

AN INTRODUCTION TO QED & QCD

By Dr G Zanderighi
University of Oxford

Lecture notes by Prof N Evans
University of Southampton

Lecture delivered at the School for Experimental High Energy Physics Students
Somerville College, Oxford, September 2009

Contents

1	Introduction	55
1.1	Relativity Review	56
2	Relativistic Wave Equations	58
2.1	The Klein-Gordon Equation	59
2.2	The Dirac Equation	62
2.3	Solutions to the Dirac Equation	63
2.4	Orthogonality and Completeness.....	65
2.5	Spin	66
2.6	Lorentz Covariance.....	68
2.7	Parity, charge conjugation and time reversal	71
2.8	Bilinear Covariants	74
2.9	Massless (Ultra-relativistic) Fermions	75
3	Quantum Electrodynamics.....	77
3.1	Classical Electromagnetism.....	77
3.2	The Dirac Equation in an Electromagnetic Field.....	79
3.3	$g - 2$ of the Electron.....	80
3.4	Interactions in Perturbation Theory	82
3.5	Internal Fermions and External Photons.....	84
3.6	Summary of Feynman Rules for QED	86
4	Cross Sections and Decay Rates.....	88
4.1	Transition Rate	88
4.2	Decay Rates.....	90
4.3	Cross Sections.....	91
4.4	Mandelstam Variables.....	92
5	Processes in QED and QCD.....	93
5.1	Electron-Muon Scattering	93
5.2	Electron-Electron Scattering	95
5.3	Electron-Positron Annihilation.....	96
5.4	Compton Scattering	98
5.5	QCD Processes	99
6	Introduction to Renormalization	101
6.1	Ultraviolet (UV) Singularities	101
6.2	Infrared (IR) Singularities	102
6.3	Renormalization.....	102
6.4	Regularization	105
7	QED as a Field Theory	107
7.1	Quantizing the Dirac Field	107
7.2	Quantizing the Electromagnetic Field	109
	Acknowledgements	112
	References.....	112
	Pre-School Problems	113
	Rotations, Angular Momentum and the Pauli Matrices.....	113
	Four Vectors	113
	Probability Density and Current Density	114

1 Introduction

The aim of this course is to teach you how to calculate transition amplitudes, cross sections and decay rates, for elementary particles in the highly successful theories of Quantum Electrodynamics (QED) and Quantum Chromodynamics (QCD). Most of our work will be in understanding how to compute in QED. By the end of the course you should be able to go from a Feynman diagram, such as the one for $e^-e^- \rightarrow \mu^-\mu^-$ in figure 5, to a number for the cross section. To do this we will have to learn how to cope with *relativistic*, *quantum*, particles and *anti*-particles that carry *spin*. In fact all these properties of particles will emerge rather neatly from thinking about relativistic quantum mechanics. The rules for calculating in QCD are slightly more complicated than in QED, as we will briefly review, however, the basic techniques for the calculation are very similar.

We have a lot to cover so will necessarily have to take some short cuts. Our main fudge will be to work in relativistic quantum mechanics rather than the full Quantum Field Theory (QFT) (sometimes referred to as ‘second quantization’). We will be in good company though since we will largely follow methods from Feynman’s papers and text books such as Halzen and Martin. In quantum mechanics a *classical* wave is used to describe a particle whose motion is subject to the Uncertainty Principle. In a full QFT the wave’s motion itself is subject to the Uncertainty Principle too - the quanta of that field are what we then refer to as particles. Luckily at lowest order in a perturbation theory calculation one neglects the quantum nature of the field and the two theories give the same answer. At higher orders the quantum nature of the field gives rise to virtual pair creation of particles - in the quantum mechanics version of the story these are included in a more ad hoc fashion as we will see. Luckily the simultaneous QFT course will give you a good grounding in more precise methodologies.

Thus our starting point will be ordinary Quantum Mechanics and our first goal (section 2) will be to write down a ‘relativistic version’ of Quantum Mechanics. This will lead us to look at relativistic wave equations, in particular the Dirac equation, which describes particles with spin 1/2. We will also develop a wave equation for photons and look at how they couple to our fermions (section 3) - this is the core of QED. A perturbation theory analysis will result in quantum mechanical probability amplitudes for particular processes. After this, we will work out how to go from the probability amplitudes to cross sections and decay rates (section 4). We will look at some examples of tree level QED processes. Here you will get hands-on experience of calculating transition amplitudes and getting from them to cross sections (section 5). We will restrict ourselves to calculations at *tree level* but, at the end of the course (section 6), we will also take a first look at higher order *loop* effects, which, amongst other things, are responsible for the running of the couplings. For QCD, this running means that the coupling appears weaker when measured at higher energy scales and is the reason why we can sometimes do perturbative QCD calculations. However, in higher order calculations divergences appear and we have to understand — at least in principle — how these divergences can be removed.

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

1.1 Relativity Review

An event in a reference frame S is described by the four coordinates of a four-vector (in units where $c = 1$)

$$x^\mu = (t, \vec{x}), \quad (1.1)$$

where the Greek index $\mu \in \{0, 1, 2, 3\}$. These coordinates are *reference frame dependent*. The coordinates in another frame S' are given by x'^μ , related to those in S by a Lorentz Transformation (LT)

$$x^\mu \rightarrow x'^\mu = \Lambda^\mu{}_\nu x^\nu, \quad (1.2)$$

where summation over repeated indices is understood. This transformation identifies x^μ as a *contravariant* 4-vector (often referred to simply as a *vector*). A familiar example of a LT is a boost along the z -axis, for which

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix}, \quad (1.3)$$

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

The “length” of the 4-vector $(t^2 - |\vec{x}|^2)$ is invariant to LTs. In general we define the Minkowski scalar product of two 4-vectors x and y as

$$x \cdot y \equiv x^\mu y^\nu g_{\mu\nu} \equiv x^\mu y_\mu, \quad (1.4)$$

where the metric

$$g^{\mu\nu} = g_{\mu\nu} = \text{diag}(1, -1, -1, -1), \quad g^{\mu\lambda} g_{\lambda\nu} = g^\mu{}_\nu = \delta^\mu_\nu = \begin{cases} 1 & \text{if } \mu = \nu \\ 0 & \text{if } \mu \neq \nu \end{cases}, \quad (1.5)$$

has been introduced. The last step in eq. (1.4) is the definition of a *covariant* 4-vector (sometimes referred to as a *co-vector*),

$$x_\mu \equiv g_{\mu\nu} x^\nu. \quad (1.6)$$

This transforms under a LT according to

$$x_\mu \rightarrow x'_\mu = \Lambda_\mu{}^\nu x_\nu. \quad (1.7)$$

Note that the invariance of the scalar product implies

$$\Lambda^T g \Lambda = g \Rightarrow g \Lambda^T g = \Lambda^{-1}, \quad (1.8)$$

i.e. a generalization of the orthogonality property of the rotation matrix $R^T = R^{-1}$.

▷ Exercise 1.1

Show eq. (1.8), starting from the invariance of the scalar product.

To formulate a coherent relativistic theory of dynamics we define kinematic variables that are also 4-vectors (i.e. transform according to eq. (1.2)). For example, we define a 4-velocity

$$u^\mu = \frac{dx^\mu}{d\tau}, \quad (1.9)$$

where τ is the *proper time* measured by a clock moving with the particle. Everyone will agree what the clock says at a particular event so this measure of time is Lorentz invariant and u^μ transforms as x^μ . Note

$$u^\mu = \frac{dt}{d\tau} \frac{dx^\mu}{dt} = \gamma(1, \vec{v}) \quad (1.10)$$

and has invariant length

$$u^\mu u_\mu = \gamma^2(1^2 - |\vec{v}|^2) = 1. \quad (1.11)$$

Similarly 4-momentum provides a relativistic definition of energy and momentum

$$p^\mu = mu^\mu \equiv (E, \vec{p}). \quad (1.12)$$

The invariant length provides the crucial relation

$$p^\mu p_\mu = E^2 - |\vec{p}|^2 = m^2. \quad (1.13)$$

▷ **Exercise 1.2**

Check that $dt/d\tau = \gamma$ and that our relativistic definitions of E and \vec{p} make sense in the non-relativistic limit.

The differentiation operator,

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial t}, \vec{\nabla} \right), \quad \partial_\mu x^\nu = \delta_\mu^\nu, \quad (1.14)$$

is a covariant 4-vector (i.e. according to eq. (1.7)). This means that the contravariant equivalent 4-vector will have an extra minus sign in its space-like components,

$$\partial^\mu = \left(\frac{\partial}{\partial t}, -\vec{\nabla} \right). \quad (1.15)$$

The convention for the totally antisymmetric Levi-Civita tensor is

$$\epsilon^{\mu\nu\lambda\sigma} = \begin{cases} +1 & \text{if } \{\mu, \nu, \lambda, \sigma\} \text{ an even permutation of } \{0, 1, 2, 3\} \\ -1 & \text{if an odd permutation} \\ 0 & \text{otherwise} \end{cases}. \quad (1.16)$$

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

▷ **Exercise 1.3**

Verify the above two properties of $\epsilon^{\mu\nu\lambda\sigma}$.

I will use natural units, $c = 1$, $\hbar = 1$, so mass, energy, inverse length and inverse time all have the same dimensions. Generally think of energy as the basic unit, e.g. mass has units of GeV and distance has units of GeV⁻¹.

▷ **Exercise 1.4**

Noting that E has SI unit kg.m².s⁻², c has SI unit m.s⁻¹ and \hbar has SI unit kg.m².s⁻¹, what is a mass of 1 GeV in kg and what is a cross-section of 1 GeV⁻² in microbarns?

2 Relativistic Wave Equations

Let's review how wave equations describe non-relativistic quantum particles. Experimentally we know that a particle with definite momentum \vec{p} and energy E can be associated with a plane wave

$$\psi = e^{i(\vec{k}\cdot\vec{x}-wt)}, \quad \text{with} \quad \vec{k} = \frac{\vec{p}}{\hbar}, \quad w = \frac{E}{\hbar}. \quad (2.1)$$

To extract E and \vec{p} from the wave we use *operators*

$$E\psi = i\hbar\frac{d}{dt}\psi, \quad \vec{p}\psi = -i\hbar\vec{\nabla}\psi. \quad (2.2)$$

In quantum mechanics, it is more usual to refer to the energy operator as the *Hamiltonian* H , and write (with $\hbar = 1$)

$$H\psi = i\frac{\partial\psi}{\partial t}. \quad (2.3)$$

I shall usually reserve the Greek symbol ψ for spin 1/2 fermions and ϕ for spin 0 bosons. So for pions and the like I shall write

$$H\phi = i\frac{\partial\phi}{\partial t}. \quad (2.4)$$

In non-relativistic systems, conservation of energy can be written

$$H = T + V, \quad (2.5)$$

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

$$T = \frac{\vec{p}^2}{2m}. \quad (2.6)$$

Replacing the energy and momentum operators with the forms seen in eq. (2.2), we arrive at the Schrödinger equation

$$i\hbar\frac{d}{dt}\psi = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi. \quad (2.7)$$

In this equation ψ is the wave function describing the single particle probability amplitude. For a slow moving particle $v \ll c$ (e.g. an electron in a Hydrogen atom) this is adequate, but for relativistic systems ($v \sim c$) the Hamiltonian above is incorrect.

For a free relativistic particle the total energy E is given by the Einstein equation

$$E^2 = \vec{p}^2 + m^2. \quad (2.8)$$

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

$$H^2 = \vec{p}^2 + m^2. \quad (2.9)$$

So far, so good, but how should this be implemented into the wave equation of eq. (2.3), which is expressed in terms of H rather than H^2 ? Naively the relativistic wave equation looks like

$$\sqrt{\vec{p}^2 + m^2}\psi(t) = i\frac{\partial\psi(t)}{\partial t} \quad (2.10)$$

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

1. Work with H^2 . By iterating the wave equation we have

$$H^2\phi(t) = -\frac{\partial^2\phi(t)}{\partial t^2} \left[\text{or } \left(\frac{i\partial}{\partial t} - V \right)^2 \phi(t) \right] \quad (2.11)$$

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

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

$$H_D\psi(t) = i\frac{\partial\psi(t)}{\partial t} \quad (2.12)$$

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

2.1 The Klein-Gordon Equation

Let us now take a more detailed look at the KG equation (2.11). In position space we write the energy-momentum operator as

$$p^\mu \rightarrow i\partial^\mu, \quad (2.13)$$

so that the KG equation (for zero potential V) becomes

$$(\partial^2 + m^2)\phi(x) = 0 \quad (2.14)$$

where we recall the notation,

$$\partial^2 = \partial_\mu\partial^\mu = \partial^2/\partial t^2 - \nabla^2 \quad (2.15)$$

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

The operator ∂^2 is Lorentz invariant, so the Klein-Gordon equation is relativistically covariant (that is, transforms into an equation of the same form) if ϕ is a scalar function. That is to say, under a Lorentz transformation $(t, \vec{x}) \rightarrow (t', \vec{x}')$,

$$\phi(t, \vec{x}) \rightarrow \phi'(t', \vec{x}') = \phi(t, \vec{x}) \quad (2.16)$$

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

The Klein-Gordon equation has plane wave solutions:

$$\phi(x) = Ne^{-i(Et - \vec{p}\cdot\vec{x})} \quad (2.17)$$

where N is a normalization constant and $E = \pm\sqrt{\vec{p}^2 + m^2}$. Thus, there are both positive and negative energy solutions. The negative energy solutions pose a severe problem if we try to interpret ϕ as a wave function (as indeed we are trying to do). The spectrum is no longer bounded from below, and we can extract arbitrarily large amounts of energy from

the system by driving it to ever more negative energy states. Any external perturbation capable of pushing a particle across the energy gap of $2m$ between the positive and negative energy continuum of states can uncover this difficulty. Furthermore, we cannot just throw away these solutions as unphysical since they appear as Fourier modes in any realistic solution of (2.14). Note that if one interprets ϕ as a quantum field there is no problem, as you will see in the field theory course. The positive and negative energy modes are just associated with operators which create or destroy particles.

A second problem with the wave function interpretation arises when trying to find a probability density. Since ϕ is Lorentz invariant, $|\phi|^2$ does not transform like a density (i.e. as the time component of a 4-vector) so we will not have a Lorentz covariant continuity equation $\partial\rho + \vec{\nabla} \cdot \vec{J} = 0$. To search for a candidate we derive such a continuity equation. Defining ρ and \vec{J} by

$$\rho \equiv i \left(\phi^* \frac{\partial\phi}{\partial t} - \phi \frac{\partial\phi^*}{\partial t} \right), \quad \left[\text{or } \phi^* \left(i \frac{\partial}{\partial t} - V \right) + \phi \left(-i \frac{\partial}{\partial t} - V \right) \phi^* \right], \quad (2.18)$$

$$\vec{J} \equiv -i (\phi^* \vec{\nabla} \phi - \phi \vec{\nabla} \phi^*), \quad (2.19)$$

we obtain (see problem) a covariant conservation equation

$$\partial_\mu J^\mu = 0, \quad (2.20)$$

where J is the 4-vector (ρ, \vec{J}) . It is thus natural to interpret ρ as a probability density and \vec{J} as a probability current. However, for a plane wave solution (2.17), $\rho = 2|N|^2 E$, so the negative energy solutions also have a negative probability!

▷ **Exercise 2.5**

Derive the continuity equation (2.20). Start with the Klein-Gordon equation multiplied by ϕ^* and subtract the complex conjugate of the KG equation multiplied by ϕ .

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

Before turning to discuss what Dirac did, let us put things in context. We have found that the Klein-Gordon equation, a candidate for describing the quantum mechanics of spinless particles, admits unacceptable negative energy states when ϕ is interpreted as the single particle wave function. We could solve all our problems here and now, and restore our faith in the Klein-Gordon equation, by simply re-interpreting ϕ as a quantum field. However we will not do that. There is another way forward (this is the way followed in the textbook of Halzen & Martin) due to Feynman and Stückelberg. Causality forces us to ensure that positive energy states propagate forwards in time, but if we force the negative energy states to propagate only backwards in time then we find a theory that is consistent with the requirements of causality and that has none of the aforementioned

problems. In fact, the negative energy states cause us problems only so long as we think of them as real physical states propagating forwards in time. Therefore, we should interpret the emission (absorption) of a negative energy particle with momentum p^μ as the absorption (emission) of a positive energy antiparticle with momentum $-p^\mu$.

In order to become more familiar with this picture, consider a process with a π^+ and a photon in the initial state and final state. In figure 1(a) the π^+ starts from the point A and at a later time t_1 emits a photon at the point \vec{x}_1 . If the energy of the π^+ is still positive, it travels on forwards in time and eventually will absorb the initial state photon at t_2 at the point \vec{x}_2 . The final state is then again a photon and a (positive energy) π^+ .

There is another process however, with the same initial and final state, shown in figure 1(b). Again, the π^+ starts from the point A and at a later time t_2 emits a photon at the point \vec{x}_1 . But this time, the energy of the photon emitted is bigger than the energy of the initial π^+ . Thus, the energy of the π^+ becomes negative and it is forced to travel backwards in time. Then at an *earlier* time t_1 it absorbs the initial state photon at the point \vec{x}_2 , thereby rendering its energy positive again. From there, it travels forward in time and the final state is the same as in figure 1(a), namely a photon and a (positive energy) π^+ .

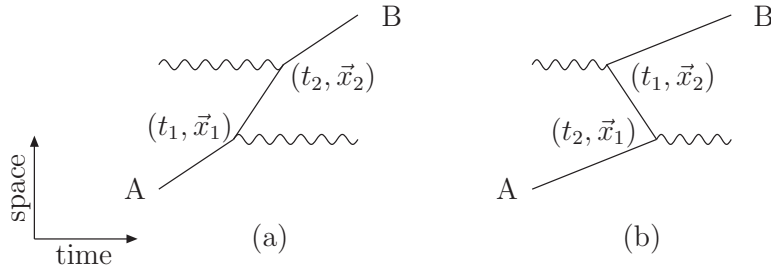


Figure 1: Interpretation of negative energy states

In today's language, the process in figure 1(b) would be described as follows: in the initial state we have a π^+ and a photon. At time t_1 and at the point \vec{x}_2 the photon creates a $\pi^+\pi^-$ pair. Both propagate forwards in time. The π^+ ends up in the final state, whereas the π^- is annihilated at (a later) time t_2 at the point \vec{x}_1 by the initial state π^+ , thereby producing the final state photon. To someone observing in real time, the negative energy state moving backwards in time looks to all intents and purposes like a negatively charged pion with positive energy moving forwards in time.

▷ **Exercise 2.6**

Consider a wave incident on the potential step shown in figure 2. Show that if the step size $V > m + E_p$, where $E_p = \sqrt{p^2 + m^2}$ then one cannot avoid using the negative square root $\vec{k} = -\sqrt{(E_p - V)^2 + m^2}$, resulting in negative currents and densities. Hint: use the continuity of $\phi(x)$ and $\partial\phi(x)/\partial x$ at $x = 0$, and ensure that the group velocity $v_g = \partial E/\partial k$ is positive for $x > 0$. Interpret the solution.

