring our theory to be consistent with the laws of relativity and quantum mechanics.

The charge operator is

$$Q = \mathrm{N}\int d^3\mathbf{x}\; j_0(x) = \int d^3\mathbf{x}\; \psi^\dagger\psi$$

which, in terms of the creation and annihilation operators, is

$$Q = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{1}{2E} E \sum_{\alpha=1,2} [b^\dagger_\alpha(\mathbf{k}) b_\alpha(\mathbf{k}) - d^\dagger_\alpha(\mathbf{k}) d_\alpha(\mathbf{k})] \tag{4.21}$$

which shows that $b^\dagger$ creates fermions while $d^\dagger$ creates the associated anti-fermions of opposite charge.

Using the techniques discussed in Dave's field theory course, we can go ahead and compute the propagator for a Dirac particle:

$$\begin{aligned} S_F(x-y) &= \langle 0|\psi(x)\overline{\psi}(y)|0\rangle \qquad x^0 > y^0, \\ &= -\langle 0|\overline{\psi}(y)\psi(x)|0\rangle \qquad x^0 < y^0. \end{aligned} \tag{4.22}$$

Skipping details (see section 4 of Dave Dunbar's course), this is

$$S_F(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i(\not{p}+m)}{p^2-m^2+i\epsilon} e^{-ip\cdot(x-y)}. \tag{4.23}$$

4.2 Quantising the Electromagnetic Field

Dave showed us how to derive the Maxwell equations from the Lagrangian density

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} - j_\mu A^\mu \tag{4.24}$$

where the field strength tensor is

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{4.25}$$

(See the pre-school problems for an introduction to this way of formulating classical electrodynamics.) He also highlighted the gauge invariance, i.e. that Maxwell's equations don't change under the transformation

$$A_\mu(x) \to A_\mu(x) + \partial_\mu \Lambda(x) \tag{4.26}$$

where $\Lambda(x)$ is some scalar field. This gauge invariance allows us to impose the Lorentz gauge condition, i.e. without loss of generality we can fix

$$\partial_\mu A^\mu = 0. \tag{4.27}$$

Note that, even after fixing the Lorentz gauge, we can perform another gauge transformation on A_μ, i.e. $A_\mu(x) \to A_\mu(x) + \partial_\mu\chi(x)$ where $\chi(x)$ must satisfy the wave equation, $\partial_\mu\partial^\mu\chi = 0$.

Immediately we try to quantise the electromagnetic field we hit a problem. To see this note that the canonically conjugate field to A_μ is

$$\Pi^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_0 A_\mu)} = F^{\mu 0} \tag{4.28}$$

and from this it follows that $\Pi^0 = 0$. This means we have no possibility to impose a non-zero commutation relation between Π^0 and A^0, which we would need if we are to quantise the field.

Fortunately, all is not lost. Let us incorporate the gauge condition into the Lagrangian density, i.e. write

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - j_\mu A^\mu - \frac{1}{2\xi}(\partial_\mu A^\mu)^2. \tag{4.29}$$

The new term fixes the gauge and ξ is a dimensionless Lagrange multiplier (as such it can take on any value we choose). At first sight it doesn't look like we've done anything useful since $\Pi^\mu = F^{\mu 0} - (1/\xi)g^{\mu 0}(\partial_\nu A^\nu)$ and so $\Pi^0 = 0$. This is certainly true classically however, we need to be a bit more careful with the quantum theory. How are we supposed to interpret the Lorentz gauge condition? If we assume it means that the operator $\partial_\mu A^\mu$ vanishes then we can't quantise. However, this is too restrictive. We need only ensure that the gauge condition holds for matrix elements of $\partial_\mu A^\mu$ and now we can impose non-zero commutation relations. The quantisation condition then leads to

$$[A^\mu(\mathbf{x},t), \partial_0 A^\mu(\mathbf{y},t)] = -ig^{\mu\nu}\delta^3(\mathbf{x}-\mathbf{y}) \tag{4.30}$$

with all other commutators vanishing. The Heisenberg operator corresponding to the photon field is (putting $\xi = 1$)[2]

$$A_\mu(x) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{1}{2E}\sum_{\lambda=0}^{3}\left[\epsilon_\mu^\lambda(k)a_\lambda(k)e^{-ik\cdot x} + \epsilon_\mu^\lambda(k)^* a_\lambda(k)^\dagger e^{ik\cdot x}\right] \tag{4.31}$$

where ϵ_μ^λ are a set of four linearly independent basis vectors ($\lambda = 0,1,2,3$). For example we might choose $\epsilon^0 = (1,0,0,0)$, $\epsilon^1 = (0,1,0,0)$, $\epsilon^2 = (0,0,1,0)$ and $\epsilon^3 = (0,0,0,1)$. If $\mathbf{k}$ is along the z-axis then ϵ^1 and ϵ^2 are polarisation vectors for transverse polarisations whilst ϵ^0 is referred to as the timelike polarisation vector and ϵ^3 is referred to as the longitudinal polarisation vector.

The commutation relation (4.30) implies that

$$\left[a_\lambda(k), a_{\lambda'}^\dagger(k')\right] = -g_{\lambda\lambda'}\,2E\,(2\pi)^3\delta^3(\mathbf{k}-\mathbf{k}'). \tag{4.32}$$

At a glance this looks fine, i.e. we interpret $a_\lambda^\dagger(k)$ as an operator which creates quanta of the electromagnetic field (photons) with polarisation λ and momentum k. However, for $\lambda = 0$ we have a problem since the sign on the RHS of (4.32) is opposite to that of the other 3 polarisations. This shows up in the fact that these timelike photons make a negative contribution to the energy:

$$H = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{1}{2E}E\left(a_1^\dagger a_1 + a_2^\dagger a_2 + a_3^\dagger a_3 - a_0^\dagger a_0\right). \tag{4.33}$$

Fortunately, although we might not realize it yet, we have already solved the problem! The Lorentz gauge condition implies that, for all physical observables, the contributions from the timelike and longitudinal photons always cancels. More explicitly, by demanding that

$$\langle\psi|\partial_\mu A^\mu|\psi\rangle = 0 \tag{4.34}$$

it follows that

$$\langle\psi|a_3^\dagger a_3 - a_0^\dagger a_0|\psi\rangle = 0. \tag{4.35}$$

This is nice because it is in accord with our knowledge that free photons are transversely polarised.

[2]This is often referred to as the Feynman gauge.

Here is a nice proof that the classical electromagnetic field is polarised transversely. The classical field can be expanded thus

$$A_\mu(x) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{1}{2E}\left[a_\mu(k)e^{-ik\cdot x} + a_\mu(k)^* e^{ik\cdot x}\right] \tag{4.36}$$

and I have absorbed the sum over polarisation vectors into the Fourier coefficients, a_μ. The Lorentz gauge condition says that $k \cdot a = 0$, i.e. $a_0(k) = \hat{\mathbf{k}} \cdot \mathbf{a}(k)$ and hence that the time component of a_μ equals the longitudinal component. So we are already down to 3 independent components. The next step comes on realizing that we have still the residual freedom to shift $A_\mu \to A_\mu + \partial_\mu \chi$ for any χ that satisfies the wave equation. Since χ satisfies the wave equation, we can perform a mode expansion just like we did for the A_μ field. Thus we are free to perform the replacement $a_\mu \to a_\mu + \lambda k_\mu$. For $k_\mu = (k, 0, 0, k)$ and $a_\mu = (a_0, a_1, a_2, a_0)$ the choice $\lambda = -a_0/k$ transforms $a_\mu \to (0, a_1, a_2, 0)$ and we have finished.

