Relativistic Quantum Mechanics

J.Pearson

November 18, 2009

Abstract

These are a set of notes I have made, based on lectures given by G.Shaw at the University of Manchester Jan-May '09. Please e-mail me with any comments/corrections: jon@jpoffline.com. These notes may be found at www.jpoffline.com.

Contents

Non-relativistic quantum mechanics is concerned with systems for whom the kinetic energy is a lot less than the rest mass, so that

$$T \approx E - mc^2 = \frac{1}{2}mv^2 = \frac{p^2}{2m}.$$

Under this, no particles are created, and everything moves slowly. That is, if a system starts out with one particle, it will always have one particle.

Now, relativistic quantum mechanics is concerned with systems for whom the energy is more than the rest mass,

$$T \neq mv^2.$$

Following this, one can make two distinctions:

$$T < mc^2, \quad T > mc^2.$$

The former case does not have enough energy to create particles, the latter is able to produce particles. The former case requires relativity, but not particle creation. The latter needs the full formalism of quantum field theory.

In discussing a relativistic quantum theory, we are able to understand electron spin, the electrons magnetic moment, predict the existence of anti-particles, the spin-statistics theorem etc.

1 Preliminaries

Let us introduce the notation we shall be adopting.

1.1 Relativistic Notation

Let us introduce the contravariant 4-vector

$$x^\mu = (ct, x, y, z) = (ct, \mathbf{x}).$$

Let us also introduce the metric tensor, such that

$$g_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1);$$

where all off-diagonal components are zero. Using this, we can define a covariant 4-vector,

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

so that

$$x_\mu = (ct, -\mathbf{x}).$$

We shall use the Einstein convention of implied summation over repeated indices.

We further define the inverse metric, as

$$g^{\lambda\mu}g_{\mu\nu} = \delta^{\lambda}_{\nu}.$$

Hence, using these results, we can see

$$g^{\sigma\mu}x_{\mu} = g^{\sigma\mu}g_{\mu\nu}x^{\nu} = \delta^{\sigma}_{\nu}x^{\nu} = x^{\sigma},$$

which is

$$x^{\sigma} = g^{\sigma\mu}x_{\mu}.$$

Therefore, we see that we can use the metric to lower raised indices, and the inverse metric to raise lower indices.

By the form of our metric and inverse, we see that they have the same components,

$$g^{\mu\nu} = g_{\mu\nu}.$$

We will commonly write expressions such as

$$px \equiv p^{\mu}x_{\mu} = Et - \mathbf{p}\cdot\mathbf{x},$$

and $\phi(x) \equiv \phi(t, \mathbf{x})$. Thus, by writing x, we mean the 4-vector (i.e. all four components), and by writing $\mathbf{x}$, only the 3-spatial components.

1.2 The Dirac δ-function

As we shall use the δ-function regularly, it is worth our noting a few of its properties. One of its definitions is that

$$\int_{-\infty}^{\infty} dx\, f(x)\delta(x - x') = f(x'). \tag{1.1}$$

That is, one can think about the δ-function as only returning a value when its argument is zero. In the context of the integral above, the integral "sweeps" over the real line, making the variable x take all possible values. Only when $x = x'$ does the δ-function return a non-zero value. Hence, the only contribution from the integral is at $x = x'$, so that $f(x')$ is returned. Thinking about the δ-function as a filtering mechanism can be very useful. For the δ-function to be used directly in this way, its argument must be of the form $(x - x')$, where x is the integration variable and x' some constant. If instead we have a functional argument, such as $\delta\left(g(x)\right)$, then we must employ a certain tactic to get it into a useable form.

Suppose we have, as a simple case, $\delta(a(x - x'))$, where a is a constant. Then, let us work out how to evaluate

$$\int_{-\infty}^{\infty} dx\, f(x)\delta(a(x - x')). \tag{1.2}$$

Let us first change variables, to

$$y = ax, \quad y' = ax',$$

then, upon substitution into (1.2), we see that

$$\int_{-\infty\times a}^{\infty\times a} \frac{dy}{a} f(\tfrac{y}{a})\delta(y - y').$$

Because we always want this form of the integral (i.e. of the limits), we must take the modulus of a; hence, we have

$$\frac{1}{|a|}\int_{-\infty}^{\infty} dy f(\tfrac{y}{a})\delta(y - y') = \frac{1}{|a|} f(\tfrac{y'}{a}),$$

where we have just used the defining property of the δ-function (1.1). An immediate consequence of this is that the δ-function is even, $\delta(-x) = \delta(x)$. Let us then consider a full functional argument, $\delta(g(x))$. Now, the δ-function only returns a non-zero value when its argument is zero (before this was just at $x = x'$). Hence, we must find the zeros of the function $g(x)$. Supposing that we now have the zeros, at, say, x_i, so that $g(x_i) = 0$, let us make a Taylor expansion of $g(x)$ about the zero,

$$g(x_i + \epsilon) = g(x_i) + (x - x_i)g'(x_i) + \ldots.$$

The first term is zero, by definition of the zero. Hence,

$$g(x) \approx (x - x')g'(x_i).$$

Therefore,

$$\delta(g(x)) = \delta\left(g'(x_i)(x - x_i)\right).$$

Now, as $g'(x_i)$ is "just a number", we can treat it exactly as our a before, so that

$$\delta(g(x)) = \delta\left(g'(x_i)(x - x_i)\right) = \frac{1}{|g'(x_i)|}\delta(x - x_i).$$

If there are multiple zeros of the function $g(x)$, we must sum over all zeros. Hence,

$$\delta(g(x)) = \sum_{x_0} \frac{1}{|g'(x_i)|}\delta(x - x_i), \quad g(x_i) = 0. \tag{1.3}$$

Therefore, we have a (very useful) relation which allows us to recast a δ-function with functional argument, into a useable form.

We may have a multi-dimensional δ-function $\delta^{(n)}(x - x')$, which we define

$$\delta^{(n)}(x - x') = \prod_{i=1}^{n} \delta(x_i - x_i').$$

So, for example,

$$\begin{aligned}\delta^{(4)}(x-x') &= \delta(x^0-x'^0)\delta(x^1-x'^1)\delta(x^2-x'^2)\delta(x^3-x'^3)\\ &= \delta(t-t')\delta^{(3)}(\mathbf{x}-\mathbf{x}').\end{aligned}$$

We can also define the δ-function to be the Fourier transform of unity:

$$\delta^{(4)}(x-x') = \int \frac{d^4k}{(2\pi)^4} e^{-ik(x-x')}.$$

Of course, the inverse Fourier transform is then

$$\delta^{(4)}(k-k') = \int d^4x e^{ix(k-k')}.$$

At this stage we have merely given a few properties of an abstract function; these properties will prove invaluable when discussing quantum fields.

1.3 Homogeneous Lorentz Transformations

Such transformations (between frames that have coinciding origins) are of the general form

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

where $\Lambda^\mu{}_\nu \in \mathbb{R}$ and depend on the relative velocity and relative orientation of the axes. Using previous relations, we see that

$$x_\mu \mapsto x'_\mu = \Lambda_\mu{}^\nu x_\nu,$$

where

$$\Lambda_\mu{}^\nu = g_{\mu\sigma} g^{\nu\tau} \Lambda^\sigma{}_\tau.$$

Now, Lorentz transformations (LTs) leave the interval $s^2 = c^2t^2 - \mathbf{x}^2$ invariant. That is,

$$s^2 = x^\mu x_\mu = x'^\mu x'_\mu.$$

Using these relations, it is simple to see that

$$\Lambda^\mu{}_\nu \Lambda_\mu{}^\sigma x^\nu x_\sigma = x^\mu x_\mu,$$

from which we read off that we require

$$\Lambda^\mu{}_\nu \Lambda_\mu{}^\sigma = \delta^\sigma_\nu.$$

The interval s^2 is an example of a Lorentz scalar – that is, something unchanged by a LT.

A four component object A^μ which transforms like x^μ under an LT, is, by definition, a contravariant 4-vector. That is, if it transforms like

$$A^\mu \mapsto A'^\mu = \Lambda^\mu{}_\nu A^\nu.$$

Similarly, one can define a covariant 4-vector if it transforms as

$$A_\mu \mapsto A'_\mu = \Lambda_\mu{}^\nu A_\nu.$$

The scalar product of any two 4-vectors can be written in a variety of ways:

$$AB = A^\mu B_\mu = A^0 B_0 - \mathbf{A} \cdot \mathbf{B},$$

and also

$$AB = A^\mu B_\mu = g^{\mu\nu} A_\mu B_\nu = g_{\mu\nu} A^\mu A^\nu = A_\mu B^\mu.$$

The scalar product is a Lorentz invariant, since one can easily show that

$$A'^\mu B'_\mu = A^\mu B_\mu.$$

For example, the energy-momentum in the 4-momentum

$$p^\mu = (E/c, \mathbf{p})$$

is a 4-vector. The scalar product with itself is just

$$p^2 = p^\mu p_\mu = \frac{E^2}{c^2} - \mathbf{p}^2 = m^2 c^2,$$

where $m^2 c^2$ is the thing we will call the Lorentz invariant. Hence,

$$p'^2 = p'^\mu p'_\mu = m^2 c^2$$

is a Lorentz invariant.

All equations written consistently in terms of 4-vectors and scalars are automatically Lorentz invariant. That is, those equations for whom indices balance automatically satisfies the principles of special relativity.

So, what about derivatives? Consider a Lorentz scalar function, whereby its value at a point in one coordinate system is the same at the same point in another coordinate system,

$$\phi'(x') = \phi(x).$$

Then,

$$\delta\phi = \frac{\partial \phi}{\partial x^\mu} \delta x^\mu.$$

So, the term on the LHS is a Lorentz scalar (as we defined it to be so). Also, the term on the far RHS is a contravariant 4-vector. So, in order that the LHS is a scalar, we must have that $\frac{\partial \phi}{\partial x^\mu}$ is a covariant 4-vector. Hence,

$$\partial_\mu \phi \equiv \frac{\partial \phi}{\partial x^\mu} = \left(\frac{1}{c} \frac{\partial \phi}{\partial t}, \nabla \phi \right)$$

is a covariant vector. It thus transforms as

$$\partial_\mu \longmapsto \partial'_\mu = \Lambda_\mu{}^\nu \partial_\nu.$$

Similarly, we define the contravariant differential 4-vector

$$\partial^\mu \phi \equiv \frac{\partial \phi}{\partial x_\mu} = \left(\frac{1}{c}\frac{\partial \phi}{\partial t}, -\nabla \phi \right).$$

We further define the D'Alembertian,

$$\Box \equiv \partial_\mu \partial^\mu = \left(\frac{1}{c}\frac{\partial^2}{\partial t^2}, -\nabla^2 \right),$$

which is a Lorentz scalar operator. We can obtain the Lorentz covariant form of the wave equation, if ϕ is a Lorentz scalar;

$$\Box \phi = \partial_\mu \partial^\mu \phi = 0 \quad \Rightarrow \quad \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = \nabla^2 \phi.$$

1.4 Equation of Continuity

Suppose we have some charge density $\rho(\mathbf{x}, t)$, such that

$$Q_\Omega = \int_\Omega dV \rho(\mathbf{x}, t) \tag{1.4}$$

is unchanged under a Lorentz transformation (i.e. is a Lorentz scalar); where Ω is some angular region of space, bounded by the surface S. This could be, for example, mass or electric charge. Then, the charge Q is conserved, provided

$$\partial_\mu j^\mu(x) = 0 \tag{1.5}$$

is satisfied; where

$$j^\mu(x) = (c\rho, \mathbf{j}).$$

To see this continuity equation, rewrite (1.5) using the definition of ∂_μ,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0. \tag{1.6}$$

Then, by (1.4),

$$\begin{aligned} \frac{dQ}{dt} &= \frac{\partial}{\partial t}\int_\Omega dV \rho(\mathbf{x}, t) \\ &= -\int_\Omega dV \nabla \cdot \mathbf{j} \\ &= -\int_S \mathbf{j} \cdot d\mathbf{S}. \end{aligned}$$

The first step follows by (1.6) and the second by the divergence theorem. Now, note that as (1.5) is a Lorentz invariant equation, then $j^\mu(x)$ is a contravariant 4-vector.

1.4.1 Electromagnetic Fields

We shall take space to be "free", so that the relative permittivities and permeabilities are $\varepsilon_r = \mu_r = 1$. Now, Maxwell's equations, in SI units, read

$$\nabla \cdot \mathbf{E} = \frac{1}{\varepsilon_0}\rho, \qquad \nabla \cdot \mathbf{B} = 0,$$
$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \qquad \nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}.$$

We note the standard relation

$$\mu_0 \varepsilon_0 = \frac{1}{c^2}, \quad \mu_0 = 4\pi \times 10^{-7}\mathrm{NA}^{-2}.$$

Now, the equations in SI units, are not used practically. Instead, we use the *rationalised Gaussian units*,

$$\varepsilon_0 = 1 \quad \Rightarrow \quad \mu_0 = \frac{1}{c^2},$$

and we redefine the magnetic field,

$$\mathbf{B} \equiv \frac{\mathbf{B}_{\mathrm{SI}}}{c}.$$

Hence, under these rationalised SI units, the four Maxwell equations are

$$\nabla \cdot \mathbf{E} = \rho, \qquad \nabla \cdot \mathbf{B} = 0,$$
$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \qquad \nabla \times \mathbf{B} = \frac{1}{c}\mathbf{j} + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}.$$

We introduce the electromagnetic potentials ϕ and $\mathbf{A}$, so that the fields are related to the potentials via

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t},$$
$$\mathbf{B} = \nabla \times \mathbf{A}.$$

Notice that upon subsitution of these expressions into each of the Maxwell equations, they are satisfied (for example, the curl of grad is zero). Note that upon substitution of these equations, into Gauss' and Amperes' law (i.e the first and fourth of Maxwell's equations), one finds

$$\nabla^2\phi + \frac{1}{c}\frac{\partial}{\partial t}\nabla \cdot \mathbf{A} = -\rho,$$
$$\nabla^2\mathbf{A} - \frac{1}{c}\frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla\left(\nabla \cdot \mathbf{A} + \frac{1}{c}\frac{\partial \phi}{\partial t}\right) = -\frac{1}{c}\mathbf{j},$$

which may be written in the Lorentz covariant (and indeed, more succinct) form

$$\Box A^\mu(x) - \partial^\mu \partial_\nu A^\nu(x) = \frac{1}{c}j^\mu,$$

where the 4-vector potential A^μ is

$$A^\mu(x) = (\phi, \mathbf{A}).$$

1.5 Quantum Mechanics

1.5.1 Interpretation of the Wavefunction

We tend to define that

$$\rho(\mathbf{x},t)d^3x = |\psi(\mathbf{x},t)|^2\, d^3x$$

is the probability of finding a particle in volume d^3x, at position $\mathbf{x}$, and at time t. Now, this requires that probabilities are positive;

$$\rho(\mathbf{x},t) = |\psi(\mathbf{x},t)|^2 \geq 0,$$

and that the wavefunction is normalised;

$$\int \rho(\mathbf{x},t)d^3x = \int |\psi(\mathbf{x},t)|^2\, d^3x = 1,$$

where integrals are taken over all space. Notice that this last statement means that probabilities are conserved, as the derivative of unity is zero;

$$\frac{\partial}{\partial t}\int \rho(\mathbf{x},t)d^3x = \frac{\partial}{\partial t}\int |\psi(\mathbf{x},t)|^2\, d^3x = 0.$$

Now, we can get to the equation of continuity, with a little work.

Consider the Schrodinger equation,

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V(x)\psi, \tag{1.7}$$

and its Hermitian conjugate,

$$-i\hbar\frac{\partial\psi^*}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi^* + V(x)\psi^*. \tag{1.8}$$

Now, take the Schrodinger equation (1.7), multiply it by ψ^*, and subtract from it the conjugated equation (1.8) multiplied by ψ, to give

$$i\hbar\left(\psi^*\frac{\partial\psi}{\partial t} + \psi\frac{\partial\psi^*}{\partial t}\right) = -\frac{\hbar^2}{2m}\left(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*\right). \tag{1.9}$$

The LHS of this is just

$$i\hbar\left(\psi^*\frac{\partial\psi}{\partial t} + \psi\frac{\partial\psi^*}{\partial t}\right) = i\hbar\frac{\partial}{\partial t}|\psi|^2 = i\hbar\frac{\partial}{\partial t}\psi^*\psi.$$

And the RHS of (1.9) can be trivially written

$$-\frac{\hbar^2}{2m}\nabla\cdot\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right).$$

Hence, bringing this together, we have that

$$\frac{\partial}{\partial t}|\psi|^2 = \frac{i\hbar}{2m}\nabla\cdot(\psi^*\nabla\psi - \psi\nabla\psi^*)\,.$$

Hence, this is of the form of a continuity equation,

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{j} = 0, \tag{1.10}$$

where

$$\rho(\mathbf{x},t) = |\psi(\mathbf{x},t)|^2\,, \quad \mathbf{j} = \frac{i\hbar}{2m}(\psi\nabla\psi^* - \psi^*\nabla\psi)\,. \tag{1.11}$$

This is the Born interpretation of quantum mechanics; that the modulus squared of the wavefunction is a probability density. Hence, we see that the continuity equation is satisfied in the Schrodinger equation, but we shall have to recheck this when we come to discuss relativistic quantum mechanics.

1.5.2 Minimal Electromagnetic Interactions

The classical Hamiltonian, for a point particle of mass m, charge q, in an EM field, is given by

$$H = \frac{1}{2m}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)^2 + q\phi. \tag{1.12}$$

Let us compute Hamilton's equations from this. That is, let us consider

$$\dot{x}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial x_i}. \tag{1.13}$$

So, the first gives

$$\dot{x}_i = \frac{1}{m}\left(p_i - \frac{q}{c}A_i\right),$$

which easily rearranges into

$$p_i = m\dot{x}_i + \frac{q}{c}A_i,$$

or,

$$\mathbf{p} = m\dot{\mathbf{x}} + \frac{q}{c}\mathbf{A}. \tag{1.14}$$

The second of (1.13) provides us with

$$m\ddot{x}_i = m\dot{x}_i + qE_i,$$

or generally with

$$m\ddot{\mathbf{x}} = q\mathbf{E} + \frac{q}{c}\mathbf{v}\times\mathbf{B}.$$

Notice that if we insert (1.14) into (1.12), then we have

$$H(x, \dot{x}) = E(x, \dot{x}) = \frac{1}{2}m\dot{x}^2 + q\phi,$$

which is just the total energy of the system. Notice that the total energy is independent of the magnetic vector potential $\mathbf{A}$; this is a consequence of the magnetic field being perpendicular to motion.

We call (1.14) the *conjugate momentum*; notice that it is the "usual" expression of momentum, with an additive term due to the magnetic vector potential.

Let us take the Schrodinger equation,

$$H\psi = i\hbar\frac{\partial\psi}{\partial t},$$

and use the "correspondence principle" (also known as first quantisation, or canonical quantisation) to turn the momentum vector $\mathbf{p}$ in (1.12) into the momentum operator $\hat{\mathbf{p}} = -i\hbar\nabla$, so that the Schrodinger equation reads (using the Hamiltonian),

$$\frac{1}{2m}\left(-i\hbar\nabla - \frac{q}{c}\mathbf{A}\right)^2\psi + q\phi\psi = i\hbar\frac{\partial\psi}{\partial t},$$

which rearranges fairly simply into

$$-\frac{\hbar^2}{2m}\left(\nabla - \frac{iq}{\hbar c}\mathbf{A}\right)^2\psi = i\hbar\left(\frac{\partial}{\partial t} + \frac{iq}{\hbar}\phi\right)\psi.$$

Now, this is related back to the "free equation" (i.e. the Schrodinger equation with no EM fields) by the substitution of space and time derivatives,

$$\nabla \longmapsto \nabla - \frac{iq}{\hbar c}\mathbf{A}, \quad \frac{\partial}{\partial t} \longmapsto \frac{\partial}{\partial t} + \frac{iq}{\hbar}\phi.$$

Indeed, this may be written as

$$\partial_\mu \longmapsto \partial_\mu + \frac{iq}{\hbar}A_\mu, \tag{1.15}$$

where

$$A_\mu = (\phi, -\mathbf{A}), \quad \partial_\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, \nabla\right).$$

This substitution, (1.15) is called minimal substitution.

1.6 Relativistic Case

By a similar argument as above, the relativistic Hamiltonian is

$$H = \sqrt{(c\mathbf{p} - q\mathbf{A})^2 + m^2c^4} + q\phi,$$

or, by the minimal substitution,

$$H = \sqrt{-\hbar^2c^2\left(\nabla - \frac{iq}{\hbar c}\mathbf{A}\right)^2 + m^2c^4} + q\phi.$$

Now, the problem is: how do we interpret this square root? For example, the free particle Schrodinger equation reads

$$\sqrt{-\hbar^2c^2\nabla^2 + m^2c^4}\psi = i\hbar\frac{\partial\psi}{\partial t}. \tag{1.16}$$

How should we interpret this? This is the main thrust of all we look at.

1.7 Natural Units

From now on, we use natural units, whereby

$$\hbar = 1, \quad c = 1.$$

To get back to practical units, we use dimensions to restore $\hbar$'s and c's; and then use

$$\hbar = 6.582 \times 10^{-22}\text{MeV s}, \quad \hbar c = 1.973 \times 10^{-13}\text{MeV m}.$$

2 The Klein-Gordon Equation

Let us work in free space, so that $(\phi, \mathbf{A}) = 0$; and in the natural units previously discussed.

So, the relativistic Schrodinger equation (1.16) reads

$$H\phi(\mathbf{x}, t) = i\frac{\partial \phi(\mathbf{x}, t)}{\partial t}, \tag{2.1}$$

where the Hamiltonian is

$$H = \sqrt{-\nabla^2 + m^2}.$$

Now, we can avoid this square-root, by noting that the Hamiltonian is independent of time t. So, if we act H upon (2.1), we get

$$\begin{aligned} H^2\phi &= Hi\frac{\partial \phi}{\partial t} \\ &= i\frac{\partial}{\partial t}H\phi \\ &= i\frac{\partial}{\partial t}i\frac{\partial \phi}{\partial t} \\ &= -\frac{\partial^2 \phi}{\partial t^2}. \end{aligned}$$

Therefore,

$$H^2\phi = \left(-\nabla^2 + m^2\right)\phi = -\frac{\partial^2 \phi}{\partial t^2}.$$

Hence,

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2\right)\phi(\mathbf{x}, t) = 0. \tag{2.2}$$

This is the Klein-Gordon (KG) equation. Using the D'Alembertian notation, this is just

$$\left(\Box + m^2\right)\phi(\mathbf{x}, t) = 0, \quad \Box \equiv \partial_\mu \partial^\mu.$$

This KG equation is mathematically correct, and is indeed Lorentz invariant (for a scalar ϕ). But, what does it mean? What is the interpretation of ϕ?

2.1 Negative Energies

Substituting a plane wave solution

$$\phi(\mathbf{x}, t) = e^{i(\mathbf{p}\cdot\mathbf{x} - Et)}$$

into the KG equation, we find the relation

$$E^2 = p^2 + m^2,$$

so that the allowed energies are

$$E = +\sqrt{p^2 + m^2} > m, \quad E = -\sqrt{p^2 + m^2} < -m.$$

Hence, the spectrum (i.e. allowed energies of the system) is discontinuous. There is an energy gap between $mc^2 < E < mc^2$ (i.e. for energies below the rest mass, and above the negative of the rest mass). In classical physics, such a gap is a bit of a problem. In quantum mechanics, interactions allow quantum jumps over this gap, releasing quanta $E_\gamma > mc^2$, as radiation or other particles; such transitions carry on to give an infinite energy release.

2.2 Conserved Current

Let us compute the conserved current and charge $\mathbf{j}, \rho$ associated with the KG equation. Let us write the KG equation (2.2) as

$$\frac{\partial^2 \phi}{\partial t^2} = \left(\nabla^2 - m^2\right) \phi, \tag{2.3}$$

also taking its complex conjugate,

$$\frac{\partial^2 \phi^*}{\partial t^2} = \left(\nabla^2 - m^2\right) \phi^*. \tag{2.4}$$

Taking the first equation (2.3), multiplying it by ϕ^*, and subtracting from it (2.4) multiplied by ϕ, gives

$$\frac{\partial}{\partial t}\left(\phi^* \frac{\partial \phi}{\partial t} - \phi \frac{\partial \phi^*}{\partial t}\right) = \nabla \cdot \left(\phi^* \nabla \phi - \phi \nabla \phi^*\right).$$

Multiplying through by i, and bringing the LHS over to the RHS,

$$i\frac{\partial}{\partial t}\left(\phi^* \frac{\partial \phi}{\partial t} - \phi \frac{\partial \phi^*}{\partial t}\right) - i\nabla \cdot \left(\phi^* \nabla \phi - \phi \nabla \phi^*\right) = 0.$$

Hence, we can compare this with the continuity equation (1.10) to identify the charge and current;

$$\rho = i\left[\phi^* \frac{\partial \phi}{\partial t} - \phi \frac{\partial \phi^*}{\partial t}\right], \tag{2.5}$$

$$\mathbf{j} = -i\left[\phi^* \nabla \phi - \phi \nabla \phi^*\right]. \tag{2.6}$$

Now, if we compare these with those found in the non-relativistic case, (1.11), we see that we cannot interpret ρ as a probability density (as we did do in the non-relativistic case). In our case, (2.5) can be negative as well as positive (thus negating its ability to be a probability). Hence, we see that ϕ cannot (generally) be interpreted as a probability amplitude (as we did for the non-relativistic ψ).

So, to summarise, we have found that the KG equation is a covariant wave equation; and with it is a conserved charge $j^\mu = (\rho, \mathbf{j})$ such that $\partial_\mu j^\mu = 0$ is satisfied. But, we have problems with negative energies, and interpreting the "wavefunction" ϕ. For the moment, let us carry on, regardless of these issues.

2.3 EM Fields and the "Hydrogen Atom"

If the 4-vector potential is non-zero, $A^\mu \neq 0$, then we make the minimal substitution

$$\partial_\mu \longmapsto \partial_\mu + iqA_\mu,$$

so that the KG equation becomes

$$\left[(\partial_\mu + iqA_\mu)(\partial^\mu + iqA^\mu) + m^2\right]\phi = 0. \tag{2.7}$$

Now, consider electrostatic fields, so that $A^\mu = (\Phi, \mathbf{0})$ only (i.e. only a static term, and no magnetic term); such that

$$\mathbf{A} = 0, \quad V(x) = qA_0(x) = q\Phi(x),$$

and V is not a function of time t. Then, substituting this into the KG equation (2.7), one easily obtains

$$\left[\left(\frac{\partial}{\partial t} + iV\right)^2 - \nabla^2 + m^2\right]\phi(\mathbf{x},t) = 0. \tag{2.8}$$

Now, if we substitute a solution of the form

$$\phi(\mathbf{x},t) = \psi(\mathbf{x})e^{-iEt}$$

into equation (2.8), we easily find that

$$\left[-(E-V)^2 + -\nabla^2 + m^2\right]\psi = 0. \tag{2.9}$$

Now, let us solve for the energy levels, where the potential is the central Coulomb

$$V = -\frac{Z\alpha}{r}, \quad \alpha = \frac{e^2}{4\pi} = \frac{1}{137}. \tag{2.10}$$

We can solve by comparing with the non-relativistic Schrodinger equation

$$\left[-\nabla^2 - \frac{2mZ\alpha}{r}\right]\psi = 2m\epsilon\psi, \tag{2.11}$$

where $\epsilon = E - mc^2$. The way to solve this, is to make the separable ansatz

$$\psi = \frac{U(r)}{r}Y_{\ell m}(\theta,\phi), \tag{2.12}$$

and use that the Laplacian in spherical polars can be written

$$\nabla^2 = \frac{1}{r}\frac{\partial^2}{\partial r^2}r - \frac{\hat{\mathbf{L}}^2}{r^2}. \tag{2.13}$$

So, substituting these into (2.11) fairly easily results in

$$\left[-\frac{d^2}{dr^2}+\frac{\ell(\ell+1)}{r^2}-\frac{2mZ\alpha}{r}\right]U(r)=2m\epsilon U(r). \tag{2.14}$$

Making the same ansatz (2.12), after putting in the central Coulomb potential (2.10), we see that we can write the KG equation (2.9) as

$$\left[-\nabla^2-\frac{(Z\alpha)^2}{r^2}-\frac{2EZ\alpha}{r}\right]\frac{U(r)}{r}Y_{\ell m}(\theta,\phi)=(E^2-m^2)\frac{U(r)}{r}Y_{\ell m}(\theta,\phi),$$

using the form of the Laplacian (2.13) gives

$$\left[-\frac{d^2}{dr^2}+\frac{\ell(\ell+1)-(Z\alpha)^2}{r^2}-\frac{2EZ\alpha}{r}\right]U(r)=(E^2-m^2)U(r). \tag{2.15}$$

Hence, upon comparison of this with the non-relativistic version (2.14), we see that we have the relations

$$\begin{aligned} 2m &\longmapsto 2E, \\ 2m\epsilon &\longmapsto E^2-m^2, \\ \ell(\ell+1) &\longmapsto \ell'(\ell'+1)=\ell(\ell+1)-(Z\alpha)^2. \end{aligned} \tag{2.16}$$

Notice that (2.16) may be solved to give

$$\ell'=-\frac{1}{2}+\sqrt{\left(\ell+\frac{1}{2}\right)^2-(Z\alpha)^2}. \tag{2.17}$$

Now, the energy levels of the non-relativistic Schrodinger equation (2.14) are given by

$$2m\epsilon=-\frac{2m(Z\alpha)^2}{2n^2}m,$$

where the principle quantum number is given in terms of the number of radial modes and angular quantum number, as

$$n=n_{\rm r}+\ell+1, \quad n_{\rm r}=0,1,2\ldots.$$

Hence, using our relations between the solutions of the relativistic and non-relativistic equations, we see that the equivalent expression for the relativistic KG equation is

$$E^2-m^2=-\frac{2E(Z\alpha)^2}{2n'^2}E, \tag{2.18}$$

where

$$n'=n_{\rm r}+\ell'+1, \quad n_{\rm r}=0,1,2,\ldots.$$

We can solve (2.18) for E,

$$E=m\left(1+\frac{(Z\alpha)^2}{n'^2}\right)^{-1/2}. \tag{2.19}$$

We can consider two cases: the strong and weak potential, in order that we may expand this energy.

