Chapter 2

QED: Quantum Electrodynamics

2.1 Negative-Energy States: Antiparticles

2.1.1 Setting the Stage: Non-Relativistic Quantum Mechanics

In non-relativistic Quantum Mechanics, it was seen that (in the "standard" position representation) essentially everything can be derived by the substitution

$$E \rightarrow i\frac{\partial}{\partial t} \tag{2.1}$$

$$\vec{p} \rightarrow -i\vec{\nabla} \tag{2.2}$$

(remember that we have set $\hbar = 1$). This substitution directly converts the classical Hamiltonian

$$H = \frac{\vec{p}^2}{2m} + V(\vec{x})$$

into the Schrödinger equation

$$i\frac{\partial}{\partial t}\psi(\vec{x}) = \left(-\frac{\vec{\nabla}^2}{2m} + V(x)\right)\psi(\vec{x})$$

acting on the wave function $\psi(\vec{x})$.

Once we have found a wave function $\psi(\vec{x})$ satisfying the Schrödinger equation, we can also take the complex conjugate expression:

$$-i\frac{\partial}{\partial t}\psi^*(\vec{x}) = \left(-\frac{\vec{\nabla}^2}{2m} + V(x)\right)\psi^*(\vec{x})$$

We then multiply the original Schrödinger equation by $\psi^*(\vec{x})$, and its conjugate by $\psi(\vec{x})$. Subtracting the two yields

$$i\left(\psi^*(\vec{x})\frac{\partial}{\partial t}\psi(\vec{x}) + \psi(\vec{x})\frac{\partial}{\partial t}\psi^*\right) = -\frac{1}{2m}\left(\psi^*(\vec{x})\vec{\nabla}^2\psi(\vec{x}) - \psi(\vec{x})\vec{\nabla}^2\psi^*(\vec{x})\right).$$

It is easily seen that this can be written alternatively as

$$i\frac{\partial}{\partial t}|\psi(\vec{x})|^2 = -\frac{1}{2m}\vec{\nabla}\cdot\left(\psi^*(\vec{x})\vec{\nabla}\psi(\vec{x}) - \psi(\vec{x})\vec{\nabla}\psi^*(\vec{x})\right).$$

Thus, this leads us to the *continuity equation*

$$\frac{\partial}{\partial t}\rho(\vec{x}) + \vec{\nabla}\cdot\vec{j}(\vec{x}) = 0,$$

with

$$\begin{aligned}\rho(\vec{x}) &= |\psi(\vec{x})|^2 \quad \text{and} \\ \vec{j}(\vec{x}) &= \frac{-i}{2m}\left(\psi^*(\vec{x})\vec{\nabla}\psi(\vec{x}) - \psi(\vec{x})\vec{\nabla}\psi^*(\vec{x})\right)\end{aligned}$$

The quantity $\rho(\vec{x})$ occurring in this equation is *positive definite*, making the interpretation of $|\psi(\vec{x})|^2$ as the probability density of finding a particle at the position $\vec{x}$ a proper one.

2.1.2 Translation to the Relativistic Case

The approach in the case of *relativistic* Quantum Mechanics is exactly the same; however, this time it must be applied to the "Hamiltonian" of special relativity. Restricting ourselves to free particles, $V(x) = 0$, the basic classical equation is then

$$p_\mu p^\mu = m^2 \quad \text{or} \quad E^2 = \vec{p}^2 + m^2. \tag{2.3}$$

When we again make the substitutions of Eqn. 2.2, and make the resulting equation act on a wave function $\phi(x)$ (this notation combines the spatial and temporal dependence), the result is

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

or, in explicitly covariant form:

$$(\partial_\mu\partial^\mu + m^2)\phi(x) \equiv (\Box + m^2)\phi(x) = 0. \tag{2.4}$$

This is the *Klein-Gordon equation*.

Unsurprisingly, for our case of free particles, this equation is easily solved to yield plane waves just like in the non-relativistic case:

$$\phi(x) = Ne^{-ip\cdot x} = Ne^{-i(Et - \vec{p}\cdot\vec{x})}, \tag{2.5}$$

with N an *a priori* arbitrary normalization constant, and the four-momentum components E and $\vec{p}$ satisfying our original classical Eqn. 2.3.

But here we are in trouble! For the solution to Eqn. 2.3 is

$$E = \pm\sqrt{\vec{p}^2 + m^2}.$$

While the solution with the $+$ sign gives us a "standard" picture, the solution with the $-$ sign cannot be ignored. As a consequence, the system has no ground state (it is unbounded from below), and hence no meaningful physical interpretation seems possible.

To make things worse, also the continuity equation becomes problematic. As in the non-relativistic case, it is obtained by taking also the complex conjugate of the Klein-Gordon equation and multiplying it with $\phi(x)$, and combining it with the original equation multiplied with $\phi^*(x)$. However, due to the fact that the Klein-Gordon equation involves a second order rather than a first order time derivative, this time we have to *subtract* the two. The result is

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

which can again be considered as a continuity equation, but with

$$\begin{aligned}\rho(x) &= i(\phi^*(x)\frac{\partial\phi(x)}{\partial t}-\phi(x)\frac{\partial\phi^*(x)}{\partial t}),\\ \vec{j}(x) &= -i(\phi^*(x)\vec{\nabla}\phi(x)-\phi(x)\vec{\nabla}\phi^*(x)).\end{aligned} \tag{2.6}$$

This can again be cast into explicitly Lorentz-covariant form:

$$\partial_\mu j^\mu(x)=0, \quad \text{with} \quad j^\mu(x)=i(\phi^*(x)\partial^\mu\phi(x)-\phi(x)\partial^\mu\phi^*(x)). \tag{2.7}$$

When we now substitute the free-particle solution of Eqn. 2.5 in Eqn. 2.6, we find that

$$j^\mu(x)=2p^\mu|N|^2. \tag{2.8}$$

In particular, we have $\rho(x)=2E|N|^2$. So in the case of a negative-energy solution, we also find that $\rho(x)$ becomes negative, *i.e.*, it can no longer be interpreted as a probability density.

Finally, there is another problem with negative-energy solutions. Consider some localized spatial wavefunction at some time t. It is then straightforward to determine its Fourier spectrum, and in general it will be seen that this will contain both positive- and negative-energy components, which will have opposite time evolutions. Constructing the norm of the wavefunction would then contain oscillating terms; the corresponding "zitterbewegung" is not observed in reality. (This issue is discussed in more detail in Section 2-2-2 of Ref. [1].)

2.1.3 Field-theoretical Interpretation

A proper interpretation can only be given in the context of Quantum Field Theory. In that context, ϕ is a *field* rather than a wave function, and its plane-wave expansion leads to particle creation operators for the positive energies combined with antiparticle annihilation operators for the negative energies:

$$\phi(x)=\int\frac{\mathrm{d}^3p}{(2\pi)^3}\frac{1}{2E_{\vec{p}}}\left(a(\vec{p})e^{-ip\cdot x}+b^\dagger(\vec{p})e^{ip\cdot x}\right)$$

The case of a real-valued classical field (we will discuss classical fields in a bit more detail in Section 2.1.4) then translates into a hermitian quantum field, *i.e.*, with $b(\vec{p}) = a(\vec{p})$. So in terms of plane-wave solutions, the action of the field is either to *create* a particle with four-momentum p^μ or to annihilate one with four-momentum p^μ (where it can be shown, although doing so is outside the scope of this course, that the positive-energy solution is associated with the *annihilation* of a particle, while the negative-energy solution is associated with the *creation* of a particle). Note that the four-momentum here is in both cases the *physical* (positive-energy) four-momentum.