Having convinced ourselves that we can go ahead and quantise the Maxwell field, we can now proceed to look for the photon propagator. Again I'm going to skip the details and refer back to Dave Dunbar's course:

$$\begin{aligned} iD_F^{\mu\nu}(x-y) &= \langle 0|\mathrm{T}(A^\mu(x)A^\nu(y))|0\rangle \\ &= -ig^{\mu\nu}\int \frac{d^4k}{(2\pi)^4}\frac{e^{-ik\cdot(x-y)}}{k^2+i\epsilon}. \end{aligned} \tag{4.37}$$

This is the Feynman propagator ($\xi = 1$). Generalising away from $\xi = 1$ is a bit more tricky, it gives

$$iD_F^{\mu\nu}(x-y) = -i\int \frac{d^4k}{(2\pi)^4}\frac{e^{-ik\cdot(x-y)}}{k^2+i\epsilon}\left(g^{\mu\nu} + (\xi - 1)\frac{k^\mu k^\nu}{k^2}\right). \tag{4.38}$$

4.3 Feynman Rules of QED

We are now ready to let our fermions and photons interact with each other. The interaction is described by the Lagrangian

$$\mathcal{L}_{\text{int}} = -e\,\overline{\psi}\gamma^\mu A_\mu \psi. \tag{4.39}$$

Such an interaction may be introduced by the concept of "minimal substitution" familiar from classical electrodynamics:

$$\mathbf{p} \to \mathbf{p} - e\mathbf{A},$$
$$E \to E - e\phi.$$

In four vector notation:

$$p^\mu \to p^\mu - eA^\mu.$$

Applying this classical concept of minimal substitution to the Dirac equation gives

$$(i\not{D} - m)\psi = 0 \tag{4.40}$$

where we have introduced the covariant derivative

$$D_\mu \equiv \partial_\mu + ieA_\mu.$$

The QED Lagrangian describing electrons, photons and their interactions is then given by

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}(\partial_\mu A^\mu)^2 + \overline{\psi}(i\not{D} - m)\psi, \tag{4.41}$$

where $(\partial{\cdot}A)^2/2$ is the gauge fixing term for the Feynman gauge.

The QED Lagrangian is invariant under a symmetry called *gauge symmetry*, which consists of the simultaneous gauge transformations of the photon field:

$$A_\mu \to A_\mu + \partial_\mu \Lambda \tag{4.42}$$

and a phase transformation on the electron field:

$$\psi \to e^{-ie\Lambda}\psi. \tag{4.43}$$

I refer to Dave Dunbar's course for more details on gauge symmetries.

For us, the important thing is that we have got our hands on the QED Lagrangian density. From it we can figure out the Feynman rules. We have already written down the fermion and photon propagators and can more or less write down the other Feynman rules without further ado (I refer back to Dave Dunbar's course, Section 4, for the detailed discussion of how to get Feynman rules from the Lagrangian density). The Feynman rules for QED are summarised in Table 4.1.

The rule for the vertex can be obtained directly from $\mathcal{L}_{\text{int}}$. The external line factors are easily derived by considering simple matrix elements in the operator formalism. They

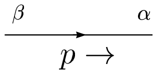
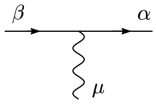
For every ...	draw ...	write ...
Internal photon line		$\dfrac{-ig^{\mu\nu}}{q^2 + i\epsilon}$
Internal fermion line		$\dfrac{i(\not{p} + m)_{\alpha\beta}}{p^2 - m^2 + i\epsilon}$
Vertex		$-ie\gamma^\mu_{\alpha\beta}$
Outgoing electron		$\overline{u}_s(p)$
Incoming electron		$u_s(p)$
Outgoing positron		$v_s(p)$
Incoming positron		$\overline{v}_s(p)$
Outgoing photon		$\epsilon^{*\mu}$
Incoming photon		ϵ^μ
• Attach a directed momentum to every internal line		
• Conserve momentum at every vertex		

Table 4.1 Feynman rules for QED. μ, ν are Lorentz indices and α, β are spinor indices.

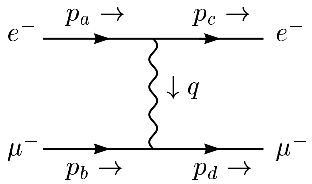


Figure 4.1 Lowest order Feynman diagram for electron–muon scattering.

are what is left behind from the expansions of fields in terms of annihilation and creation operators, after the operators have all been (anti-)commuted until they annihilate the vacuum.

The spinor indices in the Feynman rules are such that matrix multiplication is performed in the opposite order to that defining the flow of fermion number. The arrow on the fermion line itself denotes the fermion number flow, *not* the direction of the momentum associated with the line: I will try always to indicate the momentum flow separately as in Table 4.1. This will become clear in the examples which follow. We have already met the Dirac spinors u and v. I will say more about the photon polarisation vector ϵ when we need to use it.

4.4 Electron–Muon Scattering

To lowest order in the electromagnetic coupling, just one diagram contributes to this process. It is shown in Figure 4.1. The amplitude obtained by applying the Feynman rules to this diagram is

$$i\mathcal{M}_{fi} = (-ie)\,\overline{u}(p_c)\gamma^\mu u(p_a)\left(\frac{-ig_{\mu\nu}}{q^2}\right)(-ie)\,\overline{u}(p_d)\gamma^\nu u(p_b). \tag{4.44}$$

Note that, for clarity, I have dropped the spin label on the spinors, I'll restore it when I need to. In constructing this amplitude we have followed the fermion lines backwards with respect to fermion flow when working out the order of matrix multiplication (which makes sense if you think of an unbarred spinor as a column vector and a barred spinor as a row vector and remember that the amplitude carries no spinor indices).

The cross section involves the squared modulus of the amplitude, which is

$$|\mathcal{M}_{fi}|^2 = \frac{e^4}{q^4}\, L^{\mu\nu}_{(e)} L_{(\mu)\,\mu\nu},$$

where the subscripts e and μ refer to the electron and muon respectively and

$$L^{\mu\nu}_{(e)} = \overline{u}(p_c)\gamma^\mu u(p_a)\overline{u}(p_a)\gamma^\nu u(p_c),$$

with a similar expression for $L^{\mu\nu}_{(\mu)}$.

▷ **Exercise 4.1**
Verify the expression for $|\mathcal{M}_{fi}|^2$.

Usually we have an unpolarised beam and target and do not measure the polarisation of the outgoing particles. Thus we calculate the squared amplitudes for each possible spin combination, then average over initial spin states and sum over final spin states.

Note that we square and then sum since the different spin configurations are in principle distinguishable. In contrast, if several Feynman diagrams contribute to the same process, you have to sum the amplitudes first. We will see examples of this below.

The spin sums are made easy by the following results (I temporarily restore spin labels on spinors):

$$\begin{aligned} \sum_r u_r(p)\,\overline{u}_r(p) &= \not{p} + m, \\ \sum_r v_r(p)\,\overline{v}_r(p) &= \not{p} - m. \end{aligned} \tag{4.45}$$

Don't forget that by m I really mean m times the unit 4×4 matrix.

▷ **Exercise 4.2**
Derive the spin sum relations in equation (4.45).