Weak potential If we take

$$Z\alpha \ll 1,$$

then (2.19) can be expanded in $Z\alpha$

$$E = m\left(1 - \frac{(Z\alpha)^2}{2n'^2} + \frac{3}{8}\frac{(Z\alpha)^4}{n'^4} + \dots\right).$$

Now, by (2.17),

$$\ell' = -\frac{1}{2} + \left(\ell + \frac{1}{2}\right) - \frac{(Z\alpha)^2}{2\left(\ell + \frac{1}{2}\right)} + \dots$$

So, to first order in $Z\alpha$, $\ell' = \ell$, which means that to first order, $n' = n$. To second order,

$$\ell' = \ell - \frac{(Z\alpha)^2}{2\left(\ell + \frac{1}{2}\right)},$$

so that

$$\begin{aligned} n' &= n_{\rm r} + \ell' + 1 \\ &= n_{\rm r} + \ell - \frac{(Z\alpha)^2}{2\left(\ell + \frac{1}{2}\right)} + 1 \\ &= n - \frac{(Z\alpha)^2}{2\left(\ell + \frac{1}{2}\right)}, \end{aligned}$$

and hence that

$$n'^4 = n^4\left(1 - \frac{4(Z\alpha)^2}{2n\left(\ell + \frac{1}{2}\right)} + \dots\right).$$

One can then show that

$$E = m\left[1 - \frac{(Z\alpha)^2}{2n^2} - \frac{(Z\alpha)^4}{2n^4}\left(\frac{n}{\ell + \frac{1}{2}} - \frac{3}{4}\right) + \dots\right].$$

The first term is the rest mass of the electron; the second the non-relativistic result. The third term gives a splitting for modes with different ℓ, and is thus the fine-splitting term. This fine-splitting term does not agree with experimental results for Hydrogen, and thus cannot be used as a description of the Hydrogen atom. However, the derivation for the term is correct (infact, the result is correct for pionic atoms, spin-zero particles). Hence, this scalar KG equation described spin-0 particles.

Strong potential Let us take

$$Z\alpha > \ell + \frac{1}{2},$$

so that for $\ell = 0$, this corresponds to $Z \geq 69$. Notice that in this regime, (2.17) is complex, which will give complex energies! To understand this, we shall take a look at a simpler problem.

2.3.1 The Klein Paradox

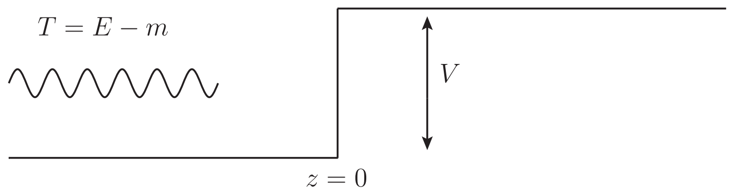


Figure 2.1: A beam with kinetic energy T travels towards a barrier which has height V, centred at $z = 0$, along the $\hat{\mathbf{z}}$-axis,

Consider a beam as in Figure (2.1). Then, for $z < 0$, the KG equation (2.9) is just

$$\left[-E^2 - \frac{d^2}{dz^2} + m^2\right]\psi(z) = 0,$$

which is easily solved to find

$$\psi(z) = Ae^{ipz} + Be^{-ipz}, \quad p^2 = E^2 - m^2.$$

Notice that A is the amplitude for the incident wave, and B for some reflected wave. For $z > 0$, the KG equation is just

$$\left[-(E-V)^2 - \frac{d^2}{dz^2} + m^2\right]\psi(z) = 0,$$

which is solved to give

$$\psi(z) = Ce^{ip'z}, \quad p'^2 = (E-V)^2 - m^2;$$

where we have imposed that there should be no imaginary wave from the right.

Now, if $p'^2 > 0$, then

$$(E-V)^2 > m^2 \quad \Rightarrow \quad (E-m)^2 > V^2.$$

Similarly, if $p'^2 < 0$, then

$$(E-m)^2 < V^2,$$

where we can thus write $p' = i|p'|$, i.e. that p' is complex. Thus, if p' is complex, then the argument of the exponential $e^{ip'z}$ is real (and negative), and thus we have exponential decay "inside the barrier" (this is called "evanescence"); a somewhat expected result.

If we impose continuity of the wavefunction at $z = 0$, where ψ and $d\psi/dz$ are continuous, then one finds the relation

$$B = \frac{p - p'}{p + p'}A, \quad C = \frac{2p}{p + p'}A.$$

Now, in 1D, the conserved current is

$$j = -i\left(\psi^* \frac{d\psi}{dz} - \psi \frac{d\psi^*}{dz}\right).$$

So, the current for the incident wave Ae^{ipz} is

$$j_{\mathrm{I}} = 2p|A|^2,$$

for the reflected Be^{-ipz} is

$$j_{\mathrm{R}} = -2p|B|^2,$$

and for the transmitted wave $Ce^{ip'z}$

$$j_{\mathrm{T}} = 2p'|C|^2.$$

Notice that

$$j_{\mathrm{I}} = j_{\mathrm{R}} + j_{\mathrm{T}},$$

somewhat as expected. Hence, we have propagation through the barrier if $(E-V)^2 > m^2$ and evanescence if $(E-V)^2 < m^2$ (these cases correspond to p' being real and imaginary, respectively).

Let us then consider two cases, as before.

Weak potential Consider $V < E$. Then, propagation if $E - V > m$, or equivalently, $E - m > V$. Then, this requires the kinetic energy to be above the barrier, which is an expected result.

Strong potential Consider $V > E$. Then, propagation can occur if $E - V < -m$. That is, $V > E+m$, which just gives $(E-V)^2 < m^2$. This result says that propagation is allowed, for large potential, if the kinetic energy is below the barrier. This is the Klein paradox.

Consider waves on the right (i.e. for $z > 0$). For $V > E + m$, then

$$p'^2 = (V-E)^2 - m^2,$$

which we can differentiate with respect to p', to give

$$2p' = 2(V-E)\left(-\frac{dE}{dp'}\right).$$

Hence, the group velocity,

$$v_{\mathrm{g}} = \frac{dE}{dp'} = -\frac{p'}{V-E}.$$

Thus, we see that the group velocity (which is the speed of the movement of particles) is "left", whereas the wave is traveling "right". If we substitute p' into the expression for the conserved current, then one finds

$$z < 0 \quad \Rightarrow \quad |j_{\mathrm{R}}| > |j_{\mathrm{I}}|,$$

and

$$z > 0 \quad \Rightarrow \quad j_{\mathrm{T}} < 0.$$

That is, current is flowing to the left, even though waves are traveling to the right.

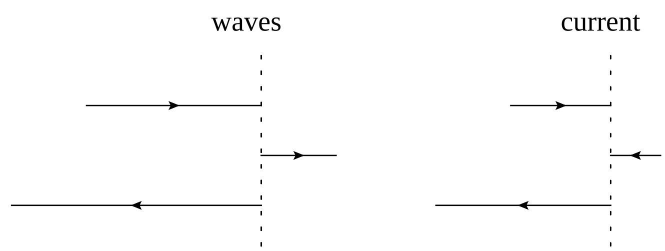


Figure 2.2: The relative amplitudes of the waves and current at the barrier.

The suggested interpretation of this, is that particle/anti-particle pairs are created at the barrier, by an interaction with the barrier. Particles move to the left, in the reflected wave, and anti-particles move to the right in the transmitted wave (with opposite charge). This suggests a similar effect in atoms, if Z is large.

So, for a full description for strong potentials, we must abandon a single particle description, and we will need a field theory – which we will come back to later.

3 The Dirac Equation

Let us reconsider the Schrodinger equation,

$$H\psi = i\frac{\partial \psi}{\partial t}, \tag{3.1}$$

where the relativistic Hamiltonian is

$$H = \sqrt{-\nabla^2 + m^2},$$

which was derived using canonical quantisation. Recall that we got the Klein-Gordon equation by multiplying the Schrodinger equation through by H,

$$HH\psi = (-\nabla^2 + m^2)\psi = -\frac{\partial^2 \psi}{\partial t^2}. \tag{3.2}$$

This made the KG equation second order (in time), which led to negative conserved quantities. To get around this, Dirac looked for an equation that would be first order in time, of the form (3.1). So, he proposed a Hamiltonian

$$H = -i\boldsymbol{\alpha} \cdot \nabla + \beta m, \tag{3.3}$$

so that (3.1) becomes

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

This is known as the *Dirac equation*. The four constant coefficients $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ & β are chosen such that the Dirac equation is consistent with (3.2), so that $E^2 = p^2 + m^2$ is recovered. They are also chosen such that the Hamiltonian H is Hermitian, resulting in real energy eigenvalues.

To be consistent with the KG equation, we thus require

$$HH\psi = \left(-i\sum_{j=1}^{3}\alpha_j\frac{\partial}{\partial x_j} + \beta m\right)\left(-i\sum_{k=1}^{3}\alpha_k\frac{\partial}{\partial x_k} + \beta m\right)\psi = -\frac{\partial^2 \psi}{\partial t^2}.$$

So, expanding out,

$$\begin{aligned}
HH\psi &= -\sum_{j,k}\alpha_j\alpha_k\frac{\partial^2\psi}{\partial x_j \partial x_k} - im\sum_j \alpha_j\beta\frac{\partial\psi}{\partial x_j} - im\sum_k \beta\alpha_k\frac{\partial\psi}{\partial x_k} + \beta^2 m^2\psi \\
&= -\sum_{j,k}\alpha_j\alpha_k\frac{\partial^2\psi}{\partial x_j \partial x_k} - im\sum_j \left(\alpha_j\beta + \beta\alpha_j\right)\frac{\partial\psi}{\partial x_j} + \beta^2 m^2\psi \\
&= -\sum_j \alpha_j^2\frac{\partial^2\psi}{\partial x_j^2} - \sum_{j\neq k}\left(\alpha_j\alpha_k + \alpha_k\alpha_j\right)\frac{\partial^2\psi}{\partial x_j \partial x_k} \\
&\quad -im\sum_j \left(\alpha_j\beta + \beta\alpha_j\right)\frac{\partial\psi}{\partial x_j} + \beta^2 m^2\psi.
\end{aligned}$$

Hence,

$$\begin{aligned} HH\psi &= -\sum_j \alpha_j^2 \frac{\partial^2\psi}{\partial x_j^2} - \sum_{j\neq k} (\alpha_j\alpha_k + \alpha_k\alpha_j) \frac{\partial^2\psi}{\partial x_j \partial x_k} \\ &\quad -im\sum_j (\alpha_j\beta + \beta\alpha_j) \frac{\partial\psi}{\partial x_j} + \beta^2 m^2\psi = -\frac{\partial^2\psi}{\partial t^2}. \end{aligned}$$

Now, upon comparison of this with (3.2), we see that we require

$$\alpha_j^2 = 1, \quad \alpha_j\alpha_k + \alpha_k\alpha_j = 0 \quad (j \neq k), \quad \alpha_j\beta + \beta\alpha_j = 0, \quad \beta^2 = 1.$$

Thus, we have the anti-commutation relations

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij}, \tag{3.4}$$

$$\{\beta, \alpha_i\} = 0, \tag{3.5}$$

$$\beta^2 = 1. \tag{3.6}$$

Thus, we see that α_i, β cannot be numbers, but they can be matrices.

3.0.2 Properties of α_i, β & Dirac Representation

Now, one can easily see that α_i and β must be Hermitian, in order that the Hamiltonian is Hermitian.

Consider the eigenvalue equation

$$\alpha_i\psi = \lambda\psi,$$

and further

$$\alpha_i^2\psi = \lambda^2\psi,$$

but, from (3.4), we see that we have $\alpha_i^2 = 1$. Hence, $\lambda^2 = 1$, which means that the eigenvalues of α_i are $\lambda = \pm 1$. This result also holds for β, by (3.6).

We shall state, then prove, that

$$\text{Tr } \beta = 0, \quad \text{Tr } \alpha_i = 0. \tag{3.7}$$

That is, the sum over diagonal components is zero. Now, consider that

$$\text{Tr } (ABC) = \text{Tr } (CAB) = \text{Tr } (BCA),$$

which is easily provable by

$$\text{Tr } (ABC) = \sum_{i,j,k} A_{ij}B_{jk}C_{ki} = \sum_{i,j,k} C_{ki}A_{ij}B_{jk} = \text{Tr } (CAB). \tag{3.8}$$

Then, consider that (3.5) is

$$\beta\alpha_i = -\alpha_i\beta,$$

multiplying from the left by β gives

$$\beta^2\alpha_i = \alpha_i = -\beta\alpha_i\beta,$$

where the first equality comes from (3.6). Hence,

$$\text{Tr } \alpha_i = -\text{Tr } (\beta\alpha_i\beta) = -\text{Tr } (\beta^2\alpha_i) = -\text{Tr } \alpha_i,$$

where the second equality follows from (3.8) and the third by (3.6). Hence,

$$2\text{Tr } (\alpha_i) = 0 \quad \Rightarrow \quad \text{Tr } \alpha_i = 0,$$

which is the second of (3.7). The proof for Tr $\beta = 0$ is analogous.

The final property of α_i and β is that they must be of even dimension. To see this, consider that the trace of a matrix is the sum of its eigenvalues. But, we found that the eigenvalues of both α_i and β are $\lambda = \pm 1$. We also showed that the trace of α_i and β is zero. Hence, the only way to add such eigenvalues to give zero, is to have an even number of them. Hence, α_i and β are of even dimension.

Let us consider how to represent the four matrices α_i and β.

In dimension $N = 2$, there are only 3 linearly independent matrices of the form

$$\begin{pmatrix} a_1 & a_2 + ia_3 \\ a_2 - ia_3 & -a_1 \end{pmatrix}, \quad a_i \in \mathbb{R}.$$

These are

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \tag{3.9}$$

notice that they satisfy

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}, \tag{3.10}$$

as required. They also satisfy the commutation relation

$$[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\sigma_l \tag{3.11}$$

If $N = 4$, there are a few ways in which we can represent the matrices. The way we shall is called the *Dirac representation*, where

$$\beta = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}, \quad \alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}. \tag{3.12}$$

Just to make this block-notation clear, we shall write some of these out, long hand.

$$\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \alpha_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}.$$

As we previously said, there are other representations. Infact, a specific representation is unneeded, and one can just work with the commutation relations. That we use the Dirac representation is a mere convenience for the non-relativistic limits, and observable predictions are not affected by the choice of representation.

Hence, we have that the Dirac equation

$$i\frac{\partial \psi}{\partial t} = -i\boldsymbol{\alpha}\cdot\nabla\psi + \beta m\psi \tag{3.13}$$

is a matrix equation with four components. That is, the wavefunction has four components

$$\psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix}.$$

Since α_i, β are Hermitian, the adjoint equation of (3.13) is

$$-i\frac{\partial \psi^\dagger}{\partial t} = i\nabla\cdot\left(\psi^\dagger\boldsymbol{\alpha}\right) + m\psi^\dagger\beta. \tag{3.14}$$

Notice that one can fairly easily see the anti-commutation relations,

$$\{\beta, \alpha_i\} = 0, \quad \{\alpha_i, \alpha_j\} = 2\delta_{ij};$$

the second of which can be seen just by (3.10).

3.1 Conserved Current

We can compute the conserved quantities by the usual argument. We multiply (3.13) by $\psi^\dagger$ *from the left* (we must be careful about this now), to give

$$i\psi^\dagger\frac{\partial \psi}{\partial t} = -i\psi^\dagger\boldsymbol{\alpha}\cdot\nabla\psi + m\psi^\dagger\beta\psi. \tag{3.15}$$

In a similar way, we multiply ψ *from the right* onto the adjoint equation (3.14), to give

$$-i\frac{\partial \psi^\dagger}{\partial t}\psi = i\nabla\cdot\left(\psi^\dagger\boldsymbol{\alpha}\right)\psi + m\psi^\dagger\beta\psi. \tag{3.16}$$

Then, subtracting (3.16) from (3.15), we see that

$$2i\frac{\partial}{\partial t}\left(\psi^\dagger\psi\right) = -2i\nabla\cdot\left(\psi^\dagger\boldsymbol{\alpha}\psi\right),$$

or, upon comparison with the continuity equation

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

then we see that

$$\rho = \psi^\dagger\psi, \quad \mathbf{j} = \psi^\dagger\boldsymbol{\alpha}\psi. \tag{3.17}$$

In particular, notice that

$$\rho = \psi^\dagger\psi = (\psi_1^*\,\psi_2^*\,\psi_3^*\,\psi_4^*)\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{pmatrix} = \sum_{i=1}^{4}|\psi_i(x)|^2 \geq 0.$$

Hence, we see that we can treat ψ as a probability density function for position, in the usual way (which we could not do for the KG equation).

3.2 Angular Momentum & Spin

Consider the orbital angular momentum operator,

$$\mathbf{L} = \mathbf{x}\times\mathbf{p},$$

or, in index notation,

$$L_i = \epsilon_{ijk}x_jp_k.$$

Let us ask: is orbital angular momentum conserved? If it is, then we require that

$$[H,\mathbf{L}] = 0.$$

So, using the Dirac Hamiltonian (3.3),

$$\begin{aligned} H &= -i\boldsymbol{\alpha}\cdot\nabla + \beta m \\ &= \boldsymbol{\alpha}\cdot\mathbf{p} + \beta m, \end{aligned}$$

and using the usual commutation relation

$$[x_j,p_k] = i\delta_{jk},$$

we consider the commutator

$$\begin{aligned}[H, L_k] &= [\alpha_j p_j, \epsilon_{knm} x_n p_m] \\ &= \epsilon_{knm} \alpha_j [p_j, x_n] p_m \\ &= -i\epsilon_{knm} \alpha_j \delta_{jn} p_m \\ &= -i\epsilon_{kjm} \alpha_j p_m,\end{aligned}$$

where the last equality allows us to see that

$$[H, \mathbf{L}] = -i\boldsymbol{\alpha} \times \mathbf{p}, \tag{3.18}$$

which is non-zero. If this were zero, then we could have said that the Hamiltonian commuted with the orbital angular momentum operator. However, it does not, in which case we must say that the Hamiltonian does not commute with orbital angular momentum, and hence orbital angular momentum is not conserved.

Now, we only considered orbital angular momentum. Instead, let us suppose that the total angular momentum is conserved (i.e. orbital plus something-else).

Let us consider some matrix operators

$$\Sigma_j = \begin{pmatrix} \sigma_j & 0 \\ 0 & \sigma_j \end{pmatrix}. \tag{3.19}$$

Then,

$$\begin{aligned}\alpha_j \Sigma_k - \Sigma_k \alpha_j &= \begin{pmatrix} 0 & \sigma_j \sigma_k \\ \sigma_j \sigma_k & 0 \end{pmatrix} - \begin{pmatrix} 0 & \sigma_k \sigma_j \\ \sigma_k \sigma_j & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & [\sigma_j, \sigma_k] \\ [\sigma_j, \sigma_k] & 0 \end{pmatrix} \\ &= 2i\epsilon_{jkn} \begin{pmatrix} 0 & \sigma_n \\ \sigma_n & 0 \end{pmatrix} \\ &= 2i\epsilon_{jkn} \alpha_n.\end{aligned}$$

Hence, we have used (3.11) to show that

$$[\alpha_j, \Sigma_k] = 2i\epsilon_{jkn} \alpha_n. \tag{3.20}$$

In a much simpler, but analogous manner, one can also show that

$$[\beta, \Sigma_k] = 0. \tag{3.21}$$

Hence, the commutator of the Hamiltonian with this Σ_i operator,

$$\begin{aligned}[H, \Sigma_k] &= [\alpha_j p_j + \beta m, \Sigma_k] \\ &= [\alpha_j, \Sigma_k] p_j \\ &= 2i\epsilon_{jkn} \alpha_n p_j \\ &= 2i\epsilon_{knj} \alpha_n p_j,\end{aligned}$$

thus,

$$[H, \mathbf{\Sigma}] = 2i\boldsymbol{\alpha} \times \mathbf{p}. \tag{3.22}$$

Then, consider the combination

$$\mathbf{J} = \mathbf{L} + \frac{1}{2}\mathbf{\Sigma}, \tag{3.23}$$

so that

$$[H, \mathbf{J}] = -i\boldsymbol{\alpha} \times \mathbf{p} + \frac{1}{2}2i\boldsymbol{\alpha} \times \mathbf{p} = 0$$

from (3.22) and (3.18). Therefore, we see that the operator $\mathbf{J}$, defined by (3.23) is conserved, as it commutes with the Hamiltonian. We can define some spin operator,

$$\mathbf{S} = \frac{1}{2}\mathbf{\Sigma} = \frac{1}{2}\begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix},$$

in Dirac representation. Notice further, since $\sigma_i^2 = 1$, then $\Sigma_i^2 = 1$, and hence the eigenvalues of S_i are $\pm\frac{1}{2}$. Thus, we see that the particle has spin$-\frac{1}{2}$.

3.3 Plane Wave States

Let us consider solutions to the Dirac equation, where the solutions are of the form

$$\psi(x) = Ne^{-ipx}u(\mathbf{p}), \tag{3.24}$$

where $px = p^\mu x_\mu = Et - \mathbf{p} \cdot \mathbf{x}$, N is some normalisation and $u(\mathbf{p})$ some 4-component spinor, whose form is to be determined. We also require that $p^2 = E^2 - m^2$, to satisfy the KG equation.

We use Dirac representation, whereby

$$\beta = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}, \quad \boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \tag{3.25}$$

and we write

$$u = \begin{pmatrix} \xi \\ \eta \end{pmatrix}, \tag{3.26}$$

where both ξ and η are two component spinors (thus making u a 4-component spinor). We will substitute into the Dirac equation,

$$i\frac{\partial \psi}{\partial t} = -i\boldsymbol{\alpha} \cdot \nabla\psi + \beta m\psi, \tag{3.27}$$

and use the useful, easily derivable

$$\partial_\mu e^{-ipx} = -i\,(E, -\mathbf{p})\,e^{-ipx} = -ip_\mu e^{-ipx}.$$

Notice that the Dirac equation (3.27) can be written as

$$\begin{aligned}\beta m\psi &= i\left(\frac{\partial}{\partial t} + \boldsymbol{\alpha}\cdot\nabla\right)\psi \\ &= i\left(\partial_0 + \alpha_j\partial_j\right) Ne^{-ipx}u(\mathbf{p}) \\ &= i\left(-iE + i\alpha_j p_j\right) Ne^{-ipx}u(\mathbf{p}),\end{aligned}$$

indeed as,

$$\beta m u(\mathbf{p}) = (E - \boldsymbol{\alpha}\cdot\mathbf{p})\,u(\mathbf{p}).$$

Hence, using the matrices (3.25) and (3.26)

$$m\begin{pmatrix}\mathbf{1} & 0\\ 0 & -\mathbf{1}\end{pmatrix}\begin{pmatrix}\xi\\ \eta\end{pmatrix} = E\begin{pmatrix}\xi\\ \eta\end{pmatrix} + \begin{pmatrix}0 & \boldsymbol{\sigma}\cdot\mathbf{p}\\ \boldsymbol{\sigma}\cdot\mathbf{p} & 0\end{pmatrix}\begin{pmatrix}\xi\\ \eta\end{pmatrix},$$

rearranging

$$E\begin{pmatrix}\xi\\ \eta\end{pmatrix} = \begin{pmatrix}m\mathbf{1} & \boldsymbol{\sigma}\cdot\mathbf{p}\\ \boldsymbol{\sigma}\cdot\mathbf{p} & -m\mathbf{1}\end{pmatrix}\begin{pmatrix}\xi\\ \eta\end{pmatrix}.$$

So, expanding out this matrix, we get two coupled simultaneous equations,

$$E\xi = m\xi + \boldsymbol{\sigma}\cdot\mathbf{p}\eta, \tag{3.28}$$

$$E\eta = \boldsymbol{\sigma}\cdot\mathbf{p}\xi - m\eta. \tag{3.29}$$

So, solving (3.29) for η, gives

$$\eta = \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\xi, \tag{3.30}$$

so that upon substitution into (3.28), one finds

$$E\xi = m\xi + \boldsymbol{\sigma}\cdot\mathbf{p}\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\xi,$$

which rearranges easily into

$$(E-m)(E+m)\xi = (\boldsymbol{\sigma}\cdot\mathbf{p})^2\,\xi. \tag{3.31}$$

Now, notice that

$$\begin{aligned}(\boldsymbol{\sigma}\cdot\mathbf{p})^2 = &\quad (\sigma_1 p_1 + \sigma_2 p_2 + \sigma_3 p_3)(\sigma_1 p_1 + \sigma_2 p_2 + \sigma_3 p_3)\\ = &\quad \sigma_1^2 p_1^2 + \sigma_2^2 p_2^2 + \sigma_3^2 p_3^2\\ &\quad +(\sigma_1\sigma_2 + \sigma_2\sigma_1)p_1 p_2\\ &\quad +(\sigma_1\sigma_3 + \sigma_3\sigma_1)p_1 p_3\\ &\quad +(\sigma_2\sigma_3 + \sigma_3\sigma_2)p_2 p_3.\end{aligned}$$

The final three terms are all zero, by the anti-commutation relation (3.10), and also $\sigma_i^2 = 1$. Hence, we see that

$$(\boldsymbol{\sigma}\cdot\mathbf{p})^2 = p_1^2 + p_2^2 + p_3^2 = \mathbf{p}^2. \tag{3.32}$$

Thus, substituting this into (3.31) – also expanding out the brackets on the LHS – gives

$$(E^2 - m^2)\xi = \mathbf{p}^2\xi. \tag{3.33}$$

Now, recall the positive energy solution, whereby

$$E(\mathbf{p}) = +\sqrt{m^2 + \mathbf{p}^2},$$

we then see that there are two independent solutions to (3.33), which we label

$$\xi_s = \sqrt{E+m}\chi_s, \tag{3.34}$$

where the χ_s are such that

$$\chi_{s=\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_{s=-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

In writing (3.34), we have somewhat inserted the factor $\sqrt{E+m}$ by forethought. Later on we shall be writing the normalisation of the spinors, where the inserted factor will make expressions a lot simpler. Hence, using this solution in (3.30), we see that (3.26) is

$$\begin{aligned} u_s(\mathbf{p}) &= \begin{pmatrix} \xi \\ \eta \end{pmatrix} \\ &= \begin{pmatrix} \xi \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\xi \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{E+m}\chi_s \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\sqrt{E+m}\chi_s \end{pmatrix} \\ &= \sqrt{E+m}\begin{pmatrix} \chi_s \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi_s \end{pmatrix}. \end{aligned}$$

Therefore, we have derived that the *positive energy solution* is

$$u_s(\mathbf{p}) = \sqrt{E+m}\begin{pmatrix} \chi_s \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi_s \end{pmatrix}. \tag{3.35}$$

Notice that

$$u_s^\dagger(\mathbf{p}) = \sqrt{E+m}\left(\chi_s \quad \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})^*}{E+m}\chi_s\right),$$

moreover that

$$u_s^\dagger(\mathbf{p})u_s(\mathbf{p}) = (E+m)\left(\chi_s^2 + \frac{(\boldsymbol{\sigma}\cdot\mathbf{p})^2}{(E+m)^2}\chi_s^2\right),$$

which becomes, using (3.32),

$$\begin{aligned} u_s^\dagger(\mathbf{p})u_s(\mathbf{p}) &= (E+m)\left(1+\frac{\mathbf{p}^2}{(E+m)^2}\right) \\ &= \frac{(E+m)^2+\mathbf{p}^2}{E+m} \\ &= \frac{(E+m)^2+E^2-m^2}{E+m} \\ &= 2E. \end{aligned} \tag{3.36}$$

If we recall the solution we were seeking, (3.24)

$$\psi(x) = Ne^{-ipx}u(\mathbf{p}),$$

and if we require its normalisation,

$$\begin{aligned} \int d^3x\,\psi^\dagger(x)\psi(x) &= |N|^2\int d^3x\,u_s^\dagger(\mathbf{p})u_s(\mathbf{p}) \\ &= |N|^2 2EV \\ &= 1, \end{aligned}$$

where we normalise in some box of volume V, then we see that

$$N = \frac{1}{\sqrt{2EV}}.$$

Therefore, the full *positive energy solution* is

$$\psi^{(+)}_{\mathbf{p},s}(x) = \frac{u_s(\mathbf{p})e^{-ipx}}{\sqrt{2EV}}, \tag{3.37}$$

where $u_s(\mathbf{p})$ is given by (3.35). This solution describes a particle with momentum $\mathbf{p}$, energy $E > m$ and spin $S_z = s = \pm 1/2$.

The *negative energy solution* is written as

$$\psi^{(-)}_{\mathbf{p},s}(x) = Ne^{ipx}v_s(\mathbf{p}).$$

One finds that the equivalent of (3.28) and (3.29) are

$$\begin{aligned} -E\xi &= m\xi + \boldsymbol{\sigma}\cdot\mathbf{p}\eta, \\ -E\eta &= \boldsymbol{\sigma}\cdot\mathbf{p}\xi - m\eta. \end{aligned}$$

Then, one can easily rearrange to find that

$$\xi = -\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\eta,$$

which results in the equivalent of (3.33) being

$$p^2\eta = (E^2 - m^2)\eta.$$

Hence, with a negative energy, one still has the relation $E^2 = p^2 + m^2$. This equation has similar solutions, and so one arrives at

$$v_s(\mathbf{p}) = \sqrt{E+m}\begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi_{-s} \\ \chi_{-s} \end{pmatrix}, \tag{3.38}$$

and thus the wavefunction

$$\psi^{(-)}_{\mathbf{p},s}(x) = \frac{v_s(\mathbf{p})e^{ipx}}{\sqrt{2EV}}. \tag{3.39}$$

This corresponds to a particle with momentum $(-\mathbf{p})$, energy $E = -\sqrt{p^2+m^2}$ and spin $(-s)$, in volume V.