However, here we have to watch the other desired properties of this field. Consider the case of a field representing a *charged* particle like the electron. The action of the field must be to change the charge by the *same* one unit, irrespective of whether creation or annihilation is concerned (motivating this requirement further would bring us too far in the realm of Quantum Field Theory; a construction can *e.g.* be found in the book by Peskin and Schroeder [2]). This means that in this case, the equality $b(\vec{p}) = a(\vec{p})$ cannot hold anymore (and hence also that ϕ cannot be a hermitian field). This can be achieved by making $b^\dagger(p)$ represent the creation of an *anti*-particle with four-momentum p^μ; we should therefore expect the existence of a particle much like the electron, but with opposite properties such of its charge. This particle is called the positron.

Of course, one would hope for experimental evidence of the existence of the positron. It was first observed in 1932, in a cloud chamber exposed to cosmic rays (see Fig. 2.1). Its discovery earned Anderson [3] the 1936 Nobel Prize. (The discovery followed the prediction of the positron by Dirac by only a year. Dirac used a different interpretation of negative-energy states, though, which is not appropriate for the description of bosons.)

So what about the continuity equation, and the fact that there doesn't appear to be a conserved quantity (*i.e.*, one occurring in a continuity equation) that can be associated with a probability density? The fact of the matter is that the (conserved) probability density is a concept that is useful in non-relativistic quantum mechanics (non-conservation would correspond to the creation or disappearance of particles). However, in a relativistic context, *it is perfectly acceptable for (anti-)particles to be created or annihilated* (and the operator nature of quantum fields allows to describe such processes). So it doesn't make sense to ask for a conserved probability density.

2.1.4 Principle of Least Action and Euler-Lagrange Equations; Noether Theorem

Accepting that we need a field-theoretical interpretation (per Sect. 2.1.3), we can now also use a different starting point for our computations than the Klein-Gordon equation. Going back to a classical single-particle system of a single degree of freedom $q(t)$, we can express the *action* S as $S = \int_{t_0}^{t_1} \mathrm{d}t L(q, \dot{q})$, where L represents the Lagrangian. Demanding that S be stationary under arbitrary but small changes of $q(t)$ at each t results in the requirement

$$\delta S = \int_{t_0}^{t_1} \mathrm{d}t \left(\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} \right) = 0.$$

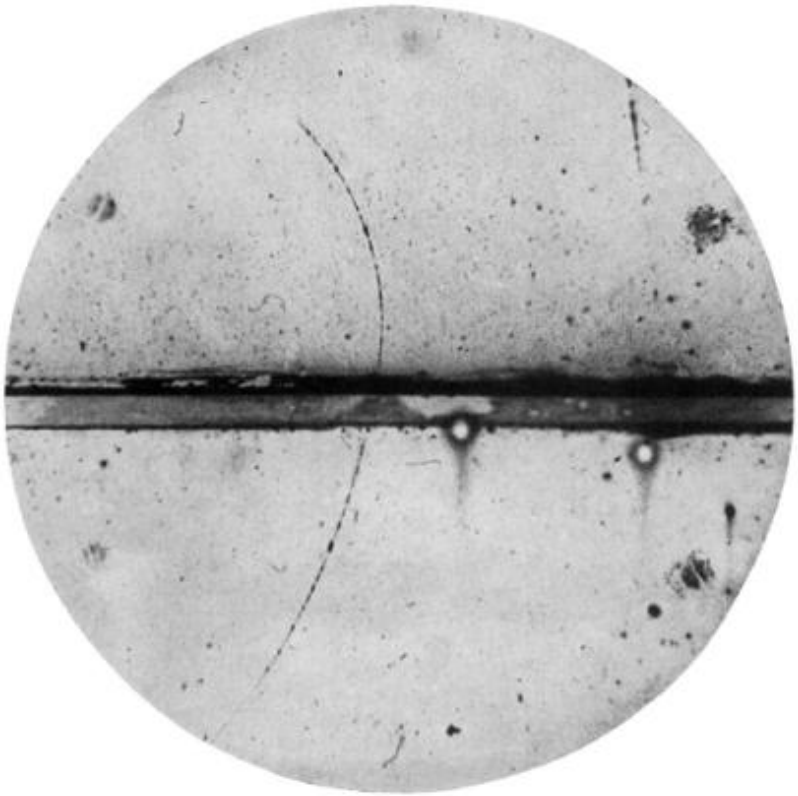

FIG. 1. A 63 million volt positron ($H\rho = 2.1 \times 10^5$ gauss-cm) passing through a 6 mm lead plate and emerging as a 23 million volt positron ($H\rho = 7.5 \times 10^4$ gauss-cm). The length of this latter path is at least ten times greater than the possible length of a proton path of this curvature.

Figure 2.1: Photograph made of a positron bent in a magnetic field and traversing (and losing energy in) a Pb plate. The positron hypothesis follows from (1) the *sign* of the curvature, indicating a positively charged particle; and (2) the track length after having traversed the plate and before being stopped, indicating a particle much lighter than a proton.

Interchanging the order of the time derivative and the δ operation and carrying out an integration by parts then results in the condition

$$\int_{t_0}^{t_1} \mathrm{d}t \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}} \right) \right) \delta q = 0.$$

If this equality is to hold for arbitrary $\delta q(t)$, then we immediately arrive at the *Euler-Lagrange* equation

$$\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}} \right) = 0.$$

In a field-theoretical setting, things work in much the same way. The essential difference is that the Lagrangian L is obtained as the spatial integral of the *Lagrange density* $\mathscr{L}(\phi(x), \partial_\mu \phi(x))$, where $\partial_\mu \phi \equiv \frac{\partial}{\partial(\frac{\partial \phi}{\partial x^\mu})}$ refers to the time as well as spatial derivatives of ϕ. The action therefore

becomes a four-dimensional integral – convenient since this allows us to express it in a covariant form. The arbitrary changes are then in the field $\phi(x)$, and the principle of least action becomes

$$\delta S = \int \mathrm{d}^4x \left(\frac{\partial \mathscr{L}}{\partial \phi} \delta\phi(x) + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)} \delta\partial_\mu \phi(x) \right) = 0. \tag{2.9}$$

The same manipulations as for the above single degree of freedom then lead to the Euler-Lagrange equation for the field:

$$\frac{\partial \mathscr{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathscr{L}}{\partial_\mu \phi} \right) = 0. \tag{2.10}$$

We will make use of this equation, as well as of properties of the Lagrange density, later in this and in other chapters. Note that in these lecture notes we will follow the common particle physicists' sloppiness and simply call $\mathscr{L}$ the Lagrangian. For now, suffice it to say that the Klein-Gordon equation can be recovered from the following choice of Lagrangian:

$$\mathscr{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2. \tag{2.11}$$

The *Noether theorem* is related to so-called *internal symmetries*, which we will cover later in more detail, but which for now we can illustrate using the relativistic wavefunction ϕ of Section 2.1.2, which we subsequently concluded should really be treated as a quantum field. In the wavefunction picture, Quantum Mechanics dictates that the physics should not depend on any complex phase of ϕ. Now in the field theoretical context, it is quite well possible to posit a real scalar field; however as an alternative we can posit a *complex* scalar field ϕ, and still make the *assumption* that the physics described by the Langrangian indeed does not depend on the phase of ϕ. This is arguably the simplest example of an internal symmetry.