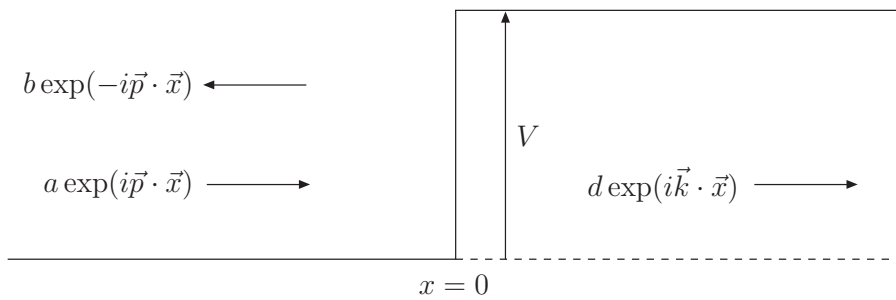


Figure 2: A potential step

2.2 The Dirac Equation

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

$$H_D = \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta m, \quad (2.21)$$

where p_i are the three components of the momentum operator \vec{p} , and α_i and β are some unknown quantities, which we will show must be interpreted as 4×4 matrices. Substituting the expressions for the operators eq. (2.13) into the Dirac Hamiltonian of eq. (2.21) results in the equation

$$i \frac{\partial \psi}{\partial t} = (-i \vec{\alpha} \cdot \vec{\nabla} + \beta m) \psi \quad (2.22)$$

which is the position space Dirac equation.

If ψ is to describe a free particle it must satisfy the Klein-Gordon equation so that it has the correct energy-momentum relation. This requirement imposes relationships among $\alpha_1, \alpha_2, \alpha_3$ and β . To see this, apply the Hamiltonian operator to ψ *twice*, to give

$$-\frac{\partial^2 \psi}{\partial t^2} = [-\alpha^i \alpha^j \nabla^i \nabla^j - i(\beta \alpha^i + \alpha^i \beta) m \nabla^i + \beta^2 m^2] \psi, \quad (2.23)$$

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

$$-\frac{\partial^2 \psi}{\partial t^2} = [-\nabla^i \nabla^i + m^2] \psi. \quad (2.24)$$

It is clear that we cannot recover the KG equation from the Dirac equation if the α^i and β are normal numbers. Insisting that the terms linear in ∇^i vanish independently would require either β to vanish or *all* the α^i to vanish. This would remove either $\nabla^i \nabla^j$ term or the m^2 term, both of which are unacceptable. Instead we must insist that the terms linear in ∇^i vanish in their sum *without* any of α^i or β vanishing, i.e. we must assume that α^i and β *anti-commute*. We recover the KG equation only if

$$\alpha_i \alpha_j + \alpha_j \alpha_i = 2\delta_{ij}$$

$$\begin{aligned}\beta\alpha_i + \alpha_i\beta &= 0 \\ \beta^2 &= 1\end{aligned}\tag{2.25}$$

for $i, j = 1, 2, 3$. In principle, these equations *define* α^i and β , and any objects which obey these relations are good representations of them. However, in practice, we will represent them by matrices. In this case, ψ is a multi-component *spinor* on which these matrices act.

▷ **Exercise 2.7**

Prove that any matrices $\vec{\alpha}$ and β satisfying eq. (2.25) are traceless with eigenvalues ± 1 . Hence argue that they must be even dimensional.

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

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.\tag{2.26}$$

However, there is no other independent 2×2 matrix with the right properties for β , so we must use a higher dimensional form. The smallest number of dimensions for which the Dirac matrices can be realized is four. One choice is the *Dirac representation*:

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.\tag{2.27}$$

Note that each entry above denotes a two-by-two block and that the 1 denotes the 2×2 identity matrix. The spinor ψ therefore has four components.

There is a theorem due to Pauli that states that all sets of matrices obeying the relations in eq. (2.25) are equivalent. Since the hermitian conjugates $\vec{\alpha}^\dagger$ and β^\dagger clearly obey the relations, you can, by a change of basis if necessary, assume that $\vec{\alpha}$ and β are hermitian. All the common choices of basis have this property. Furthermore, we would like α_i and β to be hermitian so that the Dirac Hamiltonian (2.42) is hermitian.

If we define

$$\rho = J^0 = \psi^\dagger\psi, \quad \vec{J} = \psi^\dagger\vec{\alpha}\psi,\tag{2.28}$$

then it is a simple exercise using the Dirac equation to show that this satisfies the continuity equation $\partial_\mu J^\mu = 0$. We will see in section 2.8 that (ρ, \vec{J}) transforms, as it must, as a 4-vector. Note that ρ is now also positive definite.

2.3 Solutions to the Dirac Equation

We look for plane wave solutions of the form

$$\psi = \begin{pmatrix} \chi(\vec{p}) \\ \phi(\vec{p}) \end{pmatrix} e^{-i(Et - \vec{p}\cdot\vec{x})}\tag{2.29}$$

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

$$E \begin{pmatrix} \chi \\ \phi \end{pmatrix} = \begin{pmatrix} m & \vec{\sigma} \cdot \vec{p} \\ \vec{\sigma} \cdot \vec{p} & -m \end{pmatrix} \begin{pmatrix} \chi \\ \phi \end{pmatrix}.\tag{2.30}$$

There are two simple cases for which eq. (2.30) can readily be solved, namely

1. $\vec{p} = 0$, $m \neq 0$, which might represent an electron in its rest frame.
2. $m = 0$, $\vec{p} \neq 0$, which describes a massless particle or a particle in the ultra-relativistic limit ($E \gg m$).

For case (1), an electron in its rest frame, the equations (2.30) decouple and become simply,

$$E\chi = m\chi, \quad E\phi = -m\phi. \quad (2.31)$$

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

The negative energy solutions persist for an electron with $\vec{p} \neq 0$ for which the solutions to equation (2.30) are

$$\phi = \frac{\vec{\sigma} \cdot \vec{p}}{E+m} \chi, \quad \chi = \frac{\vec{\sigma} \cdot \vec{p}}{E-m} \phi. \quad (2.32)$$

▷ **Exercise 2.8**

Show that $(\vec{\sigma} \cdot \vec{p})^2 = \vec{p}^2$.

Using $(\vec{\sigma} \cdot \vec{p})^2 = \vec{p}^2$ we see that $E = \pm|\sqrt{\vec{p}^2 + m^2}|$. We write the positive energy solutions with $E = +|\sqrt{\vec{p}^2 + m^2}|$ as

$$\psi(x) = \begin{pmatrix} \chi \\ \frac{\vec{\sigma} \cdot \vec{p}}{E+m} \chi \end{pmatrix} e^{-i(Et - \vec{p} \cdot \vec{x})}, \quad (2.33)$$

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

$$\psi(x) = \begin{pmatrix} \frac{\vec{\sigma} \cdot \vec{p}}{E-m} \phi \\ \phi \end{pmatrix} e^{-i(Et - \vec{p} \cdot \vec{x})}, \quad (2.34)$$

for arbitrary constant ϕ and χ . Clearly when $\vec{p} = 0$ these solutions reduce to the positive and negative energy solutions discussed previously.

It is interesting to see how Dirac coped with the negative energy states. Dirac interpreted the negative energy solutions by postulating the existence of a “sea” of negative energy states. The vacuum or ground state has all the negative energy states full. An additional electron must now occupy a positive energy state since the Pauli exclusion principle forbids it from falling into one of the filled negative energy states. On promoting one of these negative energy states to a positive energy one, by supplying energy, an electron-hole pair is created, i.e. a positive energy electron and a hole in the negative energy sea. The hole is seen in nature as a positive energy positron. This was a radical new idea, and brought pair creation and antiparticles into physics. The problem with Dirac’s hole theory is that it does not work for bosons. Such particles have no exclusion principle to stop them falling into the negative energy states, releasing their energy.

It is convenient to rewrite the solutions, eqs. (2.33) and (2.34), introducing the spinors $u_\alpha^{(s)}(\vec{p})$ and $v_\alpha^{(s)}(\vec{p})$. The label $\alpha \in \{1, 2, 3, 4\}$ is a spinor index that often will be suppressed, while $s \in \{1, 2\}$ denotes the spin state of the fermion, as we shall see later. We take the positive energy solution eq. (2.33) and define

$$\sqrt{E+m} \begin{pmatrix} \chi_s \\ \frac{\vec{\sigma} \cdot \vec{p}}{E+m} \chi_s \end{pmatrix} e^{-ip \cdot x} \equiv u^{(s)}(p) e^{-ip \cdot x}. \quad (2.35)$$

For the negative energy solution of eq. (2.34), change the sign of the energy, $E \rightarrow -E$, and the three-momentum, $\vec{p} \rightarrow -\vec{p}$, to obtain,

$$\sqrt{E+m} \begin{pmatrix} \frac{\vec{\sigma} \cdot \vec{p}}{E+m} \chi_s \\ \chi_s \end{pmatrix} e^{ip \cdot x} \equiv v^{(s)}(p) e^{ip \cdot x}. \quad (2.36)$$

In these two solutions E is now (and for the rest of the course) always positive and given by $E = (\vec{p}^2 + m^2)^{1/2}$. The χ_s for $s = 1, 2$ are

$$\chi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (2.37)$$

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

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

2.4 Orthogonality and Completeness

Our solutions to the Dirac equation take the form

$$\psi = Nu^{(s)}e^{-ip \cdot x}, \quad \psi = Nv^{(r)}e^{ip \cdot x}, \quad \text{with } r, s = 1, 2, \quad (2.38)$$

where N is a normalization factor. We have already included a factor $\sqrt{E+m}$ in our spinors (see eqs. (2.35) and (2.36)), which results in

$$u^{(r)\dagger}(p)u^{(s)}(p) = v^{(r)\dagger}(p)v^{(s)}(p) = 2E\delta^{rs}. \quad (2.39)$$

This convention allows $u^\dagger u$ to transform as the time component of a 4-vector under Lorentz transformations, which is essential to its interpretation as a probability density (see eq. (2.28) and section 2.8). Also note that the spinors are *orthogonal*.

▷ Exercise 2.9

Check the normalization condition for the spinors in eq. (2.39).

We must further normalize the spatial part of the wave functions. In fact a plane wave is not normalizable in an infinite space so in the computations that follow where we use them we will work in a large box of volume V - such a construction is not Lorentz invariant. The number of particles in the box will be

$$\int \psi^\dagger \psi d^3x = 2E N^2 V, \quad (2.40)$$

so setting $N = 1/\sqrt{V}$ allows us to adopt the standard relativistic normalization convention of $2E$ particles per box of volume V . Most people and the books use this

convention. I frequently find it more intuitive, given we've broken Lorentz invariance, to set $N = 1/\sqrt{2EV}$ so there's one particle in the box. I'll try to be clear below when I do this.

Remember that the solutions to the wave equation form a complete set of states meaning that we can expand (like a Fourier expansion) an arbitrary function $\chi(x)$ in terms of them

$$\chi(x) = \sum_n a_n \psi_n(x) \quad (2.41)$$

The a_n are the equivalent of Fourier coefficients and if χ is a wave function in some quantum mixed state then $|a_n|^2$ is the probability of being in the state ψ_n (or $2E$ times that!).

2.5 Spin

Now it is time to justify the statements we have been making that the Dirac equation describes spin-1/2 particles. The Dirac Hamiltonian in momentum space is given in eq. (2.21) as

$$H_D = \vec{\alpha} \cdot \vec{p} + \beta m, \quad (2.42)$$

and the orbital angular momentum operator is

$$\vec{L} = \vec{R} \times \vec{p}. \quad (2.43)$$

Evaluating the commutator of \vec{L} with H_D ,

$$\begin{aligned} [\vec{L}, H_D] &= [\vec{R} \times \vec{p}, \vec{\alpha} \cdot \vec{p}] \\ &= [\vec{R}, \vec{\alpha} \cdot \vec{p}] \times \vec{p} \\ &= i\vec{\alpha} \times \vec{p}, \end{aligned} \quad (2.44)$$

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

$$\vec{J} = \vec{L} + \vec{S}, \quad [\vec{J}, H_D] = 0. \quad (2.45)$$

To this end, consider the three matrices,

$$\vec{\Sigma} \equiv \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} = -i\alpha_1\alpha_2\alpha_3\vec{\alpha}, \quad (2.46)$$

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

$$[\vec{\Sigma}, \beta] = 0, \quad [\Sigma_i, \alpha_j] = 2i\varepsilon_{ijk}\alpha_k. \quad (2.47)$$

From the relations in (2.47) we find that

$$[\vec{\Sigma}, H_D] = -2i\vec{\alpha} \times \vec{p}. \quad (2.48)$$

▷ **Exercise 2.10**

Using $\alpha_1\alpha_2\alpha_3 \equiv \frac{1}{3}\epsilon_{ijk}\alpha_i\alpha_j\alpha_k$ verify the commutation relations in eqs. (2.47) and (2.48).

Comparing eq. (2.48) with the commutator of \vec{L} with H_D in eq. (2.44), you see that

$$[\vec{L} + \frac{1}{2}\vec{\Sigma}, H_D] = 0, \quad (2.49)$$

and we can identify

$$\vec{S} = \frac{1}{2}\vec{\Sigma} \quad (2.50)$$

as the additional quantity that, when added to \vec{L} in equation (2.45), yields a conserved total angular momentum \vec{J} . We interpret \vec{S} as an angular momentum *intrinsic* to the particle. Now

$$\vec{S}^2 = \frac{1}{4} \begin{pmatrix} \vec{\sigma} \cdot \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \cdot \vec{\sigma} \end{pmatrix} = \frac{3}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (2.51)$$

and, recalling that the eigenvalue of \vec{J}^2 for spin j is $j(j+1)$, we conclude that \vec{S} represents spin-1/2 and the solutions of the Dirac equation have spin-1/2 as promised. We worked in the Dirac representation of the matrices for convenience, but the result is necessarily independent of the representation.

Now consider the u -spinor solutions $u^{(s)}(p)$ of eq. (2.35). Choose $\vec{p} = (0, 0, p_z)$ and write

$$u_{\uparrow} \equiv u^{(1)}(p) = \begin{pmatrix} \sqrt{E+m} \\ 0 \\ \sqrt{E-m} \\ 0 \end{pmatrix}, \quad u_{\downarrow} \equiv u^{(2)}(p) = \begin{pmatrix} 0 \\ \sqrt{E+m} \\ 0 \\ -\sqrt{E-m} \end{pmatrix}. \quad (2.52)$$

With these definitions, we get

$$S_z u_{\uparrow} = \frac{1}{2} u_{\uparrow}, \quad S_z u_{\downarrow} = -\frac{1}{2} u_{\downarrow}. \quad (2.53)$$

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

$$v_{\downarrow} = v^{(1)}(p) = \begin{pmatrix} \sqrt{E-m} \\ 0 \\ \sqrt{E+m} \\ 0 \end{pmatrix}, \quad v_{\uparrow} = v^{(2)}(p) = \begin{pmatrix} 0 \\ -\sqrt{E-m} \\ 0 \\ \sqrt{E+m} \end{pmatrix}, \quad (2.54)$$

where now,

$$S_z v_{\downarrow} = \frac{1}{2} v_{\downarrow}, \quad S_z v_{\uparrow} = -\frac{1}{2} v_{\uparrow}. \quad (2.55)$$

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

2.6 Lorentz Covariance

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

$$\gamma^0 = \beta, \quad \vec{\gamma} = \beta\vec{\alpha}. \quad (2.56)$$

In the Dirac representation,

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \vec{\gamma} = \begin{pmatrix} 0 & \vec{\sigma} \\ -\vec{\sigma} & 0 \end{pmatrix}. \quad (2.57)$$

In terms of these, the relations between the $\vec{\alpha}$ and β in eq. (2.25) can be written compactly as,

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \quad (2.58)$$

▷ Exercise 2.11

Prove that $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$.

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

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

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

Observe that using the γ -matrices the Dirac equation (2.22) becomes

$$(i\not{\partial} - m)\psi = 0, \quad (2.59)$$

or, in momentum space,

$$(\not{p} - m)\psi = 0. \quad (2.60)$$

The spinors u and v satisfy

$$(\not{p} - m)u^{(s)}(p) = 0, \quad (2.61)$$

$$(\not{p} + m)v^{(s)}(p) = 0, \quad (2.62)$$

since for $v^{(s)}(p)$, $E \rightarrow -E$ and $\vec{p} \rightarrow -\vec{p}$.

We want the Dirac equation (2.59) to preserve its form under Lorentz transformations eq. (1.2). We’ve just naively written the matrices in the Dirac equation as γ_μ however this does not make them a 4-vector! They are just a set of numbers in four matrices and there’s no reason they should change when we do a boost. Since ∂^μ does transform, for the equation to be Lorentz covariant we are led to propose that ψ transforms too. We know that 4-vectors get their components mixed up by LT’s, so we expect that the components of ψ might get mixed up too:

$$\psi(x) \rightarrow \psi'(x') = S(\Lambda)\psi(x) = S(\Lambda)\psi(\Lambda^{-1}x') \quad (2.63)$$

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

In order to appreciate the above it is useful to consider a vector field, where the corresponding transformation is

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

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

$$A'^\mu(x') = \Lambda^\mu{}_\nu A^\nu(x). \quad (2.64)$$

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

We now need to figure out what S is. The requirement is that the Dirac equation has the same form in any inertial frame. Thus, if we make a LT from our original frame into another ('primed') frame and write down the Dirac equation in this frame, it has to have the same form.

$$(i\gamma^\mu \partial_\mu - m)\psi(x) = 0 \quad \longrightarrow \quad (i\gamma'^\mu \partial'_\mu - m)\psi'(x') = 0, \quad (2.65)$$

where we used the fact that m is a scalar, i.e. $m' = m$.

The derivative transforms as a covector, eq. (1.7), so using the orthogonality condition of eq. (1.8), we can write $\partial_\mu = \Lambda^\sigma{}_\mu \partial'_\sigma$ and multiplying the Dirac equation in the *original* frame by S it becomes

$$S(i\gamma^\mu \Lambda^\sigma{}_\mu \partial'_\sigma - m)\psi(x) = 0. \quad (2.66)$$

On the other hand, we can use the definition of S in eq. (2.63) to rewrite the equation in the primed frame as

$$(i\gamma'^\mu \partial'_\mu - m)S\psi(x) = 0. \quad (2.67)$$

We can see that the second term (containing m) of eqs. (2.66) and (2.67) are now identical. To make the first term identical we need $S\Lambda^\sigma{}_\mu \gamma^\mu = \gamma'^\sigma S$. Thus, in order for the Dirac equation to be Lorentz invariant, $S(\Lambda)$ has to satisfy

$$\Lambda^\sigma{}_\mu \gamma^\mu = S^{-1} \gamma'^\sigma S \quad (2.68)$$

We still haven't solved for S explicitly. We need to find an S that satisfies eq. (2.68). Since S depends on the LT, we first have to find a convenient parameterization of a LT and then express $S(\Lambda)$ in terms of these parameters. For an infinitesimal LT, it can be shown that,

$$\Lambda^\mu{}_\nu = g^\mu{}_\nu + \omega^\mu{}_\nu \quad (2.69)$$

where $\omega_{\mu\nu}$ is an antisymmetric set of infinitesimal parameters. For example, a boost along the z -axis corresponds to $\omega_{03} = -\omega_{30} = -\beta$ (remember that $\omega_{0i} = \omega^0{}_i = -\omega_0{}^i$ etc) with all other entries of $\omega_{\mu\nu}$ zero,

$$\Lambda^\mu{}_\nu = g^\mu{}_\nu + \omega^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & -\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta & 0 & 0 & 1 \end{pmatrix}. \quad (2.70)$$

This corresponds to eq. (1.3) when one makes an expansion in small β , i.e. $\gamma = 1 + \mathcal{O}(\beta^2)$. Non-zero ω_{01} or ω_{02} correspond to boosts along the x and y axes respectively. The remaining combinations, non-zero ω_{23} , ω_{31} or ω_{12} , correspond to infinitesimal anti-clockwise rotations through an angle ω_{ij} about the x , y and z axes respectively. It's a nice exercise to check this out.

