6. Quantum Electrodynamics

In this section we finally get to quantum electrodynamics (QED), the theory of light interacting with charged matter. Our path to quantization will be as before: we start with the free theory of the electromagnetic field and see how the quantum theory gives rise to a photon with two polarization states. We then describe how to couple the photon to fermions and to bosons.

6.1 Maxwell's Equations

The Lagrangian for Maxwell's equations in the absence of any sources is simply

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{6.1}$$

where the field strength is defined by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{6.2}$$

The equations of motion which follow from this Lagrangian are

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \right) = -\partial_\mu F^{\mu\nu} = 0 \tag{6.3}$$

Meanwhile, from the definition of $F_{\mu\nu}$, the field strength also satisfies the Bianchi identity

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0 \tag{6.4}$$

To make contact with the form of Maxwell's equations you learn about in high school, we need some 3-vector notation. If we define $A^\mu = (\phi, \vec{A})$, then the electric field $\vec{E}$ and magnetic field $\vec{B}$ are defined by

$$\vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t} \quad \text{and} \quad \vec{B} = \nabla \times \vec{A} \tag{6.5}$$

which, in terms of $F_{\mu\nu}$, becomes

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix} \tag{6.6}$$

The Bianchi identity (6.4) then gives two of Maxwell's equations,

$$\nabla \cdot \vec{B} = 0 \quad \text{and} \quad \frac{\partial \vec{B}}{\partial t} = -\nabla \times \vec{E} \tag{6.7}$$

These remain true even in the presence of electric sources. Meanwhile, the equations of motion give the remaining two Maxwell equations,

$$\nabla \cdot \vec{E} = 0 \quad \text{and} \quad \frac{\partial \vec{E}}{\partial t} = \nabla \times \vec{B} \tag{6.8}$$

As we will see shortly, in the presence of charged matter these equations pick up extra terms on the right-hand side.

6.1.1 Gauge Symmetry

The massless vector field A_μ has 4 components, which would naively seem to tell us that the gauge field has 4 degrees of freedom. Yet we know that the photon has only two degrees of freedom which we call its polarization states. How are we going to resolve this discrepancy? There are two related comments which will ensure that quantizing the gauge field A_μ gives rise to 2 degrees of freedom, rather than 4.

- The field A_0 has no kinetic term $\dot{A}_0$ in the Lagrangian: it is not dynamical. This means that if we are given some initial data A_i and $\dot{A}_i$ at a time t_0, then the field A_0 is fully determined by the equation of motion $\nabla \cdot \vec{E} = 0$ which, expanding out, reads

$$\nabla^2 A_0 + \nabla \cdot \frac{\partial \vec{A}}{\partial t} = 0 \tag{6.9}$$

This has the solution

$$A_0(\vec{x}) = \int d^3 x' \, \frac{(\nabla \cdot \partial \vec{A}/\partial t)(\vec{x}')}{4\pi |\vec{x} - \vec{x}'|} \tag{6.10}$$

So A_0 is not independent: we don't get to specify A_0 on the initial time slice. It looks like we have only 3 degrees of freedom in A_μ rather than 4. But this is still one too many.

- The Lagrangian (6.3) has a *very* large symmetry group, acting on the vector potential as

$$A_\mu(x) \to A_\mu(x) + \partial_\mu \lambda(x) \tag{6.11}$$

for any function $\lambda(x)$. We'll ask only that $\lambda(x)$ dies off suitably quickly at spatial $\vec{x} \to \infty$. We call this a *gauge symmetry*. The field strength is invariant under the gauge symmetry:

$$F_{\mu\nu} \to \partial_\mu(A_\nu + \partial_\nu \lambda) - \partial_\nu(A_\mu + \partial_\mu \lambda) = F_{\mu\nu} \tag{6.12}$$

So what are we to make of this? We have a theory with an infinite number of symmetries, one for each function $\lambda(x)$. Previously we only encountered symmetries which act the same at all points in spacetime, for example $\psi \to e^{i\alpha}\psi$ for a complex scalar field. Noether's theorem told us that these symmetries give rise to conservation laws. Do we now have an infinite number of conservation laws?