Under an infinitesimal phase change, which we will describe more generally as a *group transformation* (see Appendix B for more details), we can then write the transformation of the field ϕ as

$$\phi \to \phi' = \phi + i\alpha T \phi,$$

where α is the infinitesimal phase change, and T is the generator of the group transformation. In the case of phase changes, we know the transformation properties:

$$\phi \to \phi' = e^{i\alpha}\phi = \phi + i\alpha\phi, \tag{2.12}$$

so we simply have $T = 1$. We now require again that the action be invariant under this transformation, so we obtain the condition

$$\begin{aligned}
\delta S &= i \int \mathrm{d}^4x \left(\frac{\partial \mathscr{L}}{\partial \phi} \alpha T \phi + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)} \partial_\mu (\alpha T \phi) \right) \\
&= i \int \mathrm{d}^4x \left(\partial_\mu \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)} \right) \alpha T \phi + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)} \partial_\mu (\alpha T \phi) \right) \\
&= i \int \mathrm{d}^4x \partial_\mu \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)} \alpha T \phi \right) = 0.
\end{aligned}$$

Requiring that this equality hold for any α and integration boundaries, we find that

$$\partial_\mu j^\mu = 0, \quad \text{with} \quad j^\mu = i\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)} T\phi\right).$$

This is the essence of the Noether theorem: *every symmetry brings with it a conserved quantity*.

2.2 Perturbation Theory and Electromagnetic Interactions

2.2.1 Perturbation Theory

A theory describing only free particles is not terribly exciting…therefore, let us see how interactions can be incorporated. The aim here is not to be entirely rigourous, but rather to provide a heuristic introduction to the computation of scattering amplitudes that can be understood as a reasonably straightforward extension of (time-dependent) non-relativistic perturbation theory.

Suppose that the Hamiltonian of a system is described by

$$H = H_0 + V(\vec{x},t)$$

and that the system corresponding to the unperturbed Hamiltonian H_0 can be solved exactly,

$$H_0\phi_n = E_n\phi_n \quad \text{with} \quad \int \mathrm{d}^3x \phi_n^*(\vec{x})\phi_m(\vec{x}) = \delta_{nm}.$$

(Here we are assuming that the system leads to a set of *discrete* eigenstates. That limitation does not affect the following argument.) We now want to know the time evolution of a system that at a time t is in the state $\psi(\vec{x})$. To this end, we decompose ψ in terms of the eigenfunctions of the unperturbed Hamiltonian:

$$\psi(\vec{x},t) = \sum_n a_n(t)\phi_n(\vec{x})e^{-iE_nt}.$$

Applying the Schrödinger equation then yields

$$\begin{aligned} i\frac{\partial\psi(\vec{x},t)}{\partial t} &= \sum_n \phi_n(\vec{x})e^{-iE_nt}\left(E_n a_n(t) + \frac{\mathrm{d}a_n(t)}{\mathrm{d}t}\right) \\ &= (H_0 + V(\vec{x},t))\psi = \sum_n (H_0 + V(\vec{x},t))a_n(t)\phi_n(\vec{x})e^{-iE_nt} \\ &= \sum_n (E_n + V(\vec{x},t))a_n(t)\phi_n(\vec{x})e^{-iE_nt} \\ \Rightarrow i\sum_n \frac{\mathrm{d}a_n(t)}{\mathrm{d}t}\phi_n(\vec{x})e^{-iE_nt} &= \sum_n V(\vec{x},t)a_n(t)\phi_n(\vec{x})e^{-iE_nt}. \end{aligned} \tag{2.13}$$

Now assume that the interaction $V(\vec{x},t)$ is switched off for large times $T \to \infty$, so that the decomposition into eigenstates of the unperturbed system is the "proper" thing to do for such large

times. Multiplying Eqn. 2.13 by $\phi_f^*(\vec{x})e^{iE_f t}$ and integrating the result over all space then yields

$$\begin{aligned}\frac{\mathrm{d}a_f(t)}{\mathrm{d}t} &= -i\sum_n a_n(t)e^{-i(E_n-E_f)t}\cdot V_{fn}(t), \quad \text{with}\\ V_{fn}(t) &= \int \mathrm{d}^3x\phi_f^*(\vec{x})V(\vec{x},t)\phi_n(\vec{x})\end{aligned}$$

This is just the well-known *Dyson series* from non-relativistic Quantum Mechanics.

Also the solution of this integro-differential equation proceeds in the same way as in non-relativstic Quantum Mechanics. In addition, suppose that before the interaction is switched on the system is in an eigenstate of the unperturbed Hamiltonian, *i.e.*, $a_n(-T)=\delta_{ni}$. Order by order, we have

$$\begin{aligned}a_f(t) &= \delta_{fi}\\ &+ (-i)\int_{-T}^{t}\mathrm{d}t' V_{fi}(t')e^{-i(E_i-E_f)t'}\\ &+ (-i)^2\sum_n\int_{-T}^{t}\mathrm{d}t' V_{fn}(t')e^{-i(E_n-E_f)t'}\\ &\quad\cdot\int_{-T}^{t'}\mathrm{d}t'' V_{ni}(t'')e^{-i(E_i-E_n)t''}\\ &+ \ldots\end{aligned}$$

At this point, we formulate the above equation in a more covariant form by setting

$$\phi_n(x)\equiv\phi_n(\vec{x})e^{-iE_n t}.$$

Retaining only the lowest-order (nontrivial) transition, we then obtain

$$\begin{aligned}a_f(t) &= -i\int_{-T}^{t}\mathrm{d}t'\int\mathrm{d}^3x\left(\phi_f(\vec{x})e^{-iE_f t}\right)^* V(\vec{x},t')\left(\phi_i(\vec{x})e^{-iE_i t}\right)\\ &= -i\int_{-T}^{t}\mathrm{d}t'\int\mathrm{d}^3x\phi_f^*(x)V(x)\phi_i(x).\end{aligned}$$

Finally, considering this quantity far after the interaction, at $t=T$, and letting $T\to\infty$, this leads to the transition amplitude

$$T_{fi}=-i\int\mathrm{d}^4x\phi_f^*(x)V(x)\phi_i(x). \tag{2.14}$$

2.2.2 Covariant Formulation of Classical Electrodynamics

Before proceeding to the implementation in Eqn. 2.14, it is useful to pay some attention to the covariant formulation of classical electrodynamics. The starting point is the Maxwell equations:

$$\vec{\nabla}\cdot\vec{E} = \rho \qquad \text{(Gauss)}, \tag{2.15}$$

$$\vec{\nabla}\times\vec{B} - \frac{\partial\vec{E}}{\partial t} = \vec{j} \qquad \text{(Ampère)}, \tag{2.16}$$

$$\vec{\nabla}\cdot\vec{B} = 0 \qquad \text{(Gauss)}, \tag{2.17}$$

$$\vec{\nabla}\times\vec{E} + \frac{\partial\vec{B}}{\partial t} = 0 \qquad \text{(Faraday)}. \tag{2.18}$$

Eqn. 2.17 indicates that $\vec{B}$ can be written as

$$\vec{B} = \vec{\nabla}\times\vec{A},$$

where $\vec{A}$ is called the *vector potential*. Combining this with Eqn. 2.18, it follows that $\vec{E}$ can be written as

$$\vec{E} = -\vec{\nabla}\Phi - \frac{\partial\vec{A}}{\partial t},$$

with Φ the scalar potential. With this notation, it then follows that Eqn. 2.16 can be written as

$$\begin{aligned}\vec{\nabla}\times\vec{B} - \frac{\partial\vec{E}}{\partial t} &= \left(-\vec{\nabla}^2\vec{A} + \vec{\nabla}\cdot(\vec{\nabla}\cdot\vec{A})\right) + \vec{\nabla}\frac{\partial\Phi}{\partial t} + \frac{\partial^2\vec{A}}{\partial t^2} \\ &= \Box\vec{A} + \vec{\nabla}(\vec{\nabla}\cdot\vec{A} + \frac{\partial\Phi}{\partial t}) = \vec{j}.\end{aligned}$$