Notice that in the non-relativistic limit $\mathbf{p} \to 0$ (and the energy E is of the order the rest mass m), (3.35) and (3.38) become

$$\begin{aligned} u_s &\longrightarrow \sqrt{2m}\begin{pmatrix} \chi_s \\ 0 \end{pmatrix}, \\ v_s &\longrightarrow \sqrt{2m}\begin{pmatrix} 0 \\ \chi_{-s} \end{pmatrix}. \end{aligned}$$

3.4 Dirac Hole Theory

As spin$-\frac{1}{2}$ particles are fermions, they obey the exclusion principle, which states that no two fermions can be in the same state. Now, imagine the "vacuum" as being a completely filled "sea" of negative energy states. That the sea is completely filled is important, as no positive energy particle can drop in there, as there is no free state for it to occupy.

If the negative energy states are full, then no positive energy particle can transition into the negative states. But, a negative energy particle can transition into the positive energy states, if it is given enough energy $E > 2m$. Then, there will be a hole in the negative energy states, which looks like an anti-particle. That is, the absence of a e^- particle with $-E, -\mathbf{p}, -s$, which looks like a e^+ particle with $E, \mathbf{p}, s$.

See Figure (3.1) for a depiction of the theory.

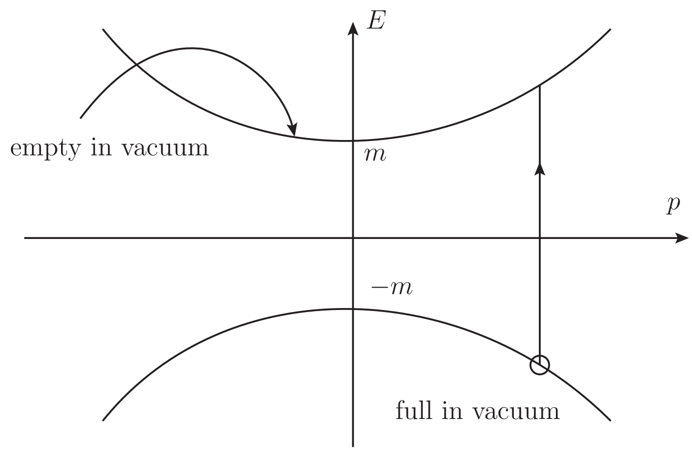


Figure 3.1: A schematic of the Dirac hole picture. In vacuum, negative energy states – given by $E = -\sqrt{p^2 + m^2}$ are full, so that no positive energy particles can occupy the negative states. However, a negative energy particle can be promoted to the positive states, and appears as an anti-particle

3.5 Relativistic Covariance of The Dirac Equation

3.5.1 Covariant Notation

The Dirac equation written in the form

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

is not obviously covariant. Now, if we multiply from the left by β, and recall that $\beta^2 = 1$, then we see that

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

writing $\beta\boldsymbol{\alpha}\cdot\nabla$ as

$$\beta\boldsymbol{\alpha}\cdot\nabla = \beta\alpha^i\frac{\partial}{\partial x^i},$$

then we see that the Dirac equation looks like

$$i\beta\frac{\partial\psi}{\partial t} = \left(-i\beta\alpha^i\frac{\partial}{\partial x^i} + m\right)\psi. \tag{3.40}$$

Now, if we make the definitions,

$$\gamma^0 \equiv \beta, \quad \gamma^i \equiv \beta\alpha^i, \tag{3.41}$$

then we see that (3.40) can be written

$$i\gamma^0 \frac{\partial \psi}{\partial t} = \left(-i\gamma^i \frac{\partial}{\partial x^i} + m \right) \psi. \tag{3.42}$$

Then, if we write

$$\gamma^\mu = (\gamma^0, \gamma^i) = (\beta, \beta\alpha^i),$$

we see that (3.42) can be written as

$$(i\gamma^\mu \partial_\mu - m)\, \psi = 0. \tag{3.43}$$

The properties of the γ^μ matrices can be deduced from the properties of β, α^i (i.e. the commutation relations). It is easy to see that $\{\alpha_i, \beta\} = 0$, so that

$$\begin{aligned}
\{\gamma^0, \gamma^i\} &= \{\beta, \beta\alpha^i\} \\
&= \beta^2\alpha^i + \beta\alpha^i\beta \\
&= \alpha^i + \beta(-\beta\alpha^i) \\
&= 0.
\end{aligned}$$

Similarly, one can see that $\{\gamma^0, \gamma^0\} = 2$. Also,

$$\begin{aligned}
\{\gamma^i, \gamma^j\} &= \{\beta\alpha^i, \beta\alpha^j\} \\
&= \beta\alpha^i\beta\alpha^j + \beta\alpha^j\beta\alpha^i \\
&= -\beta^2\alpha^i\alpha^j - \beta^2\alpha^j\alpha^i \\
&= -\{\alpha^i, \alpha^j\}.
\end{aligned}$$

Hence, one can deduce that

$$\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}. \tag{3.44}$$

Also, one can see that

$$\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0, \tag{3.45}$$

which is equivalent to saying $\gamma^{0\dagger} = \gamma^0, \gamma^{i\dagger} = -\gamma^i$. Also, as $\beta^2 = 1$, then $(\gamma^0)^2 = 1$.

It is also useful to introduce the Dirac adjoint,

$$\bar{\psi}(x) \equiv \psi^\dagger(x)\gamma^0. \tag{3.46}$$

If we recall the components of j^μ, from (3.17),

$$j^\mu = (\psi^\dagger\psi, \psi^\dagger\boldsymbol{\alpha}\psi),$$

then we see that using the Dirac adjoint, we can write j^μ as

$$j^\mu(x) = \bar{\psi}(x)\gamma^\mu\psi(x), \tag{3.47}$$

and so the equation of continuity is $\partial_\mu j^\mu = 0$.

Now, this is not obvious, but the γ^μ do not form the components of a 4-vector. Hence, it is still not obvious as to whether (3.43) is covariant or not. If (3.43) is a valid physical equation (which it is), the principle of special relativity says that it must be covariant. Hence, we must look for transformation properties of ψ so that (3.43) is covariant.

3.5.2 Proof of Covariance

Consider a Lorentz transformation from one frame to another, $O \mapsto O'$, such that

$$x \longmapsto x' = ax,$$

so that

$$x'^{\mu} = a^{\mu}{}_{\nu} x^{\nu},$$

where $a^{\mu}{}_{\nu}$ is some transformation matrix that satisfies

$$a^{\mu}{}_{\nu} a_{\mu}{}^{\sigma} = \delta^{\sigma}_{\nu}. \tag{3.48}$$

So, let us define some operator $S(a)$ by

$$\psi(x) \longmapsto \psi'(x') = S(a)\psi(x), \tag{3.49}$$

and, correspondingly,

$$\psi(x) = S^{-1}(a)\psi'(x') = S(a^{-1})\psi(x').$$

We shall state, and then prove, that the Dirac equation (3.43) and (3.47) are covariant, provided,

$$\begin{aligned} S^{-1}(a)\gamma^{\mu}S(a) &= a^{\mu}{}_{\nu}\gamma^{\nu}, && (3.50)\\ \gamma^{0}S^{\dagger}(a)\gamma^{0} &= S^{-1}(a). && (3.51) \end{aligned}$$

Those familiar with the terminology of group theory will notice that (3.50) is just the statement of equivalence. That is, we require $a^{\mu}{}_{\nu}\gamma^{\nu}$ and γ^{μ} to be equivalent, with "equivalence element" $S(a)$. Physically, we are requiring a γ-matrix to still be a γ-matrix, even after transformation (which is not too much to ask of a matrix). One will also notice that (3.51) is just the statement that $S^{-1}(a)$ is equivalent to $S^{\dagger}(a)$, with equivalence element γ^{0}.

So, prove the covariance, we want to show

$$\begin{aligned} \left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi(x) &= 0 \\ \Rightarrow \quad \left(i\gamma^{\mu}\partial'_{\mu} - m\right)\psi'(x') &= 0, \end{aligned} \tag{3.52}$$

where

$$\partial'_{\mu} = a_{\mu}{}^{\nu}\partial_{\nu}. \tag{3.53}$$

Let us substitute (3.49) and (3.53) into (3.52), to give

$$\left(i\gamma^{\mu}a_{\mu}{}^{\nu}\partial_{\nu} - m\right)S(a)\psi(x) = 0,$$

multiplying from the left by $S^{-1}(a)$,

$$\left(iS^{-1}(a)\gamma^{\mu}S(a)a_{\mu}{}^{\nu}\partial_{\nu} - m\right)\psi(x) = 0. \tag{3.54}$$

Now, if we multiply (3.50) by ${a_\mu}^\sigma$ from the right, then we see that

$$\begin{aligned} S^{-1}(a)\gamma^\mu S(a){a_\mu}^\sigma &= {a^\mu}_\nu \gamma^\nu {a_\mu}^\sigma \\ &= \gamma^\nu \delta^\sigma_\nu \\ &= \gamma^\sigma, \end{aligned}$$

after using (3.48) in the second equality. So, using this result in (3.54), gives

$$(i\gamma^\mu \partial_\mu - m)\,\psi(x) = 0,$$

as required. Thus, the Dirac equation is covariant.

We still need to show that (3.47) is covariant. That is, show that

$$j'^\mu = \bar{\psi}'\gamma^\mu\psi' = {a^\mu}_\nu \bar{\psi}\gamma^\nu\psi \tag{3.55}$$

holds.

Let us consider the transformation law of $\bar{\psi}(x)$,

$$\psi'(x') = S(a)\psi(x) \quad \Rightarrow \quad \psi'^\dagger(x') = \psi^\dagger(x)S^\dagger(a),$$

and multiplying by γ^0 from the right,

$$\psi'^\dagger\gamma^0 = \bar{\psi}' = \psi^\dagger S^\dagger \gamma^0,$$

where the first equality follows by (3.46). Now, if we insert $(\gamma^0)^2 = 1$, between $\psi^\dagger$ and $S^\dagger$, then

$$\bar{\psi}' = \psi^\dagger\gamma^0\gamma^0 S^\dagger\gamma^0 = \bar{\psi}\gamma^0 S^\dagger \gamma^0,$$

where the second equality again follows by the definition of the Dirac adjoint. If we now use (3.51) for $\gamma^0 S^\dagger \gamma^0$, then we see that

$$\bar{\psi}' = \bar{\psi}S^{-1}(a). \tag{3.56}$$

So, inserting this for the $\bar{\psi}'$ in the middle term of (3.55), gives

$$\bar{\psi}'\gamma^\mu\psi' = \bar{\psi}S^{-1}\gamma^\mu\psi',$$

using (3.49) for ψ' on the far RHS,

$$\bar{\psi}S^{-1}\gamma^\mu\psi' = \bar{\psi}S^{-1}\gamma^\mu S\psi,$$

using (3.50) for $S^{-1}\gamma^\mu S$, gives

$$\bar{\psi}S^{-1}\gamma^\mu S\psi = {a^\mu}_\nu\bar{\psi}\gamma^\nu\psi.$$

And therefore,

$$j'^\mu = {a^\mu}_\nu\bar{\psi}\gamma^\nu\psi, \tag{3.57}$$

as required.

Notice that if we multiply (3.56) by ψ' from the right,

$$\bar{\psi}'\psi' = \bar{\psi}S^{-1}\psi' = \bar{\psi}S^{-1}S\psi = \bar{\psi}\psi.$$

That is, $\bar{\psi}\psi$ is a Lorentz scalar: its value is the same in all coordinate systems.

3.5.3 Lorentz Boost

Consider a Lorentz boost along the x-axis. Then, we have a coordinate transformation

$$x' = ax,$$

where

$$a = \begin{pmatrix} \cosh\omega & -\sinh\omega & 0 & 0 \\ -\sinh\omega & \cosh\omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

so that for a frame O' moving at speed v, relative to O, we have

$$\sinh\omega = \frac{v}{\sqrt{1-v^2}} = \gamma v, \quad \cosh\omega = \frac{1}{\sqrt{1-v^2}} = \gamma.$$

The form of the $S(a)$ is

$$S(a) = e^{\omega\alpha_1/2}.$$

Now, consider that

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots, \quad \cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots,$$

and that

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots = \cosh x + \sinh x.$$

Hence,

$$\begin{aligned} \sinh\left(\frac{\omega\alpha_1}{2}\right) + \cosh\left(\frac{\omega\alpha_1}{2}\right) &= 1 + \frac{\omega\alpha_1}{2} + \left(\frac{\omega\alpha_1}{2}\right)^2\frac{1}{2!} + \left(\frac{\omega\alpha_1}{2}\right)^3\frac{1}{3!} + \left(\frac{\omega\alpha_1}{2}\right)^4\frac{1}{4!} + \ldots \\ &= 1 + \frac{\omega}{2}\alpha_1 + \left(\frac{\omega}{2}\right)^2\frac{1}{2!}\mathbf{1} + \left(\frac{\omega}{2}\right)^3\alpha_1\frac{1}{3!} + \left(\frac{\omega}{2}\right)^4\frac{1}{4!}\mathbf{1} + \ldots \\ &= \mathbf{1}\cosh\left(\frac{\omega}{2}\right) + \alpha_1\sinh\left(\frac{\omega}{2}\right), \end{aligned}$$

where we have made use of $\alpha_1^{2n} = \mathbf{1}$ and $\alpha_1^{2n+1} = \alpha_1$. Hence,

$$S(a) = e^{\omega\alpha_1/2} = \mathbf{1}\cosh\left(\frac{\omega}{2}\right) + \alpha_1\sinh\left(\frac{\omega}{2}\right), \tag{3.58}$$

for a Lorentz boost.

3.5.4 Parity

Consider the transformation

$$t \longmapsto t' = t, \quad x_i \longmapsto x'_i = -x_i,$$

so that $x'^{\mu} = a^{\mu}{}_{\nu}x^{\nu}$, where

$$a = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Notice that this matrix satisfies, as indeed it must,

$$a^{\mu}{}_{\nu}a_{\mu}{}^{\sigma} = \delta^{\sigma}_{\nu}.$$

Now, if we have that $S(a) = P$, then

$$\psi'(x') = P\psi(x).$$

Now, for covariance, we require (3.50) and (3.51) to hold,

$$\begin{aligned} P^{-1}\gamma^{\mu}P &= a^{\mu}{}_{\nu}\gamma^{\nu}, \\ P^{-1} &= \gamma^0 P^{\dagger}\gamma^0. \end{aligned}$$

Now, by the form of the matrix a, we see that $P^{-1} = P$, so that $P^2 = 1$. All of these conditions are satisfied by the choice

$$P = \eta\gamma^0, \quad \eta = \pm 1.$$

Recall that in Dirac representation,

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

Particles at Rest Consider a particle at rest. Then, the positive and negative energy solutions (3.37) and (3.39) become

$$\psi^{(+)} = \begin{pmatrix} \chi_s \\ 0 \end{pmatrix} e^{-imt}, \quad \psi^{(-)} = \begin{pmatrix} 0 \\ \chi_{-s} \end{pmatrix} e^{imt}.$$

So, acting P upon both states, gives

$$P\psi^{(+)} = \eta\psi^{(+)}, \quad P\psi^{(-)} = -\eta\psi^{(-)}.$$

Therefore, we see that both the positive and negative energy solutions are eigenstates of parity P, but particles and anti-particles have opposite "intrinsic" parity. This has been experimentally verified, and is crucial for understanding positronium e^+e^- and mesons $q\bar{q}$. We cannot measure η, and we conventionally set it $\eta = 1$, so that $P = \gamma^0$.

Particles in Motion We now consider the full wavefunctions,

$$\psi^{(+)}_{\mathbf{p},s}(x) = N\begin{pmatrix}\chi_s \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi_s\end{pmatrix}e^{-i(Et-\mathbf{p}\cdot\mathbf{x})}, \quad \psi^{(-)}_{\mathbf{p},s}(x) = N\begin{pmatrix}\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi_{-s} \\ \chi_{-s}\end{pmatrix}e^{i(Et-\mathbf{p}\cdot\mathbf{x})}.$$

Hence, using the parity transformation, $P : x \longmapsto -x$,

$$\begin{aligned}
P\psi^{(+)}_{\mathbf{p},s}(x) &= N\begin{pmatrix}\chi_s \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi_s\end{pmatrix}e^{-i(Et'+\mathbf{p}\cdot\mathbf{x}')}, \\
&= \psi^{(+)}_{-\mathbf{p},s}(x') \\
P\psi^{(-)}_{\mathbf{p},s}(x) &= -N\begin{pmatrix}\frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\chi_{-s} \\ \chi_{-s}\end{pmatrix}e^{i(Et'+\mathbf{p}\cdot\mathbf{x}')} \\
&= -\psi^{(-)}_{-\mathbf{p},s}(x').
\end{aligned}$$

Thus again, we see the intrinsic parity difference between particles and anti-particles.

3.6 Interaction With Fields

Consider interactions with an electromagnetic field, $A^\mu(x) = (A^0, \mathbf{A})$, which can be introduced via the minimal substitution

$$\partial_\mu \longmapsto \partial_\mu + iqA_\mu,$$

where q is some charge. We could also introduce interactions with a scalar field, via

$$m \longmapsto m + S(x).$$

Hence, the Dirac equation is rewritten into

$$\left[i\gamma^\mu\left(\partial_\mu + iqA_\mu\right) - m - S\right]\psi(x) = 0,$$

or, indeed,

$$i\frac{\partial\psi}{\partial t} = \left[\boldsymbol{\alpha}\cdot(-i\nabla - q\mathbf{A}) + \beta(m+S) + qA^0\right]\psi. \tag{3.59}$$

3.6.1 Magnetic Moment of Electrons

Let us consider interactions with EM fields only, so that $S = 0$, and look for energy eigenfunctions of the form

$$\psi(x) = \begin{pmatrix}\phi(\mathbf{x}) \\ \eta(\mathbf{x})\end{pmatrix}e^{-iEt}. \tag{3.60}$$

If we substitute (3.60) into (3.59), we see that

$$\boldsymbol{\sigma}\cdot(-i\nabla - q\mathbf{A})\,\eta + (qA^0 + m)\phi \;=\; E\phi, \qquad (3.61)$$
$$\boldsymbol{\sigma}\cdot(-i\nabla - q\mathbf{A})\,\phi + (qA^0 - m)\eta \;=\; E\eta; \qquad (3.62)$$

where the $-m$ in the second expression comes from the -1 in the β matrix.

We shall consider the non-relativistic weak-field limit

$$\epsilon = E - m \ll m, \quad |qA^\mu| \ll m.$$

We neglect the qA^0 term in (3.62), and take $m\eta$ to the RHS, to give

$$\boldsymbol{\sigma}\cdot(-i\nabla - q\mathbf{A})\,\phi = (E+m)\eta,$$

but, $E + m = \epsilon + 2m$, however, $\epsilon \ll m$, so that $E + m \approx 2m$. Therefore,

$$\eta = \frac{1}{2m}\boldsymbol{\sigma}\cdot(-i\nabla - q\mathbf{A})\,\phi. \qquad (3.63)$$

Now, consider (3.61); we can easily rearrange it into

$$\boldsymbol{\sigma}\cdot(-i\nabla - q\mathbf{A})\,\eta + qA^0\phi = (E-m)\phi,$$

and we use that $\epsilon = E - m$ (notice that we do not have that $qA^0 \ll \epsilon$, so that we cannot neglect the qA^0 term). So, inserting (3.63) for η,

$$\frac{1}{2m}\left[\boldsymbol{\sigma}\cdot(-i\nabla - q\mathbf{A})\right]^2\phi + qA^0\phi = \epsilon\phi. \qquad (3.64)$$

Our task now is to tidy this expression up, so that we can read off terms we can recognise.

We first use the easily provable

$$(\boldsymbol{\sigma}\cdot\mathbf{a})(\boldsymbol{\sigma}\cdot\mathbf{b}) = \mathbf{a}\cdot\mathbf{b} + i\boldsymbol{\sigma}\cdot(\mathbf{a}\times\mathbf{b}),$$

so that

$$\begin{aligned}\left[\boldsymbol{\sigma}\cdot(-i\nabla - q\mathbf{A})\right]^2\phi \;=\;& (-i\nabla - q\mathbf{A})\cdot(-i\nabla - q\mathbf{A})\,\phi \\ &+ i\boldsymbol{\sigma}\cdot(-i\nabla - q\mathbf{A})\times(-i\nabla - q\mathbf{A})\,\phi. \end{aligned}\qquad (3.65)$$

Consider the second term,

$$i\boldsymbol{\sigma}\cdot(-i\nabla - q\mathbf{A})\times(-i\nabla - q\mathbf{A})\,\phi = -q\boldsymbol{\sigma}\cdot\left[\nabla\times(\mathbf{A}\phi) + \mathbf{A}\times(\nabla\phi)\right],$$

after noting that curl grad is zero, and that $\mathbf{a}\times\mathbf{a} = 0$. Let us then use the vector identity

$$\nabla\times(\mathbf{A}\phi) = \phi\nabla\times\mathbf{A} - \mathbf{A}\times\nabla\phi,$$

so that the second term of (3.65) is just

$$-q\boldsymbol{\sigma}\cdot\left(\nabla\times\mathbf{A}\right)\phi=-q\boldsymbol{\sigma}\cdot\mathbf{B}\phi,$$

where we have used that $\mathbf{B}=\nabla\times\mathbf{A}$. We can also use that $\mathbf{S}=\frac{1}{2}\boldsymbol{\sigma}$, to see that this term is just

$$-2q\mathbf{S}\cdot\mathbf{B}\phi. \tag{3.66}$$

Let us now consider the first term of (3.65),

$$\left(-i\nabla-q\mathbf{A}\right)\cdot\left(-i\nabla-q\mathbf{A}\right)\phi=-\nabla^2\phi+iq\nabla\cdot\left(\mathbf{A}\phi\right)+iq\mathbf{A}\cdot\nabla\phi+q^2A^2\phi. \tag{3.67}$$

Now, notice that we have

$$\nabla\cdot\left(\mathbf{A}\phi\right)=\nabla\cdot\mathbf{A}\phi+\mathbf{A}\cdot\nabla\phi,$$

and if we use the Coulomb gauge,

$$\nabla\cdot\mathbf{A}=0,$$

then

$$\nabla\cdot\left(\mathbf{A}\phi\right)=\mathbf{A}\cdot\nabla\phi.$$

Hence, (3.67) becomes

$$\left(-\nabla^2+2iq\mathbf{A}\cdot\nabla+q^2A^2\right)\phi. \tag{3.68}$$

Thus far, our discussion has been very general, but, we shall now specify ourselves to a homogeneous magnetic field; which one can show is given by

$$\mathbf{A}=\frac{1}{2}\mathbf{B}\times\mathbf{x}.$$

Notice then that

$$2iq\mathbf{A}\cdot\nabla=iq\left(\mathbf{B}\times\mathbf{x}\right)\cdot\nabla=iq\left(\mathbf{x}\times\nabla\right)\cdot\mathbf{B},$$

but,

$$\mathbf{L}=\mathbf{x}\times\mathbf{p}=-i\mathbf{x}\times\nabla,$$

and hence that

$$2iq\mathbf{A}\times\nabla=-q\mathbf{L}\cdot\mathbf{B}.$$

Thus, using this in (3.68), we see that the first term of (3.65) is

$$\left(-\nabla^2-q\mathbf{L}\cdot\mathbf{B}+q^2A^2\right)\phi.$$

Hence, using this and (3.66) in the first and second terms of (3.65), we see that (3.65) becomes

$$\left[\boldsymbol{\sigma}\cdot\left(-i\nabla-q\mathbf{A}\right)\right]^2\phi=\left(-\nabla^2-2q\mathbf{S}\cdot\mathbf{B}-q\mathbf{L}\cdot\mathbf{B}+q^2A^2\right)\phi.$$

Hence, we can rewrite (3.64) with things that we recognise;

$$\left(-\frac{1}{2m}\nabla^2 + qA^0 - \boldsymbol{\mu}\cdot\mathbf{B} + \frac{q^2}{2m}A^2\right)\phi = \epsilon\phi,$$

where

$$\boldsymbol{\mu} = \frac{q}{2m}\left(\mathbf{L} + 2\mathbf{S}\right).$$

The factor of 2 infront of $\mathbf{S}$ is the so-called g-factor (gyromagnetic ratio) of the electron. This prediction has been verified by experiment.

3.7 Particles in a Spherical Potential

Consider the time independent Dirac equation,

$$\left[-i\boldsymbol{\alpha}\cdot\nabla + \beta(m + S(r)) + V(r)\right]\psi = E\psi. \tag{3.69}$$

Notice that $S(r), V(r)$ depend on $r = |\mathbf{x}|$ only. We have that

$$V(r) = qA^0(r), \quad \mathbf{A} = 0,$$

so that we have an interaction with static fields only.

Let us consider solutions with $j = 1/2, P = 1$. Now, wavefunctions in Dirac representation have the form

$$\psi(r) = \begin{pmatrix} f(r)\chi_s \\ g(r)i\boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\chi_s \end{pmatrix}, \tag{3.70}$$

so that $f(r)$ is large for positive energy solutions, and $g(r)$ is large for negative energy solutions, and both f, g are scalars. Let us then substitute (3.70) into (3.69),

$$\begin{pmatrix} m + S + V & -i\boldsymbol{\sigma}\cdot\nabla \\ -i\boldsymbol{\sigma}\cdot\nabla & -m - S + V \end{pmatrix}\begin{pmatrix} f(r)\chi_s \\ g(r)i\boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\chi_s \end{pmatrix} = E\begin{pmatrix} f(r)\chi_s \\ g(r)i\boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\chi_s \end{pmatrix}. \tag{3.71}$$

We can use that

$$(\boldsymbol{\sigma}\cdot\nabla)(\boldsymbol{\sigma}\cdot\hat{\mathbf{r}})g(r) = \frac{2g(r)}{r} + \frac{dg}{dr}.$$

and expand (3.71) to see that we get two coupled differential equations,

$$\left(\frac{d}{dr} + \frac{2}{r}\right)g + (m + S + V - E)\,f = 0, \tag{3.72}$$

$$\frac{df}{dr} + (m + S - V + E)\,g = 0. \tag{3.73}$$

These coupled differential equations can then be solved for f, g. Notice that particles have f large, and anti-particles have g being large.

3.7.1 The MIT Bag Model

This is a very crude model of hadrons. For example, the *uud* quarks are confined to make a proton p, with mass $m_{\mathrm{p}} = 940\mathrm{MeV}/c^2$ and radius $R_{\mathrm{p}} = 1\mathrm{fm}$. If we consider typical quark momenta, taking $\lambda = 2R$ and $p = h/\lambda = 600\mathrm{MeV}$, we see that quarks inside a protons are highly relativistic. We also know that quarks strongly interact at distances $r \sim 1\mathrm{fm}$, and interact weakly at $r \ll 1\mathrm{fm}$ (their so-called asymptotic freedom). So, the simple model we will discuss – the MIT bag model – binds quarks in a scalar potential

$$S(r) = \begin{cases} 0 & r < R, \\ S_0 & r > R; \end{cases}$$

where $R \sim 1\mathrm{fm}$. So, we have some scalar potential, which will allow quarks and anti-quarks to attract (note, we want anti-quarks to attract as well, as we would like to be able to allow anti-protons to exist).

We shall set $m_{\mathrm{q}} = 0$ (which is a good approximation for u, d-quarks). Notice that we are not setting the constituent mass to zero, which is the mass of the bare quark and the virtual gluons.

So, the radial equations (3.72), (3.73) become for the various limits:

Inside $r < R$

$$\begin{aligned} \left(\frac{d}{dr} + \frac{2}{r}\right) g - Ef &= 0, \\ \frac{df}{dr} + Eg &= 0, \end{aligned}$$

we can substitute the second,

$$g = -\frac{1}{E}\frac{df}{dr}, \tag{3.74}$$

into the first,

$$\frac{d^2 f}{dr^2} + \frac{2}{r}\frac{df}{dr} + E^2 f = 0. \tag{3.75}$$

We can solve this by setting

$$f(r) = \frac{u(r)}{r},$$

so that upon substitution into (3.75), one can easily find

$$\frac{d^2 u}{dr^2} = -E^2 u,$$

which has solution

$$u(r) = A_1 \sin Er + A_2 \cos Er.$$

Now, under the requirement that $f(r)$ is finite at $r = 0$, we must therefore set $A_2 = 0$, and hence

$$f(r) = \frac{A \sin Er}{r}. \tag{3.76}$$

Then, by (3.74), we see that

$$g(r) = A\left(\frac{\sin Er}{Er^2} - \frac{\cos Er}{r}\right). \tag{3.77}$$

Outside $r > R$

$$\begin{aligned} \left(\frac{d}{dr} + \frac{2}{r}\right) g + (S_0 - E) f &= 0, \\ \frac{df}{dr} + (S_0 + E) g &= 0, \end{aligned}$$

again, substituting the second into the first to give

$$\frac{d^2 f}{dr^2} + \frac{2}{r}\frac{df}{dr} + \left(E^2 - S_0^2\right) f = 0.$$

To solve this, we make a similar ansatz, and we find

$$f(r) = \frac{Be^{-\kappa r}}{r}, \quad \kappa \equiv \sqrt{S_0^2 - E^2},$$

where we used the boundary condition that f should not diverge as $r \to \infty$, and

$$g(r) = \frac{B}{S_0 + E}\left(\frac{\kappa e^{-\kappa r}}{r} - \frac{e^{-\kappa r}}{r^2}\right).$$

If we assume that $S_0 \gg E$ (i.e. strong potential), then $\kappa \approx S_0$ and

$$g(r) = \frac{Be^{-\kappa r}}{r} \approx f(r).$$

We can now apply continuity at the boundary $r = R$, so that (3.76) and (3.77) are equal from below,

$$f(R) = g(R) \quad \Rightarrow \quad \frac{\sin ER}{R} = \frac{\sin ER}{ER^2} - \frac{\cos ER}{R},$$

or,

$$\tan ER = \frac{ER}{1 - ER}.$$

This can be graphically solved to give

$$ER = 2.04.$$

Therefore,

$$E_{s\frac{1}{2}} = \frac{2.04}{R},$$

or, if we have 3 quarks,

$$3E = \frac{6.12}{R}.$$

Hence, the mass of the proton, in this crude model, is $m_{\mathrm{p}} = 3E = 1200\mathrm{MeV}/c^2$. This isnt too far off, given that we have completely ignored things like the quarks interact with nothing except the binding bag – so, ignoring the gluons around them.