Using the spin sums we find that

$$\begin{aligned} \frac{1}{4}\sum_{\text{spins}} |\mathcal{M}_{fi}|^2 &= \frac{e^4}{4q^4}\left[\gamma^{\mu}_{ij}(\not{p}_a + m_e)_{jk}\gamma^{\nu}_{kl}(\not{p}_c + m_e)_{li}\right]\left[\gamma_{\mu,ab}(\not{p}_b + m_\mu)_{bc}\gamma_{\nu,cd}(\not{p}_d + m_\mu)_{da}\right] \\ &= \frac{e^4}{4q^4}\,\mathrm{tr}\left(\gamma^{\mu}(\not{p}_a + m_e)\gamma^{\nu}(\not{p}_c + m_e)\right)\,\mathrm{tr}\left(\gamma_{\mu}(\not{p}_b + m_\mu)\gamma_{\nu}(\not{p}_d + m_\mu)\right). \end{aligned} \tag{4.46}$$

Where in the first expression, I chose to make explicit the spinor indices in order that you can see how the trace which appears in the second expression emerges. Since all calculations of cross sections or decay rates in QED require the evaluation of traces of products of gamma matrices, you will generally find a table of "trace theorems" in any quantum field theory textbook [1]. All these theorems can be derived from the fundamental anticommutation relations of the gamma matrices in equation (2.27) together with the invariance of the trace under a cyclic change of its arguments. For now it suffices to use

$$\begin{aligned} \mathrm{tr}(\not{a}\not{b}) &= 4\,a{\cdot}b, \\ \mathrm{tr}(\not{a}\not{b}\not{c}\not{d}) &= 4(a{\cdot}b\,c{\cdot}d - a{\cdot}c\,b{\cdot}d + a{\cdot}d\,b{\cdot}c), \\ \mathrm{tr}(\gamma^{\mu_1}\cdots\gamma^{\mu_n}) &= 0 \quad \text{for } n \text{ odd}. \end{aligned} \tag{4.47}$$

▷ **Exercise 4.3**
Derive the trace results in equation (4.47).

Using these results, and expressing the answer in terms of the Mandelstam variables of equation (3.21), we find

$$\frac{1}{4}\sum_{\text{spins}} |\mathcal{M}_{fi}|^2 = \frac{2e^4}{t^2}\left(s^2 + u^2 - 4(m_e^2 + m_\mu^2)(s+u) + 6(m_e^2 + m_\mu^2)^2\right). \tag{4.48}$$

This can now be used in the $2 \to 2$ cross section formula (3.20) to give, in the high energy limit $(s, |u| \gg m_e^2, m_\mu^2)$,

$$\frac{d\sigma}{d\Omega^*} = \frac{e^4}{32\pi^2 s}\,\frac{s^2 + u^2}{t^2} \tag{4.49}$$

for the differential cross section in the centre of mass frame.

▷ **Exercise 4.4**
Derive the result for the electron–muon scattering cross section in equation (4.49).

Figure 4.2 Lowest order Feynman diagrams for electron–electron scattering.

Other calculations of cross sections or decay rates will follow the same steps we have used above. You draw the diagrams, write down the amplitude, square it and evaluate the traces (if you are using spin sum/averages). There are one or two more wrinkles to be aware of, which we will meet below.

4.5 Electron–Electron Scattering

Since the two scattered particles are now identical, you can't just replace m_μ by m_e in the calculation we did above. If you look at the diagram of Figure 4.1 (with the muons replaced by electrons) you will see that the outgoing legs can be labelled in two ways. Hence we get the two diagrams of Figure 4.2.

The two diagrams give the amplitudes,

$$\begin{aligned} i\mathcal{M}_1 &= \frac{ie^2}{t}\overline{u}(p_c)\gamma^\mu u(p_a)\overline{u}(p_d)\gamma_\mu u(p_b), \\ i\mathcal{M}_2 &= -\frac{ie^2}{u}\overline{u}(p_d)\gamma^\mu u(p_a)\overline{u}(p_c)\gamma_\mu u(p_b). \end{aligned}$$

Notice the additional minus sign in the second amplitude, which comes from the anti-commuting nature of fermion fields. You should accept as part of the Feynman rules for QED that when diagrams differ by an interchange of two fermion lines, a relative minus sign must be included (you don't need to get the absolute sign of an amplitude right, just its sign relative to the other amplitudes). This is important because

$$|\mathcal{M}_{fi}|^2 = |\mathcal{M}_1 + \mathcal{M}_2|^2,$$

so the interference term will have the wrong sign if you don't include the extra sign difference between the two diagrams.

Squaring the amplitude and doing the traces yields (in the limit of negligible fermion masses),

$$\frac{1}{4}\sum_{\text{spins}} |\mathcal{M}_{fi}|^2 = 2e^4\left(\frac{s^2+u^2}{t^2} + \frac{s^2+t^2}{u^2} + \frac{2s^2}{tu}\right). \tag{4.50}$$

4.6 Electron–Positron Annihilation

4.6.1 $e^+e^- \rightarrow e^+e^-$

For this process the two diagrams are shown in Figure 4.3, with the one on the right known as the annihilation diagram. They are just what you get from the diagrams for electron–electron scattering in Figure 4.2 if you twist round the fermion lines. The fact that the diagrams are related this way implies a relation between the amplitudes. The interchange of incoming particles/antiparticles with outgoing antiparticles/particles is

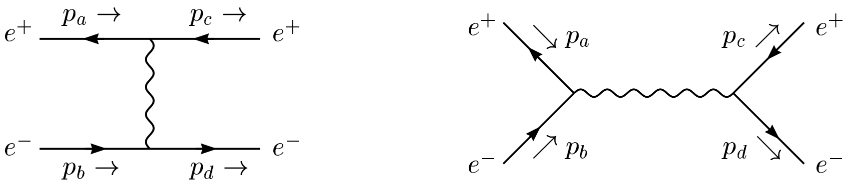


Figure 4.3 Lowest order Feynman diagrams for electron–positron scattering in QED.

called *crossing*. For our particular example, the squared amplitude for $e^+e^- \to e^+e^-$ is related to that for $e^-e^- \to e^-e^-$ by performing the interchange $s \leftrightarrow u$. Hence, squaring the amplitude and doing the traces yields (again neglecting fermion mass terms)

$$\frac{1}{4}\sum_{\text{spins}} |\mathcal{M}_{fi}|^2 = 2e^4\left(\frac{s^2+u^2}{t^2} + \frac{u^2+t^2}{s^2} + \frac{2u^2}{ts}\right). \tag{4.51}$$

4.6.2 $e^+e^- \to \mu^+\mu^-$ and $e^+e^- \to$ hadrons

If electrons and positrons collide and produce muon–antimuon or quark-antiquark pairs, then the annihilation diagram is the only one which contributes. At sufficiently high energies that the quark masses can be neglected, this immediately gives the lowest order QED prediction for the ratio of the annihilation cross section into hadrons to that into $\mu^+\mu^-$:

$$R \equiv \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} = 3\sum_f Q_f^2, \tag{4.52}$$

where the sum is over quark flavours f and Q_f is the quark's charge in units of e. The 3 comes from the existence of three colours for each flavour of quark. Historically this was important: you could look for a step in the value of R as your e^+e^- collider's CM energy rose through a threshold for producing a new quark flavour. If you didn't know about colour, the height of the step would seem too large. Incidentally, another place the number of colours enters is in the decay of a π^0 to two photons. There is a factor of 3 in the amplitude from summing over colours, without which the predicted decay rate would be one ninth of its real size.

At the energies used at LEP you have to remember the diagram with a Z replacing the photon.

▷ **Exercise 4.5**
Show that the cross section for $e^+e^- \to \mu^+\mu^-$ is equal to $4\pi\alpha^2/(3s)$, neglecting the lepton masses.