For an infinitesimal LT we are at liberty to write

$$S(\Lambda) = 1 + \frac{i}{4} \omega_{\mu\nu} \sigma^{\mu\nu}, \quad (2.71)$$

which is nothing but a definition of the set of matrices $\sigma^{\mu\nu}$. Our task is to determine these matrices. To do this, substitute the expression for S , eq. (2.71), into eq. (2.68) (and remember that $S^{-1}(\Lambda) = 1 - \frac{i}{4} \omega_{\mu\nu} \sigma^{\mu\nu}$). After some algebra, we can convince ourselves that the solution is

$$\sigma^{\mu\nu} = \frac{i}{2} [\gamma^\mu, \gamma^\nu] \quad (2.72)$$

Thus S can be written explicitly in terms of γ -matrices for a general LT by building the finite transformation out of lots of infinitesimal ones.

▷ **Exercise 2.12**

Verify that eq. (2.72) is true.

Now that we now how ψ transforms we can find quantities that are Lorentz invariant, or transform as vectors or tensors under LT's. To this end, we will find it useful to introduce the Dirac *adjoint*. The Dirac adjoint $\bar{\psi}$ of a spinor ψ is defined by

$$\bar{\psi} \equiv \psi^\dagger \gamma^0 \quad (2.73)$$

With the help of

$$S^\dagger(\Lambda) \gamma^0 = \gamma^0 S^{-1}(\Lambda) \quad (2.74)$$

we see that $\bar{\psi}$ transforms under LT's as

$$\bar{\psi} \rightarrow \bar{\psi}' = \bar{\psi} S^{-1}(\Lambda). \quad (2.75)$$

▷ **Exercise 2.13**

1. Verify that $\gamma^{\mu\dagger} = \gamma^0 \gamma^\mu \gamma^0$.
2. Prove eq. (2.74)
3. Show that $\bar{\psi}$ satisfies the equation

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

where the arrow over $\not{\partial}$ implies the derivative acts to the left.

4. Hence prove that $\bar{\psi}$ transforms as in eq. (2.75).

Combining the transformation properties of ψ and $\bar{\psi}$ in eqs. (2.63) and (2.75) we see that the bilinear $\bar{\psi}\psi$ is Lorentz invariant. In section 2.8 we will consider the transformation properties of general bilinears.

Let's close this section by recasting the spinor normalization eq. (2.39) in terms of Dirac inner products. The conditions become

$$\begin{aligned} \bar{u}^{(r)}(p)u^{(s)}(p) &= 2m\delta^{rs} \\ \bar{u}^{(r)}(p)v^{(s)}(p) &= \bar{v}^{(r)}(p)u^{(s)}(p) = 0 \\ \bar{v}^{(r)}(p)v^{(s)}(p) &= -2m\delta^{rs} \end{aligned} \quad (2.76)$$

where, in analogy to eq. (2.73), we defined $\bar{u} \equiv u^\dagger \gamma^0$ and $\bar{v} \equiv v^\dagger \gamma^0$.

▷ **Exercise 2.14**

Verify the normalization properties in the above equations (2.76).

2.7 Parity, charge conjugation and time reversal

2.7.1 Parity

We usually use LT's which are in the connected Lorentz Group, $SO(3,1)$, meaning they can be obtained by a continuous deformation of the identity transformation (i.e. by lots of little transformations)¹. This class of LT is often referred to as proper LT. However, the full Lorentz group consists not only of the proper transformations but also includes the discrete operations of *parity* (space inversion), P , and *time reversal*, T :

$$\Lambda_P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \Lambda_T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (2.77)$$

LT's satisfy $\Lambda^T g \Lambda = g$, so taking determinants shows that $\det \Lambda = \pm 1$. Proper LT's are continuously connected to the identity so must have determinant 1, but both P and T operations have determinant -1 .

Let us now find the action of parity on the Dirac wave function and determine the wave function ψ_P in the parity-reversed system. According to the discussion of the previous section, we need to find a matrix P satisfying

$$P^{-1}\gamma^0 P = \gamma^0, \quad P^{-1}\gamma^i P = -\gamma^i. \quad (2.78)$$

¹Indeed in the last section we considered LT's very close to the identity in equation (2.69)

Using some properties of the γ -matrices we see that $P = P^{-1} = \gamma^0$ is an acceptable solution (Clearly one could multiply γ^0 by a phase and still have an acceptable definition for the parity transformation.), from which it follows that the transformation is

$$\psi(t, \vec{x}) \rightarrow \psi_P(t, -\vec{x}) = P\psi(t, \vec{x}) = \gamma^0\psi(t, \vec{x}). \quad (2.79)$$

Since

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (2.80)$$

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

2.7.2 Charge Conjugation

Another discrete invariance of the Dirac equation is *charge conjugation*, which takes you from particle to antiparticle and vice versa. For scalar fields the symmetry is just complex conjugation, but in order for the charge conjugate Dirac field to remain a solution of the Dirac equation, you have to mix its components as well. The transformation on the fermion wavefunction is

$$\psi \rightarrow \psi_C = C\bar{\psi}^T, \quad (2.81)$$

where $\bar{\psi}^T = (\psi^\dagger \gamma^0)^T = \gamma^{0T} \psi^{\dagger T} = \gamma^0 \psi^*$. To find the form of C , let's take the complex conjugate of the Dirac Equation,

$$\begin{aligned} (i\gamma^\mu \partial_\mu - m)^* \psi^* &= \left(i(\gamma^{\mu\dagger})^T \partial_\mu - m \right) (\psi^\dagger)^T \\ &= \gamma^{0T} (-i\gamma^{\mu T} \partial_\mu - m) \bar{\psi}^T, \end{aligned} \quad (2.82)$$

where we have additionally used $\gamma^{\mu\dagger} = \gamma^0 \gamma^\mu \gamma^0$. Premultiply by C and the Dirac equation becomes

$$(-iC\gamma^{\mu T} C^{-1} \partial_\mu - m) \psi_c = 0. \quad (2.83)$$

In order for ψ_C to satisfy the Dirac equation we require C to be a matrix satisfying the condition

$$C\gamma_\mu^T C^{-1} = -\gamma_\mu \quad (C^{-1} = C^\dagger). \quad (2.84)$$

In the Dirac representation, a suitable choice for this operator is

$$C = i\gamma^2 \gamma^0 = \begin{pmatrix} 0 & -i\sigma^2 \\ -i\sigma^2 & 0 \end{pmatrix}. \quad (2.85)$$

The charge-conjugation transformation is then

$$\psi(t, \vec{x}) \rightarrow \psi_C(t, \vec{x}) = C\bar{\psi}^T(t, \vec{x}) = i\gamma^2 \gamma^0 \bar{\psi}^T(t, \vec{x}). \quad (2.86)$$

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

2.7.3 Time reversal

As already noted, time reversal is an improper LT, given by Λ_T in eq. (2.77). Naively one would expect to derive a time reversal operation in the same way as for parity. However, there is a subtlety that the momentum of a particle is a *rate of change*, so if we reverse the direction of time, the momentum must change direction. When we reverse the momentum \vec{p} in a plane wave we find

$$e^{-i(Et-\vec{p}\cdot\vec{x})} \longrightarrow e^{-i(Et-(-\vec{p})\cdot\vec{x})} = e^{i(E(-t)-\vec{p}\cdot\vec{x})} = \left(e^{-i(E(-t)-\vec{p}\cdot\vec{x})}\right)^*. \quad (2.87)$$

In this example, taking the complex conjugate is the equivalent of reversing the time coordinate and reversing the momentum. So once again, we must take the *complex conjugate* of the field, transforming it according to

$$\psi(t, \vec{x}) \rightarrow \psi_T(-t, \vec{x}) = T\psi^*(t, \vec{x}). \quad (2.88)$$

To find the form of T , let's take the complex conjugate of the Dirac equation, premultiply by T and interchange $t \rightarrow -t$,

$$\begin{aligned} \left(i\gamma^0 \frac{\partial}{\partial t} + i\vec{\gamma} \cdot \vec{\nabla} - m\right) \psi(t, \vec{x}) &\longrightarrow S_T \left(-i\gamma^{0*} \frac{\partial}{\partial(-t)} - i\vec{\gamma}^* \cdot \vec{\nabla} - m\right) T^{-1}T\psi^*(-t, \vec{x}) \\ &= \left(i \left[T\gamma^{0*}T^{-1}\right] \frac{\partial}{\partial t} + i \left[-T\vec{\gamma}^*T^{-1}\right] \cdot \vec{\nabla} - m\right) \psi_T(t, \vec{x}). \end{aligned} \quad (2.89)$$

For ψ_T to satisfy the Dirac equation we need

$$i \left[T\gamma^{0*}T^{-1}\right] = \gamma^0, \quad \left[-T\vec{\gamma}^*T^{-1}\right] = -\vec{\gamma}. \quad (2.90)$$

A suitable choice is

$$T = i\gamma^1\gamma^3 = \begin{pmatrix} 0 & -i\sigma_1\sigma_3 \\ -i\sigma_1\sigma_3 & 0 \end{pmatrix} = i \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad (2.91)$$

and the time reversal transformation on a fermion field is

$$\psi(t, \vec{x}) \rightarrow \psi_T(-t, \vec{x}) = T\psi^*(t, \vec{x}) = i\gamma^1\gamma^3\psi^*(t, \vec{x}) \quad (2.92)$$

2.7.4 CPT

We are now in the position to ask what is the effect of performing charge conjugation, parity and time-reversal all together on a Dirac field. The combined transformation is known as *CPT*. Using eqs. (2.79), (2.86) and (2.92), the CPT transformation is,

$$\begin{aligned} \psi(t, \vec{x}) \rightarrow \psi_{CPT}(-t, -\vec{x}) &= i\gamma^2\gamma^0\gamma^{0T} \left[\gamma^0 i\gamma^1\gamma^3\psi^*(t, \vec{x})\right]^* \\ &= i\gamma^2\gamma^0\gamma^0\gamma^0 (-i)\gamma^1\gamma^3\psi(t, \vec{x}) \\ &= \gamma^0\gamma^1\gamma^2\gamma^3\psi(t, \vec{x}) \\ &= -i\gamma^5\psi(t, \vec{x}) \end{aligned} \quad (2.93)$$

Thus, apart from the factor of γ^5 , a particle moving forward in time is equivalent to an anti-particle moving backwards in time and in the opposite direction. In fact, the extra γ^5 makes no difference to observable quantities (see the next section) so this justifies the Feynman-Stückelberg interpretation of negative energy states we used earlier.

2.8 Bilinear Covariants

Now, as promised, we will construct and classify the bilinears. These are useful for defining quantities with particular properties under Lorentz transformations, and appearing in Lagrangians for fermion field theories.

To begin, note that by forming products of the γ -matrices it is possible to construct 16 linearly independent 4×4 matrices. Any constant 4×4 matrix can then be decomposed into a sum over these basis matrices. In equation (2.72) we have defined

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

and now it is convenient to define

$$\gamma^5 \equiv \gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (2.94)$$

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

$$\gamma^{5\dagger} = \gamma^5, \quad \{\gamma^5, \gamma^\mu\} = 0, \quad (\gamma^5)^2 = 1. \quad (2.95)$$

▷ Exercise 2.15

Prove the three results in eq. (2.95) independently of the γ -matrix representation.

Now, the set of 16 matrices

$$\{1, \gamma^5, \gamma^\mu, \gamma^\mu\gamma^5, \sigma^{\mu\nu}\}$$

form a basis for γ -matrix products. There are 16 matrices since there is 1 unit matrix, 1 γ^5 matrix, 4 γ^μ matrices and 4 $\gamma^\mu\gamma^5$ matrices, and 6 $\sigma^{\mu\nu}$ matrices (see equation (2.72) for the definition of $\sigma^{\mu\nu}$).

Using the transformations of ψ and $\bar{\psi}$ from eqs. (2.63) and (2.75), together with the transformation of γ^μ in eq. (2.74), the 16 fermion bilinears and their transformation properties can be written as follows:

$$\begin{aligned} \bar{\psi}\psi &\rightarrow \bar{\psi}\psi && \text{S scalar} \\ \bar{\psi}\gamma^5\psi &\rightarrow \det(\Lambda)\bar{\psi}\gamma^5\psi && \text{P pseudoscalar} \\ \bar{\psi}\gamma^\mu\psi &\rightarrow \Lambda^\mu{}_\nu\bar{\psi}\gamma^\nu\psi && \text{V vector} \\ \bar{\psi}\gamma^\mu\gamma^5\psi &\rightarrow \det(\Lambda)\Lambda^\mu{}_\nu\bar{\psi}\gamma^\nu\gamma^5\psi && \text{A axial vector} \\ \bar{\psi}\sigma^{\mu\nu}\psi &\rightarrow \Lambda^\mu{}_\lambda\Lambda^\nu{}_\sigma\bar{\psi}\sigma^{\lambda\sigma}\psi && \text{T tensor} \end{aligned} \quad (2.96)$$

In particular we note that

$$\bar{\psi}\gamma^\mu\psi = \psi^\dagger\gamma^0\gamma^\mu\psi = (\psi^\dagger\psi, \psi^\dagger\vec{\alpha}\psi) \quad (2.97)$$

which is our previous definition eq. (2.28) of the current 4-vector J^μ , i.e. we now see that it is really a 4-vector.

▷ **Exercise 2.16**

Derive the transformation properties of the bilinears in equation (2.96) under C, P, T and CPT transformations.

2.9 Massless (Ultra-relativistic) Fermions

At very high energies we may neglect the masses of particles ($E^2 \simeq |\vec{p}|^2$). Therefore, let us look at solutions of the Dirac equation with $m = 0$, on the basis that this will be an extremely good approximation for many situations.

From equation (2.30) we have in this case

$$E\phi = \vec{\sigma} \cdot \vec{p}\chi, \quad E\chi = \vec{\sigma} \cdot \vec{p}\phi. \quad (2.98)$$

These equations can easily be decoupled by taking linear combinations and defining the two component spinors Ψ_L and Ψ_R ,

$$\Psi_{R/L} \equiv \frac{\chi \pm \phi}{2}, \quad (2.99)$$

which leads to

$$E\Psi_R = \vec{\sigma} \cdot \vec{p}\Psi_R, \quad E\Psi_L = -\vec{\sigma} \cdot \vec{p}\Psi_L. \quad (2.100)$$

The system is in fact described by two entirely separated two component spinors. If we take them to be moving in the z direction, and noting that $\sigma_3 = \text{diag}(1, -1)$, we see that there is one positive and one negative energy solution in each.

Further since $E = |\vec{p}|$ for massless particles, these equations may be written

$$\frac{\vec{\sigma} \cdot \vec{p}}{|\vec{p}|}\Psi_L = -\Psi_L, \quad \frac{\vec{\sigma} \cdot \vec{p}}{|\vec{p}|}\Psi_R = \Psi_R \quad (2.101)$$

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

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

$$\Psi_L \rightarrow e^{\frac{i}{2}\vec{\sigma} \cdot (\vec{\theta} - i\vec{\phi})}\Psi_L \quad (2.102)$$

$$\Psi_R \rightarrow e^{\frac{i}{2}\vec{\sigma} \cdot (\vec{\theta} + i\vec{\phi})}\Psi_R \quad (2.103)$$

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

However, under parity transformations $\vec{\sigma} \rightarrow \vec{\sigma}$ (like $\vec{R} \times \vec{p}$), $\vec{p} \rightarrow -\vec{p}$, therefore $\vec{\sigma} \cdot \vec{p} \rightarrow -\vec{\sigma} \cdot \vec{p}$, i.e. the spinors transform into each other:

$$\Psi_L \leftrightarrow \Psi_R. \quad (2.104)$$

So a theory in which Ψ_L has different interactions to Ψ_R (such as the standard model in which the weak force only acts on left handed particles) manifestly violates parity.

Although massless particles can be described very simply using two component spinors as above, they may also be incorporated into the four-component formalism by using the γ^5 we defined earlier. Let's define *projection operators*

$$P_{R/L} \equiv \frac{1}{2} (1 \pm \gamma^5). \quad (2.105)$$

In the Dirac representation, these are,

$$P_{R/L} = \frac{1}{2} \begin{pmatrix} 1 & \pm 1 \\ \pm 1 & 1 \end{pmatrix}, \quad (2.106)$$

where 1 denotes the 2×2 identity matrix. Acting these projection operators on a general Dirac field of the form eq. (2.29) projects onto right- or left-handed eigenstates. To see this, first note that

$$P_{R/L} \begin{pmatrix} \chi \\ \phi \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & \pm 1 \\ \pm 1 & 1 \end{pmatrix} \begin{pmatrix} \chi \\ \phi \end{pmatrix} = \begin{pmatrix} \Psi_{R/L} \\ \Psi_{R/L} \end{pmatrix}. \quad (2.107)$$

The helicity operator in four-component Dirac space is given by $\vec{S} \cdot \vec{p}/|\vec{p}|$, with $\vec{S} = \frac{1}{2}\vec{\Sigma}$, where $\vec{\Sigma}$ is defined in equation (2.46). Acting this operator on the projected state gives

$$\frac{1}{2} \begin{pmatrix} \frac{\vec{\sigma} \cdot \vec{p}}{|\vec{p}|} & 0 \\ 0 & \frac{\vec{\sigma} \cdot \vec{p}}{|\vec{p}|} \end{pmatrix} \begin{pmatrix} \Psi_{R/L} \\ \Psi_{R/L} \end{pmatrix} = \pm \frac{1}{2} \begin{pmatrix} \Psi_{R/L} \\ \Psi_{R/L} \end{pmatrix}, \quad (2.108)$$

indicating that the projected states are indeed right- or left-handed eigenstates with helicity $\pm \frac{1}{2}$.

This can be made more explicit by using a different representation for the γ -matrices. In the *chiral representation* (sometimes called the *Weyl representation*) we define the γ -matrices to be

$$\gamma^0 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \vec{\gamma} \equiv \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}, \quad (2.109)$$

so that, with $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ as before, the projection operators eq. (2.105) become

$$P_R = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad P_L = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \quad (2.110)$$

Now, the left-handed Weyl spinor sits in the upper two components of the Dirac spinor, while the right-handed Weyl spinor sits in the lower two components of the Dirac spinor. The projection operators pick out only the upper or lower component, e.g.

$$P_R \begin{pmatrix} \Psi_L \\ \Psi_R \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Psi_L \\ \Psi_R \end{pmatrix} = \begin{pmatrix} 0 \\ \Psi_R \end{pmatrix}, \quad (2.111)$$

so the projected states are once again helicity eigenstates.