3.7.2 Hydrogen-like Atoms

Now consider a system with

$$S(r) = 0, \quad V(r) = -\frac{Z\alpha}{r}$$

Then, in a similar way to above, one can solve the radial equations, and find

$$E_n = m\left[1 + \frac{Z^2\alpha^2}{n - \left(j + \frac{1}{2}\right) + \sqrt{\left(j + \frac{1}{2}\right)^2 - Z^2\alpha^2}}\right]^{-1/2}. \tag{3.78}$$

Notice that ℓ does not appear in this expression, only j (and n, obviously), and this is because the Dirac equation "knows about" spin. Also notice that for $Z^2\alpha^2 > j + \frac{1}{2}$, the energy becomes imaginary (for $j = 1/2$, this corresponds to atoms with $Z > 137$) – this is another instance of the Klein paradox for strong potentials.

Consider a hydrogen atom, with $Z = 1$, then, expanding in powers of α, one finds

$$\epsilon_n = E_n - m = -\frac{m\alpha^2}{2n^2}\left[1 + \frac{\alpha^2}{n}\left(\frac{1}{j + \frac{1}{2}} - \frac{3}{4n}\right) + \mathcal{O}(\alpha^4)\right].$$

The first term gives the Rydberg energy, and the α^2 -term is the first relativistic correction and is $\mathcal{O}(10^{-5})$.

With reference to Figure (3.2), we see the spectrum of the $n = 2$ level, and the splitting. In the non-relativistic case, the spin-orbit interaction is added in manually, as is $g_{\mathrm{s}} = 2$. Notice that the $2s_{1/2}$ level splits into $2p_{1/2}$ and $2p_{3/2}$, and those two levels are separated by an energy gap Δ.

Notice that in the relativistic spectrum, the splitting gives a degeneracy between $2s_{1/2}$ and $2p_{1/2}$; this is because in (3.78), there is no mention of ℓ, and all energy-information is

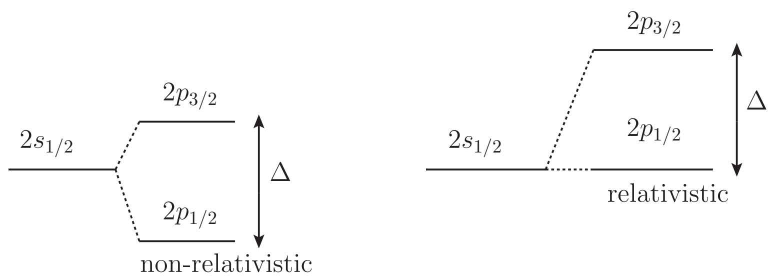


Figure 3.2: The spectrum of splittings in the H-atom. In the non-relativistic case, the spin-orbit coupling is put in "by hand", but in the relativistic case, the Dirac equation "knows about" spin inherently.

contained in j. This degeneracy is exact to all orders of expansion. The states have the same splitting Δ as in the non-relativistic case. In this relativistic case, the Dirac equation already "knew about" spin, and orbital angular momentum, and one can predict the magnetic moment $g_s = 2$, so that one does not need to add them in "by hand", as one did for the non-relativistic case.

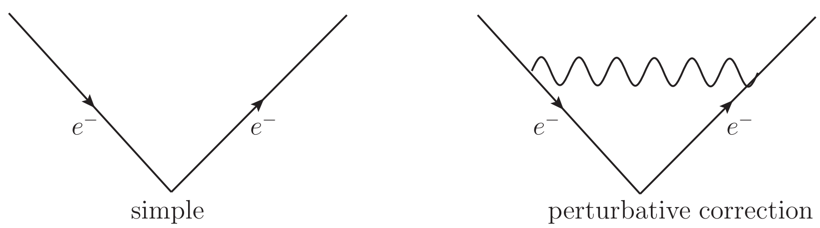


Figure 3.3: A difference in the simple Dirac model and a more complicated – but needed – perturbative correction.

There was fantastic agreement in the relativistic prediction of these splittings, until 1947, when it was observed that the $2s_{1/2}$-level was shifted down by $\mathcal{O}(\alpha^3)$. This shift is called the *Lamb shift*, and was promptly awarded a Nobel prize. The reason behind the failure is that one needs to go further than the rather simple Dirac equation, and quantum fluctuations come into play; for example, as in Figure (3.3). Also, one must think that problems are present when one gets complex energies for strong potentials!

Also, the magnetic moment is not exactly $g_s = 2$ in the more complicated model;

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

which is again confirmed by experiment.

This highlights the need for a more complicated description of relativistic quantum systems.

4 Quantum Fields

We now want to move on to a many-body theory, in which particle numbers can change. An idea of how to do this, lies in that photons are quanta of the electromagnetic field, and that the KG and Dirac equations are wave equations. But, how do we quantise waves? The general idea is to expand a field in terms of modes of harmonic oscillators, where the modes are operators. It is the technicalities of doing this, and some of the results of doing this, which we shall consider in this section. The idea of assigning an operator to a field, or to a quantity in general, is called secondary quantisation.

We use fields, where a field is some continuous dynamical variable which has a value at each point in space. For example, the displacement field of a string, or density of a compressible fluid, or the electromagnetic field. Fields can support waves, and can transmit energy, momentum and forces; and have infinitely many degrees of freedom.

4.1 Quantum Mechanics of a String

Consider a string, length L, with tension T having transverse displacement $\phi(t, z)$. The string will satisfy the wave equation

$$\frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = \frac{\partial^2 \phi}{\partial z^2},$$

where the speed of sound is

$$c = \sqrt{\frac{T}{\mu}},$$

where μ is the mass per unit length of the string. The energy of the string is given by

$$E = \int_0^L dz \left[\frac{1}{2}\mu \left(\frac{\partial \phi}{\partial t}\right)^2 + \frac{1}{2}T\left(\frac{\partial \phi}{\partial z}\right)^2 \right]. \tag{4.1}$$

For simplicity, we shall set $T = \mu = 1$, so that $c = 1$ and so the wave equation and energy become

$$\frac{\partial^2 \phi}{\partial t^2} = \frac{\partial^2 \phi}{\partial z^2}, \quad E = \int_0^L dz \left[\frac{1}{2}\left(\frac{\partial \phi}{\partial t}\right)^2 + \frac{1}{2}\left(\frac{\partial \phi}{\partial z}\right)^2 \right].$$

The first terms is a kinetic energy, and the second a potential energy density. Notice that the wave equation is like a KG equation, with $m = 0$: massless.

Now, we want to impose travelling waves. We let our theory do so by imposing periodic boundaries,

$$\phi(t, z) = \phi(t, z + L),$$

where we could let $L \to \infty$, which we will do at the end of the calculation. Now, we can expand the solution in terms of normal modes,

$$\phi_n = e^{\pm i(\omega_n t - k_n z)},$$

where

$$k_n = \frac{2\pi n}{L},$$

and n is an integer. Then, with this choice of k_n, the periodic boundary condition is satisfied. Upon substitution of the normal mode into the wave equation, we find the requirement,

$$\omega_n = |k_n|,$$

a dispersion relation. Notice that at each position z along the string, the string performs simple harmonic motion. We can expand an arbitrary ϕ in terms of the normal modes ϕ_n,

$$\phi(t,z) = \sum_{n=-\infty}^{\infty} \frac{1}{\sqrt{2\omega_n L}} \left[a_n e^{-i(\omega_n t - k_n z)} + a_n^* e^{i(\omega_n t - k_n z)} \right]. \tag{4.2}$$

The prefactor $1/\sqrt{2\omega_n L}$ is there purely for later convenience – if we did not put it in, we would have found that equations become more complicated than if we did. The first term in the brackets bears resemblance with the positive energy solutions, and the second term to negative energy solutions.

We now want to work out the Hamiltonian – i.e. the energy – in terms of the coefficients a_n.

So, the kinetic energy, where we evaluate at $t = 0$,

$$\begin{aligned}
\frac{1}{2}\int_0^L dz \left(\frac{\partial\phi}{\partial t}\right)^2 &= \frac{1}{2}\sum_{n,m}\int_0^L dz \, \frac{1}{2L\sqrt{\omega_n\omega_m}} \left\{ \left[-i\omega_n a_n e^{ik_n z} + i\omega_n a_n^* e^{-ik_n z} \right] \right. \\
&\qquad \left. \times \left[-i\omega_m a_m e^{ik_m z} + i\omega_m a_m^* e^{-ik_m z} \right] \right\} \\
&= \sum_{n,m}\int_0^L dz \, \frac{1}{4L\sqrt{\omega_n\omega_m}} \omega_n\omega_m \left\{ -a_n a_m e^{i(k_n+k_m)z} + a_n a_m^* e^{i(k_n-k_m)z} \right. \\
&\qquad \left. \times + a_n^* a_m e^{i(k_m-k_n)z} - a_n^* a_m^* e^{-i(k_n+k_m)z} \right\} \\
&= -\sum_{n,m} \frac{\sqrt{\omega_n\omega_m}}{4L} L \left\{ a_n a_m \delta_{n,-m} - a_n a_m^* \delta_{nm} - a_n^* a_m \delta_{mn} + a_n^* a_m^* \delta_{n,-m} \right\},
\end{aligned}$$

where in the last equality, we made use of the orthogonality relation

$$\int_0^L e^{i(k_n-k_m)z} dz = \delta_{nm} L.$$

Hence, performing the sum over m, we find an expression for the kinetic energy,

$$\frac{1}{2}\int_0^L dz \left(\frac{\partial\phi}{\partial t}\right)^2 = -\sum_n \frac{\omega_n}{4} \left(a_n a_{-n} - 2a_n^* a_n + a_n^* a_{-n}^* \right).$$

In a very similar fashion, we compute the potential energy,

$$\frac{1}{2}\int_0^L dz \left(\frac{\partial\phi}{\partial z}\right)^2 = \sum_n \frac{\omega_n}{4}\left(a_n a_{-n} + 2a_n^* a_n + a_n^* a_{-n}^*\right),$$

where we have used $|k_{-n}| = k_n$ and $\omega_{-n} = \omega_n$. Notice that we have not been careful with allowing the coefficients a_n to commute – this is because they are purely classical in nature, and therefore commute. We then add the potential to the kinetic energy to find the total energy,

$$H = \sum_n \omega_n a_n^* a_n. \tag{4.3}$$

To make this look more familiar, we could introduce the real variables,

$$q_n \equiv \frac{a_n + a_n^*}{\sqrt{2\omega_n}}, \quad p_n \equiv -i\sqrt{\frac{\omega}{2}}\left(a_n - a_n^*\right),$$

so that upon rearrangement and insertion into the total energy (4.3), we find

$$H = \sum_n \frac{1}{2}\left(p_n^2 + \omega_n^2 q_n^2\right).$$

We can then compare this with the usual expression for the energy of a simple harmonic oscillator,

$$H = \frac{p^2}{2m} + \frac{1}{2}\omega^2 x^2,$$

to see that (4.3) is actually equivalent to an infinite discrete sum over modes of the harmonic oscillator.

So far, our analysis has been completely classical in nature. To quantise the theory, we promote q_n, p_n to Hermitian operators, and a_n, a_n^* to the $a_n, a_n^\dagger$ operators. We also impose the canonical commutation relations

$$[p_n, q_m] = -i\delta_{nm}, \quad [p_n, p_m] = [q_n, q_m] = 0; \tag{4.4}$$

and hence,

$$\left[a_n, a_m^\dagger\right] = \delta_{nm}, \quad [a_n, a_m] = [a_n^\dagger, a_m^\dagger] = 0. \tag{4.5}$$

To find the Hamiltonian of this theory (i.e. replacement of quantities with non-commuting operators), we repeat the procedure that resulted in (4.3), but being careful with ordering; and we find

$$H = \sum_n \frac{\omega_n}{2}\left(a_n^\dagger a_n + a_n a_n^\dagger\right), \tag{4.6}$$

or, using the commutator $[a_n, a_n^\dagger] = 1$,

$$H = \sum_n \omega_n \left(a_n^\dagger a_n + \frac{1}{2} \right). \tag{4.7}$$

The identity

$$[AB, C] = A[B, C] + [A, C]B \tag{4.8}$$

is very useful in computing

$$[H, a_n^\dagger] = \omega_n a_n^\dagger, \quad [H, a_n] = -\omega_n a_n. \tag{4.9}$$

4.1.1 Phonons

Let us consider the state $|\psi\rangle$, and suppose that the Hamiltonian acting upon the state gives the energy,

$$H |\psi\rangle = E |\psi\rangle .$$

Now, consider

$$\begin{aligned} H \left(a_n^\dagger |\psi\rangle \right) &= H a_n^\dagger |\psi\rangle \\ &= \left(\omega_n a_n^\dagger + a_n^\dagger H \right) |\psi\rangle \\ &= \left(E + \omega_n \right) a_n^\dagger |\psi\rangle . \end{aligned}$$

The second equality follows after using the first of the commutation relations (4.9). Similarly, one can compute that

$$H a_n |\psi\rangle = \left(E - \omega_n \right) a_n |\psi\rangle .$$

Therefore, we see that $a_n^\dagger$ is an eigenstate of the Hamiltonian, with eigenvalue $E + \omega_n$, and that a_n is an eigenstate with eigenvalue $E - \omega_n$. Therefore, we say that $a_n^\dagger$ increases the energy of a system by ω_n, and a_n decreases the energy by ω_n.

We define the *ground state* of a system to be the state with lowest possible energy, denoted by $|0\rangle$. Hence, as it is the state with lowest energy, any attempt to lower the energy with a_n results in zero;

$$a_n |0\rangle = 0, \quad \forall n. \tag{4.10}$$

We can then construct *excited states* from the ground state, by applying $a_n^\dagger$ to the ground state as many times as we wish,

$$|\{N_n\}\rangle = \prod_n \frac{\left(a_n^\dagger\right)^{N_n}}{\sqrt{N_n!}} |0\rangle , \tag{4.11}$$

so that we will have N_n quanta in the mode n. The energy eigenvalues are then given by

$$E(\{N_n\}) = \sum_{n=-\infty}^{\infty} \left(N_n + \frac{1}{2} \right) \omega_n.$$

Notice that in this expression for the energy, we have an infinity; the zero-point energy,

$$H|0\rangle = E_{\text{ZP}} = \frac{1}{2} \sum_{n=-\infty}^{\infty} \omega_n,$$

this sum is over an infinite range, and hence gives an infinite value. This infinity can be "taken care of", by redefining the energy,

$$E \longmapsto E - E_{\text{ZP}},$$

or, equivalently, redefining the Hamiltonian,

$$H \longmapsto H = \sum_n \omega_n a_n^\dagger a_n.$$

In practice, this infinite energy is not a problem: we measure energies relative to some value – we measure energy differences, as opposed to absolute values. Hence, in this definition, the energy is just the sum over the occupation number of each mode,

$$E = \sum_n N_n \omega_n.$$

This has the interpretation of looking like a state with N_n particles (or phonons) with energy ω_n. We say that $a_n^\dagger$ creates a particle, and a_n destroys a particle. Just as we say that photons are the quanta of electromagnetic waves, we say that phonons are the quanta of elastic waves.

So, what about momentum, as waves carry momentum. We will make a guess – and confirm later – that the momentum operator in the z-direction is

$$P_z = \sum_n k_n a_n^\dagger a_n. \tag{4.12}$$

Recall that $k_n = 2\pi/\lambda_n$, so that we recover the de Broglie relation $P = k$, with $\hbar = 1$. By analogy with (4.9), we can easily see that

$$[P_z, a_n^\dagger] = k_n a_n^\dagger, \quad [P_z, a_n] = -k_n a_n. \tag{4.13}$$

Thus, we see that $a_n^\dagger$ increases, and a_n decreases, momenta by units of k_n, and energy by units of ω_n. That is, the action of $a_n^\dagger$ upon a system, is to create a massless particle with $\omega_n = k_n$. The particle is associated with the wavefunction $e^{-i(\omega_n t - k_n z)}$, as usual, but, this is not the probability amplitude. The states $|\psi\rangle$ are the probability amplitudes.

Finally, the states are symmetric under interchange of particles; since, for example,

$$|k_n, k_m\rangle = a_n^\dagger a_m^\dagger |0\rangle = a_m^\dagger a_n^\dagger |0\rangle = |k_m, k_n\rangle ,$$

where the second equality follows from the commutator $[a_n^\dagger, a_m^\dagger] = 0$, (4.5). We can also place any number of particles in a given state, so that $N_n = 0, 1, 2, \ldots$ is allowed, with identical properties k_n, ω_n.

Hence, the particles described by this formalism (infact, it all stems from the usage of commutators, rather than anti-commutators) are bosons.

4.1.2 Field Operators

Recall (4.2),

$$\phi(t,z) = \sum_{n=-\infty}^{\infty} \frac{1}{\sqrt{2\omega_n L}} \left[a_n e^{-i(\omega_n t - k_n z)} + a_n^\dagger e^{i(\omega_n t - k_n z)} \right] ,$$

where we have now promoted the coefficients to operators. Hence, the field ϕ is a *field operator*. Infact, the operator is time dependent, but the states $|\{N_n\}\rangle$ are independent of time. Therefore, we are in the Heisenberg picture of quantum mechanics – operators evolve, but states do not. Following from the commutators (4.9), we thus have

$$[H, \phi] = \sum_n \sqrt{\frac{\omega_n}{2L}} \left[-a_n e^{-i(\omega_n t - k_n z)} + a_n^\dagger e^{i(\omega_n t - k_n z)} \right] ,$$

and therefore, by inspection,

$$[H, \phi] = -i \frac{\partial \phi}{\partial t}, \tag{4.14}$$

which is the Heisenberg equation of motion. Hence, we have H in terms of ϕ. So, we have the (verifiable)

$$P_z = \sum_n k_n a_n^\dagger a_n = -\int_0^L dz \frac{\partial \phi}{\partial t} \frac{\partial \phi}{\partial z},$$

and, using the commutators (4.13), we can easily verify that

$$[P_z, \phi] = i \frac{\partial \phi}{\partial z}. \tag{4.15}$$

Now, consider some wave equation,

$$\frac{\partial^2 \phi}{\partial t^2} = \frac{\partial^2 \phi}{\partial z^2},$$

where ϕ is now a Lorentz scalar, and using a unit speed of light. That is, this is just the invariant Klein-Gordon equation, for a massless particle,

$$\Box \phi = \partial_\mu \partial^\mu \phi = 0,$$

in (1 + 1)-dimensions; where $x^\mu = (t, z)$ and $x_\mu = (t, -z)$. So, let us say that $H = E$ must then be a component of the energy-momentum vector $P^\mu = (H, P_z)$, and the invariant generalisation of the Heisenberg equation of motion is

$$i\frac{\partial\phi}{\partial x^\mu} = [\phi, P_\mu]; \tag{4.16}$$

that is, it has two components,

$$i\frac{\partial\phi}{\partial t} = [\phi, H], \quad i\frac{\partial\phi}{\partial z} = -[\phi, P_z],$$

which is in accordance with (4.14) and (4.15), after noting that $[\phi, H] = -[H, \phi]$.

4.2 The Klein-Gordon Field

Let us consider quantising fields that satisfy the Klein-Gordon equation. We shall work in a box, with volume $V = L^3$, under periodic boundaries. Hence, the normal modes are discrete (as a finite sized box), and are of the form

$$e^{\pm ipx}, \quad px = p^\mu x_\mu = Et - \mathbf{p}\cdot\mathbf{x}.$$

The momentum and energy are given by

$$\mathbf{p} = \frac{2\pi}{L}\left(n_x, n_y, n_z\right), \quad E_p \equiv E(\mathbf{p}) = \left|\sqrt{\mathbf{p}^2 + m^2}\right|.$$

We shall consider complex scalar fields, so that the field operator is

$$\phi(x) = \sum_{\mathbf{p}} \frac{1}{\sqrt{2VE_{\mathbf{p}}}}\left(a(\mathbf{p})e^{-ipx} + c^\dagger(\mathbf{p})e^{ipx}\right) \tag{4.17}$$

We have that $a(\mathbf{p}) \neq c(\mathbf{p})$, as we do not want $\phi = \phi^\dagger$ (in general, at least). We have that the coefficients $a, a^\dagger, c, c^\dagger$ are operators, and we assume the usual commutation relations

$$\left[a(\mathbf{p}), a^\dagger(\mathbf{p}')\right] = \delta_{\mathbf{p}\mathbf{p}'} = \left[c(\mathbf{p}), c^\dagger(\mathbf{p}')\right]. \tag{4.18}$$

This holds on a component-by-component basis, so that $\delta_{\mathbf{p}\mathbf{p}'} = \delta_{p_1 p_1'}\delta_{p_2 p_2'}\delta_{p_3 p_3'}$. All other commutators vanish.

So, what about the Hamiltonian? The essential thing is that we require the Heisenberg equation of motion,

$$i\frac{\partial\phi}{\partial t} = [\phi, H],$$

to hold. Consider the Hamiltonian for a 1D classical massless real string, (4.1),

$$H = \int_0^L dz\left[\frac{1}{2}\left(\frac{\partial\phi}{\partial t}\right)^2 + \frac{1}{2}\left(\frac{\partial\phi}{\partial z}\right)^2\right].$$

Notice that we have appropriately chosen units to make the propagation speed unity. The generalisation of this, to a massive complex 3D field, is simple,

$$H = \int d^3x \left[\frac{\partial \phi^\dagger}{\partial t}\frac{\partial \phi}{\partial t} + (\nabla\phi)^\dagger \cdot (\nabla\phi) + m^2\phi^\dagger\phi \right]. \tag{4.19}$$

Notice where the "mass term" comes in (i.e. the prefactor of $\phi^\dagger\phi$); also notice that there is no factor of one-half (this is because we treat ϕ and $\phi^\dagger$ independently). We then expand this Hamiltonian by inserting our field operator (4.17).

We will make use of the orthogonality relations

$$\int d^3x e^{-ipx} e^{ip'x} = V\delta_{pp'}, \quad \int d^3x e^{ipx} e^{ip'x} = V\delta_{p,-p'}.$$

So, we have

$$\begin{aligned}
\frac{\partial \phi^\dagger}{\partial t}\frac{\partial \phi}{\partial t} = \sum_{p,p'} \frac{E_p E_{p'}}{2V\sqrt{E_p E_{p'}}} \quad & \Big(a^\dagger(p)a(p')e^{ipx}e^{-ip'x} \\
& -a^\dagger(p)c^\dagger(p')e^{ipx}e^{ip'x} \\
& -c(p)a(p')e^{-ipx}e^{-ip'x} \\
& +c(p)c^\dagger(p')e^{-ipx}e^{ip'x} \Big), \\
\left(\nabla\phi^\dagger\right)\cdot(\nabla\phi) = \sum_{p,p'} \frac{p_i p_i'}{2V\sqrt{E_p E_{p'}}} \quad & \Big(a^\dagger(p)a(p')e^{ipx}e^{-ip'x} \\
& -a^\dagger(p)c^\dagger(p')e^{ipx}e^{ip'x} \\
& -c(p)a(p')e^{-ipx}e^{-ip'x} \\
& +c(p)c^\dagger(p')e^{-ipx}e^{ip'x} \Big);
\end{aligned}$$

and hence, using the orthogonality condition,

$$\begin{aligned}
\int d^3x \frac{\partial \phi^\dagger}{\partial t}\frac{\partial \phi}{\partial t} &= \sum_p \left\{ \frac{E_p}{2}\left[a^\dagger(p)a(p) + c(p)c^\dagger(p)\right] \right. \\
& \left. - \frac{E_p E_{-p}}{2\sqrt{E_p E_{-p}}}\left[a^\dagger(p)c^\dagger(-p) + c(p)a(-p)\right] \right\},
\end{aligned} \tag{4.20}$$

$$\begin{aligned}
\int d^3x \left(\nabla\phi^\dagger\right)\cdot(\nabla\phi) &= \frac{1}{2}\sum_p \left\{ \frac{\mathbf{p}^2}{E_p}\left[a^\dagger(p)a(p) + c(p)c^\dagger(p)\right] \right. \\
& \left. + \frac{\mathbf{p}^2}{\sqrt{E_p E_{-p}}}\left[a^\dagger(p)c^\dagger(-p) + c(p)a(-p)\right] \right\}.
\end{aligned} \tag{4.21}$$

We can also compute that

$$\int d^3x\, m^2\phi^\dagger\phi = \frac{1}{2}\sum_p m^2 \left\{ \frac{1}{E_p}\left[a^\dagger(p)a(p) + c(p)c^\dagger(p)\right] + \frac{1}{\sqrt{E_pE_{-p}}}\left[a^\dagger(p)c^\dagger(-p) + c(p)a(-p)\right] \right\}. \tag{4.22}$$

Now, notice that

$$E_p = \left|\sqrt{\mathbf{p}^2 + m^2}\right| \quad \Rightarrow \quad E_p = E_{-p},$$

which can be used to see that upon adding (4.21) and (4.22), we have a factor of E_p which cancels from the second term of (4.20). And hence, using these last two results, we can write that the Hamiltonian is

$$\begin{aligned} H &= \int d^3x \left[\frac{\partial\phi^\dagger}{\partial t}\frac{\partial\phi}{\partial t} + (\nabla\phi)^\dagger\cdot(\nabla\phi) + m^2\phi^\dagger\phi\right] \\ &= \sum_p E_p\left(a^\dagger a + c^\dagger c + 1\right), \end{aligned}$$

where in the last step we used the commutator (4.18). Hence, the Hamiltonian is

$$H = \sum_{\mathbf{p}} E(\mathbf{p})\left(a^\dagger(\mathbf{p})a(\mathbf{p}) + c^\dagger(\mathbf{p})c(\mathbf{p}) + 1\right), \tag{4.23}$$

Thus, this is like two infinite sets of simple harmonic oscillators – this corresponds to the two degrees of freedom in a complex field. We shall ignore the zero-point energy in what follows.

Thus, we can then easily compute the following commutation relations, using (4.18),

$$\begin{aligned} \left[H, a^\dagger(\mathbf{p})\right] &= E(\mathbf{p})a^\dagger(\mathbf{p}), \qquad [H, a(\mathbf{p})] = -E(\mathbf{p})a(\mathbf{p}), \\ \left[H, c^\dagger(\mathbf{p})\right] &= E(\mathbf{p})c^\dagger(\mathbf{p}), \qquad [H, c(\mathbf{p})] = -E(\mathbf{p})c(\mathbf{p}); \end{aligned} \tag{4.24}$$

which are similar to (4.9). Hence, we can use these to show that

$$i\frac{\partial\phi}{\partial t} = [\phi, H]$$

holds. That is, our formalism of a quantised complex scalar field, with the assumed commutators, is consistent with the KG equation of motion. By covariance of the theory, we have that the 4-momentum is $P^\mu = (H, \mathbf{P})$, where

$$\mathbf{P} = \sum_{\mathbf{p}} \mathbf{p}\left(a^\dagger a + c^\dagger c\right);$$

and we have a set of commutators exactly like (4.24), except H is replaced by $\mathbf{P}$ and E_p by $\mathbf{p}$, for example $\left[\mathbf{P}, a^\dagger(\mathbf{p})\right] = \mathbf{p}a^\dagger(\mathbf{p})$.

So, what we see, is that $a^\dagger(\mathbf{p})$ and $c^\dagger(\mathbf{p})$ both create spin-0 particles, with 4-momentum $(E(\mathbf{p}), \mathbf{p})$ and mass m; and correspondingly, $a(\mathbf{p})$ and $c(\mathbf{p})$ both destroy spin-0 particles, with 4-momentum $(E(\mathbf{p}), \mathbf{p})$ and mass m. Then, we may ask, what is the difference between a and c. To answer this, consider the conserved 4-current for the KG equation,

$$j^\mu(x) = iq \left(\phi^\dagger \left(\partial^\mu \phi\right) - \left(\partial^\mu \phi^\dagger\right) \phi\right),$$

where we have merely inserted a constant q. Now, corresponding to this, is a conserved charge,

$$\begin{aligned} Q &= \int d^3x \, j^0(x) && (4.25) \\ &= iq \int d^3x \left(\phi^\dagger \frac{\partial \phi}{\partial t} - \frac{\partial \phi^\dagger}{\partial t} \phi\right). && (4.26) \end{aligned}$$

We can then insert the fields, and expand just as we did for the Hamiltonian (and we still ignore the vacuum energy part); and we find

$$Q = q \sum_{\mathbf{p}} \left(a^\dagger(\mathbf{p}) a(\mathbf{p}) - c^\dagger(\mathbf{p}) c(\mathbf{p})\right). \tag{4.27}$$

Hence, we see that the first term is the number operator for particles with charge q, and the second term is the number operator for particles with charge $-q$. We can now use this conserved charge to construct a set of commutators (which we derive either directly or by analogy – they are effectively trivial to see, given (4.24));

$$\begin{aligned} \left[Q, a^\dagger(\mathbf{p})\right] &= q a^\dagger(\mathbf{p}), \qquad & [Q, a(\mathbf{p})] &= -q a(\mathbf{p}), \\ \left[Q, c^\dagger(\mathbf{p})\right] &= -q c^\dagger(\mathbf{p}), \qquad & [Q, c(\mathbf{p})] &= q c(\mathbf{p}). \end{aligned} \tag{4.28}$$

Therefore, we can see that $a^\dagger$ and c increase the charge by q, and $a, c^\dagger$ decrease the charge by q.