Finally, we have

$$\vec{\nabla}\cdot\vec{E} = -\vec{\nabla}^2\Phi - \frac{\partial}{\partial t}\left(\vec{\nabla}\cdot\vec{A}\right) = \rho.$$

When we add and subtract here a term $\frac{\partial^2\Phi}{\partial t^2}$, this last equation can be rewritten as

$$\Box\Phi - \frac{\partial}{\partial t}\left(\vec{\nabla}\cdot\vec{A} + \frac{\partial\Phi}{\partial t}\right) = \rho.$$

The two rewritten inhomogeneous equations now have a very similar form; defining

$$A^\mu = (\Phi, \vec{A}) \quad \text{and} \quad j^\mu = (\rho, \vec{j})$$

allows us to finally put the inhomogeneous equations into a manifestly covariant form:

$$\Box A^\mu - \partial^\mu(\partial_\nu A^\nu) = j^\mu,$$

which can also be written as

$$\partial_\mu F^{\mu\nu} = j^\nu, \quad \text{with} \quad F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{2.19}$$

The quantity $F_{\mu\nu}$ is called the *electromagnetic field tensor*, and it turns out that its elements are just $\vec{E}$ and $\vec{B}$.

(Of course, putting these equations into a nicely covariant-looking form does not guarantee the right -known- behaviour of $\vec{E}$ and $\vec{B}$ under Lorentz transformations. But that can be verified explicitly and turns out to be in good order.)

Even this nice formula can be simplified further. The field tensor $F_{\mu\nu}$ encodes the *physical* information. Therefore, a change in A^μ

$$A^\mu \to A'^\mu = A^\mu + \partial^\mu \chi, \tag{2.20}$$

with χ an arbitrary function, does not affect the physics. This is the *gauge* freedom of electromagnetism.

As a consequence, we can choose χ such that $\partial_\nu A^\nu = 0$: this is called the *Lorentz condition*. So finally

$$\Box A^\mu = j^\mu. \tag{2.21}$$

This choice for A^μ is also called the *Lorentz gauge*. It is to be emphasised again that the choice of gauge does *not* affect the physics of the system (and other choices are indeed used, such as the *Coulomb gauge*, in which $\vec{\nabla}\cdot\vec{A} = 0$).

A last ingredient that will be extremely useful in the following is the fact that the interaction of particles with an electromagnetic field can be described simply by the *minimal substitution*[1]:

$$p^\mu \to p^\mu + eA^\mu. \tag{2.22}$$

The usefulness of this substitution is that we can use it instead of a "proper" field theoretical treatment of gauge symmetries: the so-called *covariant derivative* corresponding to the $U(1)$ symmetry group relevant for this treatment of QED yields precisely the same result.

2.2.3 The covariant derivative, and implications of U(1) symmetry

As discussed in the exercises, the use of the minimal substitution allows for a derivation of the Lorentz force in *classical* electrodynamics. If we are to extend this validity to the realm of (non-relativistic) quantum mechanics, this results in a Schrödinger equation

$$\left(\frac{1}{2m}(-i\vec{\nabla} - q\vec{A})^2 + qV\right)\psi(\vec{x},t) = i\frac{\partial\psi(\vec{x},t)}{\partial t} \tag{2.23}$$

(where we have replaced $-e$ in the minimal substitution with the more general charge q). However, the requirement that the gauge transformation of Eqn. 2.20 should not affect the *physics* (*i.e.*, should leave the *form* of eqn. 2.23 invariant) now has a nontrivial consequence. For it can be shown that this invariance is only achieved if simultaneously with Eqn. 2.20, also the *wavefunction* transforms:

$$\psi(\vec{x},t) \to \psi'(\vec{x},t) = e^{-iq\chi(\vec{x},t)}\psi(\vec{x},t). \tag{2.24}$$

[1] The derivation of this property is lengthy and we will not venture into it here. More details can be found *e.g.* in Jackson [4], Chapter 12. Also one of the exercises offers a partial justification.

Although the above is done in the framework of non-relativistic quantum mechanics, exactly the same conclusion (Eqn. 2.24) holds in the relativistic case. In conclusion, we end up with a space- and time-dependent phase transformation of the wavefunction, which does not affect any physics. In group theoretical terms, the $U(1)$ symmetry group can be identified exactly with all possible phase transformations – hence the statement that QED implements a $U(1)$ symmetry.

But having drawn this conclusion, matters can in fact be turned around: let us suppose that we *require* that Eqn. 2.24 does not affect any physics. Then it can be shown that the quantum mechanical analogue of Eqn. 2.22,

$$i\partial_\mu \to iD_\mu \equiv i\partial_\mu - qA_\mu, \tag{2.25}$$

precisely achieves this. The quantity D_μ is called the *covariant derivative*.

Note that this phase change looks a lot like the one encountered in Eqn. 2.12. An essential difference is that rather than merely requiring invariance under *global* phase changes, we now impose this requirement even for *local* (*i.e.*, space and time dependent) phase changes. Another important difference, although we will not prove it here, is that we now require not merely the action to be invariant under the transformation, but also the Lagrangian itself!

Of course, all of the above hinges on the known properties of QED. However, it turns out that the *gauge principle* (starting here with the assumed phase transformation property of the wavefunction – or field – and constructing the appropriate covariant derivative, which then ultimately describes the interaction of charged particles with the electromagnetic field) is very powerful. The same principle will be used later to describe the strong and weak interactions.

Finally, note that while Eqn. 2.20 does not depend on the charge q of the fermion involved, the covariant derivative and the phase transformation so. This means that we can use the same principle (and with the same electromagnetic field!) for particles of different charge. In group theoretical terms, this means that different *representations* of the underlying phase symmetry are possible. This is a fact that will be exploited later on.

2.2.4 Transition Amplitudes

We now have all the required ingredients in hand to proceed further. In the Klein-Gordon equation, we make the minimal substitution of Eqn. 2.22; the resulting equation can be recast as

$$(\Box + m^2)\psi = -V\psi, \tag{2.26}$$

with the "potential" V given by

$$V\psi = -ie(\partial_\mu A^\mu + A^\mu \partial_\mu)\psi - e^2 A^2 \psi$$

(note the operator character of the derivative: it acts on ψ as well as on A). We will neglect the last term in this equation, on account of the fact that e is small. Retaining only the first two terms, we then have

$$\begin{aligned} T_{fi} &= -i\int \mathrm{d}^4x \phi_f^*(x) V(x) \phi_i(x) \\ &= i\int \mathrm{d}^4x \phi_f^*(x)(ie)(A^\mu \partial_\mu + \partial_\mu A^\mu)\phi_i(x). \end{aligned}$$

The last term is amenable to integration by parts, and neglecting the resulting surface integral the result becomes

$$T_{fi} = -i \int \mathrm{d}^4x j^{\mu}_{fi}(x) A_{\mu}(x) \quad \text{with} \quad j^{\mu}_{fi}(x) = -ie\left(\phi^*_f(x)\partial^{\mu}\phi_i(x) - (\partial^{\mu}\phi^*_f(x))\phi_i(x)\right). \quad (2.27)$$

Note that the quantity $j^{\mu}_{fi}(x)$ looks almost exactly like the quantity $j^{\mu}(x)$ in Eqn. 2.7. There is however a difference in that $j^{\mu}_{fi}(x)$ involves two *different* wavefunctions, those of both the initial and final states. The proper interpretation of $j^{\mu}_{fi}(x)$ is that of the current involved in the *interaction of a microscopic particle*. This is relevant in that the absorption or emission of a photon (we'll see later that this picture is appropriate) may affect the particle noticeably.