3 Quantum Electrodynamics

3.1 Classical Electromagnetism

So far, we have only considered relativistic wave equations for free particles. Now we want to include electromagnetic interactions, so let's start by reviewing Maxwell's Equations in differential form:

$$\begin{aligned}\vec{\nabla} \cdot \vec{E} &= \rho, & \vec{\nabla} \cdot \vec{B} &= 0, \\ \vec{\nabla} \times \vec{E} &= -\frac{\partial \vec{B}}{\partial t}, & \vec{\nabla} \times \vec{B} &= \vec{J} + \frac{\partial \vec{E}}{\partial t}.\end{aligned}\tag{3.1}$$

Note here that I'm using Heaviside Lorentz units - I've used my freedom to choose the unit of charge to set $\epsilon_0 = 1$. Then, since in natural units $c = 1$, $\mu_0 = 1$ too. When one plays these games the value of the electron charge changes but the dimensionless quantity $\alpha = e^2/4\pi\epsilon_0\hbar c$ remains unchanged - $\alpha = 1/137$.

We can rewrite the Maxwell equations in terms of a scalar potential ϕ , and a vector potential \vec{A} . Writing

$$\begin{aligned}\vec{E} &= -\frac{\partial \vec{A}}{\partial t} - \vec{\nabla}\phi, \\ \vec{B} &= \vec{\nabla} \times \vec{A},\end{aligned}\tag{3.2}$$

we automatically have solutions of two of the Maxwell equations,

$$\vec{\nabla} \cdot \vec{B} = \vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) \equiv 0\tag{3.3}$$

and

$$\begin{aligned}\vec{\nabla} \times \vec{E} &= \vec{\nabla} \times \left(-\frac{\partial \vec{A}}{\partial t} - \vec{\nabla}\phi \right) \\ &= -\frac{\partial(\vec{\nabla} \times \vec{A})}{\partial t} - \vec{\nabla} \times (\vec{\nabla}\phi) \\ &= -\frac{\partial \vec{B}}{\partial t}.\end{aligned}\tag{3.4}$$

This simplifies things greatly since now there are only two Maxwell equations to solve.

Let's write them out in terms of the potentials,

$$\vec{\nabla} \cdot \vec{E} = -\nabla^2\phi - \frac{d(\vec{\nabla} \cdot \vec{A})}{dt} = \rho,\tag{3.5}$$

and (since $\vec{\nabla} \times \vec{\nabla} \times \vec{A} \equiv -\nabla^2\vec{A} + \vec{\nabla} \cdot (\vec{\nabla} \cdot \vec{A})$),

$$\vec{\nabla}(\vec{\nabla} \cdot \vec{A}) - \nabla^2\vec{A} = \vec{J} + \frac{\partial}{\partial t} \left(-\frac{\partial \vec{A}}{\partial t} - \vec{\nabla}\phi \right).\tag{3.6}$$

or rearranging,

$$-\nabla^2 \vec{A} + \frac{\partial^2 \vec{A}}{\partial t^2} = \vec{J} - \vec{\nabla}(\vec{\nabla} \cdot \vec{A} + \frac{\partial \phi}{\partial t}). \quad (3.7)$$

Unfortunately these two equations we are left with are quite complicated. To simplify them up we note that we can redefine our potentials,

$$\begin{aligned} \vec{A} &\rightarrow \vec{A} + \vec{\nabla}\psi, \\ \phi &\rightarrow \phi - \frac{\partial \psi}{\partial t}, \end{aligned} \quad (3.8)$$

without changing \vec{E} and \vec{B} . This redefinition of the potentials is known as a *gauge transformation*.

▷ **Exercise 3.17**

Check that \vec{E} and \vec{B} are invariant under the gauge transformation in eq. (3.8).

We can choose a gauge transformation such that

$$\vec{\nabla} \cdot \vec{A} = -\frac{\partial \phi}{\partial t}. \quad (3.9)$$

In this gauge (the Lorentz gauge) Maxwell's equations simplify to

$$-\nabla^2 \phi + \frac{\partial^2 \phi}{\partial t^2} = \rho, \quad (3.10)$$

$$-\nabla^2 \vec{A} + \frac{\partial^2 \vec{A}}{\partial t^2} = \vec{J}. \quad (3.11)$$

As well as being prettier, these equations also have a very suggestive form. They suggest we should define the 4-vectors,

$$J^\mu = (\rho, \vec{J}), \quad A^\mu = (\phi, \vec{A}), \quad (3.12)$$

so the Maxwell equations may be written in a manifestly covariant form,

$$\partial^2 A^\mu = J^\mu. \quad (3.13)$$

The $\mu = 0$ equation is the ϕ eq. (3.10) and the $\mu = 1, 2, 3$ equations give the components of the eq. (3.11) for \vec{A} . The gauge condition, eq. (3.9), becomes

$$\partial^\mu A_\mu = 0. \quad (3.14)$$

Eq. (3.13) is the classical wave equation for the electromagnetic field. In free space we have eq. (3.13) with no source, i.e.

$$\partial^2 A^\mu = 0, \quad (3.15)$$

which has plane wave solutions,

$$A^\mu = \epsilon^\mu e^{iq \cdot x}, \quad (3.16)$$

where ϵ^μ is the polarization tensor and $q^2 = 0$.

The Lorentz condition, eq. (3.14), enforces

$$q^\mu \epsilon_\mu = 0, \quad (3.17)$$

which removes one degree of freedom. Even after enforcing this condition, there is still room to make more gauge transformations,

$$A^\mu \rightarrow A^\mu + \partial^\mu \chi \quad \text{where} \quad \partial^2 \chi = 0. \quad (3.18)$$

This can be used to remove one extra degree of freedom from ϵ^μ . There are therefore two physical degrees of freedom, the normal polarizations of a photon.

3.2 The Dirac Equation in an Electromagnetic Field

We will now treat A^μ as a quantum mechanical wave function for photons. In the limit of a large number of photons the wave function is interpreted as a number density and produces the classical wave theory. But so far we have no interactions; to allow electrons to interact with electromagnetism we have to include the photon field into our Dirac equation.

The 'obvious' thing to do is to just be led by Lorentz invariance. The field A^μ is a vector field so we need to 'soak up' its free index with a γ -matrix. We therefore include it into the Dirac equation as

$$(i\gamma^\mu \partial_\mu - e\gamma_\mu A^\mu - m)\psi = 0, \quad (3.19)$$

where the factor of e is a free constant which quantifies how strongly the electron couples to the photon (the charge of the electron is $-e$).

It is convenient to incorporate this extra term into a new definition of a *covariant derivative*²,

$$D^\mu \equiv \partial^\mu + ieA^\mu. \quad (3.20)$$

Our interacting Dirac equation was therefore obtained from the free Dirac equation by the *minimal substitution* $\partial^\mu \rightarrow D^\mu$, and the Dirac equation becomes

$$(i\not{D} - m)\psi = 0. \quad (3.21)$$

There is a much nicer and theoretically much more appealing way to get the interaction term. That is if we require the QED Lagrangian to be invariant under a *local gauge symmetry* consisting of the transformations

$$\psi \rightarrow e^{-ie\Lambda(x)}\psi, \quad A^\mu \rightarrow A^\mu - \partial^\mu \Lambda(x). \quad (3.22)$$

then we are forced to the wave equation in eq. (3.21). For more details, I refer you to the Standard Model course.

We must also allow the electrons to enter into the photon wave equation but here the classical theory already tells us how a current density enters. We expect

$$\partial^2 A^\mu = J^\mu \quad (3.23)$$

where J^μ is just given by the charge times the Dirac equation number density ($-e\bar{\psi}\gamma^\mu\psi$).

²Conventions for the covariant derivative vary. *Halzen and Martin*, and *Mandl and Shaw* both use $D^\mu \equiv \partial^\mu - ieA^\mu$ whereas *Peskin and Schroeder* both use eq. (3.20). Both conventions define the electron charge to be $-e$ but differ by a sign in the definition of the photon field, A^μ .

3.3 $g - 2$ of the Electron

We now have a wave equation which describes how an electron behaves in an electromagnetic field, i.e. eq. (3.19). We will immediately put this to use by investigating the interaction between the spin of a non-relativistic electron and a magnetic field.

Writing the electron field in the form of eq. (2.29), we see that eq. (3.19) gives

$$\begin{pmatrix} \chi \\ \phi \end{pmatrix} = \begin{pmatrix} m & \vec{\sigma} \cdot (-i\vec{\nabla} - e\vec{A}) \\ \vec{\sigma} \cdot (-i\vec{\nabla} - e\vec{A}) & -m \end{pmatrix} \begin{pmatrix} \chi \\ \phi \end{pmatrix} \quad (3.24)$$

Substituting the equation from the second row into the that from the first leads to,

$$\left(E - m + \frac{[\vec{\sigma} \cdot (-i\vec{\nabla} - e\vec{A})]^2}{E + m} \right) \chi = 0. \quad (3.25)$$

We can simplify this somewhat by using to relation

$$\sigma_i \sigma_j = \delta_{ij} + i\epsilon_{ijk} \sigma_k, \quad (3.26)$$

to show

$$[\vec{\sigma} \cdot (-i\vec{\nabla} - e\vec{A})]^2 = |-i\vec{\nabla} - e\vec{A}|^2 - e(\vec{\nabla} \times \vec{A} + \vec{A} \times \vec{\nabla}) \cdot \vec{\sigma}, \quad (3.27)$$

and note

$$\vec{\nabla} \times \vec{A} \psi + \vec{A} \times \vec{\nabla} \psi = (\vec{\nabla} \times \vec{A}) \psi = \vec{B} \psi. \quad (3.28)$$

Putting all this together we find,

$$\left(E - m + \frac{|\vec{p} - e\vec{A}|^2 - e\vec{B} \cdot \vec{\sigma}}{E + m} \right) \chi = 0. \quad (3.29)$$

In the non-relativistic limit we can write $E \approx m$ and observe that the lower 2-component spinor is

$$\phi \approx \frac{\vec{\sigma} \cdot (\vec{p} - e\vec{A})}{2m} \chi \ll \chi. \quad (3.30)$$

This allows us to write, for the 4-component spinor ψ ,

$$\frac{1}{2m} |\vec{p} - e\vec{A}|^2 \psi - \frac{e\vec{B} \cdot \vec{\Sigma}}{2m} \psi = 0. \quad (3.31)$$

Notice that we have a coupling between the magnetic field \vec{B} and the spin of the electron $\vec{S} = \frac{1}{2}\vec{\Sigma}$. This is known as a *magnetic moment interaction* and takes the form

$$-\vec{\mu} \cdot \vec{B}. \quad (3.32)$$

Our Dirac equation in an electromagnetic field has predicted

$$\vec{\mu} = -\frac{e}{2m} \vec{\Sigma}. \quad (3.33)$$

In classical physics the *magnetic moment* of an orbiting charge is written

$$\vec{\mu}_{\text{orb}} = -\frac{e}{2mc}\vec{L}. \quad (3.34)$$

This is the magnetic moment associated with orbital angular momentum. By analogy we define the magnetic moment due to intrinsic angular momentum (i.e. spin) as

$$\vec{\mu}_{\text{spin}} = -g\frac{e}{2m}\vec{S} = -\frac{g}{2}\frac{e}{2m}\vec{\Sigma} \quad (3.35)$$

where g is the *gyromagnetic ratio* of the particle. The Dirac equation predicts

$$g = 2. \quad (3.36)$$

Experimentally one finds for the electron that

$$g = 2.0023193043738 \pm 0.0000000000082, \quad (3.37)$$

so the Dirac equations prediction is pretty close. It is not exactly correct, as we can see from the incredible precision with which this quantity has been measured. The discrepancy is due to us not yet including quantum corrections to our prediction. The interaction of an electron with a photon (and thus the gyromagnetic ratio) will be changed by processes of the form seen in fig. 3, and processes involving yet more particle loops. When

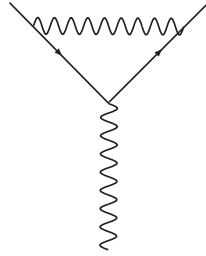


Figure 3: Quantum corrections to the electron-photon interaction.

one performs a more careful analysis, including these quantum effects, one predicts the deviation from 2 to be

$$\frac{g-2}{2} = 1 + \frac{\alpha}{2\pi} - 0.328 \left(\frac{\alpha}{\pi}\right)^2 + 1.181 \left(\frac{\alpha}{\pi}\right)^3 - 1.510 \left(\frac{\alpha}{\pi}\right)^4 + \dots + 4.393 \times 10^{-12}, \quad (3.38)$$

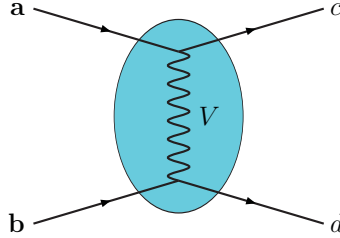
and comparing this prediction with experiment:

$$\begin{aligned} \text{Theory : } \quad \frac{g-2}{2} &= 1159652140(28) \times 10^{-12}, \\ \text{Experiment : } \quad \frac{g-2}{2} &= 1159652186.9(4.1) \times 10^{-12}. \end{aligned} \quad (3.39)$$

The figure in brackets denotes the error on the last significant figure. We can see that the experimental measurement matches the theoretical prediction to 8 significant figures, making this prediction of QED the most precisely tested prediction in physics.

3.4 Interactions in Perturbation Theory

The principle technique for computations of particle scatterings is perturbation theory - in other words we assume that the coupling $e \ll 1$ and expand about $e = 0$. We will be interested in processes such as



Outside the shaded interaction region we assume the particles are free. We will use the plane wave particle solutions derived in section 2.3 which, as noted in section 2.4, can only be normalized in a box of volume V . The shaded region is a sketch of this box - if we take a very large box then we expect it's presence to vanish from the answer for the scattering which is dominated when the particles are close and at the centre of the box. This will indeed be the case for our final cross-section results but we will need to keep track of factors of V for a while to see that result. I find it intuitive to have just one of each of the incoming and outgoing states in the box and to calculate the probability of that scatter occurring - I therefore pick the normalization $N = 1/\sqrt{2EV}$ from section 2.4. None of this analysis is Lorentz invariant but as the volume will factor out of our final results we will finally recover the Lorentz invariant forms for cross-sections.

To begin let's write the Dirac equation in a way that displays the smallness of the interaction

$$i\gamma^0 \frac{\partial \psi}{\partial t} + i\gamma^i \partial_i \psi - m\psi + \gamma^0 \delta V \psi = 0 \quad (3.40)$$

so for the electromagnetic interaction

$$\delta V = -e\gamma^0 \gamma^\mu A_\mu \quad (3.41)$$

Note that $(\gamma^0)^2 = 1$ so the γ^0 have been included simply for notational convenience. We will assume that the scattering particles begin in a pure \vec{p} state but the interaction then scatters them to another \vec{p} state with some (small) probability. In general we can write

$$\psi = \sum_n \kappa_n \phi_n(x) e^{iE_n t} \quad (3.42)$$

The $\phi_n(x)$ are the free Dirac equation solutions with n labelling the spinor state and the \vec{p} state. The κ_n are the probability amplitudes for the given state n . Before the interaction all the κ_n will be zero except one but during the interaction ($-T/2 < t < T/2$) we allow κ_n to change - $\kappa_n(t)$. If we now substitute the solution into the perturbed Dirac equation above then, at leading order, we obtain zero since we have expanded in solutions of the unperturbed equation. At next order we find

$$i\gamma_0 \sum_n \left(\frac{d\kappa_n}{dt} \right) \phi_n e^{-iE_n t} = \sum_n \gamma_0 \delta V \kappa_n \phi_n(x) e^{iE_n t} \quad (3.43)$$

Now we will make use of the orthogonality of the ϕ_n to extract the final state κ_n . We multiply through by $\int d^3x \phi_f^\dagger \gamma_0$

$$\frac{d\kappa_f}{dt} = -i \sum_n \kappa_n \int d^3x \phi_f^\dagger \delta V \phi_n e^{-i(E_n - E_f)t} \quad (3.44)$$

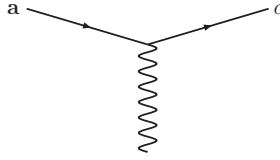
For a discussion of normalization of the spinors see section 2.4. Remembering that at $t = -T/2$ $\kappa_i = 1$ and $\kappa_{i \neq n} = 0$ at leading order we have

$$\frac{d\kappa_f}{dt} = -i \int \psi_f^\dagger \delta V \psi_i d^3x \quad (3.45)$$

and integrating with respect to t we find the important result

$$\kappa_f(T/2) = -i \int \psi_f^\dagger \delta V \psi_i d^4x \quad (3.46)$$

Now lets use our explicit form for δV in QED and concentrate on the scattering of a particle $a \rightarrow c$ by a photon A^μ



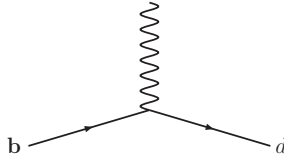
$$\begin{aligned} \kappa_{ca} &= -i \int \bar{\psi}^c (-e \gamma_\mu A^\mu) \psi^a d^4x \\ &= -i \int J_\mu^{ca} A^\mu d^4x \end{aligned} \quad (3.47)$$

where

$$J_\mu^{ca} = -e \bar{\psi}^c \gamma_\mu \psi^a = -e N_a N_c \bar{u}^c \gamma_\mu u^a e^{i(p_c - p_a) \cdot x} \quad (3.48)$$

The N s here are the normalizations of the spatial wave functions ψ again from section 2.5.

We're really interested in two particles scattering off each other so we'd better compute the A^μ field produced when another particle scatters from state $b \rightarrow d$



$$\square A^\mu = J_{db}^\mu = -e N_b N_d \bar{u}_d \gamma_\mu u_b e^{i(p_d - p_b) \cdot x} \quad (3.49)$$

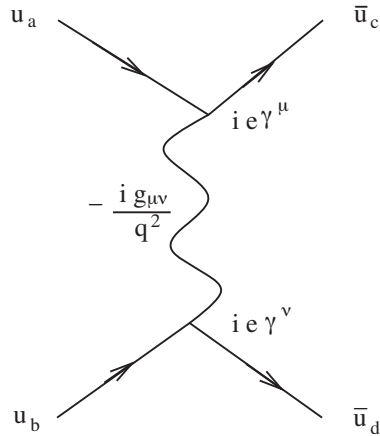
the solution is

$$A^\mu = -\frac{1}{q^2} J_{db}^\mu, \quad q = p_d - p_b \quad (3.50)$$

So finally substituting this back into our expression for κ_{ca} we find

$$\kappa_{fi} = -i N_a N_b N_c N_d \bar{u}^c (-e\gamma_\mu) u^a \left(-\frac{1}{q^2} \right) \bar{u}^d (-e\gamma^\mu) u^b \int e^{i(p_b + p_d - p_a - p_c) \cdot x} d^4x \quad (3.51)$$

Note that the integral is just a delta function that ensures 4-momentum conservation in the interaction. In order to make this result more memorable Feynman developed his famous rules that associate different parts of the expression with elements of a diagram of the scattering.



where momentum is conserved at the vertices. Multiplying these rules out gives us $-i\mathcal{M}_{fi}$ where

$$\kappa_{fi} = -i N_a N_b N_c N_d (2\pi)^4 \delta^4(p_f - p_i) \mathcal{M}_{fi} \quad (3.52)$$

▷ **Exercise 3.18**

Derive the Feynman rules for the scattering of two particles described by the Klein Gordon equation to leading order in e . You may Assume the form of the result in (3.46).