Hence, putting this all together, we say that $a^\dagger$ creates particles with charge $Q = +q$, $c^\dagger$ creates anti-particles with charge $Q = -q$; and a destroys particles with charge $Q = +q$, c destroys anti-particles with charge $Q = -q$. It is interesting to notice that if the field ϕ were Hermitian, so that we only had $a(\mathbf{p})$-type coefficients (i.e no $c(\mathbf{p})$-type), then the charge Q is zero. Hence, to have charge, we require non-Hermitian fields; with Hermitian fields being uncharged.

In deriving the charge operator Q, one computes the quantities

$$\phi^\dagger \frac{\partial \phi}{\partial t} = -i \sum_{p,p'} \frac{E_{p'}}{2V\sqrt{E_p E_{p'}}} \Big(a^\dagger(p)a(p')e^{ipx}e^{-ip'x} - a^\dagger(p)c^\dagger(p')e^{ipx}e^{ip'x} + c(p)a(p')e^{-ipx}e^{-ip'x} - c(p)c^\dagger(p')e^{-ipx}e^{ip'x} \Big), \tag{4.29}$$

$$\frac{\partial \phi^\dagger}{\partial t}\phi = i \sum_{p,p'} \frac{E_{p'}}{2V\sqrt{E_p E_{p'}}} \Big(a^\dagger(p)a(p')e^{ipx}e^{-ip'x} + a^\dagger(p)c^\dagger(p')e^{ipx}e^{ip'x} - c(p)a(p')e^{-ipx}e^{-ip'x} - c(p)c^\dagger(p')e^{-ipx}e^{ip'x} \Big). \tag{4.30}$$

Then, integrating the first and fourth terms of both expressions (where the second expression is subtracted from the first), one finds

$$-\frac{i}{2}\sum_p \left(a^\dagger(p)a(p) + a^\dagger(p)a(p) - c(p)c^\dagger(p) - c(p)c^\dagger(p) \right),$$

which one can simplify using commutation relations, to

$$-i\sum_p \left(a^\dagger(p)a(p) - c^\dagger(p)c(p) - 1 \right).$$

Then, the final terms which need integrating, are the second and third terms of both (4.29) and (4.30); these infact cancel due to the symmetry of the summation. Hence, the charge Q is found by multiplying the above by iq.

In a very similar fashion to the above, one can compute the equal-time commutator,

$$\left[\frac{\partial \phi}{\partial t}(t, \mathbf{x}), \phi(t, \mathbf{x}') \right] = -i\delta^{(3)}(\mathbf{x} - \mathbf{x}'),$$

which is essentially the statement that one cannot know the position and velocity of a field at the same time, unless they are at the same place.

It is worth noting, to be inkeeping with the literature, that the format of an operator, when we get the modes to have destruction on the right and creation on the left, is called "time ordered".

4.3 The Dirac Field

In our previous discussion on the Klein-Gordon field, we looked at fields which permit particles (and anti-particles) which are spinless. Naturally, we also want fields that allow particles which have spin. Hence, the obvious choice is some Dirac field (given that the Dirac equation permitted particles with spin).

4.3.1 Fermions

Suppose we have some operators $b_n^\dagger$ and b_n which create and destroy spin-$\frac{1}{2}$ fermions, in the state $n = \mathbf{p}, s$ (i.e. a given value of momentum and spin). That is,

$$\begin{aligned} b_n \left|0\right\rangle &= 0, \quad \forall n, \\ b_m^\dagger \left|0\right\rangle &= \left|n\right\rangle, \\ b_n^\dagger b_{n'}^\dagger \left|0\right\rangle &= \left|n, n'\right\rangle; \end{aligned} \tag{4.31}$$

where $\left|n\right\rangle$ is a single-particle state, in $n = \mathbf{p}, s$ and $\left|n, n'\right\rangle$ is a two-particle state.

Now, by the Pauli exclusion principle (which states that no two fermions can be in the same quantum state; and that fermions are anti-symmetric), we require

$$\left|n, n'\right\rangle = -\left|n', n\right\rangle = -b_{n'}^\dagger b_n^\dagger \left|0\right\rangle.$$

Now, this and (4.31) mean that the $b_n^\dagger$ operators are inconsistent with the usual commutation relation $[b_n^\dagger, b_{n'}^\dagger] = 0$, but are consistent with the anticommutation relation,

$$\left\{b_n^\dagger, b_{n'}^\dagger\right\} = b_n^\dagger b_{n'}^\dagger + b_{n'}^\dagger b_n^\dagger = 0. \tag{4.32}$$

We also require that

$$\left|n, n\right\rangle = b_n^\dagger b_n^\dagger \left|0\right\rangle = 0,$$

so that two fermions cannot be created in the same state.

Therefore, this suggests that we should quantise fermionic fields using anticommutation relations,

$$\left\{b_n, b_m^\dagger\right\} = \delta_{nm}, \quad \left\{b_n, b_m\right\} = \left\{b_n^\dagger, b_m^\dagger\right\} = 0, \tag{4.33}$$

for the creation and destruction operators. We use these anticommutation relations instead of the commutation relations we used for bosonic fields. So, we must ask: does this work? To answer this, we shall need to quantise a theory, using anticommutation relations, and see if the answers are consistent.

4.3.2 The Complex Dirac Field $\psi(x)$

Recall that the Dirac equation reads

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

and has plane wave solutions,

$$\frac{u_s(\mathbf{p})e^{-ipx}}{\sqrt{2E_pV}}, \quad \frac{v_s(\mathbf{p})e^{ipx}}{\sqrt{2E_pV}},$$

with the spinors being orthogonal;

$$u_s^\dagger(\mathbf{p})u_{s'}(\mathbf{p}) = 2E_p\delta_{ss'} = v_s^\dagger(\mathbf{p})v_{s'}(\mathbf{p}),$$

which follows from (3.36). Hence, the general solution of the Dirac equation can be written as a sum over modes;

$$\psi(x) = \sum_{s,\mathbf{p}} \frac{1}{\sqrt{2E_pV}} \left(b_s(\mathbf{p})u_s(\mathbf{p})e^{-ipx} + d_s^\dagger(\mathbf{p})v_s(\mathbf{p})e^{ipx}\right). \tag{4.34}$$

The first term (will have) the interpretation of being a positive energy – or, alternatively, frequency – solution, with destruction operator b_s, and the second term being a negative frequency solution, with creation operator $d_s^\dagger$. To quantise this field operator properly, we must specify a set of commutation relations. Specifically, in this case, as we want to describe fermions, we specify the anticommutation relations (4.33), so that

$$\left\{b_s(\mathbf{p}), b_{s'}^\dagger(\mathbf{p}')\right\} = \delta_{ss'}\delta_{\mathbf{pp}'} = \left\{d_s(\mathbf{p}), d_{s'}^\dagger(\mathbf{p}')\right\}, \tag{4.35}$$

with all other anticommutators vanishing.

Now we have a field operator, and a set of anticommutation relations, we want to check if the Hamiltonian (after being expanded in terms of the field operators modes) satisfies the Heisenberg equation of motion,

$$i\frac{\partial\psi}{\partial t} = [\psi, H]. \tag{4.36}$$

The Hamiltonian for the Dirac equation was

$$H = -i\boldsymbol{\alpha}\cdot\nabla + \beta m.$$

Hence, a good "guess", is to integrate this over all space. Therefore, we suppose that the Hamiltonian in the quantum Dirac field theory is

$$H = \int d^3x\, \psi^\dagger \left(-i\boldsymbol{\alpha}\cdot\nabla + \beta m\right)\psi. \tag{4.37}$$

Notice that this Hamiltonian is just the expectation value of the thing we "used to" call the Hamiltonian (i.e. is the energy). So, upon substitution of the field (4.34) into this Hamiltonian, one finds that

$$H = \sum_{s,\mathbf{p}} E(\mathbf{p})\left(b_s^\dagger(\mathbf{p})b_s(\mathbf{p}) + d_s^\dagger(\mathbf{p})d_s(\mathbf{p})\right), \tag{4.38}$$

after ignoring the vacuum energy. The easiest way to show that this is the Hamiltonian, is to recall that the Dirac equation is

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

so that upon contraction with $\psi^\dagger$ from the left,

$$i\psi^\dagger\frac{\partial\psi}{\partial t} = \psi^\dagger\left(-i\boldsymbol{\alpha}\cdot\nabla + \beta m\right)\psi,$$

and integrating,

$$i\int d^3x\psi^\dagger\frac{\partial\psi}{\partial t} = \int d^3x\psi^\dagger\left(-i\boldsymbol{\alpha}\cdot\nabla + \beta m\right)\psi,$$

we have on the RHS the Hamiltonian (4.37). Thus, in order that the RHS is true, the LHS must also be true, and it is easier to compute the LHS.

We can use the easily provable identity

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

to find the commutators of the Hamiltonian with the creation and destruction operators (to infact verify their interpretation). Then, one finds

$$\begin{aligned} \left[H, b_s^\dagger(\mathbf{p})\right] &= E(\mathbf{p})b_s^\dagger(\mathbf{p}), & \left[H, b_s(\mathbf{p})\right] &= -E(\mathbf{p})b_s(\mathbf{p}), \\ \left[H, d_s^\dagger(\mathbf{p})\right] &= E(\mathbf{p})d_s^\dagger(\mathbf{p}), & \left[H, d_s(\mathbf{p})\right] &= -E(\mathbf{p})d_s(\mathbf{p}). \end{aligned} \tag{4.39}$$

Hence, these can be used to show that the Hamiltonian (4.37) with the field quantised as (4.34), satisfies the Heisenberg equation of motion (4.36). Therefore, we now see that we have the interpretation of $b_s^\dagger(\mathbf{p}), d_s^\dagger(\mathbf{p})$ creating particles of energy E and momentum $\mathbf{p}$; and $b_s(\mathbf{p}), d_s(\mathbf{p})$ destroying particles of energy E and momentum $\mathbf{p}$.

Now, recall that the conserved charge for the Dirac equation was

$$\rho = \psi^\dagger(x)\psi(x).$$

Hence, the conserved charge operator for the quantised Dirac field is

$$Q = q\int d^3x\,\psi^\dagger(x)\psi(x),$$

where we have merely inserted a multiplicative charge q. Then, if we insert our field expansion (4.34), we find that

$$Q = q\sum_{s,\mathbf{p}}\left(b_s^\dagger(\mathbf{p})b_s(\mathbf{p}) - d_s^\dagger(\mathbf{p})d_s(\mathbf{p})\right). \tag{4.40}$$

And hence, we can compute the commutator of this charge operator with the creation operators;

$$\left[Q, b_s^\dagger(\mathbf{p})\right] = qb_s^\dagger(\mathbf{p}), \quad \left[Q, d_s^\dagger(\mathbf{p})\right] = -qd_s^\dagger(\mathbf{p});$$

and obvious equivalents for the destruction operators. This allows us to see what the difference is in the particles that $b_s^\dagger$ and $d_s^\dagger$ create; which we summarise below.

Therefore, we have the interpretation that $b_s^\dagger(\mathbf{p})$ creates a spin-$\frac{1}{2}$ particle with charge $Q = +q$, energy E and momentum $\mathbf{p}$; and $b_s(\mathbf{p})$ destroys that particle. We also have that $d_s^\dagger(\mathbf{p})$ creates a spin-$\frac{1}{2}$ antiparticle with charge $Q = -q$, energy E and momentum $\mathbf{p}$; and $d_s(\mathbf{p})$ destroys that antiparticle.

It is worth noting again that the vacuum state (which is the state with lowest energy) satisfies

$$b_s(\mathbf{p}) \left|0\right\rangle = d_s(\mathbf{p}) \left|0\right\rangle = 0, \quad \forall s, \mathbf{p}.$$

And also that single particle states

$$b_s^\dagger(\mathbf{p}) \left|0\right\rangle, \quad d_s^\dagger(\mathbf{p}) \left|0\right\rangle,$$

have energies $E(\mathbf{p}) = +\sqrt{\mathbf{p}^2 + m^2}$; notice that neither the particle or anti-particle have negative energies any more, relative to the vacuum.

4.4 The Feynman Propagator $G_{\mathbf{F}}(x)$

So far we have operators which create and destroy particles, but, we still need a way of propagating from one place to another. The crucial quantity which does this is the *Feynman propagator*. We define it for the complex KG field,

$$G_{\mathrm{F}}(x) = -i \left\langle 0\right| T\left(\phi(x)\phi^\dagger(0)\right) \left|0\right\rangle, \tag{4.41}$$

where the Dyson (also called the time-ordered) product is

$$T\left(\phi(x)\phi^\dagger(0)\right) = \theta(t)\phi(x)\phi^\dagger(0) + \theta(-t)\phi^\dagger(0)\phi(x), \tag{4.42}$$

and the θ-function is defined

$$\theta(t) = \begin{cases} 1 & t > 0, \\ 0 & t < 0. \end{cases} \tag{4.43}$$

Basically, what the Dyson product does, is to ensure that particles are created before they are annihilated. $\phi(x)$ destroys particles, and $\phi^\dagger(x)$ creates particles. Hence, if $t > 0$, then G_{F} represents the creation of a particle at $t = 0$ and subsequent destruction at time t. Similarly, if $t < 0$, then an antiparticle is created at t and destroyed at time $t = 0$. That is, it ensures causality is not violated.

Now,

$$\begin{aligned}\phi(x)\phi^\dagger(0) = \sum_{p,p'} \frac{1}{2V\sqrt{E_p E_{p'}}} \quad &\big(a(p)a^\dagger(p')e^{-ipx} + a(p)c(p')e^{-ipx} \\ &+ c^\dagger(p)a^\dagger(p')e^{ipx} + c^\dagger(p)c(p')e^{ipx}\big),\end{aligned}$$

so that the only term to contribute upon contracting with the vacuum state is the first,

$$\begin{aligned}\langle 0|\,\phi(x)\phi^\dagger(0)\,|0\rangle &= \sum_{p,p'} \frac{1}{2V\sqrt{E_p E_{p'}}}\,\langle 0|\,a(p)a^\dagger(p')\,|0\rangle\, e^{-ipx} \\ &= \sum_{p,p'} \frac{1}{2V\sqrt{E_p E_{p'}}}\delta_{pp'} e^{-ipx} \\ &= \sum_{\mathbf{p}} \frac{e^{-ipx}}{2VE_{\mathbf{p}}}.\end{aligned}$$

Similarly,

$$\langle 0|\,\phi^\dagger(0)\phi(x)\,|0\rangle = \sum_{\mathbf{p}} \frac{e^{ipx}}{2VE_{\mathbf{p}}}.$$

Hence, we can use these to see that the Feynman propagator (4.41) in its two limits is

$$G_{\mathrm{F}} = -i\sum_{\mathbf{p}} \frac{e^{-ipx}}{2VE_{\mathbf{p}}}, \quad t > 0, \tag{4.44}$$

$$G_{\mathrm{F}} = -i\sum_{\mathbf{p}} \frac{e^{ipx}}{2VE_{\mathbf{p}}}, \quad t < 0; \tag{4.45}$$

where

$$px = p^\mu x_\mu = Et - \mathbf{p}\cdot\mathbf{x}.$$

We can write this using the θ-function as

$$G_{\mathrm{F}}(x) = -i\sum_{\mathbf{p}} \left(\theta(t)\frac{e^{-ipx}}{2E_{\mathbf{p}}V} + \theta(-t)\frac{e^{ipx}}{2E_{\mathbf{p}}V}\right). \tag{4.46}$$

These can be represented by diagrams, see f4.1.

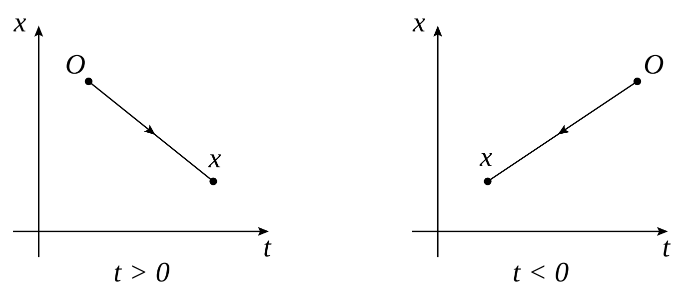


Figure 4.1: Representations of the Feynman propagator. The direction of the arrow only denotes whether the moving object is a particle or antiparticle. The arrow is sometimes emitted for bosons.

Another way of evaluating the propagator is as a single term. Let us differentiate the propagator (4.41) with respect to t, using

$$\frac{d\theta(t)}{dt} = \delta(t), \qquad \frac{d\theta(-t)}{dt} = -\delta(t).$$

So,

$$\begin{aligned} \frac{\partial G_{\rm F}(x)}{\partial t} &= -i\delta(t) \langle 0| \phi(x)\phi^\dagger(0) - \phi^\dagger(0)\phi(x) |0\rangle \\ &\quad -i \langle 0| \theta(t)\frac{\partial \phi(x)}{\partial t}\phi^\dagger(0) + \theta(-t)\phi^\dagger(0)\frac{\partial\phi(x)}{\partial t} |0\rangle . \end{aligned}$$

Now, the first term is zero as

$$\left[\phi(t,\mathbf{x}), \phi^\dagger(t,\mathbf{x}')\right] = 0. \tag{4.47}$$

Hence,

$$\frac{\partial G_{\rm F}(x)}{\partial t} = -i \langle 0| \theta(t)\frac{\partial \phi(x)}{\partial t}\phi^\dagger(0) + \theta(-t)\phi^\dagger(0)\frac{\partial\phi(x)}{\partial t} |0\rangle .$$

Let us differentiate again,

$$\begin{aligned} \frac{\partial^2 G_{\rm F}}{\partial t^2} &= -i \langle 0| \delta(t)\frac{\partial \phi(x)}{\partial t}\phi^\dagger(0) |0\rangle - i \langle 0| \theta(t)\frac{\partial^2 \phi(x)}{\partial t^2}\phi^\dagger(0) |0\rangle \\ &\quad +i \langle 0| \delta(t)\phi^\dagger(0)\frac{\partial \phi(x)}{\partial t} |0\rangle - i \langle 0| \theta(-t)\phi^\dagger(0)\frac{\partial^2 \phi(x)}{\partial t^2} |0\rangle \\ &= -i\delta(t) \langle 0| \left[\frac{\partial\phi(x)}{\partial t}, \phi^\dagger(0)\right] |0\rangle - i \langle 0| \theta(t)\frac{\partial^2 \phi(x)}{\partial t^2}\phi^\dagger(0) |0\rangle \\ &\quad -i \langle 0| \theta(-t)\phi^\dagger(0)\frac{\partial^2 \phi(x)}{\partial t^2} |0\rangle , \end{aligned}$$

where we have noticed the appearance of the commutator. Now, recall the KG equation,

$$\frac{\partial^2\phi}{\partial t^2} = (\nabla^2 - m^2)\phi,$$

which we can use in the last two terms. So, explicitly,

$$\begin{aligned} -i \langle 0| \theta(t)\frac{\partial^2 \phi(x)}{\partial t^2}\phi^\dagger(0) |0\rangle &= -i \langle 0| \theta(t)(\nabla^2 - m^2)\phi(x)\phi^+(0) |0\rangle \\ &= -i(\nabla^2 - m^2) \langle 0| \theta(t)\phi(x)\phi^+(0) |0\rangle , \\ -i \langle 0| \theta(-t)\phi^\dagger(0)\frac{\partial^2 \phi(x)}{\partial t^2} |0\rangle &= -i(\nabla^2 - m^2) \langle 0| \theta(-t)\phi^+(0)\phi(x) |0\rangle ; \end{aligned}$$

so that their sum is just the Feynman propagator (4.41). Hence,

$$\frac{\partial^2 G_{\rm F}}{\partial t^2} = -i\delta(t) \langle 0| \left[\frac{\partial\phi(x)}{\partial t}, \phi^\dagger(0)\right] |0\rangle + (\nabla^2 - m^2)G_{\rm F}.$$

Therefore, taking the far right term over to the LHS, we see that we have

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2 + m^2\right) G_{\mathrm{F}} = -i\delta(t) \langle 0| \left[\frac{\partial \phi(x)}{\partial t}, \phi^\dagger(0)\right] |0\rangle \,,$$

which is just

$$\left(\partial^\mu \partial_\mu + m^2\right) G_{\mathrm{F}} = -i\delta(t) \langle 0| \left[\frac{\partial \phi(x)}{\partial t}, \phi^\dagger(0)\right] |0\rangle \,. \tag{4.48}$$

We can show, by simply expanding in terms of modes, that

$$\left[\frac{\partial \phi(t, \mathbf{x})}{\partial t}, \phi^\dagger(t, \mathbf{x}')\right] = -i\delta^{(3)}(\mathbf{x} - \mathbf{x}'). \tag{4.49}$$

Hence, upon substitution of this into (4.48), we see that

$$\left(\partial^\mu \partial_\mu + m^2\right) G_{\mathrm{F}}(x) = -\delta^{(4)}(x). \tag{4.50}$$

This is a point-source Klein-Gordon equation, and its solution is the Feynman propagator. This can also be derived by differentiating (4.46). One should recall that such solutions are Green functions, and that Green functions can be integrated over a charge distribution to give the solution for the inhomogeneous Klein-Gordon equation. That is, we can use the Green function G_{F} (which is the solution to the above point-source equation) to solve

$$\left(\partial^\mu \partial_\mu + m^2\right) \phi(x) = -\rho(x),$$

where $\rho(x)$ is an arbitrary source.

To find the Green function $G_{\mathrm{F}}(x)$, we solve (4.50) by Fourier transforms. So, we have

$$\begin{aligned} G_{\mathrm{F}}(x) &= \int \frac{d^4k}{(2\pi)^4} e^{-ikx} \tilde{G}_{\mathrm{F}}(k) \qquad (4.51) \\ \delta^{(4)}(x) &= \int \frac{d^4k}{(2\pi)^4} e^{-ikx}, \end{aligned}$$

noting that a δ-function can be written as the Fourier tranform of unity. Now,

$$\partial^\mu \partial_\mu e^{-ikx} = -k^2 e^{-ikx},$$

so that upon substitution of these transforms into (4.50), we have

$$\int \frac{d^4k}{(2\pi)^4} \left\{-k^2 + m^2\right\} e^{-ikx} \tilde{G}_{\mathrm{F}}(k) = -\int \frac{d^4k}{(2\pi)^4} e^{-ikx}.$$

Thus, as the integrals are identical, we equate the integrands;

$$\tilde{G}_{\mathrm{F}}(k) \left(-k^2 + m^2\right) = -1,$$

and hence,

$$\tilde{G}_{\rm F}(k) = \frac{1}{k^2 - m^2}, \quad k^2 \equiv k^\mu k_\mu.$$

We then substitute this back into (4.51),

$$G_{\rm F}(x) = \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ikx}}{k^2 - m^2}. \tag{4.52}$$

This then gives us an expression for the Green function; once we have performed the integral, we have an expression which allows us to solve any inhomogeneous KG equation, by integrating over it. However, this expression has a singularity at $k^2 = m^2$;

$$\text{singularity}: \quad k^0 = \pm\omega_k \equiv \pm\sqrt{\mathbf{k}^2 + m^2}.$$

This pole is on the real axis; to get around this problem, we use the *Feynman prescription*, which is to displace the pole from the real axis, do the integral, then put everything back as it was. That is, we modify the Green function (4.52) into

$$G_{\rm F}(x) = \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ikx}}{k^2 - m^2 + i\epsilon}, \quad \epsilon > 0, \tag{4.53}$$

perform the integral, and at the end of any calculations, set $\epsilon = 0$. We evaluate the time-component,

$$I = \int_{-\infty}^{\infty} dk^0 \frac{e^{-ik^0 t}}{(k^0)^2 - \omega_k^2 + i\epsilon}, \quad \omega_k = \sqrt{\mathbf{k}^2 + m^2}, \tag{4.54}$$

by contour integration, with poles at

$$k^0 = \pm\left(\omega_k - i\epsilon\right).$$

With reference to Figure (4.2), we see the complex k^0 plane, with poles, and the contour we shall integrate over. We separate the cases $t < 0$ and $t > 0$.

$t > 0$**: Lower-half-plane** In this case, we close the contour in the lower half plane, and let the radius of the semi-circle go to infinity. There is no contribution from the semi-circle, as $k^0 < 0$. We do the contour integral by noting that Cauchy's theorem states that the integral of any analytic function, over an (anti-clockwise) closed path, is just $2\pi i$ multiplied by the sum of the residues enclosed by the path. Thus, for our contour which is closed in the lower-half plane, we only have a pole at $k^0 = \omega_k - i\epsilon$. Therefore, the value of the integral (4.54) is

$$I = -2\pi i \frac{e^{-i\omega_k t}}{2\omega_k}.$$

Notice that the minus-sign is due to our contour being clockwise, and that we have set $\epsilon = 0$.

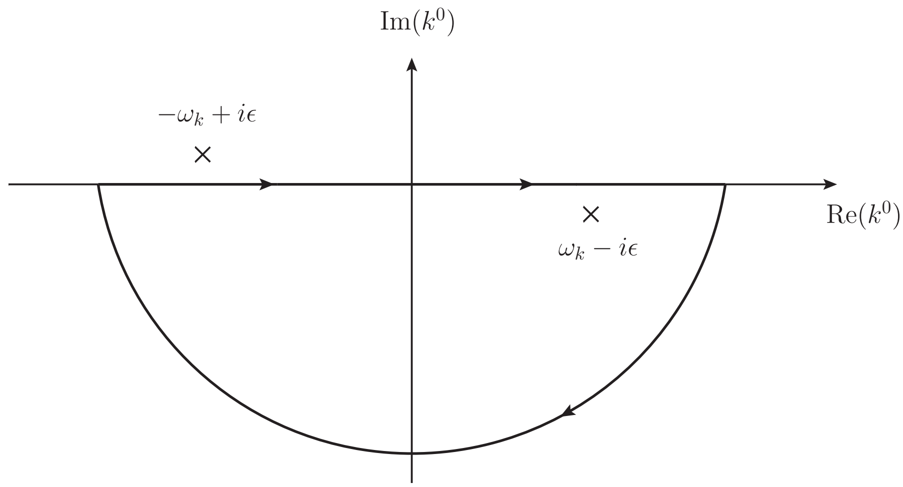


Figure 4.2: The complex plane, in k^0-space; with the poles marked on. We evaluate (4.54) over the contour shown. This is the $t > 0$ case.

$t < 0$**: Upper-half-plane** In this case, we close the contour in the upper-half-plane, so that the only contributing pole is at $k^0 = -\omega_k + i\epsilon$.

Hence, using these results, we can write down (4.53), in the two regimes,

$$G_{\mathrm{F}}(x) = -\frac{i}{(2\pi)^3}\int d^3p \frac{e^{-i(Et-\mathbf{p}\cdot\mathbf{x})}}{2E_p}, \quad t > 0, \tag{4.55}$$

$$G_{\mathrm{F}}(x) = -\frac{i}{(2\pi)^3}\int d^3p \frac{e^{i(Et-\mathbf{p}\cdot\mathbf{x})}}{2E_p}, \quad t < 0. \tag{4.56}$$

Notice that these are the same as (4.44), (4.45); only now we have moved to continuous momentum. Thus, this is the same as the causal result.

We shall leave propagators for the moment, and discuss interactions. We shall come back to them when we need a term which allows the creation and destruction of a particle within a process (i.e. when we discuss exchange bosons).

4.5 Interactions

4.5.1 The Interaction Picture

So far, we have two ways of describing the time evolution of a quantum system: the Schrodinger and Heisenberg pictures. In the Schrodinger picture, only the states evolve and not the operators, where the states obey

$$i\frac{\partial}{\partial t}|\psi(t)\rangle_{\mathrm{S}} = H|\psi(t)\rangle_{\mathrm{S}}.$$

The formal solution to this is just

$$|\psi(t)\rangle_{\rm S} = e^{-iHt}|\psi(0)\rangle_{\rm S}.$$

Hence, we can use this to find the expectation value of an operator,

$${}_{\rm S}\langle\psi(t)|\, A_{\rm S}|\psi(t)\rangle_{\rm S} = {}_{\rm S}\langle\psi(0)|\, e^{iHt}A_{\rm S}e^{-iHt}|\psi(0)\rangle_{\rm S}.$$

We can take this, and redefine a few terms,

$$|\psi\rangle_{\rm H} \equiv |\psi(0)\rangle_{\rm S}, \quad A_{\rm H}(t) \equiv e^{iHt}A_{\rm S}e^{-iHt}, \tag{4.57}$$

so that the expected value becomes just

$${}_{\rm S}\langle\psi(t)|\, A_{\rm S}|\psi(t)\rangle_{\rm S} = {}_{\rm H}\langle\psi|\, A_{\rm H}(t)|\psi\rangle_{\rm H}.$$

Hence, we see that by writing (4.57), we have preserved the form of the expectation value of an operator. What we have done, is to change the object that is evolving in time: from the state to the operator. In the Heisenberg picture, the state is constant, and all time evolution is carried out by the operators. Simply by differentiating the second of (4.57), we can derive the Heisenberg equation of motion,

$$-\, i\frac{dA_{\rm H}}{dt} = \left[H, A_{\rm H}(t)\right]. \tag{4.58}$$

For perturbation theory, we use a third picture, the interaction picture. To define the picture, we write the Hamiltonian as a sum of a free Hamiltonian and an interaction Hamiltonian,

$$H = H_0 + H_{\rm I}. \tag{4.59}$$

We also define the interaction picture states and operators as

$$|\psi(t)\rangle_{\rm I} = e^{iH_0t}|\psi(t)\rangle_{\rm S}, \quad A_{\rm I}(t) = e^{iH_0t}A_{\rm S}e^{-iH_0t}. \tag{4.60}$$

Thus, differentiating, we get

$$i\frac{d}{dt}|\psi(t)\rangle_{\rm I} = H_{\rm I}(t)|\psi(t)\rangle_{\rm I}, \quad -i\frac{dA_{\rm I}}{dt} = \left[H_0, A_{\rm I}(t)\right], \tag{4.61}$$

where

$$H_{\rm I}(t) = e^{-iH_0t}H_{\rm I,S}e^{iH_0t}. \tag{4.62}$$

If we have that $H = H_0$, then the Heisenberg and interaction pictures are identical. If $H = H_0 + H_{\rm I}$, then we have an intermediate picture in which operators (which includes field operators) evolve according to the free Hamiltonian H_0. Hence, our free-field results in the Heisenberg picture hold for interacting fields, in the interaction picture.