4.7 Compton Scattering

The diagrams which need to be evaluated to compute the Compton cross section for $\gamma e \to \gamma e$ are shown in Figure 4.4. For unpolarised initial and/or final states, the cross section calculation involves terms of the form

$$\sum_\lambda \epsilon_\lambda^{*\mu}(p)\, \epsilon_\lambda^\nu(p), \tag{4.53}$$

Figure 4.4 Feynman diagrams for Compton scattering.

where λ represents the polarisation of the photon of momentum p. Since the photon is massless, the sum is over the two transverse polarisation states, and must vanish when contracted with p_μ or p_ν. Moreover, since the photon is coupled to the electromagnetic current $J^\mu = \overline{\psi}\gamma^\mu\psi$ of equation (2.7), any term in the polarisation sum (4.53) proportional to p^μ or p^ν does not contribute to the cross section. This is because the current is conserved, $\partial_\mu J^\mu = 0$, so in momentum space $p_\mu J^\mu = 0$. The upshot is that in calculations you can make the replacement

$$\sum_\lambda \epsilon_\lambda^{*\,\mu}(p)\,\epsilon_\lambda^\nu(p) \to -g^{\mu\nu}. \tag{4.54}$$

5 Introduction to Renormalisation

5.1 Renormalisation of QED

Let's start by considering how the electric charge is defined and measured. This will bring up the question of what happens when you try to compute loop corrections. In fact, the expansion in the number of loops is an expansion in Planck's constant $\hbar$, as you can show if you put back the factors of $\hbar$.

The electric charge $\hat{e}$ can be defined as the coupling between an on-shell electron and an on-shell photon: that is, as the vertex on the left hand side of Figure 5.1 with $p_1^2 = p_2^2 = m^2$, where m is the electron mass, and $q^2 = 0$. The Lagrangian parameter e can never be measured in an experiment, since quantum fluctuations are always present. Experiment tells us that

$$\frac{\hat{e}^2}{4\pi} \approx \frac{1}{137}.$$

We call $\hat{e}$ the renormalised coupling constant of QED. We can calculate $\hat{e}$ in terms of e in perturbation theory. To one loop, the relevant diagrams are shown on the right hand side of Figure 5.1, and the result takes the form

$$\hat{e} = e + e^3\Big[a_1 \ln\frac{M^2}{m^2} + b_1\Big] + \cdots \tag{5.1}$$

where a_1 and b_1 are constants obtained from the calculation. The e^3 term is divergent, so we have introduced a cutoff M to regulate it. This is called an ultraviolet divergence since it arises from the propagation of high momentum modes in the loops. The cutoff amounts to selecting only those modes where each component of momentum is less than M in magnitude. Despite the divergence in (5.1), it still relates the measurable quantity $\hat{e}$ to the coupling e we introduced in our theory. This implies that $1/e$ itself must be divergent. For a sensible theory, in any relation between physical quantities the ultraviolet divergences must cancel leaving a relation which is independent of the method used to regulate divergences. This seems a very sensible demand of our theory. Essentially we expect that QED breaks down at very high energies (e.g. when gravitational effects start to become important), i.e. before $M \to \infty$. However, we hope that what is going on at such ultra-short (Planck length) distances does not modify physics as we know it today, e.g. Quantum Gravity does not destroy Coulombs Law! Demanding this of our theory is equivalent to saying that we want our theory to be *renormalisable.*

As an example, consider the amplitude for electron–electron scattering, which we considered at tree level in section 4.5. Some of the contributing diagrams are shown in Figure 5.2, where the crossed diagrams are understood (we showed the crossed tree level diagram explicitly in Figure 4.2). Ultraviolet divergences are again encountered when the diagrams are evaluated, and the result is of the form

$$i\mathcal{M}_{fi} = c_0 e^2 + e^4\Big[c_1 \ln\frac{M^2}{m^2} + d_1\Big] + \cdots \tag{5.2}$$

where c_0, c_1 and d_1 are constants, determined by the calculation. In order to evaluate $\mathcal{M}_{fi}$ numerically, however, we must express it in terms of the known parameter $\hat{e}$. Combining (5.1) and (5.2) yields,

$$i\mathcal{M}_{fi} = c_0\hat{e}^2 + \hat{e}^4\Big[(c_1 - 2a_1c_0)\ln\frac{M^2}{m^2} + d_1 - 2b_1c_0\Big] + \cdots \tag{5.3}$$

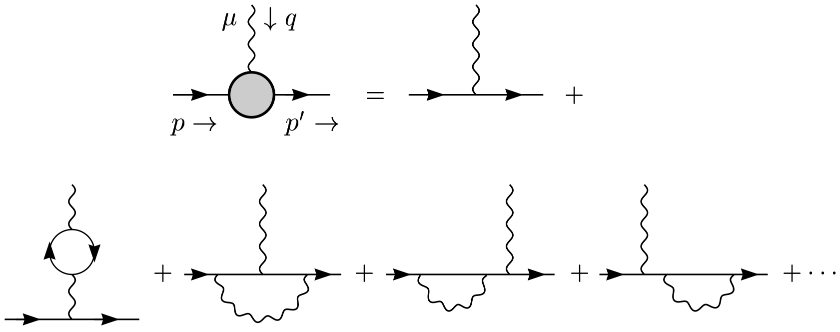


Figure 5.1 Diagrams for vertex renormalisation in QED up to one loop.

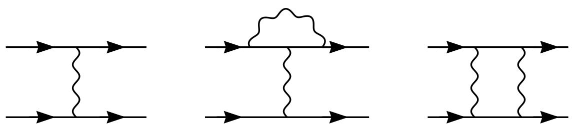


Figure 5.2 Some diagrams for electron–electron scattering in QED up to one loop.

where the ellipsis denotes terms of order $\hat{e}^6$ and above. Since $|\mathcal{M}_{fi}|^2$ is measurable, consistency (renormalisability) requires,

$$c_1 = 2a_1 c_0.$$

This result is indeed borne out by the actual calculations, and the relation between $\mathcal{M}_{fi}$ and $\hat{e}$ contains no divergences:

$$i\mathcal{M}_{fi} = c_0\hat{e}^2 + \hat{e}^4(d_1 - 2b_1c_0) + \mathcal{O}(\hat{e}^6). \tag{5.4}$$

To understand how this cancellation of divergences happened we can study the convergence properties of loop diagrams (although we shall not evaluate them). Consider the third diagram on the right hand side in Figure 5.1 and the middle diagram in Figure 5.2. These both contain a loop with one photon propagator, behaving like $1/k^2$ at large momentum k, and two electron propagators, each behaving like $1/k$. To evaluate the diagram we have to integrate over all momenta, leading to an integral,

$$I \sim \int_{\text{large } k} \frac{d^4k}{k^4}, \tag{5.5}$$

which diverges logarithmically, leading to the $\ln M^2$ terms in (5.1) and (5.2). Notice, however, that the divergent terms in these two diagrams must be the same, since the divergence is by its nature independent of the finite external momenta (the factor of two in equation (5.3) arises because there is a divergence associated with the coupling of each electron in the scattering process). In this way we can understand that at least some of the divergences are common in both (5.1) and (5.2). What about diagrams such as the third box-like one in Figure 5.2? Now we have two photon and two electron propagators, leading to

$$I \sim \int_{\text{large } k} \frac{d^4k}{k^6}.$$

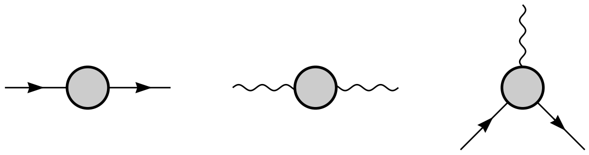


Figure 5.3 Primitive divergences of QED.