The answer is no! Gauge symmetries have a very different interpretation than the global symmetries that we make use of in Noether's theorem. While the latter take a physical state to another physical state with the same properties, the gauge symmetry is to be viewed as a redundancy in our description. That is, two states related by a gauge symmetry are to be identified: they are the same physical state. (There is a small caveat to this statement which is explained in Section 6.3.1). One way to see that this interpretation is necessary is to notice that Maxwell's equations are not sufficient to specify the evolution of A_μ. The equations read,

$$\left[\eta_{\mu\nu}(\partial^\rho\partial_\rho) - \partial_\mu\partial_\nu\right] A^\nu = 0 \tag{6.13}$$

But the operator $\left[\eta_{\mu\nu}(\partial^\rho\partial_\rho) - \partial_\mu\partial_\nu\right]$ is not invertible: it annihilates any function of the form $\partial_\mu\lambda$. This means that given any initial data, we have no way to uniquely determine A_μ at a later time since we can't distinguish between A_μ and $A_\mu + \partial_\mu\lambda$. This would be problematic if we thought that A_μ is a physical object. However, if we're happy to identify A_μ and $A_\mu + \partial_\mu\lambda$ as corresponding to the same physical state, then our problems disappear.

Since gauge invariance is a redundancy of the system, we might try to formulate the theory purely in terms of the local, physical, gauge invariant objects $\vec{E}$ and $\vec{B}$. This is fine for the free classical theory: Maxwell's equations were, after all, first written in terms of $\vec{E}$ and $\vec{B}$. But it is not possible to describe certain quantum phenomena, such as the Aharonov-Bohm effect, without using the gauge potential A_μ. We will see shortly that we also require the gauge potential to describe classically charged fields. To

Figure 29:

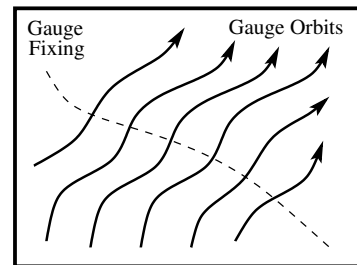describe Nature, it appears that we have to introduce quantities A_μ that we can never measure.

The picture that emerges for the theory of electromagnetism is of an enlarged phase space, foliated by gauge orbits as shown in the figure. All states that lie along a given

line can be reached by a gauge transformation and are identified. To make progress, we pick a representative from each gauge orbit. It doesn't matter which representative we pick — after all, they're all physically equivalent. But we should make sure that we pick a "good" gauge, in which we cut the orbits.

Different representative configurations of a physical state are called different *gauges*. There are many possibilities, some of which will be more useful in different situations. Picking a gauge is rather like picking coordinates that are adapted to a particular problem. Moreover, different gauges often reveal slightly different aspects of a problem. Here we'll look at two different gauges:

- **Lorentz Gauge:** $\partial_\mu A^\mu = 0$

  To see that we can always pick a representative configuration satisfying $\partial_\mu A^\mu = 0$, suppose that we're handed a gauge field A'_μ satisfying $\partial_\mu (A')^\mu = f(x)$. Then we choose $A_\mu = A'_\mu + \partial_\mu \lambda$, where

  $$\partial_\mu \partial^\mu \lambda = -f \tag{6.14}$$

  This equation always has a solution. In fact this condition doesn't pick a unique representative from the gauge orbit. We're always free to make further gauge transformations with $\partial_\mu \partial^\mu \lambda = 0$, which also has non-trivial solutions. As the name suggests, the Lorentz gauge[3] has the advantage that it is Lorentz invariant.

- **Coulomb Gauge:** $\nabla \cdot \vec{A} = 0$

  We can make use of the residual gauge transformations in Lorentz gauge to pick $\nabla \cdot \vec{A} = 0$. (The argument is the same as before). Since A_0 is fixed by (6.10), we have as a consequence

  $$A_0 = 0 \tag{6.15}$$

  (This equation will no longer hold in Coulomb gauge in the presence of charged matter). Coulomb gauge breaks Lorentz invariance, so may not be ideal for some purposes. However, it is very useful to exhibit the physical degrees of freedom: the 3 components of $\vec{A}$ satisfy a single constraint: $\nabla \cdot \vec{A} = 0$, leaving behind just 2 degrees of freedom. These will be identified with the two polarization states of the photon. Coulomb gauge is sometimes called radiation gauge.

---

[3]Named after Lorenz who had the misfortune to be one letter away from greatness.

6.2 The Quantization of the Electromagnetic Field

In the following we shall quantize free Maxwell theory twice: once in Coulomb gauge, and again in Lorentz gauge. We'll ultimately get the same answers and, along the way, see that each method comes with its own subtleties.