Eqn. 2.27 is appropriate for the description of the interaction of a particle with a general electromagnetic field. However, this is not the situation typically of interest in particle physics. Rather, our interest is in scattering particles off each other, *i.e.*, in electromagnetic fields caused by *other particles*: the field satisfies

$$\Box A^{\mu} = j^{\mu(2)}_{fi} \quad (2.28)$$

relating it to the current of the other particle (which we will also assume to be an electron).

We will also restrict the further discussion to plane-wave initial and final states (as appropriate for our discussion of scattering experiments where long before and after the scattering process, the participating particles can be considered as free particles). In this case, the current $j^{\mu(2)}_{fi}$ takes on the simple form

$$j^{\mu(2)}_{fi}(x) = -e|N|^2\left(p^{(2)}_i + p^{(2)}_f\right)^{\mu} e^{-i(p^{(2)}_i - p^{(2)}_f)\cdot x},$$

and it is not hard to see that Eqn. 2.28 is satisfied by

$$A^{\mu}(x) = -\frac{j^{\mu(2)}_{fi}(x)}{q^2} \quad \text{with} \quad q^{\mu} = (p^{(2)}_i - p^{(2)}_f)^{\mu}. \quad (2.29)$$

Therefore, the final transition amplitude is given by

$$\begin{aligned} T_{fi} &= -i\int \mathrm{d}^4x j^{\mu(1)}_{fi}(x)\frac{-g_{\mu\nu}}{q^2} j^{\nu(2)}_{fi}(x) \\ &= |N|^4 \int \mathrm{d}^4x e^{-i(p^{(1)}_i - p^{(1)}_f + p^{(2)}_i - p^{(2)}_f)\cdot x} \\ &\quad \cdot\left(ie(p^{(1)}_i + p^{(1)}_f)^{\mu}\right)\cdot\frac{-ig_{\mu\nu}}{q^2}\cdot\left(ie(p^{(2)}_i + p^{(2)}_f)^{\nu}\right) \quad (2.30) \\ &= |N|^4(2\pi)^4\delta^4(p^{(1)}_i + p^{(2)}_i - p^{(1)}_f - p^{(2)}_f)\cdot\left(ie(p^{(1)}_i + p^{(1)}_f)^{\mu}\right)\cdot\frac{-ig_{\mu\nu}}{q^2}\cdot\left(ie(p^{(2)}_i + p^{(2)}_f)^{\nu}\right). \end{aligned}$$

A few remarks are in order at this point:

1. For clarity, a label (1) has been attached to the current representing particle 1 (the particle that is scattered by the potential caused by particle 2). However, Eqn. 2.31 is clearly symmetric in the treatment of the two particles under consideration. This is in fact to be expected! For in our -now microscopic- setup, we are scattering two electrons off each other, and there really isn't any physics reason to treat them differently.

2. The factor $(2\pi)^4\delta^4(\ldots)$ arises from the integration over all of spacetime of the plane-wave exponents. Its effect is to impose conservation of four-momentum, as desirable for these scatterings. In fact, this is not at all particular to the process we are considering here, but is instead related to the assumption of asymptotically free states.

3. Implicit in Eqn. 2.31 is the assumption that the normalization N is independent of the momentum. This is in fact correct, but we will not bother with such normalization issues.

Therefore, in general we will be simplifying the discussion of the transition amplitude to that of the so-called *matrix element*, generically denoted by $\mathscr{M}$. Their relation is defined by

$$T_{fi} = -i(2\pi)^4\delta^4(p_i^{(1)} + p_i^{(2)} - \sum_j p_j)N\mathscr{M}, \tag{2.31}$$

where the sum is over all particles in the final state, and N takes care of the above normalization. In this case, $\mathscr{M}$ is given by

$$-i\mathscr{M} = \left(ie(p_i^{(1)} + p_f^{(1)})^\mu\right) \cdot \frac{-ig_{\mu\nu}}{q^2} \cdot \left(ie(p_i^{(2)} + p_f^{(2)})^\nu\right). \tag{2.32}$$

Limitation

The thing that makes the above derivation heuristic is Eqn. 2.26, in which a "potential" term is added to the *equation of motion* for a free particle (and not to the free particle Hamiltonian). Clearly this is not a proper thing to do. Fortunately, it turns out that in a proper quantum field-theoretical context, we can use the actual Hamiltonian for a complex scalar field (which we lack the formalism to construct explicitly), and it can be shown that the expression for the transition amplitude is correct.

2.2.5 Feynman Diagrams and Feynman Rules

The transition amplitude of Eqn. 2.31 is our way to Feynman diagrams. Apart from the delta function and normalization factors, it contains three ingredients:

- two terms originating from the currents involving the two particles (and which are called the *couplings*);

- and one term that represents the electromagnetic field, as per Eqn. 2.29.

In addition, that same equation shows us that the four-momentum q^μ occurring in the term corresponding to the electromagnetic field corresponds precisely to the difference between the initial- and final-state particles, or in other words, their *momentum transfer*. This leads us to a very simple picture, especially given that we are aware of the particle nature of the photon: in this process, *a photon is exchanged* between the two electrons, absorbing four-momentum from one electron and transferring it to the other. The $-g_{\mu\nu}/q^2$ term is called the *photon propagator*.

This picture can in fact be translated easily to a graphical equivalent, as shown in Fig. 2.2, called the *Feynman diagram* corresponding to this amplitude. In it, the exchanged photon is clearly recognizable, as is its coupling to the electrons. The corresponding *Feynman rules* (given

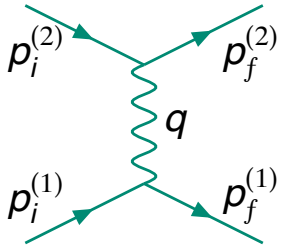

Figure 2.2: Graphical representation of the matrix element of Eqn. 2.32.

without proof – that is rather a topic for a course on Quantum Field Theory) then tell us how to go back from the diagram to the matrix element:

1. Each Feynman diagram consists of external and internal lines (in Fig. 2.2, the electron and photon lines, respectively) and of *vertices*, which are associated with the couplings of particles to each other.

2. Each vertex involves a factor

$$(2\pi)^4\delta^4(\sum_i k_i)\cdot\left(ie(p_i+p_f)^\mu\right)$$

   where the delta function expresses four-momentum conservation *at each vertex* (all the k_i are taken to be incoming; this is a generic feature of all Feynman diagrams) and in the coupling $e(p_i+p_f)^\mu$ the electron four-momenta "follow the arrows", as in Fig. 2.2.

3. Each internal photon (*i.e.*, each photon propagator) is represented by a "wavy" line and corresponds to a term

$$\int \frac{\mathrm{d}^4q}{(2\pi)^4}\frac{-ig_{\mu\nu}}{q^2},$$

   where q^μ is the photon's four-momentum (meaning that each internal four-momentum is integrated over).

4. The result contains a factor $(2\pi)^4$ times a delta function expressing overall four-momentum conservation. This factor is discarded (but of course is to be kept in mind when doing actual computations); the result is equal to $-i\mathscr{M}$.