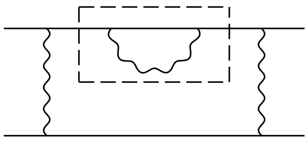


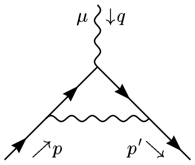
Figure 5.4 Diagram containing a primitive divergence.

This time the integral is convergent.

Detailed study like this reveals that, for QED, ultraviolet divergences always cancel in relations between physically measurable quantities. We discussed above the definition of the physical electric charge $\hat{e}$. A similar argument applies for the electron mass: the Lagrangian bare mass parameter m is divergent, but we can define a finite physical mass $\hat{m}$.

In fact you find that all ultraviolet divergences in QED stem from graphs of the type shown in Figure 5.3 and are known as the *primitive divergences*. Any divergent graph will be found on inspection to contain a divergent subgraph of one of these basic types. For example, Figure 5.4 shows a graph where the divergence comes from the primitive divergent subgraph inside the dashed box. Furthermore, the primitive divergences are always of a type that would be generated by a term in the initial Lagrangian with a divergent coefficient. Hence by rescaling the fields, masses and couplings in the original Lagrangian we can make all physical quantities finite (and independent of the exact details of the adjustment such as how we regulate the divergent integrals). This is what we mean by renormalisability. Put slightly differently, a renormalisable theory is one which needs as input only the values of those observable (i.e. renormalised) parameters which have bare (unobservable) counterparts sitting in the original Lagrangian density. This means that a renormalisable theory has real predictive power.

This should be made clearer by an example. Consider calculating the vertex correction in QED to one loop,



$$= \overline{u}(p')\Big[A\gamma^{\mu} + B\sigma^{\mu\nu}q_{\nu} + Cq^2\gamma^{\mu} + \cdots\Big]u(p).$$

The calculation shows that A is divergent. However, we can absorb this by adding a cancelling divergent coefficient to the $\overline{\psi}\not{A}\psi$ term in the QED Lagrangian (4.41). The B and C terms are finite and unambiguous. This is just as well, since an infinite part of B, for example, would need to be cancelled by an infinite coefficient of a term of the form

$$\overline{\psi}\sigma^{\mu\nu}F_{\mu\nu}\psi,$$

which is not available in (4.41).

In fact, the B term gives the QED correction to the magnetic dipole moment, g, of the electron or muon (see page 160 of the textbook by Itzykson and Zuber [1]). These are predicted to be 2 at tree level. You can do the one-loop calculation (it was first done by Schwinger between September and November 1947 [2]) with a few pages of algebra to find

$$g = 2\left(1 + \frac{\alpha}{2\pi}\right).$$

This gives $g/2 = 1.001161$, which is already impressive compared to the experimental values:

$$\begin{aligned}(g/2)_{\text{electron}} &= 1.001159652193(10),\\ (g/2)_{\text{muon}} &= 1.001165923(8).\end{aligned}$$

Higher order calculations show that the electron and muon magnetic moments differ at two loops and above. Kinoshita and collaborators have devoted their careers to these calculations and are currently at the four loop level. Theory and experiment agree for the electron up to the 11th decimal place!

The C term gives the splitting between the $2s_{1/2}$ and $2p_{1/2}$ levels of the hydrogen atom, known as the Lamb shift. Bethe's calculation [3] of the Lamb shift, done during a train ride to Schenectady in June 1947, was an early triumph for quantum field theory. Here too, the current agreement between theory and experiment is impressive.

In discussing the vertex correction in QED, we said that the divergent part of the A term could be absorbed by adding a cancelling divergent coefficient to the $\overline{\psi}\not{A}\psi$ term in the QED Lagrangian (4.41). When a theory is renormalisable, *all* divergences can be removed in this way. Thus, for QED, if the original Lagrangian is (ignoring the gauge-fixing term),

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\overline{\psi}\not{\partial}\psi - e\overline{\psi}\not{A}\psi - m\overline{\psi}\psi,$$

then redefine everything by:

$$\begin{aligned}\psi &= Z_2^{1/2}\psi_R, & A^\mu &= Z_3^{1/2}A_R^\mu,\\ e = Z_e\hat{e} &= \frac{Z_1}{Z_2 Z_3^{1/2}}\,\hat{e}, & m &= Z_m\hat{m},\end{aligned}$$

where the subscript R stands for "renormalised". In terms of the renormalised fields

$$\mathcal{L} = -\frac{1}{4}Z_3 F_{R\,\mu\nu}F_R^{\mu\nu} + iZ_2\overline{\psi}_R\not{\partial}\psi_R - Z_1\hat{e}\overline{\psi}_R\not{A}_R\psi_R - Z_m Z_2\hat{m}\overline{\psi}_R\psi_R.$$

Writing each Z as $Z = 1 + \delta Z$, re-express the Lagrangian one more time as

$$\mathcal{L} = -\frac{1}{4}F_{R\,\mu\nu}F_R^{\mu\nu} + i\overline{\psi}_R\not{\partial}\psi_R - \hat{e}\overline{\psi}_R\not{A}_R\psi_R - \hat{m}\overline{\psi}_R\psi_R + (\delta Z \text{ terms}).$$

Now it looks like the old Lagrangian, but written in terms of the renormalised fields, with the addition of the δZ *counterterms*. Now when you calculate, the counterterms give you new vertices to include in your diagrams. The divergences contained in the counterterms cancel the infinities produced by the loop integrations, leaving a finite answer.

The old A and ψ are called the *bare* fields, and e and m are the bare coupling and mass.

Note that to maintain the original form of $\mathcal{L}$, you want $Z_1 = Z_2$, so that the $\partial\!\!\!/$ and $\hat{e}A\!\!\!/$ terms combine into a covariant derivative term. This relation does hold, and is a consequence of the electromagnetic gauge symmetry: it is known as the *Ward identity.*

Let me stress again that renormalisation is not about sweeping infinities under the carpet. It is about saying that we don't need to understand physics at the Planck scale in order to interpret LEP data.

5.2 Renormalisation of Quantum Chromodynamics

QCD is a theory of interactions between spin-1/2 quarks and spin-1 gluons. It is a non-Abelian gauge theory based on the group $SU(3)$, with Lagrangian

$$\mathcal{L} = -\frac{1}{4} G^a_{\mu\nu} G^{a\,\mu\nu} + \sum_f \overline{\psi}_f (i D\!\!\!/ - m_f)\psi_f + \begin{array}{l}\text{gauge fixing and}\\ \text{ghost terms}\end{array}. \tag{5.6}$$

Here, a is a colour label, taking values from 1 to 8 for $SU(3)$, and f runs over the quark flavours. The covariant derivative and field strength tensor are given by

$$\begin{aligned} D_\mu &= \partial_\mu - igA^a_\mu T^a, \\ G^a_{\mu\nu} &= \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu, \end{aligned} \tag{5.7}$$

where the f^{abc} are the structure constants of $SU(3)$ and the T^a are a set of eight independent Hermitian traceless 3×3 matrix generators in the fundamental or defining representation (see the pre school problems and the quantum field theory course).

As in QED gauge fixing terms are needed to define the propagator and ensure that only physical degrees of freedom propagate. The gauge fixing procedure is more complicated in the non-Abelian case and necessitates, for certain gauge choices, the appearance of Faddeev–Popov ghosts to cancel the contributions from unphysical polarisation states in gluon propagators. However, the ghosts first appear in loop diagrams, which we will not compute in this course.

There are no Higgs bosons in pure QCD. The only relic of them is in the masses for the fermions which are generated via the Higgs mechanism, but in the electroweak sector of the standard model.