The first of these subtleties is common to both methods and comes when computing the momentum π^μ conjugate to A_μ,

$$
\pi^0 = \frac{\partial \mathcal{L}}{\partial \dot{A}_0} = 0
$$

$$
\pi^i = \frac{\partial \mathcal{L}}{\partial \dot{A}_i} = -F^{0i} \equiv E^i \tag{6.16}
$$

so the momentum π^0 conjugate to A_0 vanishes. This is the mathematical consequence of the statement we made above: A_0 is not a dynamical field. Meanwhile, the momentum conjugate to A_i is our old friend, the electric field. We can compute the Hamiltonian,

$$
H = \int d^3x \; \pi^i \dot{A}_i - \mathcal{L}
$$

$$
= \int d^3x \; \tfrac{1}{2}\vec{E} \cdot \vec{E} + \tfrac{1}{2}\vec{B} \cdot \vec{B} - A_0(\nabla \cdot \vec{E}) \tag{6.17}
$$

So A_0 acts as a Lagrange multiplier which imposes Gauss' law

$$
\nabla \cdot \vec{E} = 0 \tag{6.18}
$$

which is now a constraint on the system in which $\vec{A}$ are the physical degrees of freedom. Let's now see how to treat this system using different gauge fixing conditions.

6.2.1 Coulomb Gauge

In Coulomb gauge, the equation of motion for $\vec{A}$ is

$$
\partial_\mu \partial^\mu \vec{A} = 0 \tag{6.19}
$$

which we can solve in the usual way,

$$
\vec{A} = \int \frac{d^3p}{(2\pi)^3} \; \vec{\xi}(\vec{p}) \, e^{ip \cdot x} \tag{6.20}
$$

with $p_0^2 = |\vec{p}|^2$. The constraint $\nabla \cdot \vec{A} = 0$ tells us that $\vec{\xi}$ must satisfy

$$
\vec{\xi} \cdot \vec{p} = 0 \tag{6.21}
$$

which means that $\vec{\xi}$ is perpendicular to the direction of motion $\vec{p}$. We can pick $\vec{\xi}(\vec{p})$ to be a linear combination of two orthonormal vectors $\vec{\epsilon}_r$, $r = 1, 2$, each of which satisfies $\vec{\epsilon}_r(\vec{p}) \cdot \vec{p} = 0$ and

$$\vec{\epsilon}_r(\vec{p}) \cdot \vec{\epsilon}_s(\vec{p}) = \delta_{rs} \qquad r, s = 1, 2 \tag{6.22}$$

These two vectors correspond to the two polarization states of the photon. It's worth pointing out that you can't consistently pick a continuous basis of polarization vectors for every value of $\vec{p}$ because you can't comb the hair on a sphere. But this topological fact doesn't cause any complications in computing QED scattering processes.

To quantize we turn the Poisson brackets into commutators. Naively we would write

$$[A_i(\vec{x}), A_j(\vec{y})] = [E^i(\vec{x}), E^j(\vec{y})] = 0$$
$$[A_i(\vec{x}), E^j(\vec{y})] = i\delta_i^j \, \delta^{(3)}(\vec{x} - \vec{y}) \tag{6.23}$$

But this can't quite be right, because it's not consistent with the constraints. We still want to have $\nabla \cdot \vec{A} = \nabla \cdot \vec{E} = 0$, now imposed on the operators. But from the commutator relations above, we see

$$[\nabla \cdot \vec{A}(\vec{x}), \nabla \cdot \vec{E}(\vec{y})] = i\nabla^2 \, \delta^{(3)}(\vec{x} - \vec{y}) \neq 0 \tag{6.24}$$

What's going on? In imposing the commutator relations (6.23) we haven't correctly taken into account the constraints. In fact, this is a problem already in the classical theory, where the Poisson bracket structure is already altered[4]. The correct Poisson bracket structure leads to an alteration of the last commutation relation,

$$[A_i(\vec{x}), E_j(\vec{y})] = i\left(\delta_{ij} - \frac{\partial_i \partial_j}{\nabla^2}\right) \delta^{(3)}(\vec{x} - \vec{y}) \tag{6.25}$$

To see that this is now consistent with the constraints, we can rewrite the right-hand side of the commutator in momentum space,

$$[A_i(\vec{x}), E_j(\vec{y})] = i\int \frac{d^3p}{(2\pi)^3} \left(\delta_{ij} - \frac{p_i p_j}{|\vec{p}|^2}\right) e^{i\vec{p} \cdot (\vec{x} - \vec{y})} \tag{6.26}$$

which is now consistent with the constraints, for example

$$[\partial_i A_i(\vec{x}), E_j(\vec{y})] = i\int \frac{d^3p}{(2\pi)^3} \left(\delta_{ij} - \frac{p_i p_j}{|\vec{p}|^2}\right) ip_i \, e^{i\vec{p} \cdot (\vec{x} - \vec{y})} = 0 \tag{6.27}$$

---

[4]For a nice discussion of the classical and quantum dynamics of constrained systems, see the small book by Paul Dirac, "*Lectures on Quantum Mechanics*"