4.5.2 The S-matrix

For time evolution, it is useful to write

$$|\psi(t)\rangle_{\rm I} = U(t,t_0)\,|\psi(t_0)\rangle\,,$$

where

$$i\frac{d}{dt}U(t,t_0) = H_{\rm I}(t)U(t,t_0), \tag{4.63}$$

under the boundary condition $U(t_0,t_0)=1$. We define the S-matrix via

$$S = \lim_{\tau\to\infty} U(\tfrac{\tau}{2},-\tfrac{\tau}{2}) = U(\infty,-\infty). \tag{4.64}$$

Hence, we start long before particles get close enough to interact, and finish longer after particles have finished interacting. The initial state (i.e. at $t=-\infty$) is just

$$|\psi(-\infty)\rangle_{\rm I} = |\psi_i\rangle_{\rm I},$$

which is a free-particle state (i.e. an eigenfunction of the free Hamiltonian H_0). At finite t, the state looks like

$$|\psi(t)\rangle_{\rm I} = U(t,-\infty)|\psi_i\rangle_{\rm I},$$

and is a complicated state of interacting particles. The final state (i.e. at $t=\infty$), is just

$$|\psi(\infty)\rangle_{\rm I} = S|\psi_{t=-\infty}\rangle_{\rm I},$$

after using the definition of the S-matrix, (4.64). Thus, we can write the final state as a linear sum of free particle states,

$$|\psi(\infty)\rangle_{\rm I} = |\psi_f\rangle_{\rm I} = \sum_i \mathcal{S}_{\rm fi}|\psi_i\rangle_{\rm I}.$$

We can write the final state like this, as we have stated that the S-matrix ends a long time after the particles have finished interacting. Hence, we can easily see that the amplitude to end up in the final state f is

$$\mathcal{S}_{\rm fi} = \langle\psi_f|\,S\,|\psi_i\rangle\,, \tag{4.65}$$

which occurs with probability

$$P_{fi} = |\mathcal{S}_{\rm fi}|^2\,. \tag{4.66}$$

These matrix elements are the only observables of the system, and summarise the physics of a set of interactions.

One can think about the S-matrix as a "mixing matrix", which describes the transitions from the initial state that took place to give the final state. Those familiar with statistical physics can think about this matrix as being like the Markov chain transition matrix. The matrix elements $\mathcal{S}_{\rm fi}$ describe the amplitude for a particular initial and final state.

The Unitarity of the S-matrix Recall that

$$|\psi(t)\rangle_{\mathrm{I}} = e^{iH_0 t}|\psi(t)\rangle_{\mathrm{S}}$$

from the first of (4.60). Hence, we can write

$$\begin{aligned}
|\psi(t)\rangle_{\mathrm{I}} &= e^{iH_0 t}|\psi(t)\rangle_{\mathrm{S}} \\
&= e^{iH_0 t}e^{-iHt}|\psi(t_0)\rangle_{\mathrm{S}} \\
&= e^{-H_0 t}e^{-iHt}e^{iHt_0}|\psi(t_0)\rangle_{\mathrm{I}} \\
&= e^{iH_0 t}e^{-iH(t-t_0)}|\psi(t_0)\rangle_{\mathrm{I}} \\
&= U(t,t_0)|\psi(t_0)\rangle_{\mathrm{I}}.
\end{aligned}$$

Now, as H and H_0 are both Hermitian, we thus have that U is unitary;

$$U^{-1}(t,t_0) = U(t_0,t) = U^{\dagger}.$$

Hence, taking the limit $t \to \infty$ and $t_0 \to -\infty$, we arrive at the statement that the S-matrix is unitary,

$$S^{\dagger}S = 1. \tag{4.67}$$

The consequence of this, is that probabilities are conserved. For example,

$$\begin{aligned}
\sum_f P_{fi} &= \sum_f |\langle\psi_f|\, S\, |\psi_i\rangle|^2 \\
&= \sum_f \langle\psi_i|\, S^{\dagger}\, |\psi_f\rangle \, \langle\psi_f|\, S\, |\psi_i\rangle \\
&= \langle\psi_i|\, S^{\dagger}S\, |\psi_i\rangle \\
&= \langle\psi_i|\psi_i\rangle.
\end{aligned}$$

The third equality follows from the second due to the completeness of eigenfunctions. Hence, we see that if a state is normalised, then the normalisation is preserved.

4.5.3 The S-matrix Expansion

Now, from (4.63), we have the formal solution

$$U(t,t_0) = 1 - i\int_{t_0}^{t} dt_1\, H_{\mathrm{I}}(t_1)U(t_1,t_0), \tag{4.68}$$

under the boundary condition $U(t_0,t_0) = 1$. The reason we say "formal", is that the thing we are trying to solve for, $U(t,t_0)$, appears within the integral itself. Now, we can form an iterative solution, by inserting the formal solution into itself,

$$U(t,t_0) = 1 - i\int_{t_0}^{t} dt_1 H_{\mathrm{I}}(t_1) + (-i)^2\int_{t_0}^{t} dt_1 H_{\mathrm{I}}(t_1)\int_{t_0}^{t_1} dt_2 H_{\mathrm{I}}(t_2)U(t_2,t_0).$$

We can imagine doing this iterative insertion many times,

$$U(t,t_0) = 1 - i\int_{t_0}^{t} dt_1 H_{\rm I}(t_1) + (-i)^2 \int_{t_0}^{t} dt_1 H_{\rm I}(t_1)\int_{t_0}^{t_1} dt_2 H_{\rm I}(t_2) + \ldots .$$

So, if we take $t \to \infty$, and $t_0 \to -\infty$, this becomes the expansion of the S-matrix, $U(\infty,-\infty) = S$, by definition. Hence, we write this as a sum of terms of incrementing orders,

$$S = S^{(0)} + S^{(1)} + S^{(2)} + \ldots ,$$

where $S^{(0)} = 1$ denotes no interaction, and for example,

$$S^{(1)} = -i\int_{-\infty}^{\infty} dt_1 H_{\rm I}(t_1), \tag{4.69}$$

$$S^{(2)} = (-i)^2\int_{-\infty}^{\infty} dt_1 H_{\rm I}(t_1)\int_{-\infty}^{t_1} dt_2 H_{\rm I}(t_2). \tag{4.70}$$

Now, as the expression stands, it is not covariant: we have singled out the time coordinate. Let us rewrite $S^{(2)}$ in a more symmetric form by interchanging t_1 and t_2,

$$S^{(2)} = \frac{(-i)^2}{2}\int_{-\infty}^{\infty} dt_1 H_{\rm I}(t_1)\int_{-\infty}^{t_1} dt_2 H_{\rm I}(t_2) + \frac{(-i)^2}{2}\int_{-\infty}^{\infty} dt_2 H_{\rm I}(t_2)\int_{-\infty}^{t_2} dt_1 H_{\rm I}(t_1).$$

We can notice that in the first term, $t_2 < t_1$ and in the second term, $t_1 < t_2$. We can then notice that in the both terms, the "earlier" Hamiltonian appears to the right of the "later" Hamiltonian;

$$\begin{aligned} t_1 > t_2 \quad &\Rightarrow \quad H_{\rm I}(t_1)H_{\rm I}(t_2), \\ t_2 > t_1 \quad &\Rightarrow \quad H_{\rm I}(t_2)H_{\rm I}(t_1). \end{aligned}$$

Hence, we have an expression which is of the time-ordered form (or Dyson product),

$$S^{(2)} = \frac{(-i)^2}{2}\int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2\, T\left\{H_{\rm I}(t_1)H_{\rm I}(t_2)\right\}. \tag{4.71}$$

We have set the upper-limit on the second integral to ∞ as we notice that the whole $t_1 - t_2$-plane will be covered by such a Dyson product. More generally, for the $n^{\rm th}$-order term,

$$S^{(n)} = \frac{(-i)^n}{n!}\int_{-\infty}^{\infty} dt_1 \ldots \int_{-\infty}^{\infty} dt_n\, T\left\{H_{\rm I}(t_1)\ldots H_{\rm I}(t_n)\right\}. \tag{4.72}$$

This now does not single out a specific time value, as the case was before we wrote the symmetric sum, but does single out time from space; hence, the expression still is not explicitly covariant. We now note that in field theory the Hamiltonian H is the spatial integral of the Hamiltonian density $\mathcal{H}$,

$$H_{\rm I}(t) = \int d^3x\,\mathcal{H}_{\rm I},$$

where we may have multiple fields which are functions of the 4-position vector x,

$$\mathcal{H}_{\mathrm{I}}(x) = \mathcal{H}_i\left(\phi_i, \partial_\mu \phi_i\right).$$

Finally, using the Hamiltonian density, rather than the Hamiltonian, the expressions (4.69) and (4.70) can be written in a covariant form,

$$S^{(1)} = (-i)\int d^4x\,\mathcal{H}_{\mathrm{I}}(x), \quad (4.73)$$

$$S^{(2)} = \frac{(-i)^2}{2}\int d^4x_1\int d^4x_2\,T\left\{\mathcal{H}_{\mathrm{I}}(x_1)\mathcal{H}_{\mathrm{I}}(x_2)\right\}, \quad (4.74)$$

$$S^{(n)} = \frac{(-i)^n}{n!}\int d^4x_1\ldots\int d^4x_n\,T\left\{\mathcal{H}_{\mathrm{I}}(x_1)\ldots\mathcal{H}_{\mathrm{I}}(x_n)\right\}. \quad (4.75)$$

Hence, we have an explicitly covariant form of the S-matrix expansion.

5 Decays

We shall illustrate calculations by considering examples.

5.1 Simple 2-body Decay: ϕ^3

Let us consider an example of a Higgs boson decaying to two tauons,

$$H^0 \longrightarrow \tau^+ + \tau^-. \tag{5.1}$$

The scalar uncharged Higgs boson H^0 has zero spin; the tauons $\tau^\pm$ are spin-$\frac{1}{2}$ leptons (i.e. fermions), with mass $m_\tau \approx 1.8\mathrm{GeV}/c^2$. In terms of momentum and spin, $(\mathbf{p}, s)$, the decay can be written

$$(\mathbf{p}, 0) \longrightarrow (\mathbf{k}', s') + (\mathbf{k}, s).$$

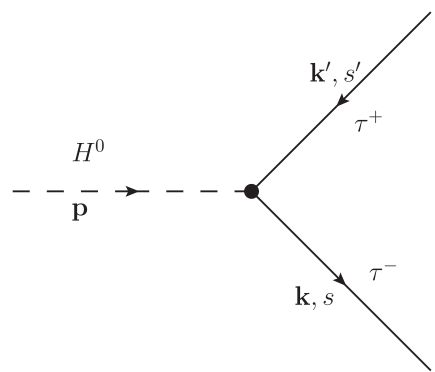


Figure 5.1: Schematic of the 2-body decay (5.1).

Now, in the standard model, the interaction is of the form

$$\mathcal{H}_{\mathrm{I}}(x) = gN\left(\bar{\psi}(x)\psi(x)\phi(x)\right), \tag{5.2}$$

where N denotes normal (or time) ordering. The choice of interaction term is effectively "the problem"; different interaction terms produce different physics. Hence, a given interaction term is "the theory" we are investigating. Thus, above, we are considering a ϕ^3-theory. Notice that $\bar{\psi}(x)\psi(x)$ is a Lorentz scalar density. As H^0 is an uncharged boson, we have that $\phi(x)$ is a Hermitian scalar field; as $\tau^\pm$ are charged particles with spin, we have that $\psi(x)$ is a Dirac spinor field. The dimensionless coupling parameter g is

$$g = m_\tau \left(\frac{\mathcal{G}}{\sqrt{2}}\right)^{1/2}, \quad \mathcal{G} \equiv 1.16 \times 10^{-5}\mathrm{GeV}^{-2}; \tag{5.3}$$

where $\mathcal{G}$ is Fermi's coupling constant. To first order, we need to evaluate the matrix element

$$\mathcal{S}_{\rm fi}^{(1)} = -i\int d^4x \left\langle f = \tau^+\tau^- \right| \mathcal{H}_{\rm I}(x) \left| i = H^0 \right\rangle ,$$

which denotes the amplitude for a decay of $H^0 \to \tau^+ + \tau^-$, with given momenta. Notice that other matrix elements are possible, such as $\langle f = \tau^+\tau^- H^0 | \, \mathcal{H}_{\rm I}(x) \, | i = 0\rangle$, which would denote particle creation. We are interested therefore in

$$\text{2 body decay :} \quad \langle \mathbf{k}, s; \mathbf{k}', s' | \, \mathcal{H}_{\rm I} \, | \mathbf{p} \rangle .$$

The neutral scalar field has its expansion being

$$\begin{aligned} \phi(x) &= \phi^{(+)}(x) + \phi^{(-)}(x) \\ &= \sum_{\mathbf{p}} \frac{1}{(2V\omega_{\mathbf{p}})^{1/2}} \left(a(\mathbf{p})e^{-ipx} + a^\dagger(\mathbf{p})e^{ipx} \right) , \end{aligned}$$

so that $\phi^{(+)}$ destroys and $\phi^{(-)}$ creates netural scalar bosons; where

$$\omega_{\mathbf{p}} = \omega(\mathbf{p}) = \sqrt{\mathbf{p}^2 + m^2}.$$

The charged fermionic field expansion is

$$\begin{aligned} \psi(x) &= \psi^{(+)}(x) + \psi^{(-)}(x) \\ &= \sum_{\mathbf{p},s} \frac{1}{(2VE_{\mathbf{p}})^{1/2}} \left(b_s(\mathbf{p})u_s(\mathbf{p})e^{-ipx} + d_s^\dagger(\mathbf{p})v_s(\mathbf{p})e^{ipx} \right) , \end{aligned}$$

so that $\psi^{(+)}$ destroys τ^- and $\psi^{(-)}$ creates τ^+. As the field is a Dirac field, we also note that $\bar{\psi}^{(-)}$ creates τ^-, and $\bar{\psi}^{(+)}$ destroys τ^+, where an over-bar denotes the Dirac adjoint, so that

$$\bar{\psi}(x) = \bar{\psi}^{(-)}(x) + \bar{\psi}^{(+)}(x).$$

So, our initial state is a single Higgs boson, with momentum $\mathbf{p}$, which we can create from the vacuum,

$$|i\rangle = \left| H^0, \mathbf{p} \right\rangle = a^\dagger(\mathbf{p}) \, |0\rangle .$$

The final state similarly is

$$|f\rangle = \left| \tau^-, \mathbf{k}, s; \tau^+, \mathbf{k}'s' \right\rangle = b_s^\dagger(\mathbf{k}) d_{s'}^\dagger(\mathbf{k}') \, |0\rangle .$$

So, to first order in $\mathcal{H}_{\rm I}$, the matrix element is

$$\mathcal{S}_{\rm fi} = -ig \int d^4x \, \langle f | \, N\left(\bar{\psi}(x)\psi(x)\phi(x) \right) | i \rangle .$$

Now, if we were to plug in the expressions for the field expansions, and expand out, we would have 8 terms. In fact, each of these terms correspond to a different process, but will only give non-zero contribution if they correspond to the process we are currently dealing with (due to the orthonormality of states, and our specific choice of initial and final states). One can see that the only contributing term is

$$\mathcal{S}_{\text{fi}} = -ig \int d^4x \, \langle f| \, \bar{\psi}^{(-)}(x)\psi^{(-)}(x)\phi^{(+)}(x) \, |i\rangle \,, \tag{5.4}$$

where the first term creates a τ^-, the second creates a τ^+ and the third destroys a H^0.

So, let us treat each field in turn. The term on the far RHS is

$$\begin{aligned} \phi^{(+)}(x) \, |i\rangle &= \phi^{(+)}(x) \, |\mathbf{p}\rangle \\ &= \phi^{(+)}(x) a^\dagger(\mathbf{p}) \, |0\rangle \,, \end{aligned}$$

but,

$$\phi^{(+)}(x) = \sum_{\mathbf{p}} \frac{1}{(2V\omega_{\mathbf{p}})^{1/2}} a(\mathbf{p}) e^{-ipx},$$

and hence

$$\phi^{(+)}(x) \, |i\rangle = \sum_{\mathbf{p}'} \frac{1}{(2V\omega_{\mathbf{p}'})^{1/2}} a(\mathbf{p}') e^{-ip'x} a^\dagger(\mathbf{p}) \, |0\rangle \,.$$

Now, we can write $a(\mathbf{p}')a^\dagger(\mathbf{p}) = a^\dagger(\mathbf{p})a(\mathbf{p}') + \delta_{\mathbf{pp}'}$, by the usual commutation relation. Hence, doing so,

$$\phi^{(+)}(x) \, |i\rangle = \sum_{\mathbf{p}'} \frac{1}{(2V\omega_{\mathbf{p}'})^{1/2}} e^{-ip'x} \left(a^\dagger(\mathbf{p})a(\mathbf{p}') + \delta_{\mathbf{pp}'} \right) |0\rangle \,.$$

The first bracketed term gives zero upon action with the vacuum state, as $a(\mathbf{p}) \, |0\rangle = 0$. The Kronecker-delta term simply filters out $\mathbf{p}' = \mathbf{p}$ in the summation to leave

$$\phi^{(+)}(x) \, |i\rangle = \frac{1}{(2V\omega_{\mathbf{p}})^{1/2}} e^{-ipx} \, |0\rangle \,. \tag{5.5}$$

In an entirely analogous manner, one can easily see that

$$\psi^{(+)}(x) \, |\mathbf{k}, s\rangle = \frac{1}{(2VE_{\mathbf{k}})^{1/2}} u_s(\mathbf{k}) e^{-ikx} \, |0\rangle \,.$$

Now, it is the adjoint of this term which appears in (5.4); hence, taking the adjoint,

$$\langle \mathbf{k}, s| \, \bar{\psi}^{(-)}(x) = \frac{1}{(2VE_{\mathbf{k}})^{1/2}} \bar{u}_s(\mathbf{k}) e^{ikx} \, \langle 0| \,. \tag{5.6}$$

And the final term can be seen by analogy again,

$$\langle \mathbf{k}', s'| \, \psi^{(-)}(x) = \frac{1}{(2VE'_{\mathbf{k}'})^{1/2}} v_{s'}(\mathbf{k}') e^{ik'x} \, \langle 0| \,. \tag{5.7}$$

Hence, putting together (5.5), (5.6) and (5.7), we see that

$$\langle f|\,\bar{\psi}^{(-)}\psi^{(-)}\phi^{(+)}\,|i\rangle = \frac{1}{(2VE_{\mathbf{k}})^{1/2}}\frac{1}{(2VE'_{\mathbf{k}'})^{1/2}}\frac{1}{(2V\omega_{\mathbf{p}})^{1/2}} \times e^{ix(k+k'-p)}\bar{u}_s(\mathbf{k})v_{s'}(\mathbf{k}'), \tag{5.8}$$

where we have been careful to maintain the order of the spinors, and noting that $\langle 0|0\rangle = 1$. Now, in (5.4), we have the integral over space of this quantity: this will only have the effect on the exponential,

$$\int d^4x\, e^{ix(k+k'-p)} = (2\pi)^4\delta^{(4)}(k+k'-p). \tag{5.9}$$

If we are not in the limit of an infinite volume/time box, we will have that

$$\int d^4x\, e^{ix(k+k'-p)} = VT\delta_{k+k',p}.$$

So, putting together (5.8) and (5.9) in (5.4), we see that

$$\mathcal{S}_{\text{fi}} = \frac{1}{(2VE_{\mathbf{k}})^{1/2}}\frac{1}{(2VE'_{\mathbf{k}'})^{1/2}}\frac{1}{(2V\omega_{\mathbf{p}})^{1/2}}(2\pi)^4\delta^{(4)}(k+k'-p)\mathcal{M}_{\text{fi}}, \tag{5.10}$$

where we have defined the Feynman amplitude

$$\mathcal{M}_{\text{fi}} = -ig\bar{u}_s(\mathbf{k})v_{s'}(\mathbf{k}'). \tag{5.11}$$

Let us now consider the three types of term in (5.10); its structure is

$$\mathcal{S}_{\text{fi}} = \text{normalisation} \times \text{energy-momentum conservation} \times \text{dynamics}.$$

The initial three factors involving energy are the normalisation, and will be common to all processes. The second term with the delta-function is infact the statement of conservation of energy-momentum. The final term $\mathcal{M}_{\text{fi}}$ is the only term that depends on the specific interaction, and is where all dynamical information is kept. We can represent $\mathcal{M}_{\text{fi}}$ by a Feynman graph, as in Figure (5.1). See Figure (5.2) for the assignment of mathematical factors to specific features of a Feynman graph. See Figure (5.3) to see the whole amplitude for a given graph.

5.2 Decay Rates & Lifetimes

We now want to calculate the measured decay rate for a given process. The probability of decay per unit time is

$$\Gamma = \sum_f \Gamma_{fi} = \sum_f \frac{|\mathcal{S}_{\text{fi}}|^2}{T}.$$

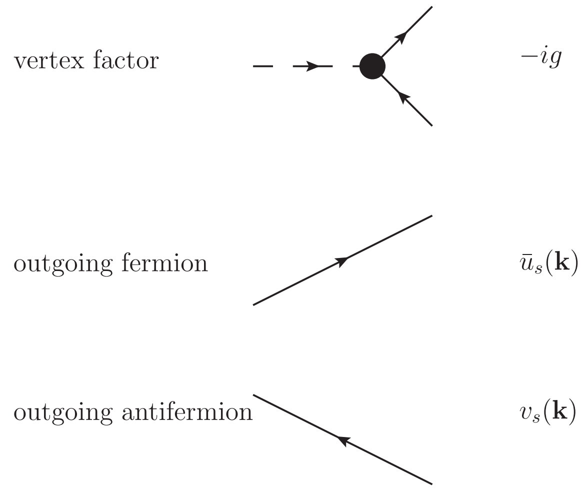


Figure 5.2: The assignment of factors to the elements of a Feynman graph.

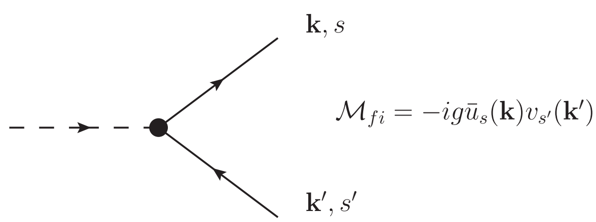


Figure 5.3: The Feynman amplitude for the given graph.

Hence, inserting (5.10), we see that

$$\Gamma = \sum_{\mathbf{k},s,\mathbf{k}',s'} \frac{1}{T} \frac{1}{8V^3 \omega_{\mathbf{p}} E_{\mathbf{k}} E'_{\mathbf{k}'}} T^2 V^2 \delta_{k+k',p} \left|\mathcal{M}_{\text{fi}}\right|^2 ,$$

where we are in the finite volume/time limit, to allow us to have a meaning in squaring the delta-function. Now we have squared the delta-function, we can take the limit $V, T \to \infty$, so that the energy-momentum conservation term becomes just

$$VT\delta_{k+k',p} \longmapsto (2\pi)^4 \delta^{(4)}(k + k' - p),$$

and the two phase-space terms

$$\frac{1}{V} \sum_{\mathbf{k}} \longmapsto \int \frac{d^3k}{(2\pi)^3}.$$

Thus, all factors that related to the volume and time of the box disappear, leaving us with

$$\Gamma = \frac{1}{2\omega_{\mathbf{p}}} \sum_{s,s'} \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} \int \frac{d^3k'}{(2\pi)^3 2E'_{\mathbf{k}'}} (2\pi)^4 \delta^{(4)}(k + k' - p) \left|\mathcal{M}_{\rm fi}\right|^2 . \tag{5.12}$$

To continue, we need to evaluate the spin sum and do the phase space integrals. Let us do the spin sum first.

We must evaluate

$$\sum_{s,s'} \left|\mathcal{M}_{\rm fi}\right|^2 = g^2 \sum_{s,s'} \left|\bar{u}_s(\mathbf{k}) v_{s'}(\mathbf{k}')\right|^2 .$$

So, to begin,

$$\begin{aligned} \sum_{s,s'} \left|\mathcal{M}_{\rm fi}\right|^2 &= \sum \mathcal{M}_{\rm fi} \mathcal{M}_{\rm fi}^* \\ &= g^2 \sum \bar{u}(\mathbf{k}) v(\mathbf{k}') \left(\bar{u}(\mathbf{k}) v(\mathbf{k}')\right)^* , \end{aligned}$$

where we drop the s, s', as they are implied by the argument of the spinor. Now, notice that the bracketed term is a scalar (indeed, as is the term directly before it). Therefore, we can treat complex conjugation exactly as Hermitian conjugation,

$$\left(\bar{u}(\mathbf{k}) v(\mathbf{k}')\right)^* = \left(\bar{u}(\mathbf{k}) v(\mathbf{k}')\right)^\dagger .$$

If we recall that the Dirac adjoint was defined to be $\bar{A} = A^\dagger \gamma^0$, we thus have

$$\left(\bar{u}(\mathbf{k}) v(\mathbf{k}')\right)^\dagger = \left(u^\dagger(\mathbf{k}) \gamma^0 v(\mathbf{k}')\right)^\dagger .$$

Hence, taking the Hermitian conjugate (which has the effect of conjugating everything, and swapping the order of all terms),

$$\left(u^\dagger(\mathbf{k}) \gamma^0 v(\mathbf{k}')\right)^\dagger = v^\dagger(\mathbf{k}') \gamma^{0\dagger} u(\mathbf{k}),$$

but, $\gamma^{0\dagger} = \gamma^0$, so that we see the Dirac adjoint appear again,

$$\left(u^\dagger(\mathbf{k}) \gamma^0 v(\mathbf{k}')\right)^\dagger = \bar{v}(\mathbf{k}') u(\mathbf{k}).$$

Therefore,

$$\sum_{s,s'} \left|\mathcal{M}_{\rm fi}\right|^2 = g^2 \sum \bar{u}(\mathbf{k}) v(\mathbf{k}') \bar{v}(\mathbf{k}') u(\mathbf{k}).$$

Now, a term such as $\bar{u}(\mathbf{k}) v(\mathbf{k}')$ is a scalar; hence, this is just two scalars multiplied together. Let us then write this with indices, so that all terms involved are numbers, and can hence be commuted at will,

$$\begin{aligned} \sum_{s,s'} \left|\mathcal{M}_{\rm fi}\right|^2 &= g^2 \sum \bar{u}_a(\mathbf{k}) v_a(\mathbf{k}') \bar{v}_b(\mathbf{k}') u_b(\mathbf{k}) \\ &= g^2 \sum u_b(\mathbf{k}) \bar{u}_a(\mathbf{k}) v_a(\mathbf{k}') \bar{v}_b(\mathbf{k}'). \end{aligned}$$

We now use two identities of the spinors,

$$\sum u_b(\mathbf{k})\bar{u}_a(\mathbf{k}) = (\not{k} + m_\tau)_{ba}\,, \quad \sum v_a(\mathbf{k}')\bar{v}_b(\mathbf{k}') = (\not{k}' - m_\tau)_{ab}\,,$$

where $\not{k} \equiv \gamma_\mu k^\mu$; and hence

$$\sum_{s,s'} |\mathcal{M}_{\text{fi}}|^2 = g^2\,(\not{k} + m_\tau)_{ba}\,(\not{k}' - m_\tau)_{ab}\,.$$

Now, an expression with indices arranged so is just a trace, so that

$$\begin{aligned}\sum_{s,s'} |\mathcal{M}_{\text{fi}}|^2 &= g^2 \text{Tr}\ [(\not{k} + m_\tau)\,(\not{k}' - m_\tau)] \\ &= g^2 \left(\text{Tr}\ (\not{k}\not{k}') - m_\tau \text{Tr}\ (\not{k}) + m_\tau \text{Tr}\ (\not{k}') - m_\tau^2 \text{Tr}\ (\mathbf{1}_4)\right).\end{aligned} \tag{5.13}$$

Now the traces of the middle two terms are zero, by noting

$$\text{Tr}\ (\not{k}) = \text{Tr}\ (\gamma_\mu k^\mu) = k^\mu \text{Tr}\ (\gamma_\mu) = 0,$$

where we first note that k^μ is not a matrix, and secondly that the γ-matrices are traceless. The last term in (5.13) is fairly obviously just Tr $(\mathbf{1}_4) = 4$. Hence, using these two results, we see that (5.13) reduces down to

$$\sum_{s,s'} |\mathcal{M}_{\text{fi}}|^2 = g^2 \left(\text{Tr}\ (\not{k}\not{k}') - 4m_\tau^2\right). \tag{5.14}$$

To evaluate the first term, we write

$$\not{k}\not{k}' = \gamma_\mu\gamma_\nu k^\mu k'^\nu = \gamma_\nu\gamma_\mu k^\nu k'^\mu.$$

We now use $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbf{1}_4$ (where we must remember that the LHS of this is a matrix equation, and hence so must the RHS – therefore the introduction of the unit matrix). Hence, we write

$$\text{Tr}\ (\gamma^\mu\gamma^\nu) + \text{Tr}\ (\gamma^\nu\gamma^\mu) = 2g^{\mu\nu}\text{Tr}\ (\mathbf{1}_4).$$

Now, due to the cyclicality of the trace, the LHS is just 2Tr $(\gamma^\mu\gamma^\nu)$. Therefore,

$$\text{Tr}\ (\gamma^\mu\gamma^\nu) = g^{\mu\nu}\text{Tr}\ (\mathbf{1}_4) = 4g^{\mu\nu}.$$

Therefore,

$$\text{Tr}\ (\not{k}\not{k}') = 4g^{\mu\nu}k_\mu k'_\nu.$$

Hence, using this in (5.14), we see that

$$\sum_{s,s'} |\mathcal{M}_{\text{fi}}|^2 = 4g^2 \left(g^{\mu\nu}k_\mu k'_\nu - m_\tau^2\right),$$

and therefore,

$$g^2 \sum_{s,s'} |\bar{u}_s(\mathbf{k}) v_{s'}(\mathbf{k}')|^2 = 4g^2 \left(kk' - m_\tau^2 \right), \tag{5.15}$$

where $kk' = k^\mu k'_\mu$.