A fundamental difference between QCD and QED is the appearance in the non-Abelian case of interaction terms (vertices) containing gluons alone. These arise from the nonvanishing commutator term in the field strength of the non-Abelian theory in equation (5.7). The photon is electrically neutral, but the gluons carry the colour charge of QCD (specifically, they transform in the adjoint representation). Since the force carriers couple to the corresponding charge, there are no multi photon vertices in QED but there are multi gluon couplings in QCD. This difference is crucial: it is what underlies the decreasing strength of the strong coupling with increasing energy scale.

In QCD, hadrons are made from quarks. Colour interactions bind the quarks, producing states with no net colour: three quarks combine to make baryons and quark–antiquark pairs give mesons. It is generally believed that the binding energy of a quark in a hadron is infinite. This property, called *confinement*, means that there is no such thing as a free quark. Because of asymptotic freedom, however, if you hit a quark with a high energy projectile it might behave in many ways as a free particle. For example, in deep inelastic

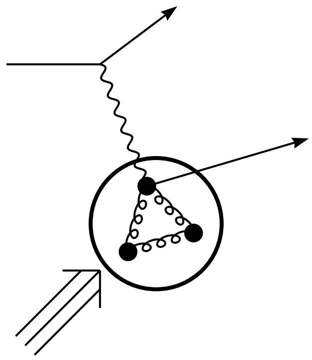


Figure 5.5 Schematic depiction of deep inelastic scattering. An incident lepton radiates a photon which knocks a quark out of a proton. The struck quark is detected indirectly only after hadronisation into observable particles.

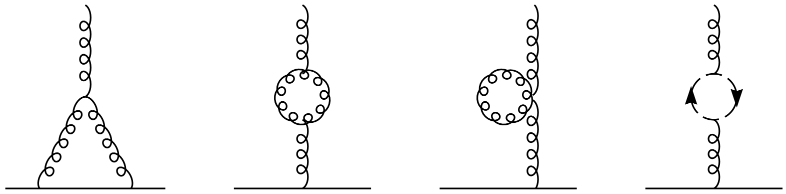


Figure 5.6 Additional diagrams for vertex renormalisation in QCD up to one loop. The dashed line denotes a ghost. For some gauge choices and some regularisation methods not all of these are required.

scattering, or DIS, a photon strikes a quark in a proton, say, imparting a large momentum to it. Some strong interaction corrections to this part of the process can be calculated perturbatively. As the quark heads off out of the proton, however, the low energy strong interactions cut in again and "hadronise" the quark into the particles you actually detect. This is illustrated schematically in Figure 5.5.

We now try to repeat the procedure we used for renormalising the coupling in QED, but this time in QCD, which is also a renormalisable theory. If we define the renormalised coupling $\hat{g}$ as the strength of the quark–gluon coupling, then in addition to the diagrams of Figure 5.1, with the photons replaced by gluons, there are more diagrams at one loop, shown in Figure 5.6. Looking at the second of these new diagrams, it is ultraviolet divergent (containing a $\ln M^2$ term) and is also infrared divergent, since there is no mass to regulate the low momentum modes. In QED all the loop diagrams contain at least one electron propagator and the electron mass provides an infrared cutoff (you still have to worry when the electron is on-shell, but this is not our concern here). In the second diagram of Figure 5.6 there is no quark in the loop. Now we can choose to define the renormalised coupling off-shell, i.e. at some non-zero q^2. The finite value of q^2 provides the infrared regulator and the diagram has a term proportional to $\ln(M^2/q^2)$.

Thus in QCD we can't define a physical coupling constant from an on-shell vertex. This is not really a serious restriction since it's up to us how we define the coupling. We could equally well have defined the QED coupling off-shell, it's just that the value of $1/137$ is easy to extract from low energy experiments which are close to the on-shell limit. Now the renormalised coupling depends on how we define it and therefore on at least one momentum scale (in almost all practical cases, only one momentum scale). The renormalised strong coupling is thus written

$$\hat{g}(q^2).$$

When physical quantities are expressed in terms of $\hat{g}(q^2)$ the coefficients of the perturbation series are finite.

You can define counterterms for QCD in the same way as was demonstrated for QED. Now the gauge coupling g enters in many terms where it has the potential to get renormalised in different ways. This would be a disaster. In fact, the gauge symmetry imposes a set of relations between the renormalisation constants, known as the *Slavnov–Taylor* identities, which generalise the Ward identities of QED.

We have just seen that the renormalised coupling in QCD, $\hat{g}(q^2)$, depends on the momentum at which it is defined. We say it depends on the *renormalisation scale*, and commonly refer to $\hat{g}$ as the "running coupling constant." We would clearly like to know just how $\hat{g}$ depends on q^2, so we calculate the diagrams in Figures 5.1 and 5.6, to get the first terms in a perturbation theory expansion:

$$\hat{g}(\mu) = g + g^3\Big[a_1 \ln\frac{M^2}{\mu^2} + b_1\Big] + \cdots \tag{5.8}$$

where a_1 and b_1 are constants and g is the "bare" coupling from the Lagrangian (5.6). I have switched to using μ^2 in place of q^2, and have written $\hat{g}$ as a function of μ for convenience. From this equation it follows that

$$\mu\,\frac{\partial \hat{g}}{\partial \mu} \equiv \beta(\hat{g}) = -2a_1\,\hat{g}^3 + \cdots \tag{5.9}$$

The discovery by Politzer and by Gross and Wilczek, in 1973, that $a_1 > 0$ led to the possibility of using perturbation theory for strong interaction processes, since it implies that the strong interactions get weaker at high momentum scales: $\hat{g}(\infty) = 0$ is a stable solution of the differential equation (5.9). Keeping just the $\hat{g}^3$ term, we can solve (5.9) to find

$$\alpha_s(\mu) \equiv \frac{\hat{g}^2(\mu)}{4\pi} = \frac{4\pi}{\beta_0 \ln(\mu^2/\Lambda^2)}, \tag{5.10}$$

where Λ is a constant of integration and $\beta_0 = 32\pi^2 a_1$. Thus $\alpha_s(\mu)$ decreases logarithmically with the scale at which it is renormalised, as shown in Figure 5.7. If for some process the natural renormalisation scale is large, there is a chance that perturbation theory will be applicable. The value of β_0 is,

$$\beta_0 = 11 - \frac{2}{3}n_f, \tag{5.11}$$

where n_f is the number of quark flavours. The crucial discovery was the appearance of the "11" coming from the self-interactions of the gluons via the extra diagrams of Figure 5.6. Quarks, and other non-gauge particles, always contribute negatively to β_0. Non-Abelian gauge theories are the only ones we know where you can have asymptotic freedom (providing you don't have too much "matter", e.g. providing that the number of flavours is less than or equal to 16 for QCD).

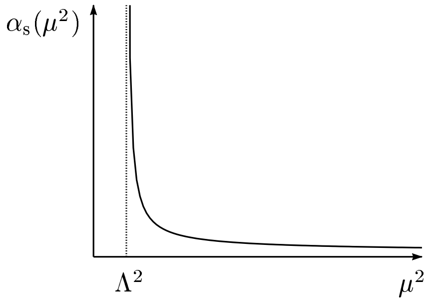


Figure 5.7 Running of the strong coupling constant with renormalisation scale.