3.5 Internal Fermions and External Photons

We concentrated above on a scattering with external fermions interacting by the exchange of a photon. We can also imagine processes where there are external photon fields or internal virtual fermions. What are the Feynman rules for these cases? Given time constraints, rather than derive them, I'll present some simple arguments to motivate the rules. If we have an external photon interacting with a fermion in some way, then the vertex rule is still $-ie\gamma^\mu$. Since the amplitude we wish to calculate is Lorentz invariant we can not allow a stray μ index to survive but must soak it up with a 4-vector. The obvious 4-vector associated with external photon is its polarization vector ϵ^μ and indeed this is the appropriate factor for an external photon. Compare this to the way an external fermion closes the gamma matrix space indices, to give a number, with the external spinor.

We have seen that an internal photon (satisfying $\square A^\mu = 0$) generates a Feynman rule (or *propagator*)

$$\square A^\mu = 0 \quad \rightarrow \quad \frac{-ig_{\mu\nu}}{p^2} \quad (3.53)$$

Since a photon is just a collection of four scalar fields we can deduce that a massless, scalar field (which satisfies the KG equation $\square\phi = 0$) will have a Feynman rule

$$\square\phi = 0 \quad \rightarrow \quad \frac{i}{p^2} \quad (3.54)$$

It turns out that the sign is that of a space-like photon degree of freedom. To find the propagator of a massive scalar field we can treat the mass as a perturbing interaction of the free particle. Writing the KG equation as

$$\square\phi = -\delta V\phi = -m^2\phi \quad (3.55)$$

will generate a Feynman rule for the scalar self interaction

$$\begin{array}{c} -im \\ \text{---} \bullet \text{---} \end{array}$$

Now we can consider the set of perturbation theory diagrams that contribute to the full scalar propagator

$$\begin{array}{c} \text{---} + \text{---} \bullet \text{---} + \text{---} \bullet \text{---} \bullet \text{---} + \dots \\ \frac{i}{p^2} + \frac{i}{p^2}(-im)\frac{i}{p^2} + \frac{i}{p^2}(-im)\frac{i}{p^2}(-im)\frac{i}{p^2} + \dots \end{array}$$

$$\frac{i}{p^2} \rightarrow \frac{i}{p^2} + \frac{i}{p^2}(-im)\frac{i}{p^2} + \frac{i}{p^2}(-im)\frac{i}{p^2}(-im)\frac{i}{p^2} + \dots \quad (3.56)$$


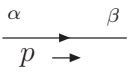
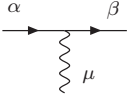
Pleasingly we can resum this series

$$\frac{i}{p^2} \left(1 + \frac{m^2}{p^2} + \frac{m^4}{p^4} + \dots \right) = \frac{i}{p^2} \left(\frac{1}{1 - \frac{m^2}{p^2}} \right) = \frac{i}{p^2 - m^2} \quad (3.57)$$

and this is indeed the full propagator in the massive case. By this point we can see that the propagator is basically just $-i$ times the inverse of the free field equation operator in momentum space. A sensible guess for the fermionic field is

$$(i\not{p} - m)\psi = 0 \quad \rightarrow \quad \frac{i}{\not{p} - m} = \frac{i}{\not{p} - m} \frac{\not{p} + m}{\not{p} + m} = \frac{i(\not{p} + m)}{p^2 - m^2} \quad (3.58)$$

This is in fact the correct answer. You will receive more insight into these results from the Field Theory course.

For every ...	draw ...	write ...
Internal photon line		$\frac{-ig^{\mu\nu}}{p^2 + i0^+}$
Internal fermion line		$\frac{i(\not{p} + m)_{\alpha\beta}}{p^2 - m^2 + i0^+}$
Vertex		$-ie\gamma_{\alpha\beta}^{\mu}$
Outgoing electron		$\bar{u}_{\alpha}(s, p)$
Incoming electron		$u_{\alpha}(s, p)$
Outgoing positron		$v_{\alpha}(s, p)$
Incoming positron		$\bar{v}_{\alpha}(s, p)$
Outgoing photon		$\varepsilon^{*\mu}(\lambda, p)$
Incoming photon		$\varepsilon^{\mu}(\lambda, p)$

- Attach a directed momentum to every internal line
- Conserve momentum at every vertex, i.e. include $\delta^{(4)}(\sum p_i)$
- Integrate over all internal momenta

Table 1: Feynman rules for QED. μ, ν are Lorentz indices, α, β are spinor indices and s and λ fix the polarization of the electron and photon respectively.

3.6 Summary of Feynman Rules of QED

The Feynman rules for computing the amplitude \mathcal{M}_{fi} for an arbitrary process in QED are summarized in Table 1.

The spinor indices in the Feynman rules are such that matrix multiplication is performed in the opposite order to that defining the flow of fermion number. The arrow on the fermion line itself denotes the fermion number flow, *not* the direction of the momentum associated with the line: I will try always to indicate the momentum flow separately as in Table 1. This will become clear in the examples which follow. We have already met the Dirac spinors u and v . I will say more about the photon polarization vector ε when we need to use it.

To summarize, the procedure for calculating the amplitude for any process in QED is the following:

1. Draw all possible distinct diagrams
2. Associate a directed 4-momentum with all lines
3. Apply the Feynman rules for the propagators, vertices and external legs
4. Ensure 4-momentum conservation at each vertex by adding $(2\pi)^4\delta^4(k_i - k_f)$, where k_i and k_f are the total incoming and outgoing 4-momenta of the vertex respectively

5. Perform the integration over all internal momenta with the measure $\int d^4k/(2\pi)^4$

It is also part of the Feynman rules for QED that when diagrams differ by an interchange of two fermion lines, a relative minus sign must be included. This is a reflection of Pauli's exclusion principle or equivalently of the anticommutation of the fermion operators discussed in the appendix. Note, however, that you don't need to get the absolute sign of an amplitude right, just its sign relative to the other amplitudes, since it is the modulus of the amplitude squared that we need ultimately. This sounds rather complicated. In particular there seem to be an awful lot of integrations to be done. However, at tree-level, i.e. if there are no loop diagrams, the delta functions attached to the vertices together with the integration over the internal momenta simply result in an overall 4-momentum conservation, i.e. in a factor $(2\pi)^4\delta^4(P_i - P_f)$, where P_i and P_f are the total incoming and outgoing 4-momenta of the process. Thus at tree-level, no 'real' integration has to be done. At one loop, however, there is one non trivial integration to be done. Generally, the calculation of an n -loop diagram involves n non trivial integrations. Even worse, these integrals very often are divergent. Still, we can get perfectly reasonable theoretical predictions at any order in QED. The procedure to get these results is called renormalization and will be the topic of section 6. At this point, some remarks concerning step 1, i.e. drawing all possible distinct Feynman diagrams, might be useful. In order to establish whether two diagrams are distinct, we have to try to convert one into the other. If this is possible without cutting lines and without gluing lines – that is solely by twisting and stretching the lines and rotating the whole figure – then the two diagrams are identical. It should be noted that the external lines are labeled in this process. Therefore, the two diagrams shown in figure 6 are different. Finally, let me mention that the diagrams shown in figure 1 are not Feynman diagrams. When drawing Feynman diagrams we are only interested in what particles are incoming and which ones are outgoing and there is no time direction involved.

4 Cross Sections and Decay Rates

Before explicitly calculating some transition amplitudes let's see how to connect those amplitudes to physical observables such as cross sections and particle widths.

4.1 Transition Rate

Consider an arbitrary scattering process with an initial state i with total 4-momentum P_i and a final state f with total 4-momentum P_f . Let's assume we computed the scattering amplitude for this process in QED, i.e. we know the matrix element

$$-i \prod_{f=1}^N N_f \prod_{\text{in}} N_i \mathcal{M}_{fi} (2\pi)^4 \delta^4(P_f - P_i) \quad (4.1)$$

Our task in this section is to convert this into a scattering cross section (relevant if there is more than 1 particle in the initial state) or a decay rate (relevant if there is just 1 particle in the initial state), see figure 4.

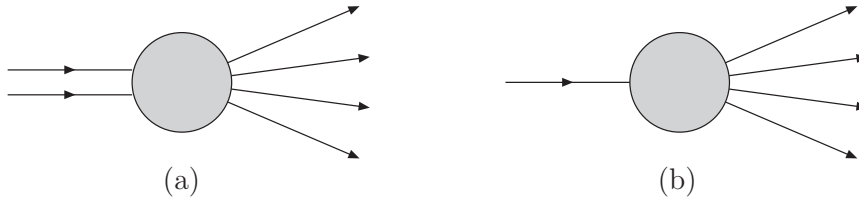


Figure 4: Scattering (a) and decay (b) processes.

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

$$\text{Probability} = \left| -i \prod_{f=1}^N N_f \prod_{\text{in}} N_i \mathcal{M}_{fi} (2\pi)^4 \delta^4(P_f - P_i) \right|^2. \quad (4.2)$$

Attempting to take the squared modulus of the amplitude produces a meaningless square of a delta function. This is a technical problem because our amplitude is expressed between plane wave states. These states are states of definite momentum and so extend throughout all of space-time. In a real experiment the incoming and outgoing states are localized (e.g. they might leave tracks in a detector). To deal with this properly we would have to construct normalized wave packet states which do become well separated in the far past and the far future. A sloppier derivation is to maintain that our interaction is occurring in a box of volume $V = L^3$ and over a time of order T . The final answers will come out independent of V and T , reproducing the ones we would get if we worked with localized wave packets. Using

$$(2\pi)^4 \delta^4(P_f - P_i) = \int e^{i(P_f - P_i)x} d^4x \quad (4.3)$$

we get in our space-time box the result

$$|(2\pi)^4 \delta^4(P_f - P_i)|^2 \simeq (2\pi)^4 \delta^4(P_f - P_i) \int e^{i(P_f - P_i)x} d^4x \simeq VT (2\pi)^4 \delta^4(P_f - P_i). \quad (4.4)$$

We must also use the explicit expressions for the wave function normalizations from section 2.4. Above we used the normalization $N = 1/\sqrt{2EV}$. So putting everything together, we find for the transition rate W , i.e. the probability per unit time

$$W = \frac{1}{T} |\mathcal{M}_{fi}|^2 VT (2\pi)^4 \delta^4(P_f - P_i) \prod_{f=1}^N \left[\frac{1}{2E_f V} \right] \prod_{\text{in}} \left[\frac{1}{2E_i V} \right]. \quad (4.5)$$

As expected, the dependence on T cancelled. Usually we are interested in much more detailed information than just the total transition rate. We want to know the differential transition rate dW , i.e. the transition rate into a particular element of the final state phase space. To get dW we have to multiply by the number of available states in the (small) part of phase space under consideration. For a single particle final state, the number of available states dn in some momentum range \vec{k} to $\vec{k} + d\vec{k}$ is, in the box normalization,

$$dn = V d^3 \vec{k} \quad (4.6)$$

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

$$dn = dn_1 dn_2$$

where

$$dn_1 = V d^3 \vec{k}_1, \quad dn_2 = V d^3 \vec{k}_2,$$

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

$$dn = \prod_{f=1}^N \frac{V d^3 \vec{k}_f}{(2\pi)^3}. \quad (4.7)$$

The transition rate for transitions into a particular element of final state phase space is thus given by, using equations (4.7) and (4.5),

$$\begin{aligned} dW &= |\mathcal{M}_{fi}|^2 (2\pi)^4 \delta^4(P_f - P_i) V \prod_{f=1}^N \left[\frac{1}{2E_f V} \right] \prod_{\text{in}} \left[\frac{1}{2E_i V} \right] \prod_{f=1}^N \frac{V d^3 \vec{k}_f}{(2\pi)^3} \\ &= |\mathcal{M}_{fi}|^2 V \prod_{\text{in}} \left[\frac{1}{2E_i V} \right] \times \text{LIPS}(N), \end{aligned} \quad (4.8)$$

where in the second step we defined the Lorentz invariant phase space with N particles in the final state

$$\text{LIPS}(N) \equiv (2\pi)^4 \delta^4(P_f - P_i) \prod_{f=1}^N \frac{d^3 \vec{k}_f}{(2\pi)^3 2E_f}. \quad (4.9)$$

Observe that everything in the transition rate is Lorentz invariant save for the initial energy factor and the factors of V .

▷ **Exercise 4.19**

Show that $d^3 k/2E$ is a Lorentz-invariant element of phase space. (Hint: Think how you would write the phase space in a 4-dimensional, integral but with the particle on-shell, i.e. $E = (\vec{k}^2 + m^2)^{1/2}$).

4.2 Decay Rates

We turn now to the special case where we have only one particle with mass m in the initial state i , i.e. we consider the decay of this particle into some final state f . In this case, the transition rate is called the partial decay rate and denoted by Γ_{if} . First of all, we observe that the dependence on V cancels, as advertised above. In the rest frame of the particle the partial decay rate is given by

$$\Gamma_{if} = \frac{1}{2m} \int |\mathcal{M}_{fi}|^2 \times \text{LIPS} \quad (4.10)$$

The important special case of two particles in the final state deserves further consideration. Consider the partial decay rate for a particle i of mass m into two particles f_1 and f_2 . The Lorentz-invariant phase space is

$$\text{LIPS}(N) = (2\pi)^4 \delta^4(p_i - p_1 - p_2) \frac{d^3\vec{p}_1}{(2\pi)^3 2E_1} \frac{d^3\vec{p}_2}{(2\pi)^3 2E_2}. \quad (4.11)$$

In the rest frame the four-vectors of each particle are

$$p_i = (m, 0), \quad p_1 = (E_1, \vec{p}), \quad p_2 = (E_2, -\vec{p}). \quad (4.12)$$

Therefore we can eliminate one three-momentum in the phase space

$$\text{LIPS}(N) = \frac{1}{(2\pi)^2} \delta(m - E_1 - E_2) \frac{d^3\vec{p}_2}{4E_1 E_2}. \quad (4.13)$$

Hence the partial decay rate becomes

$$\Gamma_{if} = \frac{1}{8m(2\pi)^2} \int |\mathcal{M}_{fi}|^2 \delta(m - E_1 - E_2) \frac{d|\vec{p}_f|^2 |\vec{p}_f| d\Omega^*}{E_1 E_2} \quad (4.14)$$

where $d\Omega^*$ is the solid angle element for the angle of one of the outgoing particles with respect to some fixed direction, and \vec{p}_f is the momentum of one of the final state particles. But from the on-shell condition $E_1 = (\vec{p}_1^2 + m_1^2)^{1/2}$, we have $dE_1 = |\vec{p}_f|/E_1 d|\vec{p}_f|$ and similarly for particle 2 and so

$$d(E_1 + E_2) = |\vec{p}_f| d|\vec{p}_f| \frac{E_1 + E_2}{E_1 E_2},$$

therefore

$$|\vec{p}_f|^2 d|\vec{p}_f| \frac{1}{E_1 E_2} = \frac{|\vec{p}_f|}{E_1 + E_2} d(E_1 + E_2). \quad (4.15)$$

Using this in eq. (4.14) and integrating over $(E_1 + E_2)$ we obtain the final result

$$\Gamma_{i \rightarrow f_1 f_2} = \frac{1}{32\pi^2 m^2} \int |\mathcal{M}_{fi}|^2 |\vec{p}_f| d\Omega^*. \quad (4.16)$$

The total decay rate of particle i is obtained by summation of the partial decay rates into all possible final states

$$\Gamma_{\text{tot}} = \sum_f \Gamma_{if} \quad (4.17)$$

The total decay rate is related to the mean life time τ via $(\Gamma_{\text{tot}})^{-1} = \tau$. For completeness I also give the definition of the branching ratio for the decay into a specific final state f

$$B_f \equiv \frac{\Gamma_{if}}{\Gamma_{\text{tot}}} \quad (4.18)$$

In an arbitrary frame we find, $W = (m/E)\Gamma_{\text{tot}}$, which has the expected Lorentz dilation factor. In the master formula (equation (4.8)) this is what the product of $1/2E_i$ factors for the initial particles does.

4.3 Cross Sections

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

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

particles per unit area per unit time. Thus the differential cross section is related to the differential transition rate in equation (4.8) by

$$d\sigma = \frac{dW}{N_0} = dW \times \frac{V}{v}. \quad (4.19)$$

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

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

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

$$d\sigma = \frac{dW}{N_0} = \frac{1}{|\vec{v}_1 - \vec{v}_2|} \frac{1}{4E_1E_2} |\mathcal{M}_{fi}|^2 \times \text{LIPS} \quad (4.20)$$

where we have used equation (4.8) for the transition rate, and the box volume V has again canceled. The amplitude-squared and phase space factors are manifestly Lorentz invariant. What about the initial velocity and energy factors? Observe that

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

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

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

and the last expression is manifestly Lorentz invariant.

▷ **Exercise 4.20**

Prove that $|E_2\vec{p}_1 - E_1\vec{p}_2|^2 = (p_1 \cdot p_2)^2 - m_1^2 m_2^2$ in a frame where the momenta are collinear. Hence we can define a Lorentz invariant differential cross section. The total cross section is obtained by integrating over the final state phase space:

$$\sigma = \frac{1}{|\vec{v}_1 - \vec{v}_2|} \frac{1}{4E_1 E_2} \sum_{\text{final states}} \int |\mathcal{M}_{fi}|^2 \times \text{LIPS}. \quad (4.21)$$

A slight word of caution is needed in deciding on the limits of integration to get the total cross section. If there are identical particles in the final state then the phase space should be integrated so as not to double count. An important special case is $2 \rightarrow 2$ scattering

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

▷ **Exercise 4.21**

Show that in the centre-of-mass frame the differential cross section for the scattering $a(p_a) + b(p_b) \rightarrow c(p_c) + d(p_d)$ is

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2 s} \frac{|\vec{p}_c|}{|\vec{p}_a|} |\mathcal{M}_{fi}|^2. \quad (4.22)$$

4.4 Mandelstam Variables

Invariant $2 \rightarrow 2$ scattering amplitudes are frequently expressed in terms of the *Mandelstam variables*. These are defined by

$$\begin{aligned} s &\equiv (p_a + p_b)^2 = (p_c + p_d)^2, \\ t &\equiv (p_a - p_c)^2 = (p_b - p_d)^2, \\ u &\equiv (p_a - p_d)^2 = (p_b - p_c)^2. \end{aligned} \quad (4.23)$$

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

▷ **Exercise 4.22**

Show that

$$s + t + u = m_a^2 + m_b^2 + m_c^2 + m_d^2. \quad (4.24)$$

▷ **Exercise 4.23**

Show that, for two body scattering of particles of equal mass m ,

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

(Hint: since all variables are invariant work in the centre of mass frame.)

5 Processes in QED and QCD

5.1 Electron–Muon Scattering

This is as simple a process as one can find since at lowest order in the electromagnetic coupling, just one diagram contributes. It is shown in figure 5. The amplitude obtained by applying the Feynman rules to this diagram is

$$i\mathcal{M}_{fi} = ie \bar{u}(p_c)\gamma^\mu u(p_a) \left(\frac{-ig_{\mu\nu}}{q^2} \right) ie \bar{u}(p_d)\gamma^\nu u(p_b), \quad (5.1)$$

where $q^2 = (p_a - p_c)^2$. Note that, for clarity, I have dropped the spin label on the spinors. I will restore it when I need to. In constructing this amplitude we have followed the fermion lines backwards with respect to fermion flow when working out the order of matrix multiplication (which makes sense if you think of an unbarred spinor as a column vector and a barred spinor as a row vector and remember that the amplitude carries no spinor indices).