5. The complete matrix element for a given process (*i.e.*, for given -completely specified- initial and final states) in general corresponds to multiple Feynman diagrams, the individual matrix elements of which have to be summed. (In fact, to obtain the complete matrix element *all* possible Feynman diagrams need to be summed. This is a consequence of the Dyson series: we have restricted ourselves to the computation of the first term in perturbation theory, and ideally we would like to compute higher-order terms as well.)

It may be noted that the photon's four-momentum q^μ does *not* in general satisfy $q^2 = 0$. On the one hand this is good (as otherwise the transition amplitude would diverge), but on the other hand the question is how this relates to the masslessness of the photon!

The resolution of this issue rests on the fact that the interaction (*i.e.*, the exchange of the photon) takes place on very short timescales. On such timescales, the Heisenberg uncertainty principle dictates that a photon of (squared) "mass" q^2 may exist for an amount of time $\sim 1/\sqrt{|q^2|}$. Such photons are called *virtual* (since they cannot propagate over macroscopic distances) or *off-shell*. In fact we will also encounter many examples of other off-shell particles being exchanged in interactions.

On a more practical note, while the process under consideration here is the elastic scattering of two *particles*, we could have equally well chosen to consider the scattering of a particle and an *anti-particle* instead (*e.g.*, electron-positron scattering). Now recall that in the Feynman-Stückelberg approach, anti-particles are (loosely speaking) considered as particles moving backward in time, and are associated with the negative-energy solutions. In Feynman diagrams, this difference between particles and anti-particles is expressed by *reversing the direction of the arrows*; so for anti-particles the direction of the arrows is always opposite the physical propagation in time. As a corollary, the conservation of (electrical) current implies that the arrows in a single "current line" (the external and internal lines featuring electrons and/or positrons) must always be in the same direction along the line.

Returning now to our computation of electron-electron scattering, it is not too hard to realize that the above Feynman rules give rise to *another* diagram, even at the lowest order in perturbation theory. Both of them are shown in Fig. 2.3. The second diagram arises because we are dealing with indistinguishable particles (this is why it is not immediately obvious that we did not find it straight from our original treatment of this process, in which we started out not making any assumptions as to the nature of the "other" particle). This process is called Møller scattering.

2.3 The Dirac Equation

2.3.1 Dirac's Attempt

As mentioned in Sect. 2.1, in a quantum mechanical setting there are two problems with the Klein-Gordon equation (perceived problems, as they are addressed by a proper field-theoretic treatment):

1. it involves a *second order* time derivative, giving rise to negative-energy states and a system that has no ground state;

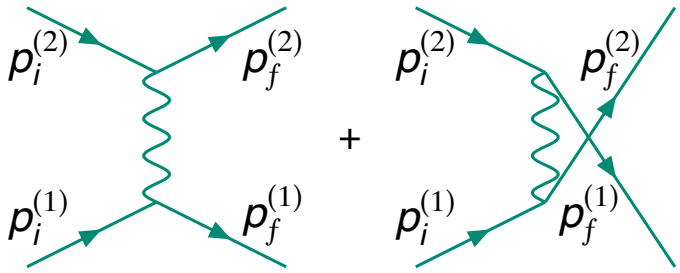

Figure 2.3: Diagrams contributing (in lowest order) to the Møller scattering process $e^- + e^- \to e^- + e^-$.

2. and these same negative-energy states lead to a continuity equation that is not amenable to a probability interpretation.

Even if in the context of field theory there is no direct problem, Dirac's attempt to address the above "issues" by constructing an equivalent equation that only involves a *first order* time derivative has proven to be of great importance, as it leads us to a proper description of spin-1/2 particles (the discussion above has not mentioned spin at all, but of course we know that electrons are spin-1/2 particles).

The Dirac equation for free spin-1/2 particles (like the Schrödinger equation, in the position representation) is

$$(i\partial_\mu \gamma^\mu - m)\psi(x) = 0, \tag{2.33}$$

with the quantities γ^μ satisfying the anticommutation relation

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}. \tag{2.34}$$

Clearly this equation cannot be satisfied by ordinary numbers, and therefore a (four-dimensional) *matrix representation* is used. Multiple conventions are possible, but the one most often used (the *Björken and Drell* convention) is

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \qquad \gamma^j = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \tag{2.35}$$

where the σ^i represent the Pauli matrices (so also the right-hand side of Eqn. 2.34 formally needs to be multiplied by the 4×4 unit matrix $\mathbb{1}$). Also $\psi(x)$ cannot be a "simple" scalar-valued wavefunction anymore; instead it becomes a column vector of dimension four, called a *bi-spinor*.

(That the Dirac equation is sufficient can be seen by multiplying it from the left by $(i\partial_\nu\gamma^\nu + m)$. This simply yields the Klein-Gordon equation, so we have proven that it is a sufficient condition for the Dirac equation to be satisfied.)

Clearly, given that we are again considering free particles here, it is to be expected that the solutions to the Dirac equation are plane waves. Now in particular, let us consider those plane-wave solutions corresponding to a particle at rest. Given the 2×2 block form of the gamma

matrices, write

$$\psi = \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix}.$$

In this case, the Dirac equation can be rewritten as

$$\begin{aligned} (i\frac{\partial}{\partial t} - m)\psi_A &= 0, \\ (-i\frac{\partial}{\partial t} - m)\psi_B &= 0. \end{aligned}$$

Clearly the solution $\psi_A \sim e^{-imt}$ corresponds to a "normal" positive-energy solution; however, $\psi_B \sim e^{+imt}$ again corresponds to a negative-energy solution. By now, however, aware of the antiparticle interpretation of $E < 0$ states, we proceed undeterred.

Like in the case of the Klein-Gordon equation, we take the hermitian conjugate of the Dirac equation. The result is

$$-i\partial_\mu \psi^\dagger(x)\gamma^{\mu\dagger} - m\psi^\dagger(x) = 0.$$

We manipulate this by noting, from Eqn. 2.35, that $\gamma^{0\dagger} = \gamma^0$ and $\gamma^{i\dagger} = -\gamma^i$ (since the Pauli matrices are hermitian). Using Eqn. 2.34, this can be written concisely as $\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0$. So we have

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

Next, we multiply the whole equation by -γ^0 from the right; the result is

$$i\partial_\mu \overline{\psi}(x)\gamma^\mu + m\overline{\psi}(x) = 0, \quad \text{with} \quad \overline{\psi}(x) \equiv \psi^\dagger(x)\gamma^0.$$

With this conjugate equation in hand, we proceed to construct again a continuity equation. This is easily done by multiplying Eqn. 2.33 by $\overline{\psi}(x)$ from the left, the conjugate equation by $\psi(x)$ from the right, and summing the result. This yields

$$i\partial_\mu(\overline{\psi}(x)\gamma^\mu\psi(x)) = 0.$$

Considering in particular the time component, we therefore find that we have

$$\overline{\psi}(x)\gamma^0\psi(x) = \psi^\dagger(x)\psi(x),$$

So we have found a solution where a probability interpretation makes sense! However, again because of the antiparticle interpretation we will not make further attempts in this direction, but instead consider this as a conserved *electric* current:

$$j_\mu = -e\overline{\psi}(x)\gamma_\mu\psi(x). \tag{2.36}$$

2.3.2 Spin-1/2 Particles

The virtue of the Dirac equation is that it allows for a description of spin-1/2 particles. This is perhaps to be expected already simply from the presence of the gamma matrices containing Pauli matrices (which also in non-relativistic Quantum Mechanics are associated with the spin operators for spin-1/2 particles). However, it can also be seen in more detail from considering the general Dirac equation, and again writing it in its 2×2 block form,

$$\psi(x) = \begin{pmatrix} u_A \\ u_B \end{pmatrix} e^{-ip\cdot x},$$

i.e., splitting off the plane-wave piece from the *spinors* u_A and u_B (at this stage we haven't yet specified whether the solution involves positive or negative energies). For nonzero momenta, we obtain *coupled* equations for the spinors:

$$\begin{aligned} (\vec{\sigma}\cdot\vec{p})u_B &= (E-m)u_A, \\ (\vec{\sigma}\cdot\vec{p})u_A &= (E+m)u_B. \end{aligned} \tag{2.37}$$