We now specialise to the rest-frame of the Higgs boson, $P^\mu = (m_\mathrm{H}, \mathbf{0})$. Therefore, conservation of 4-momentum implies that

$$\mathbf{k} = -\mathbf{k}', \quad E(\mathbf{k}) = E(\mathbf{k}') = \frac{m_\mathrm{H}}{2}. \tag{5.16}$$

Hence,

$$\begin{aligned} E^2(\mathbf{k}) = \mathbf{k}^2 + m_\tau^2 \quad \Rightarrow \quad \mathbf{k}^2 &= E^2(\mathbf{k}) - m_\tau^2 \\ &= \frac{m_\mathrm{H}^2}{4} - m_\tau^2. \end{aligned} \tag{5.17}$$

Therefore, noting that the kk' term in (5.15) can be written

$$kk' = k^\mu k'_\mu = E^2(\mathbf{k}) + \mathbf{k}^2 = 2\mathbf{k}^2 + m_\tau^2,$$

where we have taken careful note of the change in sign. Hence, inserting (5.17) into this expression, we see that

$$kk' = \frac{m_\mathrm{H}^2}{2} - m_\tau^2,$$

which we can insert into (5.15) to see that

$$\sum_{s,s'} |\mathcal{M}_\mathrm{fi}|^2 = g^2 \sum_{s,s'} |\bar{u}_s(\mathbf{k}) v_{s'}(\mathbf{k}')|^2 = 8g^2 \left(\frac{m_\mathrm{H}^2}{4} - m_\tau^2 \right). \tag{5.18}$$

Let us now do the phase-space integral (5.12), in the rest-frame of the Higgs. So, we must compute

$$I = \int \frac{d^3k}{2E_\mathbf{k}} \int \frac{d^3k'}{2E'_{\mathbf{k}'}} \delta\left(E(\mathbf{k}) + E(\mathbf{k}') - m_\mathrm{H} \right) \delta^3(\mathbf{k} + \mathbf{k}'),$$

where we have ignored the factors of 2π (to be put back in later), and we have also split the delta-function into its space and time parts. The $\delta^{(3)}(\mathbf{k} + \mathbf{k}')$ term, whilst evaluating the d^3k'-integral, filters out $\mathbf{k} = -\mathbf{k}'$ as the only non-zero contribution, so that

$$I = \int \frac{d^3k}{4E_\mathbf{k} E_{-\mathbf{k}}} \delta\left(E(\mathbf{k}) + E(-\mathbf{k}) - m_\mathrm{H} \right).$$

We now use the second of (5.16) to write

$$I = \int \frac{d^3k}{4E_\mathbf{k}^2} \delta\left(2E(\mathbf{k}) - m_\mathrm{H} \right),$$

and then $E(\mathbf{k}) = \sqrt{\mathbf{k}^2 + m_\tau^2}$, to see that

$$I = \int \frac{d^3k}{4E_{\mathbf{k}}^2} \delta\left(2\sqrt{\mathbf{k}^2 + m_\tau^2} - m_{\mathrm{H}}\right). \tag{5.19}$$

To evaluate the delta-function, we use the usual relation for a functional argument,

$$\delta\left(f(x)\right) = \sum_{x_0} \frac{\delta(x - x_0)}{|f'(x_0)|}, \quad f(x_0) = 0.$$

Now, the argument of the delta-function in (5.19) is zero when

$$\mathbf{k} = \mathbf{k}_0 = \sqrt{\frac{m_{\mathrm{H}}^2}{4} - m_\tau^2},$$

and the value of the derivative of the argument, at $\mathbf{k}_0$ is

$$f'(\mathbf{k}_0) = \frac{2\mathbf{k}}{E(\mathbf{k})}.$$

Therefore, using these results, we have that (5.19) reads

$$I = \int \frac{d^3k}{4E_{\mathbf{k}}^2} \frac{E(|\mathbf{k}|)}{2|\mathbf{k}|} \delta\left(|\mathbf{k}| - \sqrt{\frac{m_{\mathrm{H}}^2}{4} - m_\tau^2}\right),$$

or, if we assume no angular dependence,

$$\begin{aligned} I &= \int \frac{4\pi|\mathbf{k}|^2}{4E_{\mathbf{k}}^2} d|\mathbf{k}| \frac{E(|\mathbf{k}|)}{2|\mathbf{k}|} \delta\left(|\mathbf{k}| - \sqrt{\frac{m_{\mathrm{H}}^2}{4} - m_\tau^2}\right) \\ &= \frac{4\pi|\mathbf{k}_0|}{8E(\mathbf{k}_0)}, \quad \mathbf{k}_0 \equiv \sqrt{\frac{m_{\mathrm{H}}^2}{4} - m_\tau^2}. \end{aligned}$$

Then, using the second of (5.16), this cancels down to

$$I = \frac{\pi|\mathbf{k}_0|}{m_{\mathrm{H}}}, \quad \mathbf{k}_0 \equiv \sqrt{\frac{m_{\mathrm{H}}^2}{4} - m_\tau^2}.$$

Hence, using this result and (5.18), we see that the decay rate (5.12) becomes

$$\begin{aligned} \Gamma &= \frac{1}{2m_{\mathrm{H}}} \frac{1}{(2\pi)^2} 8g^2 \left(\frac{m_{\mathrm{H}}^2}{4} - m_\tau^2\right) \frac{\pi|\mathbf{k}_0|}{m_{\mathrm{H}}} \\ &= \frac{g^2 \left(\frac{m_{\mathrm{H}}^2}{4} - m_\tau^2\right)^{3/2}}{\pi m_{\mathrm{H}}^2}, \end{aligned} \tag{5.20}$$

after using $\omega_{\mathbf{p}} = m_{\mathrm{H}}$ in the rest-frame of the Higgs. Now, in the limit $m_{\mathrm{H}} \gg m_\tau$ (which is a very good approximation),

$$\Gamma\left(H^0 \to \tau^+\tau^-\right) = \frac{g^2 m_{\mathrm{H}}}{8\pi}.$$

We can use g from (5.3) to see that

$$\Gamma\left(H^0 \to \tau^+\tau^-\right) = \mathcal{K} m_{\mathrm{H}} m_\tau^2, \quad \mathcal{K} \equiv \frac{\mathcal{G}}{8\sqrt{2}\pi}. \tag{5.21}$$

Therefore, we have found that the decay rate for the process $H^0 \to \tau^+\tau^-$ increases with the mass of the Higgs particle (infact, it will increase for the larger product mass as well, as there is a larger phase space to "decay into"). We can somewhat generalise this decay rate to any lepton $\ell = e^-, \mu^-, \tau^-$, to give the decay rate

$$\Gamma\left(H^0 \to \ell^+\ell^-\right) = \mathcal{K} m_{\mathrm{H}} m_\ell^2. \tag{5.22}$$

For quarks, we have to multiply the answer by 3, as there are 3 colour states that could be "decayed into", hence increasing the phase space further;

$$\Gamma\left(H^0 \to q\bar{q}\right) = 3\mathcal{K} m_{\mathrm{H}} m_q^2.$$

As the top t quark has the highest mass, the dominant decay channel is $H^0 \to t\bar{t}$.

5.3 Example: Decay of a Scalar to two Scalars

Let us consider a second example, of a neutral scalar particle decaying into two charged scalars. We will compute the matrix element $\mathcal{M}_{\mathrm{fi}}$, the general decay rate Γ and the decay rate of a stationary particle.

So, consider

$$\sigma \longrightarrow \phi^- + \phi^+ \quad \mathbf{p} \longrightarrow \mathbf{k} + \mathbf{k}'.$$

The interaction Hamiltonian we are considering is

$$\mathcal{H}_{\mathrm{I}} = g\phi^\dagger(x)\phi(x)\sigma(x),$$

where the field expansions are

$$\begin{aligned}
\sigma(x) &= \sum_{\mathbf{p}} \frac{1}{(2VE_{\mathbf{p}})^{1/2}} \left(a(\mathbf{p})e^{-ipx} + a^\dagger(\mathbf{p})e^{ipx}\right), \\
\phi(x) &= \sum_{\mathbf{k}} \frac{1}{(2V\omega_{\mathbf{k}})^{1/2}} \left(b(\mathbf{k})e^{-ikx} + d^\dagger(\mathbf{k})e^{ikx}\right);
\end{aligned}$$

and

$$\begin{aligned}
px &= p^\mu x_\mu = Et - \mathbf{p}\cdot\mathbf{x}, \\
E_{\mathbf{p}} &= \sqrt{\mathbf{p}^2 + m_\sigma^2}, \quad \omega_{\mathbf{k}} = \sqrt{\mathbf{k}^2 + m_\phi^2}.
\end{aligned}$$

The important features of these expansions are: the σ-field is uncharged, which is modeled by being Hermitian (i.e. we use a and $a^\dagger$). In contrast, the ϕ-field is charged, modeled by not being Hermitian, so we use b and $d^\dagger$.

We need to first compute $\mathcal{S}_{\text{fi}}$,

$$\mathcal{S}_{\text{fi}} = -i \int d^4x \, \langle f| \mathcal{H}_{\text{I}}(x) |i\rangle \, ,$$

where the initial and final states are just

$$|i\rangle = |\sigma, \mathbf{p}\rangle \, , \quad |f\rangle = \left|\phi^-, \mathbf{k}; \phi^+, \mathbf{k}'\right\rangle .$$

These states can be created from the vacuum via application of the creation operators;

$$|i\rangle = a^\dagger(\mathbf{p}) |0\rangle \, , \quad |f\rangle = b^\dagger(\mathbf{k}) d^\dagger(\mathbf{k}') |0\rangle \, .$$

Then, using our interaction Hamiltonian, we want to compute

$$g \langle f| \phi^\dagger \phi \sigma |i\rangle = g \langle 0| d(\mathbf{k}') b(\mathbf{k}) \phi^\dagger \phi \sigma a^\dagger(\mathbf{p}) |0\rangle \, .$$

The only non-zero terms in the field expansion can then be deduced by requiring those which "null out" those operators which are already present. The term from σ will be $a(\mathbf{p})e^{-ipx}$, as that will destroy the particle the already present $a^\dagger(\mathbf{p})$-term has created. Similarly, we also need $d^\dagger(\mathbf{k}')$- and $b^\dagger(\mathbf{k})$-terms to create the particles that are destroyed by the already present $d(\mathbf{k}')b(\mathbf{k})$-terms. We can get $d^\dagger(\mathbf{k}')$ simply from the field expansion as it stands (also picking up a factor of $e^{ik'x}$). We get the $b^\dagger(\mathbf{k})$-term by conjugating the field expansion, in which case we pick up a factor of e^{ikx} as well. Hence, we have (also inserting the factors of energy and volume we have neglected to mention thus far),

$$\begin{aligned} g \langle f| \phi^\dagger \phi \sigma |i\rangle &= g \langle 0| d(\mathbf{k}') b(\mathbf{k}) b^\dagger(\mathbf{k}) d^\dagger(\mathbf{k}') a(\mathbf{p}) a^\dagger(\mathbf{p}) |0\rangle \\ &\quad \times \frac{1}{(2VE_{\mathbf{p}})^{1/2}} \frac{1}{(2V\omega_{\mathbf{k}})^{1/2}} \frac{1}{(2V\omega_{\mathbf{k}'})^{1/2}} \\ &\quad \times e^{ix(k+k'-p)}. \end{aligned}$$

The entire "state" part contracts to unity (by construction), to leave just

$$g \langle f| \phi^\dagger \phi \sigma |i\rangle = g \frac{1}{(2VE_{\mathbf{p}})^{1/2}} \frac{1}{(2V\omega_{\mathbf{k}})^{1/2}} \frac{1}{(2V\omega_{\mathbf{k}'})^{1/2}} e^{ix(k+k'-p)}.$$

Thus, the integral of this is just $\mathcal{S}_{\text{fi}}$, so that

$$\begin{aligned} \mathcal{S}_{\text{fi}} &= -i \int d^4x \, g \langle f| \phi^\dagger \phi \sigma |i\rangle \\ &= \int d^4x \, \frac{1}{(2VE_{\mathbf{p}})^{1/2}} \frac{1}{(2V\omega_{\mathbf{k}})^{1/2}} \frac{1}{(2V\omega_{\mathbf{k}'})^{1/2}} e^{ix(k+k'-p)} \mathcal{M}_{\text{fi}}, \end{aligned}$$

where we have defined the Feynman amplitude to be

$$\mathcal{M}_{\text{fi}} \equiv -ig.$$

This rather simple amplitude is just because every particle involved is scalar, and hence we only have a contribution from the vertex itself. We now use the property that

$$\int d^4x\, e^{ix(k+k'-p)} = (2\pi)^4\delta^{(4)}(k+k'-p),$$

so that we can do the integral in $\mathcal{S}_{\text{fi}}$;

$$\mathcal{S}_{\text{fi}} = \frac{1}{(2VE_{\mathbf{p}})^{1/2}}\frac{1}{(2V\omega_{\mathbf{k}})^{1/2}}\frac{1}{(2V\omega_{\mathbf{k}'})^{1/2}}(2\pi)^4\delta^{(4)}(k+k'-p)\mathcal{M}_{\text{fi}}.$$

Hence, we have computed an expression for the matrix element for neutral scalar decay into charged scalar particles.

Let us now compute the decay rate. To do so, we perform a summation over all possible final states;

$$\Gamma = \sum_f \frac{|\mathcal{S}_{\text{fi}}|^2}{T}.$$

To be able to compute $|\mathcal{S}_{\text{fi}}|^2$, (infact, more precisely, to compute the square of the delta-function), we go into the finite-box limit, so that

$$(2\pi)^4\delta^{(4)}(k+k'-p) = VT\delta_{k+k',p}.$$

Hence, squaring this:

$$\left((2\pi)^4\delta^{(4)}(k+k'-p)\right)^2 = V^2T^2\left(\delta_{k+k',p}\right)^2,$$

but

$$\left(\delta_{k+k',p}\right)^2 = \delta_{k+k',p}.$$

So,

$$\left((2\pi)^4\delta^{(4)}(k+k'-p)\right)^2 = V^2T^2\delta_{k+k',p}.$$

Hence, upon computing the square of the matrix element, we easily see that

$$|\mathcal{S}_{\text{fi}}|^2 = \frac{1}{(2VE_{\mathbf{p}})}\frac{1}{(2V\omega_{\mathbf{k}})}\frac{1}{(2V\omega_{\mathbf{k}'})}V^2T^2\delta_{k+k',p}|\mathcal{M}_{\text{fi}}|^2.$$

Let us cancel off one of the factors of V, and divide by T, so that

$$\frac{|\mathcal{S}_{\text{fi}}|^2}{T} = \frac{1}{2E_{\mathbf{p}}}\frac{1}{(2V\omega_{\mathbf{k}})}\frac{1}{(2V\omega_{\mathbf{k}'})}VT\delta_{k+k',p}|\mathcal{M}_{\text{fi}}|^2,$$

and we now send $VT\delta_{k+k',p} \to (2\pi)^4\delta^{(4)}(k+k'-p)$, so that

$$\frac{|\mathcal{S}_{\rm fi}|^2}{T} = \frac{1}{2E_{\mathbf{p}}}\frac{1}{(2V\omega_{\mathbf{k}})}\frac{1}{(2V\omega_{\mathbf{k}'})}(2\pi)^4\delta^{(4)}(k+k'-p)|\mathcal{M}_{\rm fi}|^2.$$

To compute the decay rate, we must sum this expression over all possible final states, so that

$$\begin{aligned}\Gamma &= \sum_f \frac{|\mathcal{S}_{\rm fi}|^2}{T} \\ &= \sum_{\mathbf{k},\mathbf{k}'} \frac{1}{2E_{\mathbf{p}}}\frac{1}{(2V\omega_{\mathbf{k}})}\frac{1}{(2V\omega_{\mathbf{k}'})}(2\pi)^4\delta^{(4)}(k+k'-p)|\mathcal{M}_{\rm fi}|^2.\end{aligned}$$

To perform the summation, we take the infinite-box limit, so that

$$\sum_{\mathbf{k}} \frac{1}{V} \longmapsto \int \frac{d^3k}{(2\pi)^3}.$$

Hence,

$$\Gamma = \frac{1}{2E_{\mathbf{p}}}\int \frac{d^3k}{(2\pi)^3 2\omega_{\mathbf{k}}}\int \frac{d^3k'}{(2\pi)^3 2\omega_{\mathbf{k}'}}(2\pi)^4\delta^{(4)}(k+k'-p)|\mathcal{M}_{\rm fi}|^2.$$

This is a very general expression for the decay rate. We now specialise to the rest frame of the decaying particle; that is, we will compute the decay rate of a σ-particle at rest. This will allow us to compute the delta-function.

In this frame, we have $\mathbf{p} = 0$, so that

$$E_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m_\sigma^2} = m_\sigma.$$

Now, we can always split up the 4-dimensional delta-function, thus

$$\delta^{(4)}(k+k'-p) = \delta(\omega_{\mathbf{k}} + \omega_{\mathbf{k}'} - E_{\mathbf{p}})\delta^{(3)}\left(\mathbf{k}+\mathbf{k}'-\mathbf{p}\right).$$

In the rest frame of the decaying particle, the second 3-D delta-function just becomes $\delta^{(3)}(\mathbf{k}+\mathbf{k}')$. Hence, using this, the decay rate becomes

$$\begin{aligned}\Gamma &= \frac{1}{2E_{\mathbf{p}}}\int \frac{d^3k}{(2\pi)^3 2\omega_{\mathbf{k}}}\int \frac{d^3k'}{(2\pi)^3 2\omega_{\mathbf{k}'}} \\ &\quad \left\{(2\pi)^4\delta(\omega_{\mathbf{k}} + \omega_{\mathbf{k}'} - E_{\mathbf{p}})\delta^{(3)}(\mathbf{k}+\mathbf{k}')|\mathcal{M}_{\rm fi}|^2\right\}.\end{aligned}$$

The first d^3k'-integral will only return a non-zero value if $\mathbf{k}' = -\mathbf{k}$, due to the 3-D delta-function. Hence, after doing this integral, we have

$$\Gamma = \frac{1}{2E_{\mathbf{p}}}\int \frac{d^3k}{(2\pi)^3 2\omega_{\mathbf{k}}}\frac{1}{(2\pi)^3 2\omega_{-\mathbf{k}}}(2\pi)^4\delta(\omega_{\mathbf{k}} + \omega_{-\mathbf{k}} - E_{\mathbf{p}})|\mathcal{M}_{\rm fi}|^2.$$

By inspection of $\omega_{\mathbf{k}} = \sqrt{\mathbf{k}^2 + m_\phi^2}$, we see that $\omega_{-\mathbf{k}} = \omega_{\mathbf{k}}$. Hence, the remaining delta-function becomes just

$$\delta(2\omega_{\mathbf{k}} - E_{\mathbf{p}}) = \delta(2\omega_{\mathbf{k}} - m_\sigma).$$

Inserting $\omega_{\mathbf{k}}$ just gives

$$\delta(2\omega_{\mathbf{k}} - m_\sigma) = \delta\left(2\sqrt{\mathbf{k}^2 + m_\phi^2} - m_\sigma\right).$$

To make this delta-function useable, we must get it into the form $\delta(\mathbf{k} - \mathbf{k}_0)$, so that upon integration, only $\mathbf{k} = \mathbf{k}_0$ will give a non-zero contribution. To do this, we use the property

$$\delta(f(x)) = \sum_{x_0} \frac{\delta(x - x_0)}{|f'(x_0)|}, \quad f(x_0) = 0.$$

Hence, given our delta-function, we see that its argument is zero when

$$|\mathbf{k}| = |\mathbf{k}_0| = \sqrt{\frac{m_\sigma^2}{4} - m_\phi^2}.$$

We also see that

$$f'(|\mathbf{k}|) = \frac{2|\mathbf{k}|}{\omega_{\mathbf{k}}}.$$

Hence, using this,

$$\delta\left(2\sqrt{\mathbf{k}^2 + m_\phi^2} - m_\sigma\right) = \frac{\omega_{\mathbf{k}_0}}{2|\mathbf{k}_0|}\delta\left(|\mathbf{k}| - |\mathbf{k}_0|\right).$$

Therefore, we can insert this expression for the delta-function in the decay rate, to see that

$$\Gamma = \frac{1}{2E_{\mathbf{p}}} \int \frac{d^3k}{(2\pi)^3 4\omega_{\mathbf{k}}^2} \frac{1}{(2\pi)^3} (2\pi)^4 \frac{\omega_{\mathbf{k}_0}}{2|\mathbf{k}_0|} \delta\left(|\mathbf{k}| - |\mathbf{k}_0|\right) |\mathcal{M}_{\mathrm{fi}}|^2.$$

Let us rearrange slightly,

$$\Gamma = \frac{1}{2E_{\mathbf{p}}} \frac{\omega_{\mathbf{k}_0}}{2|\mathbf{k}_0|} \int \frac{d^3k}{(2\pi)^2 4\omega_{\mathbf{k}}^2} \delta\left(|\mathbf{k}| - |\mathbf{k}_0|\right) |\mathcal{M}_{\mathrm{fi}}|^2.$$

We now make the integral spherically symmetric, so that

$$\int d^3k \longmapsto \int d|\mathbf{k}|\, |\mathbf{k}|^2 4\pi.$$

Hence,

$$\Gamma = \frac{1}{2E_{\mathbf{p}}} \frac{\omega_{\mathbf{k}_0}}{2|\mathbf{k}_0|} \int \frac{d|\mathbf{k}|\, |\mathbf{k}|^2 4\pi}{(2\pi)^2 4\omega_{\mathbf{k}}^2} \delta\left(|\mathbf{k}| - |\mathbf{k}_0|\right) |\mathcal{M}_{\mathrm{fi}}|^2,$$

where integration will now just filter out $|\mathbf{k}| = |\mathbf{k}_0|$, so that

$$\begin{aligned} \Gamma &= \frac{1}{2E_{\mathbf{p}}} \frac{\omega_{\mathbf{k}_0}}{2|\mathbf{k}_0|} \frac{|\mathbf{k}_0|^2 \pi}{(2\pi)^2 \omega_{\mathbf{k}_0}^2} |\mathcal{M}_{\mathrm{fi}}|^2 \\ &= \frac{1}{2E_{\mathbf{p}}} \frac{|\mathbf{k}_0|}{8\pi\omega_{\mathbf{k}_0}} |\mathcal{M}_{\mathrm{fi}}|^2. \end{aligned}$$

We can tidy this expression up slightly, by recalling that we are in the frame where the decaying particle is at rest; so that $E_{\mathbf{p}} = m_\sigma$. Now, we have that

$$\omega_{\mathbf{k}_0}^2 = \mathbf{k}_0^2 + m_\phi^2,$$

where we defined

$$\mathbf{k}_0^2 = \frac{m_\sigma^2}{4} - m_\phi^2.$$

Hence,

$$\omega_{\mathbf{k}_0} = \frac{m_\sigma}{2}.$$

Thus,

$$\frac{|\mathbf{k}_0|}{8\pi\omega_{\mathbf{k}_0}} = \frac{1}{8\pi}\frac{2}{m_\sigma}\sqrt{\frac{m_\sigma^2}{4} - m_\phi^2}.$$

Therefore, we can use this in the decay rate, to see that

$$\Gamma = \frac{1}{4\pi m_\sigma^2}\frac{1}{2}\sqrt{\frac{m_\sigma^2}{4} - m_\phi^2}|\mathcal{M}_{\text{fi}}|^2.$$

Due to the simple form of our $\mathcal{M}_{\text{fi}} = -ig$, we easily have

$$|\mathcal{M}_{\text{fi}}|^2 = g^2.$$

Hence, we have computed that the complete expression for the decay rate of a scalar σ-particle to a pair of scalar charged ϕ-particles is

$$\Gamma\left(\sigma \to \phi^+\phi^-\right) = \frac{g^2}{8\pi m_\sigma^2}\sqrt{\frac{m_\sigma^2}{4} - m_\phi^2}. \tag{5.23}$$

Let us recall the equivalent expression we derived for a scalar decaying to two fermions, (5.20),

$$\Gamma(\sigma \to \tau^+\tau^-) = \frac{g^2}{\pi m_\sigma^2}\left(\frac{m_\sigma^2}{4} - m_\tau^2\right)^{3/2}.$$

We can now make a comparison of the two expressions. First, we note that the overall structure of the decay rate is the same. One of the differences is that the bracketed factor is raised to a different power. This is a general feature of fermionic decays. Secondly, the scalar decay rate has a factor 1/8 suppression relative to the fermionic. This is because there are more fermionic states in the phase space to decay to, than in the scalar phase space (that is, each fermion has two degrees of freedom).

6 Scattering Processes

6.1 An Outline of Particle Exchange

Let us extend our previous interaction Hamiltonian, to

$$\mathcal{H}_{\mathrm{I}} = g_\tau \bar{\psi}_\tau(x)\psi_\tau(x)\phi(x) + g_e \bar{\psi}_e(x)\psi_e(x)\phi(x). \tag{6.1}$$

Hence, we have two types of lepton, τ^- and e^-, which can scatter by boson exchange in the process

$$e^-(\mathbf{p}_1, s_1) + \tau^-(\mathbf{p}_2, s_2) \longrightarrow e^-(\mathbf{p}_3, s_3) + \tau^-(\mathbf{p}_4, s_4).$$

So, the initial state is just

$$|i\rangle = b^\dagger_{e,s_1}(\mathbf{p}_1) b^\dagger_{\tau,s_2}(\mathbf{p}_2) |0\rangle\,,$$

and the final state

$$|f\rangle = b^\dagger_{e,s_3}(\mathbf{p}_3) b^\dagger_{\tau,s_4}(\mathbf{p}_4) |0\rangle\,.$$

So that we can describe the process, we obviously need two electron and two tauon field operators. The lowest order term in the S-matrix expansion that can do this is the second order,

$$S^{(2)} = \frac{(-i)^2}{2} \int d^4x \int d^4x' \, \langle f| \, T\left(\mathcal{H}_{\mathrm{I}}(x)\mathcal{H}_{\mathrm{I}}(x')\right) |i\rangle\,. \tag{6.2}$$

From this, we need the operators $\bar{\psi}^{(-)}_e(x), \bar{\psi}^{(-)}_\tau(x'), \psi^{(+)}_e(x)$ and $\psi^{(+)}_\tau(x')$, which create electrons/tauons and destroy electrons/tauons; notice that we have electron destruction/creation at a different place to tauon destruction/creation. Hence, the term we want is

$$2 \langle f| \, \bar{\psi}^{(-)}_e(x)\bar{\psi}^{(-)}_\tau(x')\psi^{(+)}_e(x)\psi^{(+)}_\tau(x') \, |i\rangle\,,$$

where we have multiplied by 2 as the choice of which is x and which x' is arbitrary. Therefore, the matrix element to be computed is

$$\begin{aligned} S^{(2)} = (-i)^2 \quad & \int d^4x \int d^4x' g_e g_\tau \\ & \times \langle f| \, \bar{\psi}^{(-)}_e(x)\bar{\psi}^{(-)}_\tau(x')\psi^{(+)}_e(x)\psi^{(+)}_\tau(x') T\left(\phi(x)\phi(x')\right) |i\rangle\,. \end{aligned} \tag{6.3}$$

The fields act upon the initial and final states in an identical manner to our discussion on decays, so that for example,

$$\psi^{(+)}_e(x) \left|e^-, \mathbf{p}_1, s_1\right\rangle = \frac{u_{e,s_1}(\mathbf{p}_1) e^{-ip_1x}}{\sqrt{2VE(\mathbf{p}_1)}} |0\rangle\,.$$

Then, plugging all of the relevant terms into (6.3), we end up with

$$\begin{aligned}\mathcal{S}_{\text{fi}} &= (-ig_e)\bar{u}_{e,s_3}(\mathbf{p}_3)u_{e,s_1}(\mathbf{p}_1)(-ig_\tau)\bar{u}_{\tau,s_4}(\mathbf{p}_4)u_{\tau,s_2}(\mathbf{p}_2) \\ &\times \prod_{i=1}^{4}\frac{1}{(2VE_i)^{1/2}}\int d^4x\int d^4x' e^{ip_3x}e^{-ip_1x}e^{ip_4x'}e^{-ip_2x'} \\ &\times \langle 0|\, T\left(\phi(x)\phi(x')\right)|0\rangle\,. \end{aligned} \tag{6.4}$$

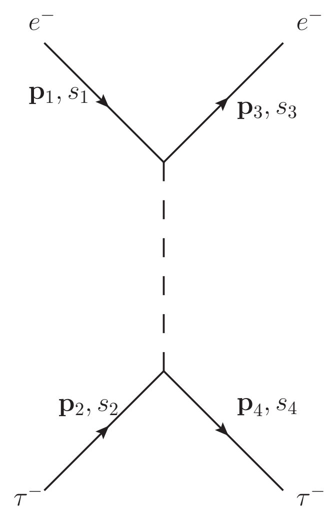


Figure 6.1: Feynman graph for $e^-\tau^- \to e^-\tau^-$ scattering.