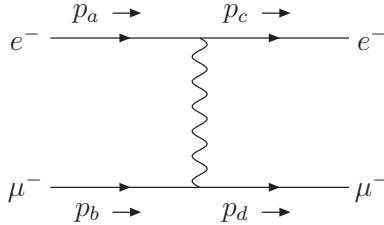


Figure 5: Lowest order Feynman diagram for $e^- \mu^- \rightarrow e^- \mu^-$ scattering.

The cross section involves the squared modulus of the amplitude, $|\mathcal{M}_{fi}|^2$. Let us see how we obtain a neat form for this. The hermitian conjugate of a ‘spinor sandwich’ is the same as its hermitian conjugate,

$$(\bar{u}(p_c)\gamma^\mu u(p_a))^* = (\bar{u}(p_c)\gamma^\mu u(p_a))^\dagger$$

since it is just a number. Using rules of matrix algebra we see that this is

$$\begin{aligned} (u(p_c)^\dagger \gamma^0 \gamma^\mu u(p_a))^\dagger &= (u(p_a)^\dagger \gamma^{\mu\dagger} \gamma^{0\dagger} u(p_c)) \\ &= (u(p_a)^\dagger \gamma^{\mu\dagger} \gamma^0 u(p_c)). \end{aligned} \quad (5.2)$$

But in section 2.6 we saw that $\gamma^0 \gamma^{\mu\dagger} \gamma^0 = \gamma^\mu$, and so this becomes

$$(\bar{u}(p_c)\gamma^\mu u(p_a))^* = \bar{u}(p_a)\gamma^\mu u(p_c). \quad (5.3)$$

▷ Exercise 5.24

If Γ represents a string of γ -matrices (not including γ^5) and Γ_R is its reverse (i.e. the same γ -matrices in reverse order), show that,

$$[\bar{u}(k')\Gamma u(k)]^* = \bar{u}(k)\Gamma_R u(k').$$

Using this result in the expression for $|\mathcal{M}_{fi}|^2$ we obtain

$$\begin{aligned} |\mathcal{M}_{fi}|^2 &= \frac{e^4}{q^4} \bar{u}(p_c) \gamma^\mu u(p_a) \bar{u}(p_d) \gamma_\mu u(p_b) \bar{u}(p_a) \gamma^\nu u(p_c) \bar{u}(p_b) \gamma_\nu u(p_d) \\ &= \frac{e^4}{q^4} L_{(e)}^{\mu\nu} L_{(\mu)\mu\nu}, \end{aligned} \quad (5.4)$$

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

$$L_{(e)}^{\mu\nu} = \bar{u}(p_c) \gamma^\mu u(p_a) \bar{u}(p_a) \gamma^\nu u(p_c),$$

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

Usually we have an unpolarized beam and target and do not measure the polarization of the outgoing particles. Thus we calculate the squared amplitudes for each possible spin combination, then average over initial spin states and sum over final spin states. Note that we square and then sum since the different spin configurations are in principle distinguishable. In contrast, if several Feynman diagrams contribute to the same process, you have to sum the amplitudes first. We will see examples of this below.

The spin sums are made easy by the results

$$\begin{aligned} \sum_s u^{(s)}(p) \bar{u}^{(s)}(p) &= \not{p} + m, \\ \sum_s v^{(s)}(p) \bar{v}^{(s)}(p) &= \not{p} - m. \end{aligned} \quad (5.5)$$

Do not forget that by m , we really mean m times the unit 4×4 matrix.

▷ **Exercise 5.25**

Prove eq. (5.5).

Using the spin sums we find that

$$\begin{aligned} \sum_{\text{spins}} L_{(e)}^{\mu\nu} &= \sum_{s_a, s_c} \bar{u}_{\alpha}^{(s_c)}(p_c) \gamma_{\alpha\beta}^{\mu} u_{\beta}^{(s_a)}(p_a) \bar{u}_{\rho}^{(s_a)}(p_a) \gamma_{\rho\sigma}^{\nu} u_{\sigma}^{(s_c)}(p_c) \\ &= \gamma_{\alpha\beta}^{\mu} [\not{p}_a + m_e]_{\beta\rho} \gamma_{\rho\sigma}^{\nu} [\not{p}_c + m_e]_{\sigma\alpha} \\ &= \text{Tr}(\gamma^{\mu} (\not{p}_a + m_e) \gamma^{\nu} (\not{p}_c + m_e)). \end{aligned} \quad (5.6)$$

where in the first line, we have make explicit the spinor indices in order to show how the trace emerges. Since all calculations of cross sections or decay rates in QED require the evaluation of traces of products of γ -matrices, you will generally find a table of “trace theorems” in any quantum field theory textbook [1]. All these theorems can be derived from the fundamental anti-commutation relations of the γ -matrices in eq. (2.58) together with the invariance of the trace under a cyclic change of its arguments. For now it suffices to use

$$\begin{aligned} \text{Tr}(\gamma^{\mu_1} \dots \gamma^{\mu_n}) &= 0 \quad \text{for } n \text{ odd} \\ \text{Tr}(\gamma^{\mu_1} \dots \gamma^{\mu_n}) &= g^{\mu_1 \mu_2} \text{Tr}(\gamma^{\mu_3} \dots \gamma^{\mu_n}) - g^{\mu_1 \mu_3} \text{Tr}(\gamma^{\mu_2} \gamma^{\mu_4} \dots \gamma^{\mu_n}) + \dots \\ &\quad + g^{\mu_1 \mu_n} \text{Tr}(\gamma^{\mu_2} \dots \gamma^{\mu_{n-1}}) \\ \text{Tr}(\not{a} \not{b}) &= 4 a \cdot b, \\ \text{Tr}(\not{a} \not{b} \not{c} \not{d}) &= 4(a \cdot b c \cdot d - a \cdot c b \cdot d + a \cdot d b \cdot c). \end{aligned} \quad (5.7)$$

▷ **Exercise 5.26**

Derive the trace results in equation (5.7). (Hint: for the first one use $(\gamma^5)^2 = 1$.)

Using these trace theorems,

$$\sum_{\text{spins}} L_{(e)}^{\mu\nu} = 4(p_a^\mu p_c^\nu - g^{\mu\nu} p_a \cdot p_c + p_a^\nu p_c^\mu) + 4g^{\mu\nu} m_e^2, \quad (5.8)$$

with a similar result for $L_{(\mu)}^{\mu\nu}$. Putting this altogether, the spin averaged/summed amplitude squared is

$$\begin{aligned} & \frac{1}{4} \sum_{\text{spins}} |\mathcal{M}_{fi}|^2 \\ &= \frac{e^4}{q^4} 4 \left(p_a^\mu p_c^\nu + p_a^\nu p_c^\mu - (p_a \cdot p_c - m_e^2) g^{\mu\nu} \right) \left(p_{b\mu} p_{d\nu} + p_{b\nu} p_{d\mu} - (p_b \cdot p_d - m_\mu^2) g_{\mu\nu} \right) \\ &= \frac{8e^4}{q^4} \left((p_c \cdot p_d)(p_a \cdot p_b) + (p_c \cdot p_b)(p_a \cdot p_d) - m_e^2(p_b \cdot p_d) - m_\mu^2(p_a \cdot p_c) + 2m_e^2 m_\mu^2 \right). \end{aligned} \quad (5.9)$$

(Notice that we have divided by 4 since we are *averaging* over initial states, and there are 4 possible initial spin configurations.)

This takes on a more compact form if expressed in terms of the Mandelstam variables of eq. (4.23),

$$\frac{1}{4} \sum_{\text{spins}} |\mathcal{M}_{fi}|^2 = \frac{2e^4}{t^2} (s^2 + u^2 - 4(m_e^2 + m_\mu^2)(s + u) + 6(m_e^2 + m_\mu^2)^2). \quad (5.10)$$

Finally, we can derive the differential cross section for this process in the centre-of-mass frame using eq. (4.22). In the high energy limit where $s, |u| \gg m_e^2, m_\mu^2$, i.e. setting the masses to zero,

$$\frac{d\sigma}{d\Omega} = \frac{e^4}{32\pi^2 s} \frac{s^2 + u^2}{t^2}. \quad (5.11)$$

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

5.2 Electron–Electron Scattering

For the scattering $e^-e^- \rightarrow e^-e^-$ we now have identical particles in the final state which may only be distinguished by their momenta. Therefore we cannot just replace m_μ by m_e in the calculation we performed above. Labeling the momenta in the process according to $e^-(p_a) + e^-(p_b) \rightarrow e^-(p_c) + e^-(p_d)$ in analogy to $e^-\mu^-$ scattering, we realize that when particle a emits a photon we do not know whether it ‘becomes’ particle c (as it did in the $e^-\mu^-$ scattering) or ‘becomes’ particle d . Since either is possible, we need to include both cases, resulting in the two diagrams of fig. 6. Applying the Feynman rules, the two

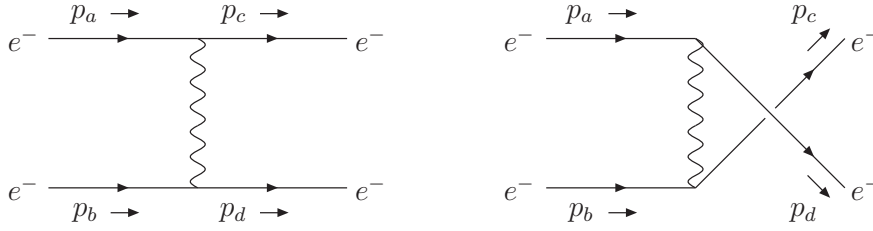


Figure 6: Lowest order Feynman diagrams for electron–electron scattering.

diagrams give the amplitudes,

$$i\mathcal{M}_1 = \frac{ie^2}{t} \bar{u}(p_c) \gamma^\mu u(p_a) \bar{u}(p_d) \gamma_\mu u(p_b), \quad (5.12)$$

$$i\mathcal{M}_2 = -\frac{ie^2}{u} \bar{u}(p_d) \gamma^\mu u(p_a) \bar{u}(p_c) \gamma_\mu u(p_b). \quad (5.13)$$

Notice the additional minus sign in the second amplitude, which comes from the anti-commuting nature of fermion fields. Remember that when diagrams differ by an interchange of two fermion lines, a relative minus sign must be included. This is important because

$$\begin{aligned} |\mathcal{M}_{fi}|^2 &= |\mathcal{M}_1 + \mathcal{M}_2|^2 \\ &= |\mathcal{M}_1|^2 + |\mathcal{M}_2|^2 + 2\text{Re}\mathcal{M}_1^* \mathcal{M}_2, \end{aligned} \quad (5.14)$$

so the interference term will have the wrong sign if you don't include the extra sign difference between the two diagrams. $|\mathcal{M}_1|^2$ and $|\mathcal{M}_2|^2$ are very similar to the previous calculation. The interference term is a little more complicated due to a different trace structure.

Performing the calculation explicitly yields (in the limit of negligible fermion masses),

$$\frac{1}{4} \sum_{\text{spins}} |\mathcal{M}_{fi}|^2 = 2e^4 \left(\frac{s^2 + u^2}{t^2} + \frac{s^2 + t^2}{u^2} + \frac{2s^2}{tu} \right). \quad (5.15)$$

▷ **Exercise 5.27**

Prove the result in eq. (5.15). It will be helpful first to prove

$$\begin{aligned} \gamma^\alpha \gamma^\mu \gamma_\alpha &= -2\gamma^\mu \\ \gamma^\alpha \gamma^\mu \gamma^\nu \gamma_\alpha &= 4g^{\mu\nu} \\ \gamma^\alpha \gamma^\mu \gamma^\nu \gamma^\rho \gamma_\alpha &= -2\gamma^\rho \gamma^\nu \gamma^\mu. \end{aligned} \quad (5.16)$$

5.3 Electron–Positron Annihilation

The two diagrams e^+e^- scattering are shown in fig. 7, with the one on the right known as the annihilation diagram. They are just what you get from the diagrams for electron–electron scattering in fig. 6 if you twist round the fermion lines. The fact that the

diagrams are related in this way implies a relation between the amplitudes. The interchange of incoming particles/antiparticles with outgoing antiparticles/particles is called *crossing*. For our particular example, the squared amplitude for $e^+e^- \rightarrow e^+e^-$ is related to that for $e^-e^- \rightarrow e^-e^-$ by performing the interchange $s \leftrightarrow u$. Hence, squaring the amplitude and doing the traces yields (again neglecting fermion mass terms)

$$\frac{1}{4} \sum_{\text{spins}} |\mathcal{M}_{fi}|^2 = 2e^4 \left(\frac{s^2 + u^2}{t^2} + \frac{u^2 + t^2}{s^2} + \frac{2u^2}{ts} \right). \quad (5.17)$$

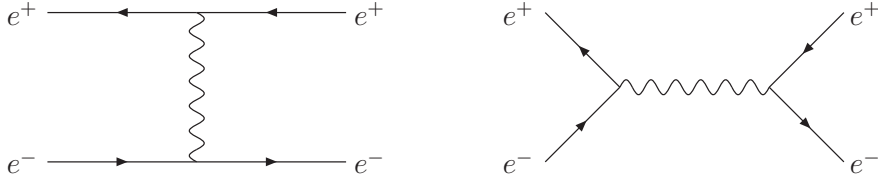


Figure 7: Lowest order Feynman diagrams for electron-positron scattering in QED.

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

$$R \equiv \frac{\sigma(e^+e^- \rightarrow \text{hadrons})}{\sigma(e^+e^- \rightarrow \mu^+\mu^-)} = 3 \sum_f Q_f^2, \quad (5.18)$$

where the sum is over quark flavours f and Q_f is the quark’s charge in units of e . The 3 comes from the existence of three colours for each flavour of quark. Historically this was important: you could look for a step in the value of R as your e^+e^- collider’s CM energy rose through a threshold for producing a new quark flavour. If you did not know about colour, the height of the step would seem too large. At the energies used at LEP you have to remember to include the diagram with a Z replacing the photon.

Finally, we compute the total cross section for $e^+e^- \rightarrow \mu^+\mu^-$, neglecting the lepton masses. Here we only have the annihilation diagram, and for the amplitude, we get

$$\begin{aligned} \mathcal{M}_{fi} &= (-ie)^2 \bar{u}(p_d) \gamma^\mu v(p_c) \frac{-ig_{\mu\nu}}{s} \bar{v}(p_a) \gamma^\nu u(p_b) \\ &= \frac{ie^2}{s} \bar{u}_d \gamma^\mu v_c \bar{v}_a \gamma_\mu u_b. \end{aligned} \quad (5.19)$$

Summing over final state spins and averaging over initial spins gives,

$$\frac{1}{4} \sum_{\text{spins}} |\mathcal{M}_{fi}|^2 = \frac{e^4}{4s^2} \text{Tr}(\gamma^\mu \not{p}_c \gamma^\nu \not{p}_d) \text{Tr}(\gamma_\mu \not{p}_b \gamma_\nu \not{p}_a),$$

where we have neglected m_e and m_μ . Using the results in equation (5.7) to evaluate the traces gives,

$$\frac{1}{4} \sum_{\text{spins}} |\mathcal{M}_{fi}|^2 = \frac{8e^4}{s^2} (p_a \cdot p_d p_b \cdot p_c + p_a \cdot p_c p_b \cdot p_d).$$

Neglecting masses we have,

$$p_a \cdot p_c = p_b \cdot p_d = -t/2, \quad (5.20)$$

$$p_a \cdot p_d = p_b \cdot p_c = -u/2. \quad (5.21)$$

Hence

$$\frac{1}{4} \sum_{\text{spins}} |\mathcal{M}_{fi}|^2 = 2e^4 \frac{t^2 + u^2}{s^2}, \quad (5.22)$$

which incidentally is what you get by applying crossing to the electron–muon amplitude of section 5.1. We can use this in eq. (4.22) to find the differential cross section in the CM frame,

$$\frac{d\sigma}{d\Omega} = \frac{e^4}{32\pi^2 s} \frac{t^2 + u^2}{s^2}. \quad (5.23)$$

You could get straight to this point by noting that the appearance of v spinors instead of u spinors in \mathcal{M}_{fi} does not change the answer since only quadratic terms in m_μ survive the Dirac algebra and we go on to neglect masses anyway. Hence you can use the result of eq. (5.11) with appropriate changes.

Neglecting masses, the CM momenta are

$$p_a = \frac{1}{2}\sqrt{s}(1, \vec{e}) \quad p_c = \frac{1}{2}\sqrt{s}(1, \vec{e}') \quad (5.24)$$

$$p_b = \frac{1}{2}\sqrt{s}(1, \vec{e}) \quad p_d = \frac{1}{2}\sqrt{s}(1, \vec{e}') \quad (5.25)$$

which gives $t = -s(1 - \cos\theta)/2$ and $u = -s(1 + \cos\theta)/2$, where $\cos\theta = \vec{e} \cdot \vec{e}'$. Hence, finally, the total cross section is,

$$\sigma = \int_{-1}^1 \frac{d\sigma}{d\Omega} 2\pi d(\cos\theta) = \frac{4\pi\alpha^2}{3s}. \quad (5.26)$$

5.4 Compton Scattering

The diagrams which need to be evaluated to compute the Compton cross section for $\gamma e \rightarrow \gamma e$ are shown in fig. 8. For unpolarized initial and/or final states, the cross section calculation involves terms of the form

$$\sum_{\lambda} \varepsilon^{*\mu}(\lambda, p) \varepsilon^{\nu}(\lambda, p), \quad (5.27)$$

where λ represents the polarization of the photon of momentum p . Since the photon is massless, the sum is over the two transverse polarization states, and must vanish when contracted with p_μ or p_ν . In principle eq. (5.27) is a complicated object. However, there is a simplification as far as the amplitude calculation is concerned. The photon is coupled

to the electromagnetic current $J^\mu = \bar{\psi}\gamma^\mu\psi$ of eq. (2.28). This is a conserved current, i.e. $\partial_\mu J^\mu = 0$, and in momentum space $p_\mu J^\mu = 0$. Hence, any term in the polarization sum, eq. (5.27), proportional to p^μ or p^ν does not contribute to the cross section. This means that in calculations one can make the replacement

$$\sum_\lambda \varepsilon^{*\mu}(\lambda, p) \varepsilon^\nu(\lambda, p) \rightarrow -g^{\mu\nu}, \quad (5.28)$$

and we have a simple, Lorentz-covariant prescription.

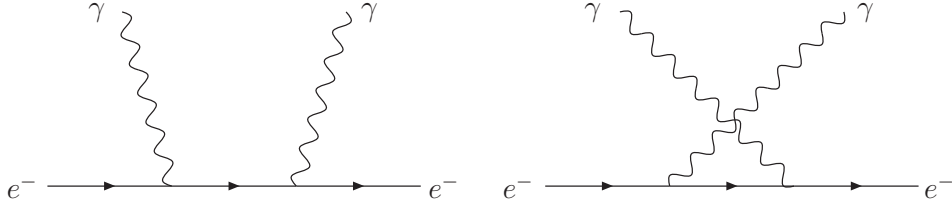


Figure 8: Lowest order Feynman diagrams for Compton scattering.