Restricting ourselves to the positive-energy solution, we can now choose two independent solutions for u_A:

$$u_A^{(1)} = \chi^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad u_A^{(2)} = \chi^{(2)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The second equation in Eqn. 2.37 then yields

$$u_B{}^{(1,2)} = \frac{(\vec{\sigma}\cdot\vec{p})}{E+m} u_A^{(1,2)}.$$

Similarly, in the case of negative-energy solutions, we choose two independent solutions for u_B, $u_B^{(1,2)} = \chi^{(1,2)}$, and find (from the first equation in Eqn. 2.37):

$$u_A{}^{(1,2)} = \frac{-(\vec{\sigma}\cdot\vec{p})}{-E+m} u_B^{(1,2)}.$$

The minus sign has been carried over to the four-momentum components here. The reason is that in this case, the *physical* four-momentum contains an extra minus sign relative to the four-momentum occurring in Eqn. 2.37. (Note that we might as well have started with the independent solutions for u_B in the case of positive-energy solutions, and vice versa. The important point is that in the non-relativistic limit, for positive-energy solutions, $u_A \gg u_B$, while for negative-energy solutions, $u_B \gg u_A$.)

Summarizing, ψ represents *four* independent degrees of freedom, two for $E > 0$ and two for $E < 0$. These two are of course nothing but the two solutions corresponding to different spin states. When doing practical calculations, the four solutions are typically written as

$$u^{(1,2)}(p) = N \begin{pmatrix} \chi^{(1,2)} \\ \frac{(\vec{\sigma}\cdot\vec{p})}{E+m}\chi^{(1,2)} \end{pmatrix}, \quad u^{(3,4)}(p) = N \begin{pmatrix} \frac{-(\vec{\sigma}\cdot\vec{p})}{-E+m}\chi^{(1,2)} \\ \chi^{(1,2)} \end{pmatrix}.$$

In addition, the negative-energy bi-spinors are usually written in terms of the *physical* four-momentum, leading to

$$v^{(1)}(p) \equiv u^{(4)}(-p) \quad \text{and} \quad v^{(2)}(p) \equiv -u^{(3)}(-p).$$

2.3.3 Perturbation Theory

The step from free to *interacting* spin-1/2 particles is made in exactly the same fashion as in the case of spin-0 particles: by means of the minimal substitution (see Sect. 2.2.2). In that case, the Dirac equation is modified to

$$(i\partial_\mu\gamma^\mu - m)\psi(x) = (i\frac{\partial}{\partial t}\gamma^0 + i\vec{\nabla}\cdot\vec{\gamma} - m)\psi(x) = -eA_\mu\gamma^\mu\psi(x). \tag{2.38}$$

The reason for separating the time and spatial components is that we can use this equation to construct explicitly a Hamiltonian suited for spin-1/2 particles. To do so, multiply (from the left) by γ^0; we then have

$$i\frac{\partial}{\partial t}\psi(x) = (-i\vec{\nabla}\cdot\gamma^0\vec{\gamma} + \gamma^0 m)\psi(x) - eA_\mu\gamma^0\gamma^\mu\psi(x),$$

the right-hand side of which nicely has the form $H = H_0 + V$, so that we can identify the term $-eA_\mu\gamma^0\gamma^\mu$ with a perturbing potential V. Inserting this in Eqn. 2.14, we obtain

$$\begin{aligned} T_{fi} &= -i\int \mathrm{d}^4x\,\psi_f^\dagger(x)(-eA_\mu\gamma^0\gamma^\mu)\psi_i(x) \\ &= -i\int \mathrm{d}^4x\, j^\mu_{fi}(x)A_\mu(x), \quad \text{with} \quad j^\mu_{fi}(x) = -e\overline{\psi}_f(x)\gamma^\mu\psi_i(x). \end{aligned}$$

Also here, we restrict ourselves to plane-wave states, and assume that the electromagnetic field is generated by another particle. This implies that we can again insert Eqn. 2.29 – this time of course with a current that is appropriate for spin-1/2 particles. From here, it is not hard to see that also the rest of the computation of the transition amplitude proceeds as for scalar particles.

2.3.4 Feynman Rules for Spin-1/2 Particles

Without further ado, we quote here the Feynman rules appropriate for the computation of matrix elements in QED:

1. The basic "building blocks" of Feynman diagrams are again propagators and vertices.
2. Each photon propagator again corresponds to a factor
$$\int \frac{\mathrm{d}^4q}{(2\pi)^4}\frac{-ig_{\mu\nu}}{q^2}.$$
3. Each fermion propagator corresponds to a factor
$$\int \frac{\mathrm{d}^4q}{(2\pi)^4}\frac{i(\not{q}+m)}{q^2-m^2}.$$
Note that we have introduced here the notation $\not{a} \equiv a_\mu\gamma^\mu$ for *any* a_μ.

4. Each vertex corresponds to a factor

$$(2\pi)^4\delta^4(\sum_i k_i)\cdot ie\gamma^\mu,$$

where all four-momenta are again taken to be towards the vertex.

5. External lines now need to be dealt with more precisely, as the fermions can be labeled by their spins, and we also allow for external photon lines corresponding to specific spin states:

incoming fermion: u	outgoing fermion: $\bar{u}$
incoming antifermion: $\bar{v}$	outgoing antifermion: v
incoming photon: ε^μ	outgoing photon: $\varepsilon^{\mu *}$

6. All appropriate Feynman diagrams should again be summed. A small refinement compared to the case of "scalar QED", however, is that when combining matrix elements that differ only in the exhange of two identical fermions, a relative minus sign must be added. (This is because wavefunctions must be fully antisymmetric under exchange of any two identical fermions.)

7. The overall $(2\pi)^4\delta^4(\ldots)$ is again discarded, and the result is again $-i\mathscr{M}$.

They are shown here mostly for completeness, as we will not attempt to perform complete calculations of Feynman diagrams; nevertheless, it is important to be aware of the differences with the "scalar QED" case.