These terms perhaps become a little more understandable if we consider Figure (6.1). The spinor factors $\bar{u}_{e,s_3}(\mathbf{p}_3)u_{e,s_1}(\mathbf{p}_1)$ can be thought about as first destroying an electron (i.e. a particle), with an un-barred spinor on the far right, then an electron created after the destruction, with a barred spinor. The exponential factors $e^{ip_3x}e^{-ip_1x}$ can be thought about equivalently. Notice that the exponential factors (and these are the only spatial terms) for the tauons have primes on them, denoting their interaction happens at a different place to the electrons interaction. The final term is the Feynman propagator, which as we saw in Section (4.4), can be written in a variety of ways:

$$\langle 0|\, T\left(\phi(x)\phi(x')\right)|0\rangle = iG_{\text{F}}(x-x') = \int \frac{d^4q}{(2\pi)^4}e^{-iq(x-x')}\frac{i}{q^2-m_{\text{H}}^2+i\epsilon}.$$

We now recall that we interpreted $\langle 0|\, T\left(\phi(x)\phi(x')\right)|0\rangle$ as creating a particle at x and destroying at x'. Therefore, this Feynman propagator term describes the exchange of a "virtual" H^0. Substituting the propagator into (6.4), we can begin to do integrals. First, integrating over x only concerns the terms

$$\int d^4x e^{ip_3x}e^{-ip_1x}e^{-iqx} = (2\pi)^4\delta^{(4)}\left(p_3-p_1-q\right),$$

and integrating over x' only

$$\int d^4x' e^{i(p_4-p_2+q)x'} = (2\pi)^4 \delta^{(4)}\left(p_4 - p_2 + q\right).$$

Finally, integrating over q,

$$\int \frac{d^4q}{(2\pi)^4}\left\{(2\pi)^4\delta^{(4)}\left(p_3-p_1-q\right)(2\pi)^4\delta^{(4)}\left(p_4-p_2+q\right)\frac{i}{q^2-m_{\mathrm{H}}^2+i\epsilon}\right\},$$

which simply gives

$$(2\pi)^4\delta^{(4)}\left(p_3+p_4-p_1-p_2\right)\frac{i}{q^2-m_{\mathrm{H}}^2+i\epsilon}, \quad q = p_3 - p_1 = p_4 - p_2.$$

Hence, putting all of this together, we have that (6.4) becomes

$$\mathcal{S}_{\mathrm{fi}}^{(2)} = \left(\prod_{i=1}^{4}\frac{1}{(2VE_i)^{1/2}}\right)(2\pi)^4\delta^{(4)}\left(p_3+p_4-p_1-p_2\right)\mathcal{M}_{\mathrm{fi}}, \tag{6.5}$$

where we have defined the Feynman amplitude to be

$$\mathcal{M}_{\mathrm{fi}} = (-ig_e)\bar{u}_{e,s_3}(\mathbf{p}_3)u_{e,s_1}(\mathbf{p}_1)\frac{i}{q^2-m_{\mathrm{H}}^2+i\epsilon}(-ig_\tau)\bar{u}_{\tau,s_4}(\mathbf{p}_4)u_{\tau,s_2}(\mathbf{p}_2). \tag{6.6}$$

Notice that (6.5) has an identical structure to the matrix element we derived for decays; normalisation times energy-momentum conservation times the dynamic parts. We can represent the amplitude (6.6) by a Feynman diagram, as in Figure (6.2).

We can note that the factors in Figure (6.2) are very similar to those we found for the decay, except we now have a factor for the virtual exchange boson. We should also note that at each vertex, charge and lepton numbers are conserved (where lepton numbers stay within their own species); Figure (6.3) has some examples.

This conservation stems from the $\bar{\psi}_e(x)\psi_e(x)$ structured-terms in the Hamiltonian; there are no terms like $\bar{\psi}_\tau(x)\psi_e(x)$ (which would create a τ^- and destroy a e^-).

We should also note that 4-momentum is conserved at each vertex (there are two vertices in Figure (6.2)), as $q = p_3 - p_1 = p_2 - p_4$, something which came about due to delta-functions in the integral. Finally, we should note that the Feynman amplitude represents two time-orderings; which comes from the Feynman propagator allowing for x or x' to come first (i.e. there is no reason which to choose to be first). Figure (6.4) has the two equivalent diagrams (the first two), which we understand to be present when we conventionally draw the third.

This highlights the fact that is dosent matter how the diagrams are drawn; and only the connections (i.e. all initial states on the left, and final on the right), and the number of connections. It is interesting to note that if we had derived the Feynman rules without the time ordering, we would get a single diagram out, which would not be Lorentz invariant (the time ordered double-diagram is Lorentz invariant).

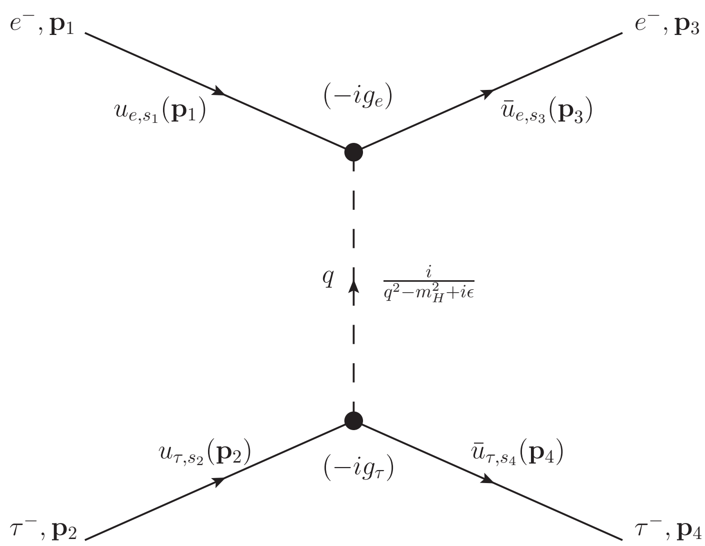
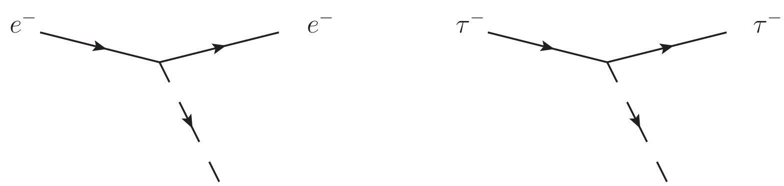
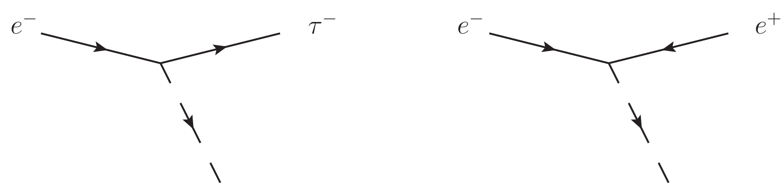


Figure 6.2: Feynman graph for $e^-\tau^- \to e^-\tau^-$ scattering, with Feynman rules.



(a) Valid diagrams.



(b) Invalid diagrams.

Figure 6.3: Examples of charge and lepton conserving diagrams, as well as invalid diagrams.

6.2 Other Scattering Processes

We have discussed non-identical particle scatterings. We could discuss identical particle scatterings as well, such as

$$\tau^-(\mathbf{p}_1) + \tau^-(\mathbf{p}_2) \longrightarrow \tau^-(\mathbf{p}_3) + \tau^-(\mathbf{p}_4).$$

There are now two possibilities. We can now have that when the particles are created, a τ^- can be created with either momentum $\mathbf{p}_3$ or $\mathbf{p}_4$.

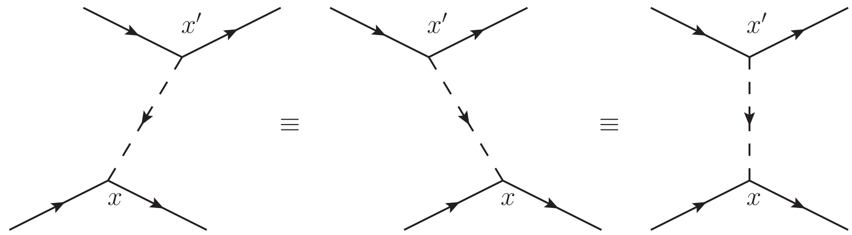


Figure 6.4: The equivalent diagrams represented by the time ordering of the Feynman propagator.

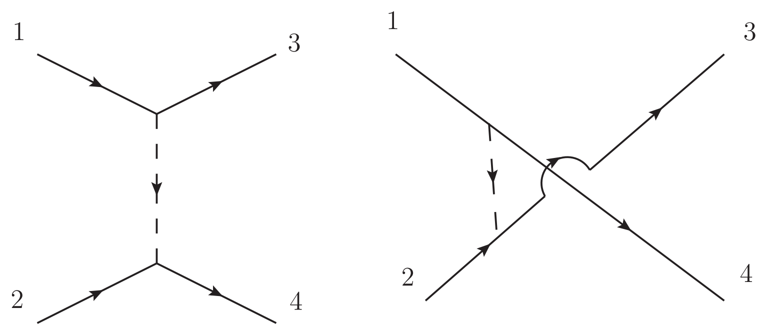


Figure 6.5: The equivalent diagrams represented by the time ordering of the Feynman propagator for scattering of identical particles. The diagram on the left is the direct diagram, on the right the exchange term.

If we do the calculation and keep track of fermion field order, we would find a relative minus sign from the anti-commutation relations (which reflects the Pauli principle). The full invariant amplitude is

$$\mathcal{M}_{\text{fi}} = \mathcal{M}_{\text{fi}}{}^{\text{direct}} + \mathcal{M}_{\text{fi}}{}^{\text{exchange}},$$

where

$$\mathcal{M}_{\text{fi}}{}^{\text{direct}}(\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3\mathbf{p}_4) = (-ig)\bar{u}(\mathbf{p}_3)u(\mathbf{p}_1)\frac{i}{q^2 - m_{\text{H}}^2 + i\epsilon}(-ig)\bar{u}(\mathbf{p}_4)u(\mathbf{p}_2),$$

as before, and

$$\mathcal{M}_{\text{fi}}{}^{\text{exchange}}(\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3\mathbf{p}_4) = -\mathcal{M}_{\text{fi}}{}^{\text{direct}}(\mathbf{p}_1\mathbf{p}_2\mathbf{p}_4\mathbf{p}_3).$$

The minus sign is simply because we are exchanging identical fermions. If we dealt with bosons, one would have the exact same thing, but without the minus sign (as bosonic field commute).

We could also have particle anti-particle scattering,

$$\tau^- + \tau^+ \longrightarrow \tau^- + \tau^+,$$

where, as in Figure (6.6), everything is conserved.

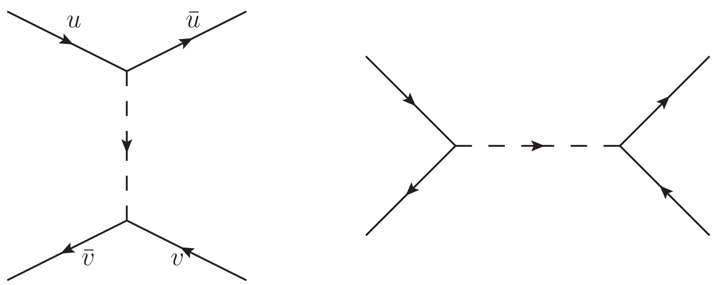


Figure 6.6: The two diagrams corresponding to particle anti-particle scattering.

6.3 Scattering Cross-sections

We now outline how to compute quantities we can actually observe. Let us consider elastic scattering, where the same particles come out as go in, but they may have different momenta and spin; see Figure (6.7).

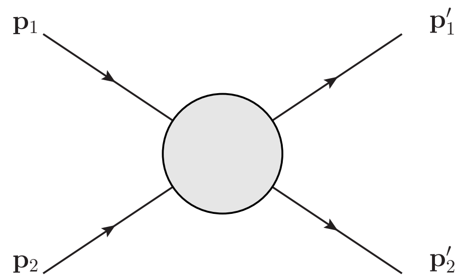


Figure 6.7: Schematic of elastic scattering.

The cross-section σ is defined as the transition rate into a given set of final states per unit flux of initial particles. Hence, in terms of dimensions,

$$[\sigma] = \frac{T^{-1}}{T^{-1}L^{-2}} = L^2.$$

Hence, the cross-section has the units of an area.

For decays, we have that the transition rate is defined as

$$\text{transition rate} = w_{fi} = \frac{|\mathcal{S}_{\text{fi}}|^2}{T}, \tag{6.7}$$

for some unique final state f.

Now, the flux for co-linear collisions (i.e. where the colliding particles hit each other head on, with 180° between them) is

$$\text{flux} \equiv f = \frac{1}{V}|\mathbf{v}_1 - \mathbf{v}_2|,$$

where V is some volume, and $\mathbf{v}_i$ are the velocities of the colliding particles. This is simply the statement that flux is the density times relative speed.

Let us consider an element of the final phase space,

$$\frac{V}{(2\pi)^3}d^3p'_1\frac{V}{(2\pi)^3}d^3p'_2,$$

so that the corresponding elemental cross-section is

$$\begin{aligned} d\sigma &= \frac{1}{\text{flux}} \times \text{transition rate} \times \text{element size} \qquad (6.8)\\ &= \frac{1}{f}\frac{|\mathcal{S}_{\text{fi}}|^2}{T}\frac{V}{(2\pi)^3}d^3p'_1\frac{V}{(2\pi)^3}d^3p'_2. \end{aligned}$$

We know $\mathcal{S}_{\text{fi}}$, by adapting (6.5) to be in the finite volume/time limit (which we do such that we can use $\delta_{ij}^2 = \delta_{ij}$, rather than having to compute the square of the delta-function),

$$\mathcal{S}_{\text{fi}} = \prod_{i=1}^{2}\left[\frac{1}{(2VE_i)^{1/2}}\frac{1}{(2VE'_i)^{1/2}}\right]TV\delta_{p_1+p_2,p'_1+p'_2}\mathcal{M}_{\text{fi}}.$$

Hence, we write the elemental cross-section

$$\begin{aligned} d\sigma &= \frac{V}{|\mathbf{v}_1-\mathbf{v}_2|}\frac{1}{T}\prod_{i=1}^{2}\left[\frac{1}{2VE_i}\frac{1}{2VE'_i}\right]T^2V^2\delta_{p_1+p_2,p'_1+p'_2}\\ &\quad\times|\mathcal{M}_{\text{fi}}|^2\frac{V}{(2\pi)^3}d^3p'_1\frac{V}{(2\pi)^3}d^3p'_2\\ &= \frac{V}{|\mathbf{v}_1-\mathbf{v}_2|}\frac{1}{T}\frac{1}{2VE_1}\frac{1}{2VE'_1}\frac{1}{2VE_2}\frac{1}{2VE'_2}T^2V^2\delta_{p_1+p_2,p'_1+p'_2}\\ &\quad\times|\mathcal{M}_{\text{fi}}|^2\frac{V}{(2\pi)^3}d^3p'_1\frac{V}{(2\pi)^3}d^3p'_2, \qquad (6.9) \end{aligned}$$

which we clean up using

$$\begin{aligned} dQ &\equiv TV\delta_{p_1+p_2,p'_1+p'_2}\frac{d^3p'_1}{(2\pi)^32E'_1}\frac{d^3p'_2}{(2\pi)^32E'_2}\\ &= (2\pi)^4\delta^{(4)}(p_1+p_2-p'_1-p'_2)\frac{d^3p'_1}{(2\pi)^32E'_1}\frac{d^3p'_2}{(2\pi)^32E'_2}, \qquad (6.10) \end{aligned}$$

where we have now taken the infinite box-size limit. Hence, using this definition, (6.9) becomes

$$d\sigma = \frac{1}{4E_1E_2|\mathbf{v}_1-\mathbf{v}_2|}|\mathcal{M}_{\text{fi}}|^2 dQ. \qquad (6.11)$$

We call (6.10) the invariant phase space factor. Now, upon comparison of this with our prototype (6.8), we see that we can write a "modified" flux factor,

$$F \equiv 4E_1E_2|\mathbf{v}_1 - \mathbf{v}_2|. \tag{6.12}$$

This can also be written as

$$F = 4\sqrt{(p_1^\mu p_{2\mu})^2 - m_1^2 m_2^2}.$$

The flux factor F and invariant phase space dQ are both Lorentz invariant, as is $\mathcal{M}_{\mathrm{fi}}$. Hence, $d\sigma$ (6.11) is Lorentz invariant under co-linear transformations. Notice that dQ (6.10) has only two independent scalar variables. The six from having two d^3p's (each has three variables) are constrained by the delta-function, eliminating four of them. Thus, the remaining two are usually chosen to be the polar angle θ, ϕ relative to the beam direction of particle 1 (say). One must note that the angles are frame dependent.

6.3.1 Centre of Momentum Frame

Let us now go into the centre of momentum frame, where $\mathbf{p}_1 + \mathbf{p}_2 = 0$ and thus $\mathbf{p}_1' + \mathbf{p}_2' = 0$. See Figure (6.8).

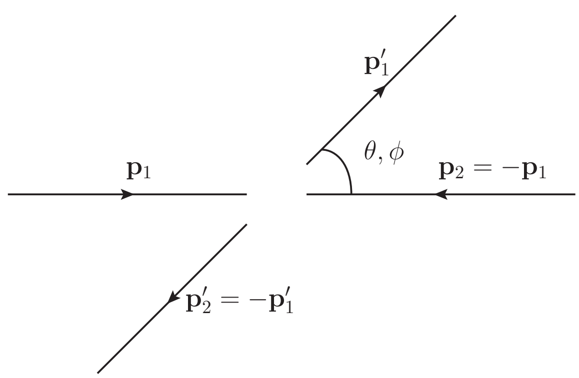


Figure 6.8: Colinear scattering in the centre of momentum frame.

Energy conservation requires that $|\mathbf{p}_1| = |\mathbf{p}_1'|$, as we have the same particles come out that go in. The flux factor (6.12) can be written as

$$F = 4E_1E_2\left(\frac{|\mathbf{p}_1|}{E_1} + \frac{|\mathbf{p}_2|}{E_2}\right) = 4(E_1 + E_2)|\mathbf{p}_1|$$

We can rewrite the phase space factor (6.10) by integrating over redundant variables. Notice that we take take out the zeroth-component of the delta-function,

$$\begin{aligned}\delta(E_1' + E_2' - E_1 - E_2) &= \delta\left(\sqrt{m_1^2 + |\mathbf{p}_1'|^2} + \sqrt{m_2^2 + |\mathbf{p}_2'|^2} - E_1 - E_2\right) \\ &= \delta(f(|\mathbf{p}|)).\end{aligned}$$

Then, we rewrite this delta-function using the usual method for the argument of the function,

$$\delta(E_1' + E_2' - E_1 - E_2) = \frac{E_1 E_2}{(E_1 + E_2)|\mathbf{p}_1|}\delta(|\mathbf{p}_1'| - |\mathbf{p}_1|).$$

Hence, doing this, we find

$$dQ = \frac{1}{(2\pi)^2}\frac{|\mathbf{p}_1|}{4(E_1 + E_2)}d\Omega_{\text{cm}},$$

so that the differential cross-section is

$$\frac{d\sigma}{d\Omega_{\text{cm}}} = \frac{|\mathcal{M}_{\text{fi}}|^2}{64\pi^2(E_1 + E_2)^2}. \tag{6.13}$$

To go further, we have to evaluate $|\mathcal{M}_{\text{fi}}|^2$, which will depend upon spins and momentum. If we "dont care about" spins (i.e. polarisation), then we average over initial spins and sum over final spins – this is done if a detector does not pick up different polarisations. Finally, if we want the total cross-section we integrate over the solid angles θ, ϕ.

6.4 Scattering From a Classical Field

Here we shall consider scattering from a static classical field – classical means that the field is so "big" that one can ignore quantum fluctuations. This could be the case, for example, if electrons are scattered from a large magnetic field, or the Coulomb field of a heavy nucleus.

Let us recall the previous interaction Hamiltonian, for the fermionic fields scattering from a bosonic field,

$$\mathcal{H}_{\text{I}} = gN\left(\bar{\psi}(x)\psi(x)\phi(x)\right).$$

Suppose we take our scalar field $\phi(x)$ to now be static, and let us denote it by $V(\mathbf{x})$. Then, the first order S-matrix element for scattering from $|i\rangle$ to $|f\rangle$ is

$$\mathcal{S}_{\text{fi}} = -ig\int d^4x\ \langle f|\, N\left(\bar{\psi}(x)\psi(x)\right)|i\rangle\, V(\mathbf{x}). \tag{6.14}$$

Now, suppose we have the initial and final states being

$$|i\rangle = |\mathbf{p}, s\rangle\,, \quad |f\rangle = |\mathbf{p}', s'\rangle\,,$$

then, the only terms in the field expansion of (6.14) to contribute are

$$\mathcal{S}_{\text{fi}} = -ig\int d^4x\ \langle f|\,\bar{\psi}^{(-)}(x)\psi^{(+)}(x)\,|i\rangle\, V(\mathbf{x}).$$

Now, we know that

$$\begin{aligned}\psi^{(+)}(x)\,|\mathbf{p}, s\rangle &= \frac{e^{-ipx}u_s(\mathbf{p})}{\sqrt{2EV}}\,|0\rangle\,,\\ \langle\mathbf{p}', s'|\,\bar{\psi}^{(-)}(x) &= \langle 0|\,\frac{e^{ip'x}\bar{u}_{s'}(\mathbf{p}')}{\sqrt{2EV}}.\end{aligned}$$

Hence, we use these in the matrix element to easily obtain

$$\mathcal{S}_{\text{fi}} = -ig \int d^4x \, V(\mathbf{x}) \bar{u}_{s'}(\mathbf{p}') u_s(\mathbf{p}) \frac{e^{ix(p'-p)}}{2V\sqrt{EE'}}. \tag{6.15}$$

The dx^0 integral is easy to compute, as the only term to contribute is the exponential,

$$\int dx^0 \, e^{x^0(E'-E)} = T\delta_{E,E'} = (2\pi)\delta(E-E').$$

The only terms in the dx^i integral are the potential and exponential, which form the integral

$$V(\mathbf{q}) \equiv \int d^3x \, V(\mathbf{x}) e^{-i\mathbf{x}\cdot(\mathbf{p}'-\mathbf{p})}, \tag{6.16}$$

where we have introduced the Fourier transform of the scattering potential, $V(\mathbf{q})$, where $\mathbf{q} \equiv \mathbf{p}' - \mathbf{p}$. So, using these, (6.15) becomes

$$\mathcal{S}_{\text{fi}} = -igV(\mathbf{q}) \bar{u}_{s'}(\mathbf{p}') u_s(\mathbf{p}) T\delta_{E,E'} \frac{1}{2V\sqrt{EE'}}.$$

Now, the transition rate is just

$$\begin{aligned} w_{fi} &= \frac{|\mathcal{S}_{\text{fi}}|^2}{T} \\ &= \frac{g^2}{4V^2EE'} \frac{T^2\delta^2_{E,E'}}{T} |\bar{u}_{s'}(\mathbf{p}') u_s(\mathbf{p})|^2 |V(\mathbf{q})|^2. \end{aligned} \tag{6.17}$$

We now take advantage of

$$T\delta^2_{E,E'} = T\delta_{E,E'} = (2\pi)\delta(E-E'),$$

so that (6.17) becomes

$$w_{fi} = |\mathcal{M}|^2 2\pi \frac{\delta(E-E')}{V^2},$$

where

$$|\mathcal{M}|^2 \equiv \frac{g^2}{4EE'} |\bar{u}_{s'}(\mathbf{p}') u_s(\mathbf{p})|^2 |V(\mathbf{q})|^2.$$

To calculate the cross-section, we need both the flux and density of states. The flux is just

$$f = \rho v = \frac{1}{V} v,$$

where v is the difference in velocities. The density of (final) states is simply

$$\frac{V d^3p'}{(2\pi)^3}.$$

Hence, following the same structure as (6.8), the elemental cross-section is

$$\begin{aligned} d\sigma &= \frac{V}{v}|\mathcal{M}|^2 2\pi \frac{\delta(E-E')}{V^2}\frac{V d^3p'}{(2\pi)^3} \\ &= \frac{1}{v}|\mathcal{M}|^2 2\pi\delta(E-E')\frac{d^3p'}{(2\pi)^3}. \end{aligned} \tag{6.18}$$

Now, let us use a trick to convert d^3p' into an integral over energy. We can easily recall that

$$E^2 = |\mathbf{p}|^2 + m^2,$$

and hence

$$\frac{dE}{dp} = \frac{p}{E},$$

where $p \equiv |\mathbf{p}|$ (i.e. not the 4-vector). So,

$$d^3p' = p'^2 dp' d\Omega' = p'E'dE'd\Omega'.$$

Hence, inserting this into the elemental cross-section (6.18), we see that

$$d\sigma = \frac{1}{v}|\mathcal{M}|^2 2\pi\delta(E-E')\frac{1}{(2\pi)^3}p'E'dE'd\Omega'.$$

If we integrate over E' now, we simply filter out $E = E'$ and $\mathbf{p} = \mathbf{p}'$, so that the differential cross-section is

$$\frac{d\sigma}{d\Omega} = \frac{Ep}{(2\pi)^2 v}|\mathcal{M}|^2, \tag{6.19}$$

where

$$|\mathcal{M}|^2 \equiv \frac{g^2}{4EE'}|\bar{u}_{s'}(\mathbf{p}')u_s(\mathbf{p})|^2|V(\mathbf{q})|^2. \tag{6.20}$$

Basically, one of the points of this, is that the differential cross-section is proportional to the modulus-squared of the Fourier transform of the scattering potential – this is very similar to the Born approximation in non-relativistic quantum mechanics.

If we consider the low energy limit, then the spinors reduce

$$u_s(\mathbf{p}\to 0) = \sqrt{2m}\begin{pmatrix}\chi_s \\ 0\end{pmatrix},$$

so that

$$\bar{u}_{s'}(\mathbf{p}')u_s(\mathbf{p}) = 2m\delta_{ss'}.$$

Thus, in the low energy limit, the spins of the outgoing particles is the same as that of the incoming particles. That is, spin is conserved; in which case the result reduces to the non-relativistic Born approximation obtained from the Schrodinger equation.

We can also get the same result from our previous completely quantised calculation of $e^-\tau^-$ scattering by one-particle exchange; where we make $V(\mathbf{q})$ proportional to the propagator. That is, the scattering potential is proportional to the Fourier transform of the propagator of the exchanged particle.

7 Feynman Rules

The Feynman rules can be used to construct the invariant Feynman amplitude $\mathcal{M}_{\text{fi}}$ for a given process. These only apply up to second order perturbation theory, and do not include any QED, weak or QCD interactions. When drawing Feynman graphs, we use the convention that the arrow only denotes whether the "particle" described is a particle or anti-particle. We also use the convention that a solid line denotes a fermion, and a dashed line a boson.

All factors are pretty self-explanatory; but it is worth noting that m denotes the mass of the exchange boson, and q the 4-momentum carried by the exchange boson. One write a vertex factor for every vertex in the diagram.

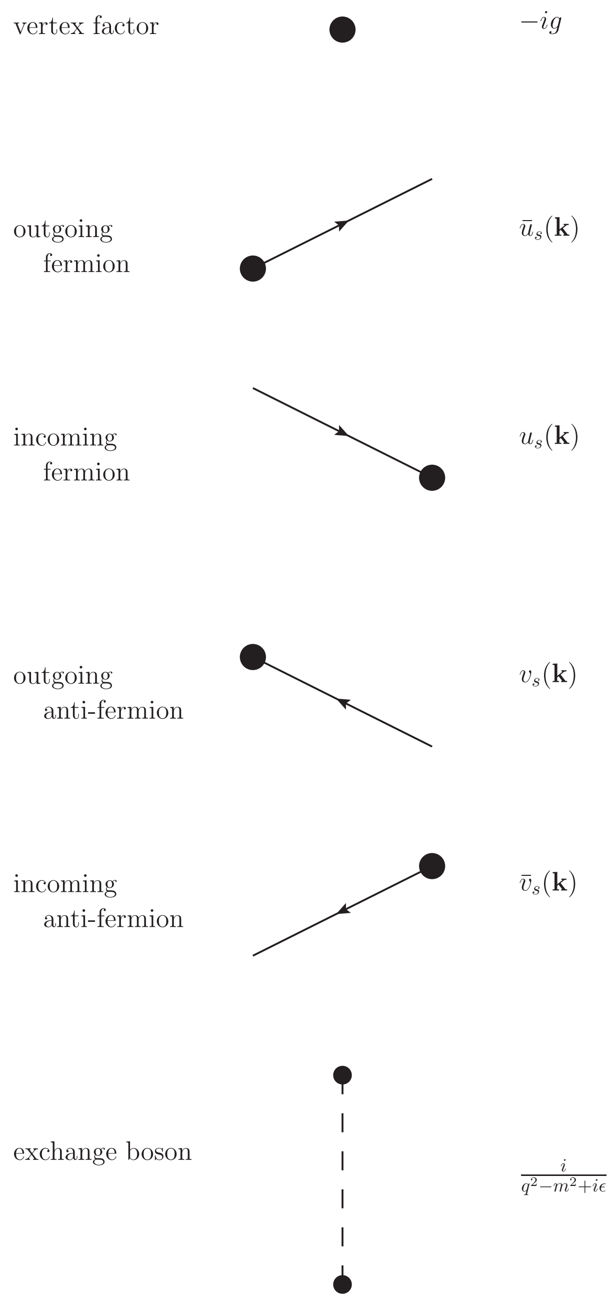


Figure 7.1: Summary of the Feynman rules we have derived.