What is the significance of the integration constant Λ? The original QCD Lagrangian (5.6) contained only a dimensionless bare coupling g (the quark masses don't matter here, since the phenomenon occurs for a pure glue theory), but now we have a dimensionful parameter. The real answer is that the radiative corrections (in all field theories except finite ones) break the scale invariance of the original Lagrangian. In QED there was an implicit choice of scale in the on-shell definition of $\hat{e}$. Lacking such a canonical choice for QCD, you have to say "measure α_s at $\mu = M_Z$" or "find the scale where $\alpha_s = 0.2$," so that a scale is necessarily involved. The phenomenon was called *dimensional transmutation* by Coleman. Λ is given by

$$\Lambda = \mu \exp\left(- \int^{\hat{g}(\mu)} \frac{dg}{\beta(g)} \right), \tag{5.12}$$

and is μ-independent. The explicit μ dependence is cancelled by the implicit μ dependence of the coupling.

We've seen that the coupling depends on the scale at which it is renormalised. Moreover, there are many ways of defining the renormalised coupling at a given scale, depending on just how you have regulated the infinities in your calculations and which momentum scales you set equal to μ. The value of $\hat{g}(\mu)$ thus depends on the *renormalisation scheme* you pick, and with it, Λ. In practice, the most popular scheme today is called modified minimal subtraction, $\overline{\text{MS}}$, in which integrals are evaluated in $4 - \epsilon$ dimensions and divergences show up as poles of the form ϵ^{-n} for positive integer n. In the particle data book you will find values quoted for $\Lambda_{\overline{\text{MS}}}$ around 200 MeV (it also depends on the number of quark flavours). Don't buy a value of Λ unless you know which renormalisation scheme was used to define it.

In Figure 5.7 you see that the coupling blows up at $\mu = \Lambda$. This is an artifact of using perturbation theory. We can't trust our calculations if $\alpha_s(\mu) > 1$. In practice, you can perhaps use scales for μ down to about 1 GeV, but not much lower, and 2 GeV is probably safer.

▷ **Exercise 5.1**

Extending the expansion of $\hat{g}$ in terms of g in (5.8) to two loops gives

$$\hat{g}(\mu) = g + g^3\Big[a_1 \ln\frac{M^2}{\mu^2} + b_1\Big] + g^5\Big[a_2 \ln^2\frac{M^2}{\mu^2} + b_2 \ln\frac{M^2}{\mu^2} + c_2\Big],$$

with a similar equation for $\hat{g}(\mu_0)$ in terms of g. Renormalisability implies that $\hat{g}(\mu)$ can be expanded in terms of $\hat{g}(\mu_0)$,

$$\hat{g}(\mu) = \sum_{n=0}^{\infty} \hat{g}^{2n+1}(\mu_0) X_n,$$

where the X_n are finite coefficients. Show that this implies that a_2 is determined once the one loop coefficient a_1 is known. In fact a_1 determines all the terms $(\alpha_s \ln \mu)^n$, called the leading logarithms: from a one loop calculation, you can sum up all the leading logarithms.

For QED there is no positive contribution to the beta function, so the electromagnetic coupling has a logarithmic increase with renormalisation scale. However the effect is small even going up to LEP energies: α goes from $1/137$ to about $1/128$. The so called Landau pole, where α blows up, is safely hidden at an enormous energy scale.

▷ **Exercise 5.2**
Any process sensitive to strong interactions is in principle able to measure α_s at any scale. However, in practise there is some optimal choice. Discuss this statement.

Acknowledgements

This lecture course has a long history. I have made a few changes to the course that Steve King gave last year. I'd like to thank Steve for supplying me with the LaTeX files. I'm also very grateful to Jonathan Flynn: Steve's 'supplier'.

Many thanks to Steve Lloyd, for inviting me to come and deliver the course and for doing such a good job as director. Thanks also to Ann Roberts, for organizing the school. I'd also like to say how much I enjoyed myself. I thank the students, tutors and other lecturers for that.

References

[1] I J R Aitchison and A J G Hey *Gauge theories in particle physics* 2nd edition Adam Hilger 1989
J D Bjorken and S D Drell *Relativistic Quantum Mechanics* McGraw-Hill 1964
F Halzen and A D Martin *Quarks and Leptons* Wiley 1984
C Itzykson and J-B Zuber *Quantum Field Theory* McGraw-Hill 1987
F Mandl and G Shaw *Quantum Field Theory* Wiley 1984
M E Peskin and D V Shroeder *An Introduction to Quantum Field Theory* Addison Wesley 1995
L H Ryder *Quantum Field Theory* Cambridge University Press 1985

[2] J Schwinger, Physical Review **73** (1948) 416

[3] H A Bethe, Physical Review **72** (1947) 339

A Pre School Problems

The main aim of this course will be to teach the techniques required for performing simple calculations of amplitudes, cross sections and decay rates, particularly in Quantum Electrodynamics but also in Quantum Chromodynamics. Some aspects of quantum mechanics, special relativity and electrodynamics will be assumed during the lectures at the school. The following problems should be helpful in consolidating your knowledge in these areas. The solutions can be found in many standard textbooks.

Probability Density and Current Density

Starting from the Schrödinger equation for the wave function $\psi(\mathbf{x}, t)$, show that the probability density $\rho = \psi^*\psi$ satisfies the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0$$

where

$$\mathbf{J} = \frac{\hbar}{2im}\left[\psi^*(\nabla\psi) - (\nabla\psi^*)\psi\right]$$

What is the interpretation of $\mathbf{J}$?

Rotations and the Pauli Matrices

Show that a 3-dimensional rotation can be represented by a 3×3 orthogonal matrix R with determinant $+1$ (Start with $\mathbf{x}' = R\mathbf{x}$, and impose $\mathbf{x}' \cdot \mathbf{x}' = \mathbf{x} \cdot \mathbf{x}$). Such rotations form the special orthogonal group, $SO(3)$.

For an *infinitesimal* rotation, write $R = \mathbb{1} + iA$ where $\mathbb{1}$ is the identity matrix and A is a matrix with infinitesimal entries. Show that A is antisymmetric (the i is there to make A hermitian).

Parameterise A as

$$A = \begin{pmatrix} 0 & -ia_3 & ia_2 \\ ia_3 & 0 & -ia_1 \\ -ia_2 & ia_1 & 0 \end{pmatrix} \equiv \sum_{i=1}^{3} a_i L_i$$

where the a_i are infinitesimal and verify that the L_i satisfy the angular momentum commutation relations

$$[L_i, L_j] = i\epsilon_{ijk} L_k$$

Note that the Einstein summation convention is used here. In general, I will switch around between different notational conventions without warning. You should be able to tell from the context what is meant: notation should be your slave, not your master.

The Pauli matrices σ_i are,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Verify that $\frac{1}{2}\sigma_i$ satisfy the same algebra as L_i. If the two-component spinor

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

transforms into $(\mathbb{1} + i\mathbf{a} \cdot \boldsymbol{\sigma}/2)\psi$ under an infinitesimal rotation, check that $\psi^\dagger\psi$ is invariant under rotations.

Raising and Lowering Operators

From the angular momentum commutation relations,

$$[L_i, L_j] = i\epsilon_{ijk}L_k$$

show that the operators

$$L_\pm = L_1 \pm iL_2$$

satisfy

$$[L_+, L_-] = 2L_3$$
$$[L_\pm, L_3] = \mp L_\pm$$

and show that

$$[L^2, L_3] = 0$$

where $L^2 = L_1^2 + L_2^2 + L_3^2$. From the last commutator it follows that there are simultaneous eigenstates of L^2 and L_3. Let ψ_{lm} be such an eigenvector of L^2 and L_3 with eigenvalues $l(l+1)$ and m respectively. Show that each of $L_\pm\psi_{lm}$ either vanishes or is an eigenstate of L^2 with eigenvalue $l(l+1)$ and of L_3 with eigenvalue $m \pm 1$.