Polarization states of spin-1 bosons

The polarization vectors ε^μ mentioned in the above merit some further discussion. Let us first discuss the case of *massive* spin-1 bosons. In this case, one can transform to the particle's rest frame, so that the polarization vectors from a non-relativistic treatment are appropriate:

$$\vec{\varepsilon}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{\varepsilon}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{\varepsilon}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

for plane polarization states, and (taking the z axis as our quantisation axis)

$$\vec{\varepsilon}_{\lambda=1} = \frac{-1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \quad \vec{\varepsilon}_{\lambda=-1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}, \quad \vec{\varepsilon}_{\lambda=0} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

for circularly polarized states. These polarization states are orthonormal:

$$\vec{\varepsilon}^*_\lambda \cdot \vec{\varepsilon}_{\lambda'} = \delta_{\lambda\lambda'}.$$

Next, we promote these polarization vectors to proper *four-vectors* and require that they remain orthonormal:

$$\begin{aligned} \varepsilon(p;\lambda)\cdot p &= 0, \\ \varepsilon^*(p;\lambda)\cdot\varepsilon(p;\lambda') &= -\delta_{\lambda\lambda'}. \end{aligned}$$

For a boost *e.g.* along the z axis, the *transverse* polarization states ($\lambda = \pm 1$) do not change under this transformation. However, the $\lambda = 0$ ("longitudinal") polarization state does change. From the orthonormality conditions it is not hard to see that for a momentum $p_\mu = (E,0,0,p)$, a vector

$$\varepsilon_\mu(p;\lambda=0) = \frac{1}{m}(p,0,0,E)$$

is required, where m is the particle mass. (Note: it is far from obvious to see how the polarization vectors transform under *general* Lorentz transformations! Suffice it to say that a proper covariant expression can be found, in the form of the so-called *Pauli-Lubanski* vector.) A final useful property of these polarization vectors is

$$\sum_\lambda \varepsilon^\mu(p;\lambda)\varepsilon^{\nu *}(p;\lambda) = -g^{\mu\nu} + p^\mu p^\nu/m^2.$$

(This can either be verified explicitly, or by realising that the result cannot depend anymore on any *specific* polarization vector, and hence only terms proportional to $g^{\mu\nu}$ and $p^\mu p^\nu$ remain. The orthonormality conditions can then be used to determine the corresponding coefficients.)

Let us now consider the case of *massless* spin-1 bosons. As discussed in Sect. 2.2.2, the QED gauge freedom allows for transformations $A_\mu \to A_\mu - \partial_\mu \chi$, with χ an arbitrary function. Specialising to plane waves

$$A^\mu \propto \varepsilon^\mu e^{-iq\cdot x},$$

these gauge transformations amount to changes of the polarization vectors

$$\varepsilon^\mu \to \varepsilon'^\mu = \varepsilon^\mu + \alpha q^\mu.$$

(Note that this does not violate the orthogonality condition $\varepsilon\cdot q = 0$: after all, for on-shell massless particles we have $q^2 = 0$.) This means that we can in fact *choose* χ such that $\varepsilon^0 = 0$. Given the Lorentz condition, this implies $\vec{\varepsilon}\cdot\vec{q} = 0$. So only the *transverse* polarization states survive (but of course this is well known from classical electrodynamics!).

2.4 The Electron's Magnetic Moment

As a final application of our manipulations involving spin-1/2 particles, consider the interaction of an electron with an external magnetic field. The non-relativistic quantum mechanical treatment of this phenomenon is to posit an interaction term $\vec{\mu}\cdot\vec{B}$ in the total Hamiltonian, leading to the *Zeeman splitting* in the presence of a (weak) static magnetic field. In this term, $\vec{\mu}$ is the

electron's magnetic moment. It is typically expressed in terms of the Bohr magneton $\mu_B \equiv e/2m$ as

$$\vec{\mu} = g\mu_B \vec{S},$$

where $\vec{S}$ is the electron's spin vector. For electrons in a quantum mechanical treatment, we have $\vec{S} = \frac{1}{2}\vec{\sigma}$, $\vec{\sigma}$ denoting the Pauli matrices as usual. In summary, we find a term in the Hamiltonian equal to

$$\frac{1}{2} g\mu_B \vec{\sigma} \cdot \vec{B}. \tag{2.39}$$

The issue is that in a "simple" quantum mechanical context, the *Landé factor* g cannot be computed from first principles. The following calculation shows that QED *does* provide a prediction for g – and a correct one at that!

We start again from the Dirac equation with the interaction with an electromagnetic field added through the minimal substitution, as in Eqn. 2.38. Writing in 2×2 block form, we have (cf. Eqn. 2.37)

$$\begin{aligned} \vec{\sigma} \cdot (\vec{p} + e\vec{A}) u_B &= (E + eA^0 - m) u_A, \\ \vec{\sigma} \cdot (\vec{p} + e\vec{A}) u_A &= (E + eA^0 + m) u_B. \end{aligned}$$

Combining these equations yields

$$\left(\vec{\sigma} \cdot (\vec{p} + e\vec{A})\right)^2 u_A = ((E + eA^0)^2 - m^2) u_A.$$

Next, we simplify the left-hand side of this equation, but *keeping in mind the operator character of $\vec{p}$!* This yields

$$\begin{aligned} \left(\vec{\sigma} \cdot (\vec{p} + e\vec{A})\right)^2 &= \sigma^i \sigma^j \left(p^i p^j + e^2 A^i A^j + e(p^i A^j + A^i p^j)\right) \\ &= (\delta_{ij} + i\varepsilon^{ijk}\sigma^k)\left(p^i p^j + e^2 A^i A^j + e(p^i A^j + A^i p^j)\right) \\ &= p^i p^i + e^2 A^i A^i + e(p^i A^i + A^i p^i) + ie(p^i A^j + A^i p^j)\varepsilon^{ijk}\sigma^k \\ &= (\vec{p} + e\vec{A})^2 + e(\vec{\nabla} \times \vec{A}) \cdot \vec{\sigma} \\ &= (\vec{p} + e\vec{A})^2 + e\vec{\sigma} \cdot \vec{B}. \end{aligned}$$

Here, repeated indices are to be summed over (from 1 to 3). Clearly, this square *almost* trivially reduces to the first term on the one-but-last line. It is precisely the operator nature of $\vec{p}$, $p^i A^j = A^j p^i - i\partial^i A^j$, which leads to the nontrivial additional term.

Next, we consider the right-hand side of the equation for u_A, in the non-relativistic limit. This implies that the kinetic energy and A^0 are small compared to m, so

$$((E + eA^0)^2 - m^2) = ((m + (E + eA^0 - m))^2 - m^2) \approx 2m(E + eA^0 - m).$$

With that approximation and dividing by $2m$, we obtain

$$(E - m) u_A = \left(\frac{(\vec{p} + e\vec{A})^2}{2m} - eA^0 + \frac{e}{2m} \vec{\sigma} \cdot \vec{B} \right) u_A. \tag{2.40}$$

The last term clearly corresponds to the interaction of a magnetic moment with an external magnetic field, with $g = 2$ (by comparison with Eqn. 2.39).

So is the equation $g = 2$ exact? Not quite, in fact. The static external magnetic field is "merely" one form of an electromagnetic field, and as such the interaction that is of important at the diagrammatic level is that of an electron with the photon, *i.e.*, a diagram consisting essentially only of the $ee\gamma$ vertex (this is possible kinematically since the external magnetic field represents *virtual* rather than real photons). But higher-order perturbative corrections, exemplified in Fig. 2.4, need to be applied.

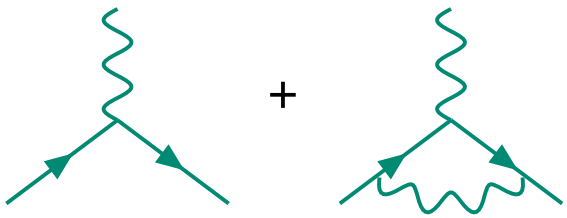

Figure 2.4: Fundamental vertex and "vertex correction" diagram describing the interaction of electrons with electromagnetic fields.

In fact, the electron's *anomalous magnetic moment* $a_\mathrm{e} \equiv (g_\mathrm{e} - 2)/2$ has been computed very accurately:

$$a_\mathrm{e} = \frac{1}{2}\left(\frac{\alpha}{\pi}\right) - 0.328478965\left(\frac{\alpha}{\pi}\right)^2 + 1.1761\left(\frac{\alpha}{\pi}\right)^3 + \ldots$$

It is one of the great achievements of QED that the measured and predicted values of a_e agree with each other, within exceedingly small uncertainties of several parts in 10^9.