▷ **Exercise 5.28**

Show that the spin summed/averaged squared matrix element for Compton scattering in the massless limit is given by

$$|\mathcal{M}_{fi}|^2 = 2e^4 \left(-\frac{u}{s} - \frac{s}{u} \right) \quad (5.29)$$

Evaluate the total cross section using the expressions in the centre-of-mass frame at the end of the last sub-section. Why does this create a problem?

5.5 QCD Processes

The theory of quarks and gluons, QCD, is in many ways very similar to QED. We have done most of the hard work to calculate tree level amplitudes already. The main difference between the theories is that QCD has three types of charges (called ‘colours’, e.g. red, green and blue). We can write a quark as a vector with the three colour states shown

$$u = \begin{pmatrix} u^R \\ u^G \\ u^B \end{pmatrix} \quad (5.30)$$

There are more possible interactions than in QED which are mediated by eight photon-like gauge fields called “gluons”. We encode the couplings of the gluons to the quarks by matrices which act on the above colour vector. For example there are two gluons with matrix “generators”

$$T^1 = \frac{1}{\sqrt{12}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad T^2 = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \quad (5.31)$$

These are just photon-like interactions with each of the two photons having different couplings to the different colours.

▷ **Exercise 5.29**

Check that the strength of a colour anti-colour quark pair scattering to itself at tree level is the same no matter which colour you pick. Show that the strength of a scattering of a colour anti-colour quark pair to a different colour pair is also the same no matter what colours you pick.

The remaining six gluons change the colour of the quark and are associated with generators of the form

$$T^3 = \frac{1}{2} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad T^4 = \frac{1}{2} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (5.32)$$

The remaining four generators are of the same form but interchange the other two colour combinations. Note these matrices are traceless and normalized so that $\text{Tr} T^a T^b = \frac{1}{2} \delta^{ab}$.

You will learn more about the origin of these fields and their couplings in the Standard Model course. From the point of view of calculating cross sections though the Feynman Rules are all we need to proceed, and these are very similar to those of QED. The generator T^a is included in the Feynman rule for the gluon–quark–anti-quark vertex as shown in fig. 9 (upper), where g is the QCD coupling constant. Also, since a gluon

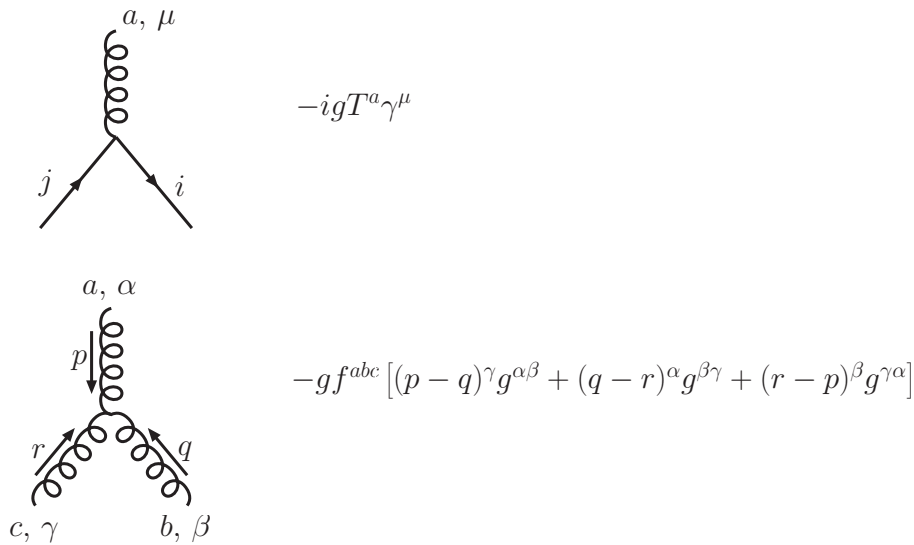


Figure 9: An example of some QCD Feynman Rules.

associated with, for example, T^3 can pair produce a red quark and an anti-green quark we see that the gluons themselves are charged. Therefore gluons can interact with other gluons, and there are multi-gluon vertices that do not occur in QED where the photon is chargeless. The Feynman rule for these vertices are given in fig. 9 (lower), where f^{abc} , $a, b, c = 1, \dots, 8$ are the QCD structure constants defined by

$$[T^a, T^b] = f^{abc} T^c, \quad (5.33)$$

The QCD Feynman rules will be discussed at greater length in the Standard Model course.

6 Introduction to Renormalization

6.1 Ultraviolet (UV) Singularities

So far, everything was computed at tree-level, that is, at the lowest nontrivial order in perturbation theory. Very often, a more precise determination of a cross section is desirable and we are thus led to consider loop diagrams. In order to illustrate this, consider the example $e^+e^- \rightarrow \mu^+\mu^-$. The perturbative expansion of the corresponding amplitude is written as

$$\mathcal{M} = \alpha\mathcal{M}_0 + \alpha^2\mathcal{M}_1 + \alpha^3\mathcal{M}_2 + \mathcal{O}(\alpha^4), \quad (6.1)$$

where $\alpha = \frac{e^2}{4\pi} \approx 1/137$. When we computed the corresponding amplitude in section 5.3 we only computed the leading order term

$$\alpha\mathcal{M}_0 = \text{[tree-level diagram]} \propto e^2 \propto \alpha \quad (6.2)$$

Using this expression for the amplitude, we will get the leading-order cross section $\sigma_0 \propto \alpha^2|\mathcal{M}_0|^2$. If we want to compute corrections of order α^3 to this result, we will have to compute the amplitude to an accuracy of order α^2 .

$$\mathcal{M} = \text{[tree-level diagram]} + \text{[loop diagrams]} + \dots \quad (6.3)$$

In fact this set of diagrams is one place where the distinction between relativistic quantum mechanics and true field theory raises its head. The diagram with an internal quark loop is naturally generated in quantum field theory but not in a perturbative expansion in quantum mechanics. In principle, a quark must also be included in this loop, but in QM you have to treat the quark as an external particle that is put there by hand. While the Feynman rules we derived are correct, you will see a much more rigorous derivation of the (scalar theory) Feynman rules in your QFT course.

The one-loop correction to the cross section is related to the interference term of \mathcal{M}_0 and \mathcal{M}_1 ,

$$\sigma_1 \propto |\alpha\mathcal{M}_0 + \alpha^2\mathcal{M}_1 + \mathcal{O}(\alpha^3)|^2 = \alpha^2|\mathcal{M}_0|^2 + 2\alpha^3\text{Re}(\mathcal{M}_0\mathcal{M}_1^*) + \mathcal{O}(\alpha^4). \quad (6.4)$$

The whole procedure looks pretty straightforward. However, if we try to compute a loop diagram, we run into trouble.

Consider as an example the vertex correction \mathcal{V} , depicted in fig. 10. Using the Feynman rules listed in section 3.6 we end up with an expression of the form

$$\mathcal{V} \propto \int \frac{d^4k}{(2\pi)^4} \frac{k \cdot k}{k^2((p_b + k)^2 - m^2)((p_a - k)^2 - m^2)} \quad (6.5)$$

where we did not bother to write down the full algebraic expression resulting from the spinor and Lorentz algebra but only the terms involving k . The two factors of k in the numerator stem from the two fermion propagators. The important point is that this integral diverges. Indeed, considering the limit $k \rightarrow \infty$ we can neglect p_a, p_b and m and find

$$\mathcal{V} \sim \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^4} \sim \int \frac{dk}{(2\pi)} \frac{1}{k} = \infty \quad (6.6)$$

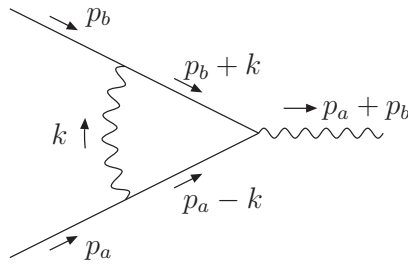


Figure 10: Vertex correction for $e^+e^- \rightarrow \mu^+\mu^-$ scattering.

where we used $d^4k \sim k^3 dk$. These singularities are called ultraviolet (UV) singularities because they come from the region $k \rightarrow \infty$.

Similar problems are encountered if we try to compute the other one-loop diagrams and our final answer for the cross section at next-to-leading order seems to be infinity.

6.2 Infrared (IR) Singularities

There is another class of singularities that shows up in QED and QCD. As we saw in section 6.1 that UV singularities are related to the region of large k . However, there is also a potential danger of singularities from the region $k \sim 0$ or more generally, from zeros in the denominators of the integrand. These singularities are called infrared (IR) singularities. These occur if some (massless) particle becomes very soft or two become very collinear. These singularities have nothing to do with the UV singularities. The solution to the problem is completely different in the two cases. In fact, you already should have encountered an IR singularity. When you tried to compute the total cross section for Compton scattering in section 5.4 you should have found that the total cross section diverges. This is due to an IR singularity. Indeed, the final state photon can become arbitrarily soft, in which case the electron-photon pair becomes indistinguishable from a single electron. One possibility to get a well defined finite answer is to require that the final state photon has some minimal energy but the general solution will be discussed in the phenomenology course.

I will not discuss the IR singularities any further and will simply ignore them, safe in the knowledge that they can be dealt with in a manner totally different to that for the UV singularities. Thus in what follows I will call a cross section finite if it has no UV singularities, but it might well have IR singularities. Strictly speaking, we should replace every ‘finite’ below by ‘UV-finite’.

6.3 Renormalization

It is important to realize that renormalization is not really about the removal of divergences, but simply an expression of the fact that in quantum field theories the value of certain parameters, e.g. the coupling constants, change with the energy scale used in a process. The infinities we encounter are then just a consequence of our ignorance of what is happening as $E \rightarrow \infty$ although we integrate up to this limit in any loop diagrams. We will demonstrate this below, and show how results do turn out to be finite after all.

To to obtain a prediction for any measurable quantity \mathcal{S} , say a cross section, we started with wave equations from which we deduced the Feynman rules, which in turn were used to compute \mathcal{S} . The wave equations of QED, eqs. (3.21) and (3.23), have some parameters. So far, we denoted them by e, m and referred to them as mass and charge of the electron. Therefore, our result \mathcal{S} will depend on these parameters. However, the parameter m in the Lagrangian is *not* the real mass of the electron, nor is e its charge. The identification of the parameter in the Lagrangian and the measurable quantity is only justified at tree level, because beyond this level the parameters themselves receive corrections, i.e. the propagator and vertex diagram which define the mass and coupling strength are themselves corrected. Therefore, from now on we will be more precise and denote the parameters in \mathcal{L} by m_0 and e_0 and call them the *bare* mass and *bare* charge respectively. Note that the bare parameters are not measurable. The (measurable) physical mass and charge of the electron will be denoted (as always) by m and e . \mathcal{L} also depends on the fields, which we denoted so far by ψ and A . From now on, we denote them by ψ_0 and A_0 and call them the bare fields.

We are now ready to reformulate the problem we encountered in section 6.1. If we try to compute a measurable quantity in terms of the unmeasurable bare quantities as a perturbative expansion in the coupling constant we generally encounter divergences. That is, if we compute

$$\mathcal{S}(e_0, m_0, \psi_0, A_0) = \mathcal{S}_0(e_0, m_0, \psi_0, A_0) + e_0^2 \mathcal{S}_1(e_0, m_0, \psi_0, A_0) + \mathcal{O}(e_0^4) \quad (6.7)$$

then we may find that $\mathcal{S}_1(e_0, m_0, \psi_0, A_0) = \infty$. In particular, this is true for two special physical quantities, namely the mass and the charge of the electron,

$$\begin{aligned} m &= m_0 + e_0^2 m_1(e_0, m_0, \psi_0, A_0) + \mathcal{O}(e_0^4) \\ e &= e_0 + e_0^2 e_1(e_0, m_0, \psi_0, A_0) + \mathcal{O}(e_0^4). \end{aligned} \quad (6.8)$$

But this is an expression for two measurable quantities in terms of unknown parameters. If the unknowns m_0 and e_0 are finite then we would get divergences in m_1 and e_1 and hence in m and e . Since m and e are finite quantities we conclude that the bare quantities are infinite. This is the root of the problem. UV divergences in our perturbative calculations show up if we try to express our results in terms of the unmeasurable, unphysical bare parameters, i.e. the parameters of the original Lagrangian.

In order to save the situation, we have to find new parameters such that the result of any physical quantity expressed in these new parameters — at any order in perturbation theory — is finite. Is this possible? Generally, the answer is no. However, for some special theories (and luckily QED is one of them) it is possible. Such theories are called renormalizable theories. The new parameters are called the *renormalized* quantities and are denoted by e_R, m_R and ψ_R, A_R . They are related to the bare quantities as follows:

$$\begin{aligned} \psi_0 &= Z_2^{1/2} \psi_R \\ A_0 &= Z_3^{1/2} A_R \\ m_0 &= Z_m^{1/2} m_R \\ e_0 &= Z_1 Z_2^{-1} Z_3^{-1/2} e_R \end{aligned} \quad (6.9)$$

This is simply a definition of the renormalization factors Z_1, Z_2, Z_3 and Z_m . Since the renormalization factors relate finite and divergent quantities, they have to be divergent themselves. More precisely, they can be written as a perturbative series with divergent coefficients.

To summarize, if we express the perturbative series for our physical quantity in terms of the renormalized quantities

$$\mathcal{S}(e_R, m_R, \psi_R, A_R) = \mathcal{S}_0(e_R, m_R, \psi_R, A_R) + e_R^2 \mathcal{S}_1(e_R, m_R, \psi_R, A_R) + \mathcal{O}(e_R^4) \quad (6.10)$$

there will be no UV-divergences at any order in perturbation theory. Some people refer to this as ‘hiding the infinities’. What is meant by this statement is that if we have a small number of input values ($m_R, e_R \dots$) and express all results in terms of these input values we get finite answers for all measurable quantities. Thus, renormalizing QED enables us to relate any measurable quantity to a small number of measurable input values.

It is a highly non-trivial exercise to show that QED is indeed a renormalizable theory. But once we know that we can find a set of renormalized parameters e_R, m_R, ψ_R, A_R such that eq. (6.10) has finite coefficients at each order, it is clear that we can find as many other sets as we like. Indeed, if we chose $e'_R, m'_R, \psi'_R, A'_R$ such that m_R and m'_R (and all other parameters) are related by a finite series, then

$$\mathcal{S}'(e'_R, m'_R, \psi'_R, A'_R) = \mathcal{S}'_0(e'_R, m'_R, \psi'_R, A'_R) + (e'_R)^2 \mathcal{S}'_1(e'_R, m'_R, \psi'_R, A'_R) + \mathcal{O}((e'_R)^4) \quad (6.11)$$

is also finite at each order in perturbation theory. In other words, the divergent pieces of the renormalization factors in eq. (6.9) are uniquely determined by requiring that the divergences cancel. However, we are completely free to fix the finite pieces to whatever we want. Choosing a particular set of renormalized quantities, that is, giving some prescription on how to fix the finite pieces of the renormalization factors, is called choosing the *renormalization scheme*. It is possible in QED that $m_R = m$ and $e_R = e$, i.e. the renormalized coupling is determined by real electron photon scattering. The renormalization scheme that satisfies these constraints is called the on-shell scheme. Alternatively, the renormalized coupling may be determined by scattering with, for example, a virtual photon. In this case the value of e_R will depend on the scale of the scattering, i.e. the coupling will “run” with the renormalization scale. To be precise let me also mention that one more constraint is needed to fix the scheme completely. Naively you would expect that four constraints are needed, since we have four renormalization factors to fix. However, two of them are related, $Z_1 = Z_2$. This identity is due to gauge invariance and is called the Ward identity. As a result, we only need three constraints to fix the renormalization scheme completely.

▷ **Exercise 6.30**

Why is it not possible in QCD to use the on-shell scheme?

Of course, the result of our calculation has to be independent of the renormalization scheme. This remark is not quite as innocuous as it looks. In fact, it is only true up to the order to which we decided to compute. If we decide to include the $\mathcal{O}(e_R^2)$ but not the higher order terms in our calculation, we have

$$\mathcal{S}(e_R, m_R, \psi_R, A_R) - \mathcal{S}'(e'_R, m'_R, \psi'_R, A'_R) = \mathcal{O}(e_R^4) \quad (6.12)$$

The numerical result for our prediction will depend on the renormalization scheme! Even though the difference is formally of higher order it still can be numerically significant, in particular in QCD.

It's worth stressing again that this ability to hide UV divergences in the couplings is not as conspiratorial as it at first seems. In the IR a theory involves long wavelength modes that are insensitive to UV physics - indeed they (like us!) don't even know what the full UV theory of nature is. The incomplete IR theory will break down (generate infinities) if extended into the UV but since we know (presumably!) that the IR theory is part of a consistent UV theory there must be a way to hide the infinities. This is fundamentally why renormalization works.

6.4 Regularization

What we have learned so far is that we have to express the result of our calculation in terms of renormalized quantities rather than the bare ones. But since the starting point of any calculation is the Lagrangian, the first step in any calculation is to get the results in terms of bare quantities. Only then, we replace the bare quantities by the renormalized quantities, using eq. (6.9) and get a finite result. In intermediate steps we will have to deal with divergent expressions.

In order to give a mathematical meaning to these intermediate expressions, we will have to regularize the integrals. That is, we have to change them in a systematic way, such that they become finite. By doing so, we change the value of the integrals. However, at the end of our calculation, we are able to undo this change. Since the final result is finite, this step will not introduce a singularity.

There are — at least in principle — many different possibilities for regularizing the integrals. To illustrate the idea of regularization I will discuss first the method of introducing a cutoff, even though in practice this method is not really used. Consider again the vertex correction in eq. (6.5). As we saw, we got the UV singularity from the region $k \rightarrow \infty$. To regularize this expression, we introduce a cutoff Λ

$$\mathcal{V} \rightarrow \mathcal{V}_{\text{reg}} \sim \int^{\Lambda} \frac{d^4k}{(2\pi)^4} \frac{k k}{k^2((p_b + k)^2 - m^2)((p_a - k)^2 - m^2)} \quad (6.13)$$

Of course, by doing so we changed the value of the integral. At the end of our calculation we will have to let $\Lambda \rightarrow \infty$. Introducing this cutoff, however, gives us the possibility to deal with such intermediate expressions.

Let me illustrate the interplay between renormalization and regularization with an oversimplified example. Assume that with the cutoff regularization we get as a result of our calculation of some physical quantity, say a cross section

$$\mathcal{S} = e_0^4 A + e_0^6 \left(B \ln \frac{\Lambda}{m} + F_S \right) + \mathcal{O}(e_0^8) \quad (6.14)$$

where A, B and F_S are some *finite* terms. The originally divergent expression for \mathcal{S} has been rendered finite by regularization. At this point we cannot let $\Lambda \rightarrow \infty$ since we would get $\mathcal{S} \rightarrow \infty$. However, we learned that we have to express our results in terms of e_R and not e_0 (For simplicity, I ignore the mass renormalization). This step is

renormalization (not regularization). Computing the relation between e_0 and e_R , using the same regularization, we would find

$$e_R = e_0 - e_0^3 \left(C \ln \frac{\Lambda}{m} + F_e \right) + \mathcal{O}(e_0^5) \quad (6.15)$$

and reversing this

$$e_0 = e_R + e_R^3 \left(C \ln \frac{\Lambda}{m} + F_e \right) + \mathcal{O}(e_R^5) \quad (6.16)$$

where C and F_e are also finite. Plugging in eq. (6.16) into eq. (6.14) we get

$$\mathcal{S} = e_R^4 A + e_R^6 \left((B + 4AC) \ln \frac{\Lambda}{m} + F_S + 4AF_e \right) + \mathcal{O}(e_R^8) \quad (6.17)$$

and we would find $(B + 4AC) = 0$. Since QED is a renormalizable theory this ‘miracle’ would happen for any measurable quantity. Finally, in the expression

$$\mathcal{S} = e_R^4 A + e_R^6 (F_S + 4AF_e) + \mathcal{O}(e_R^8) \quad (6.18)$$

we can let $\Lambda \rightarrow \infty$ and ‘undo’ the regularization.