Four Vectors

A Lorentz transformation on the coordinates $x^\mu = (ct, \mathbf{x})$ can be represented by a 4×4 matrix Λ as follows:

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu$$

For a boost along the x-axis to velocity v, show that

$$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{A.1}$$

where $\beta = v/c$ and $\gamma = (1-\beta^2)^{-1/2}$ as usual.

By imposing the condition

$$g_{\mu\nu}x'^\mu x'^\nu = g_{\mu\nu}x^\mu x^\nu \tag{A.2}$$

where

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

show that

$$g_{\mu\nu}\Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma = g_{\rho\sigma} \quad \text{or} \quad \Lambda^T g\Lambda = g$$

This is the analogue of the orthogonality relation for rotations. Check that it works for the Λ given by equation (A.1) above.

Now introduce

$$x_\mu = g_{\mu\nu}x^\nu$$

and show, by reconsidering equation (A.2) using $x^\mu x_\mu$, or otherwise, that

$$x'_\mu = x_\nu (\Lambda^{-1})^\nu{}_\mu$$

Vectors A^μ and B_μ that transform like x^μ and x_μ are sometimes called *contravariant* and *covariant* respectively. A simpler pair of names is *vector* and *covector*. A particularly important covector is obtained by letting $\partial/\partial x^\mu$ act on a scalar ϕ:

$$\frac{\partial \phi}{\partial x^\mu} \equiv \partial_\mu \phi$$

Show that ∂_μ does transform like x_μ and not x^μ.

Electromagnetism

The four Maxwell equations are:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \qquad \nabla \cdot \mathbf{B} = 0$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \qquad \nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$$

Which physical laws are represented by each of these equations? Show that

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0$$

and explain the significance of this equation. Verify that it can be written in manifestly covariant form

$$\partial_\mu J^\mu = 0$$

where $J^\mu = (c\rho, \mathbf{J})$.

Introduce scalar and vector potentials ϕ and $\mathbf{A}$ by defining $\mathbf{B} = \nabla \times \mathbf{A}$ and $\mathbf{E} = -\nabla\phi - \partial \mathbf{A}/\partial t$, and recall the gauge invariance of electrodynamics which says that $\mathbf{E}$ and $\mathbf{B}$ are unchanged when

$$\mathbf{A} \to \mathbf{A} + \nabla\Lambda \quad \text{and} \quad \phi \to \phi - \frac{\partial \Lambda}{\partial t}$$

for any scalar function Λ. Using this gauge freedom we can set

$$\nabla \cdot \mathbf{A} = -\frac{1}{c^2}\frac{\partial \phi}{\partial t}$$

Assuming that ϕ and $\mathbf{A}$ can be combined into a four vector $A^\mu = (\phi/c, \mathbf{A})$, this can be written as $\partial_\mu A^\mu = 0$, which is known as the *Lorentz gauge* condition. Defining $\Box \equiv \partial_\mu \partial^\mu$, show that with this condition Maxwell's equations are equivalent to

$$\Box A^\mu = \mu_0 J^\mu$$

The tensor $F_{\mu\nu}$ is defined by

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$$

How many independent components does $F_{\mu\nu}$ have? Rewrite $F_{\mu\nu}$ in terms of $\mathbf{E}$ and $\mathbf{B}$. Show that,

$$F_{\mu\nu}F^{\mu\nu} = -2\left(\frac{\mathbf{E}^2}{c^2} - \mathbf{B}^2\right)$$

$$\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} = -\frac{8}{c}\mathbf{E}\cdot\mathbf{B}$$

where

$$\epsilon^{\mu\nu\rho\sigma} = \begin{cases} +1 & \text{if } \mu\nu\rho\sigma \text{ is an even permutation of } 0123 \\ -1 & \text{if } \mu\nu\rho\sigma \text{ is an odd permutation of } 0123 \\ 0 & \text{otherwise} \end{cases}$$

This gives the relativistic invariants which can be constructed from $\mathbf{E}$ and $\mathbf{B}$.

Group Theory: in Particular $SU(N)$

Unitary matrices U satisfy $U^\dagger U = \mathbb{1}$. Verify that they form a group by showing that $W = UV$ is unitary if U and V are. In general, you should also show that there is an identity element and that every U has an inverse, but these are both obvious. $U(N)$ is the group of complex unitary $N \times N$ matrices and $SU(N)$ is the subgroup of matrices with determinant $+1$.

Let U be a $U(N)$ matrix close to the identity. Write

$$U = \mathbb{1} + iG$$

where G has infinitesimal entries. Show that G is hermitian. If, in addition, U has determinant 1, so $U \in SU(N)$, show that G is traceless.

Any $N \times N$ traceless hermitian matrix can be written as a linear combination of a chosen *basis set*. So, for any G we can choose infinitesimal numbers ϵ_i such that

$$G = \sum_{i=1}^{N^2-1} \epsilon_i T_i$$

where the T_i are our basis. Explain why the summation runs from 1 to $N^2 - 1$.

Show that $[T_i, T_j]$ is antihermitian and traceless, and hence can be written

$$[T_i, T_j] = i f_{ijk} T_k \tag{A.3}$$

for some constants f_{ijk}. The commutation relations between the different T_i define the *Lie algebra* of $SU(N)$. The T_i are called the *generators* and the f_{ijk} are called the *structure constants.*

Find a set of 3 independent 2×2 matrices which are generators for $SU(2)$ and a set of 8 independent 3×3 generators for $SU(3)$.

Verify the *Jacobi identity,*

$$\Big[T_i, [T_j, T_k]\Big] + \Big[T_j, [T_k, T_i]\Big] + \Big[T_k, [T_i, T_j]\Big] = 0$$

and hence show that

$$f_{jkl}f_{ilm} + f_{kil}f_{jlm} + f_{ijl}f_{klm} = 0$$

Define a new set of $(N^2 - 1) \times (N^2 - 1)$ matrices

$$(T^i_{\text{adj}})_{jk} = -i f_{ijk}$$

and show that they obey the same commutation relations as the T_i in equation (A.3). The $T^i_{\rm adj}$ define the *adjoint representation.* The W's of the weak interactions and the gluons of the strong interactions belong to the adjoint representations of $SU(2)_L$, left-handed weak $SU(2)$, and $SU(3)$, the strong interaction colour algebra, respectively.

The generators, and hence the algebra, were found by looking at group elements near the identity. Other group elements can be recovered by combining lots of these infinitesimal "rotations"

$$U = \lim_{N\to\infty} (\mathbb{1} + i\theta_i T_i/N)^N = e^{i\theta_i T_i}$$

where the θ_i are finite. This construction generates what mathematicians call a simply connected group. There is a theorem stating that every Lie algebra comes from exactly one simply connected group: $SU(N)$ and its algebra give us one example.

However, we have seen that both $SU(2)$ and the rotation group $SO(3)$ have the same, angular momentum, algebra. What is going on? It must be that $SO(3)$ is not simply connected. In fact, there is a mapping, called a *covering*, from $SU(2)$ to $SO(3)$ which preserves the group property: that is if $U \in SU(2)$ is mapped to $f(U) \in SO(3)$, then $f(UV) = f(U)f(V)$. In the $SU(2) \to SO(3)$ case, two elements of $SU(2)$ are mapped on to every element of $SO(3)$. Whenever a group G has the same Lie algebra as a simply connected group S there must be such a covering $S \to G$.

The double covering of $SO(3)$ by $SU(2)$ underlies the behaviour of spin-1/2 and other half-odd-integer spin particles under rotations: they really transform under $SU(2)$, and rotating them by 2π only gets you half way around $SU(2)$, so you pick up a minus sign. A second 2π rotation gets you back to where you started.