To summarize, regularization enables us to work with divergent intermediate expressions. In the example above, instead of writing ∞ we write $\log \Lambda$ and have in mind $\Lambda \rightarrow \infty$. Renormalization, on the other hand removes the (would be) singularities, i.e. it removes the $\log \Lambda$ terms. Therefore, after renormalization we can (and have to) undo the regularization.

Note that we could have defined a different renormalized coupling

$$\tilde{e}_R = e_0 - e_0^3 \left(C \ln \frac{\Lambda}{m} + G_e \right) + \mathcal{O}(e_0^5) \quad (6.19)$$

and this would have lead to

$$\mathcal{S} = \tilde{e}_R^4 A + \tilde{e}_R^6 (F_S + 4AG_e) + \mathcal{O}(\tilde{e}_R^8) \quad (6.20)$$

and we would have a different expression in terms of a different coupling - both equally valid, and identical up to the $\mathcal{O}(\tilde{e}_R^8)$ corrections.

As mentioned above, the method of introducing a cutoff for regularization is hardly ever used in actual calculations. The by far most popular method is to use dimensional regularization. The basic idea is to do the calculation not in 4 space-time dimensions but rather in D dimensions. Why does this help?

Consider once more our initial example of the vertex correction in eq. (6.5), which has an UV singularity in $D = 4$ space-time dimensions (see eq. (6.6)). For arbitrary D , using $d^D k \sim k^{D-1} dk$ we get

$$\mathcal{V} \sim \int \frac{d^D k}{(2\pi)^4} \frac{1}{k^4} \sim \int \frac{dk}{(2\pi)} k^{D-5} \quad (6.21)$$

and the integral is UV-finite for say $D \leq 3$. Thus changing the dimension can regulate integrals. It is important to note that this is only a technicality. There is no Physics

associated with $D \neq 4$ and at the end of the calculation we have to let $D \rightarrow 4$. If we did renormalize our theory properly this last step will not lead to UV divergences.

The reason why dimensional regularization is so popular is that it preserves gauge invariance and is technically relatively simple. Another very important issue is that this regularization not only regulates UV singularities, but also IR singularities. As mentioned in section 6.2, theories like QED or QCD are very often plagued by IR singularities. It is therefore very convenient if we do not have to introduce another regularization for IR singularities. Only after all UV and IR singularities have been removed, we can let $D \rightarrow 4$ and finally obtain a finite result.

7 QED as a Field Theory

7.1 Quantizing the Dirac Field

In this section we return to the Dirac equation and use it as the basis for a field theory, which allows the creation and annihilation of particles naturally. Quantizing a field (or second quantization) basically means that the wave function becomes an operator. The space in which this operator acts is called the Fock space. The Fock space contains states with an arbitrary number of particles and therefore we will be able to describe processes where the number of states changes.

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

$$\mathcal{L} = \bar{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x). \quad (7.1)$$

Then the Euler-Lagrange equation

$$\frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial(\partial_\mu\psi)} - \frac{\partial \mathcal{L}}{\partial\psi} = 0 \quad (7.2)$$

leads to the Dirac equation. The canonical momentum is

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial\dot{\psi}(x)} = i\psi^\dagger(x) \quad (7.3)$$

and the Hamiltonian density is

$$\mathcal{H} = \pi\dot{\psi} - \mathcal{L} = \psi^\dagger i \frac{\partial\psi}{\partial t}. \quad (7.4)$$

Now we want to regard ψ as a quantum field rather than as a wave function. In order to quantize this field, naively we would try to impose the usual equal time commutation relations, i.e.

$$\begin{aligned} [\psi_\alpha(\vec{x}, t), \pi_\beta(\vec{y}, t)] &= i\delta_{\alpha\beta}\delta^3(\vec{x} - \vec{y}), \\ [\psi_\alpha(\vec{x}, t), \psi_\beta(\vec{y}, t)] &= 0, \\ [\pi_\alpha(\vec{x}, t), \pi_\beta(\vec{y}, t)] &= 0, \end{aligned} \quad (7.5)$$

where α and β label the spinor components of ψ and π . Without proving it for the moment we note that this would lead to a disaster. In particular, the Hamiltonian is unbounded from below - there is no ground state. The only way to cure the problem is to impose anti-commutation relations (we will soon see that this leads to the desired properties for spin-1/2):

$$\begin{aligned}\{\psi_\alpha(\vec{x}, t), \pi_\beta(\vec{y}, t)\} &= i\delta_{\alpha\beta}\delta^3(\vec{x} - \vec{y}). \\ \{\psi_\alpha(\vec{x}, t), \psi_\beta(\vec{y}, t)\} &= 0, \\ \{\pi_\alpha(\vec{x}, t), \pi_\beta(\vec{y}, t)\} &= 0.\end{aligned}\tag{7.6}$$

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

The Heisenberg equations of motion for the field operators have the solution

$$\psi_\alpha(\vec{x}, t) = \int \frac{d^3\vec{k}}{(2\pi)^3} \frac{1}{2E} \sum_{s=1,2} [b(s, \vec{k})u_\alpha(s, \vec{k})e^{-ik\cdot x} + d^\dagger(s, \vec{k})v_\alpha(s, \vec{k})e^{ik\cdot x}]\tag{7.7}$$

$$\bar{\psi}_\alpha(\vec{x}, t) = \int \frac{d^3\vec{k}}{(2\pi)^3} \frac{1}{2E} \sum_{s=1,2} [b^\dagger(s, \vec{k})\bar{u}_\alpha(s, \vec{k})e^{ik\cdot x} + d(s, \vec{k})\bar{v}_\alpha(s, \vec{k})e^{-ik\cdot x}]\tag{7.8}$$

Since ψ is now an operator, so are the expansion coefficients b^\dagger, d^\dagger, b and d . They are interpreted as creation and annihilation operators for electrons and positrons respectively. The anti-commutation relations for the fields, eq. (7.6), imply that

$$\begin{aligned}\{b(r, \vec{k}), b^\dagger(s, \vec{k}')\} &= (2\pi)^3 2E \delta^3(\vec{k} - \vec{k}')\delta_{sr} \\ \{d(r, \vec{k}), d^\dagger(s, \vec{k}')\} &= (2\pi)^3 2E \delta^3(\vec{k} - \vec{k}')\delta_{sr} \\ \{b(r, \vec{k}), b(s, \vec{k}')\} &= \{b^\dagger(r, \vec{k}), b^\dagger(s, \vec{k}')\} = 0 \\ \{d(r, \vec{k}), d(s, \vec{k}')\} &= \{d^\dagger(r, \vec{k}), d^\dagger(s, \vec{k}')\} = 0\end{aligned}\tag{7.9}$$

▷ **Exercise 7.31**

Show that the anticommutation relations above lead to the correct anticommutation relations for the fields $\psi_\alpha(\vec{x}, t)$ and $\pi_\beta(\vec{x}, t)$. You will need the spinor sum relations in eq. (5.5).

The total Hamiltonian is

$$H = \int d^3\vec{x} : \mathcal{H} : \tag{7.10}$$

The symbols $: \ :$ denote normal ordering of the operator inside, i.e. we put all creation operators to the left of all annihilation operators so that $H|0\rangle = 0$ by definition, and is the way we remove the ambiguity associated with the order of operators. Note that if we move an anti-commuting (fermion) operator through another such operator then we pick up a minus sign. Using eq. (7.4) after some algebra we get

$$H = \int \frac{d^3\vec{k}}{(2\pi)^3} \frac{1}{2E} E \sum_{s=1,2} [b^\dagger(s, \vec{k})b(s, \vec{k}) + d^\dagger(s, \vec{k})d(s, \vec{k})]. \tag{7.11}$$

▷ **Exercise 7.32**

Verify the above form of the Hamiltonian. Can you see from the derivation why commutation relations for ψ and π and therefore for b and d would have led to a disaster?

The formula in eq. (7.11) has a very nice interpretation. The operator $b^\dagger b$ is nothing but the number operator for electrons and $d^\dagger d$ that for positrons. Thus, to get the total Hamiltonian, we have to count all electrons and positrons for all spin states s and momenta \vec{k} and multiply this number by the corresponding energy E .

If we had tried to impose commutation relations, the $d^\dagger d$ term would have entered with a minus sign in front, which would signal that something has gone wrong. In particular, it would mean that d^\dagger creates particles of negative energy. This is not supposed to happen in the quantized field theory. (We could try to fix the problem by simply re-labeling $d \leftrightarrow d^\dagger$ but it may be shown that this leads to acausal propagation.)

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

$$\{b^\dagger(r, \vec{k}), b^\dagger(s, \vec{k}')\} = 0$$

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

$$b^\dagger(s, \vec{k})b^\dagger(s, \vec{k})|0\rangle = 0.$$

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

Finally consider the charge operator

$$Q = \int d^3\vec{x} : j_0(x) : = \int d^3\vec{x} : \psi^\dagger \psi :$$

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

$$Q = \int \frac{d^3\vec{k}}{(2\pi)^3} \frac{1}{2E} \sum_{s=1,2} [b^\dagger(s, \vec{k})b(s, \vec{k}) - d^\dagger(s, \vec{k})d(s, \vec{k})] \quad (7.12)$$

This shows again that b^\dagger creates fermions while d^\dagger creates the associated antifermions of opposite charge.

7.2 Quantizing the Electromagnetic Field

The Maxwell equations can be derived from the Lagrangian density

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - j_\mu A^\mu \quad (7.13)$$

where the field strength tensor is

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu, \quad (7.14)$$

and j_μ is a source for the field. Maxwell's equations do not change under the gauge transformation

$$A_\mu(x) \rightarrow A_\mu(x) + \partial_\mu \Lambda(x) \quad (7.15)$$

where $\Lambda(x)$ is some scalar field. This shows that there is some redundancy, and the 4 components of $A_\mu(x)$ are more than is required to describe the electromagnetic field (there are two transverse polarizations of e.m. radiation). This leads to a problem in quantization. To see this note that the canonically conjugate field to A_μ is

$$\Pi^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_\mu)} = F^{\mu 0} \quad (7.16)$$

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

To get around this problem we recognize that gauge invariance allows us to impose an extra condition, which we use to *fix* the gauge invariance, and effectively lower the degrees of freedom. For example, we can impose the Lorentz gauge condition, i.e.

$$\partial_\mu A^\mu = 0. \quad (7.17)$$

Note that, even after fixing the Lorentz gauge, we can perform another gauge transformation on A_μ , i.e. $A_\mu(x) \rightarrow A_\mu(x) + \partial_\mu \chi(x)$ where $\chi(x)$ must satisfy the wave equation, $\partial_\mu \partial^\mu \chi = 0$, i.e. we have two unphysical degrees of freedom and the two physical fields.

We impose the constraint by noting that since $\partial_\mu A^\mu = 0$, there is no harm in adding it to the Lagrangian density as

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} - j_\mu A^\mu - \frac{1}{2\xi} (\partial_\mu A^\mu)^2. \quad (7.18)$$

Indeed what we are doing here is following the Lagrange multiplier method of imposing constraints ($1/2\xi$ being the Lagrange multiplier), and recognizing that we should find the stationary points of $S = \int d^4x \mathcal{L}$ subject to the constraint $\int d^4x (\partial_\mu A^\mu)^2 = 0$, i.e. this comes from the ‘‘equation of motion’’ $\partial \mathcal{L} / \partial(1/2\xi) = 0$.

Using the gauge-fixed Lagrangian, the equations of motion are now

$$\partial^\mu F_{\mu\nu} - j_\nu + \frac{1}{\xi} \partial_\nu (\partial^\mu A_\mu) = 0.$$

If we require that these equations are satisfied *and then* also $\partial_\mu A^\mu = 0$, we have the original equations of motion but in a fixed gauge.

In the Feynman gauge $\xi = 1$, the Lagrangian is particularly simple (after some integration by parts under $\int d^4x$):

$$\mathcal{L} = \frac{1}{2} \partial_\mu A_\nu \partial^\mu A^\nu - j_\mu A^\mu,$$

and quantization can now proceed: $\Pi^\mu = \partial_0 A^\mu$ and thus

$$[A^\mu(\vec{x}, t), \partial_0 A^\nu(\vec{y}, t)] = -ig^{\mu\nu} \delta^3(\vec{x} - \vec{y}) \quad (7.19)$$

with all other commutators vanishing. The Heisenberg operator corresponding to the photon field is

$$A_\mu(x) = \int \frac{d^3 \vec{k}}{(2\pi)^3} \frac{1}{2E} \sum_{\lambda=0}^3 \left[\varepsilon_\mu(\lambda, \vec{k}) a(\lambda, \vec{k}) e^{-ik \cdot x} + \varepsilon_\mu^*(\lambda, \vec{k}) a^\dagger(\lambda, \vec{k}) e^{ik \cdot x} \right] \quad (7.20)$$

where $\varepsilon_\mu(\lambda, \vec{k})$ are a set of four linearly independent basis 4-vectors for polarization ($\lambda = 0, 1, 2, 3$). For example, if $k = (k_0, \vec{k})$, we might choose $\varepsilon^\mu(0) = (1, 0, 0, 0)$, $\varepsilon^\mu(3) = (0, \vec{k})/k_0$, $\varepsilon^\mu(1) = (0, \vec{n}_1)$ and $\varepsilon^\mu(2) = (0, \vec{n}_2)$, where $k_0^2 = \vec{k}^2$, $\vec{n}_1 \cdot \vec{k} = 0$, $\vec{n}_2 \cdot \vec{k} = 0$ and $\vec{n}_1 \cdot \vec{n}_2 = 0$. $\varepsilon^\mu(1)$ and $\varepsilon^\mu(2)$ are therefore polarization vectors for transverse polarizations whilst $\varepsilon^\mu(0)$ is referred to as the timelike polarization vector and $\varepsilon^\mu(3)$ is referred to as the longitudinal polarization vector. For example, if $k = (k_0, 0, 0, k_0)$, $\varepsilon^\mu(0) = (1, 0, 0, 0)$, $\varepsilon^\mu(3) = (0, 0, 0, 1)$, $\varepsilon^\mu(1) = (0, 1, 0, 0)$ and $\varepsilon^\mu(2) = (0, 0, 1, 0)$.

The commutation relation (7.19) implies that

$$\left[a(\lambda, \vec{k}), a^\dagger(\lambda', \vec{k}') \right] = -g_{\lambda\lambda'} 2E (2\pi)^3 \delta^3(\vec{k} - \vec{k}'). \quad (7.21)$$

At a glance this looks fine, i.e. we interpret $a^\dagger(\lambda, \vec{k})$ as an operator that creates quanta of the electromagnetic field (photons) with polarization λ and momentum \vec{k} . However, for $\lambda = 0$ we have a problem since the sign on the RHS of (7.21) is opposite to that of the other 3 polarizations. This shows up in the fact that these timelike photons make a negative contribution to the energy:

$$H = \int \frac{d^3\vec{k}}{(2\pi)^3} \frac{1}{2E} E \left(-a^\dagger(0, \vec{k})a(0, \vec{k}) + \sum_{i=1,3} a^\dagger(i, \vec{k})a(i, \vec{k}) \right). \quad (7.22)$$

Fortunately, although we might not realize it yet, we have already solved the problem. Recall that we still have to impose $\partial_\mu A^\mu = 0$. It turns out that it is impossible to do this at the operator level, but we can do it for all physical expectation values, i.e. we can impose the correct physics. It then turns out that contributions from the timelike and longitudinal photons always cancel. More explicitly, by demanding for any state $|\chi\rangle$ that

$$\langle \chi | \partial_\mu A^\mu | \chi \rangle = 0 \quad (7.23)$$

it follows that

$$\langle \chi | a^\dagger(3, \vec{k})a(3, \vec{k}) - a^\dagger(0, \vec{k})a(0, \vec{k}) | \chi \rangle = 0. \quad (7.24)$$

and therefore $\langle \chi | H | \chi \rangle \geq 0$. This is nice because it is in accord with our knowledge that free photons are transversely polarized.

▷ **Exercise 7.33**

Show that eq. (7.24) follows from eq. (7.23).

Acknowledgements

This lecture course has a long history; David Miller and I have shared the teaching over the last 6 years. We both built upon the foundation laid by Robert Thorne, Tim Morris, Jeff Forshaw, Adrian Signer and all the previous lecturers of this course who have been responsible for its development over the years. These course notes have evolved dependent on the tastes of those lecturers and in that spirit I have only modified the L^AT_EX files which David was kind enough to provide me. Any errors which I have managed to add while doing this (as well as any errors I have missed for sections left unchanged) are, of course, entirely my responsibility.

Many thanks also to Margaret Evans for the impeccable organization and to Bill Murray for keeping us all under control. Thanks also to all the other lecturers and tutors and the students for making the school such a pleasant experience.

References

- [1] J D Bjorken and S D Drell *Relativistic Quantum Mechanics* McGraw-Hill 1964
- F Halzen and A D Martin *Quarks and Leptons* Wiley 1984
- F Mandl and G Shaw *Quantum Field Theory* Wiley 1984
- I J R Aitchison and A J G Hey *Gauge theories in particle physics* 2nd edition Adam Hilger 1989
- M E Peskin and D V Schroeder *An Introduction to Quantum Field Theory* Addison Wesley 1995
- B Hatfield *Quantum Field Theory of Point Particles and Strings* Addison Wesley 1992
- L H Ryder *Quantum Field Theory* Cambridge University Press 1985
- C Itzykson and J-B Zuber *Quantum Field Theory* McGraw-Hill 1987

Pre School Problems

Rotations, Angular Momentum and the Pauli Matrices

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

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

Parameterise A as

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

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

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

Note that the Einstein summation convention is used here. Compute $L^2 \equiv L_1^2 + L_2^2 + L_3^2$. What is the interpretation of L^2 ?

The Pauli matrices σ_i are,

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

Verify that $\frac{1}{2}\sigma_i$ satisfy the same commutation relations as L_i .

Four Vectors

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

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

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

$$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (.25)$$

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

By imposing the condition

$$g_{\mu\nu}x'^\mu x'^\nu = g_{\mu\nu}x^\mu x^\nu \quad (.26)$$

where

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

show that

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

This is the analogue of the orthogonality relation for rotations. Check that it works for the Λ given by equation (.25) above.

Now introduce

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

and show, by reconsidering equation (.26) using $x^\mu x_\mu$, or otherwise, that

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

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

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

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

Probability Density and Current Density

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

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

where

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

What is the interpretation of \vec{J} ? Verify that the continuity equation can be written in manifestly covariant form.

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

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