We now write $\vec{A}$ in the usual mode expansion,

$$\vec{A}(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2|\vec{p}|}} \sum_{r=1}^{2} \vec{\epsilon}_r(\vec{p}) \left[a_{\vec{p}}^r e^{i\vec{p}\cdot\vec{x}} + a_{\vec{p}}^{r\dagger} e^{-i\vec{p}\cdot\vec{x}} \right]$$

$$\vec{E}(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} (-i) \sqrt{\frac{|\vec{p}|}{2}} \sum_{r=1}^{2} \vec{\epsilon}_r(\vec{p}) \left[a_{\vec{p}}^r e^{i\vec{p}\cdot\vec{x}} - a_{\vec{p}}^{r\dagger} e^{-i\vec{p}\cdot\vec{x}} \right] \tag{6.28}$$

where, as before, the polarization vectors satisfy

$$\vec{\epsilon}_r(\vec{p}) \cdot \vec{p} = 0 \quad \text{and} \quad \vec{\epsilon}_r(\vec{p}) \cdot \vec{\epsilon}_s(\vec{p}) = \delta_{rs} \tag{6.29}$$

It is not hard to show that the commutation relations (6.25) are equivalent to the usual commutation relations for the creation and annihilation operators,

$$[a_{\vec{p}}^r, a_{\vec{q}}^s] = [a_{\vec{p}}^{r\dagger}, a_{\vec{q}}^{s\dagger}] = 0$$

$$[a_{\vec{p}}^r, a_{\vec{q}}^{s\dagger}] = (2\pi)^3 \delta^{rs} \delta^{(3)}(\vec{p} - \vec{q}) \tag{6.30}$$

where, in deriving this, we need the completeness relation for the polarization vectors,

$$\sum_{r=1}^{2} \epsilon_r^i(\vec{p}) \epsilon_r^j(\vec{p}) = \delta^{ij} - \frac{p^i p^j}{|\vec{p}|^2} \tag{6.31}$$

You can easily check that this equation is true by acting on both sides with a basis of vectors $(\vec{\epsilon}_1(\vec{p}), \vec{\epsilon}_2(\vec{p}), \vec{p})$.

We derive the Hamiltonian by substituting (6.28) into (6.17). The last term vanishes in Coulomb gauge. After normal ordering, and playing around with $\vec{\epsilon}_r$ polarization vectors, we get the simple expression

$$H = \int \frac{d^3p}{(2\pi)^3} |\vec{p}| \sum_{r=1}^{2} a_{\vec{p}}^{r\dagger} a_{\vec{p}}^r \tag{6.32}$$

The Coulomb gauge has the advantage that the physical degrees of freedom are manifest. However, we've lost all semblance of Lorentz invariance. One place where this manifests itself is in the propagator for the fields $A_i(x)$ (in the Heisenberg picture). In Coulomb gauge the propagator reads

$$D_{ij}^{\text{tr}}(x - y) \equiv \langle 0| T A_i(x) A_j(y) |0\rangle = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 + i\epsilon} \left(\delta_{ij} - \frac{p_i p_j}{|\vec{p}|^2} \right) e^{-ip\cdot(x-y)} \tag{6.33}$$

The *tr* superscript on the propagator refers to the "transverse" part of the photon. When we turn to the interacting theory, we will have to fight to massage this propagator into something a little nicer.

6.2.2 Lorentz Gauge

We could try to work in a Lorentz invariant fashion by imposing the Lorentz gauge condition $\partial_\mu A^\mu = 0$. The equations of motion that follow from the action are then

$$\partial_\mu \partial^\mu A^\nu = 0 \tag{6.34}$$

Our approach to implementing Lorentz gauge will be a little different from the method we used in Coulomb gauge. We choose to change the theory so that (6.34) arises directly through the equations of motion. We can achieve this by taking the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} (\partial_\mu A^\mu)^2 \tag{6.35}$$

The equations of motion coming from this action are

$$\partial_\mu F^{\mu\nu} + \partial^\nu (\partial_\mu A^\mu) = \partial_\mu \partial^\mu A^\nu = 0 \tag{6.36}$$

(In fact, we could be a little more general than this, and consider the Lagrangian

$$\mathcal{L} = -\tfrac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2\alpha} (\partial_\mu A^\mu)^2 \tag{6.37}$$

with arbitrary α and reach similar conclusions. The quantization of the theory is independent of α and, rather confusingly, different choices of α are sometimes also referred to as different "gauges". We will use $\alpha = 1$, which is called "Feynman gauge". The other common choice, $\alpha = 0$, is called "Landau gauge".)

Our plan will be to quantize the theory (6.36), and only later impose the constraint $\partial_\mu A^\mu = 0$ in a suitable manner on the Hilbert space of the theory. As we'll see, we will also have to deal with the residual gauge symmetry of this theory which will prove a little tricky. At first, we can proceed very easily, because both π^0 and π^i are dynamical:

$$\pi^0 = \frac{\partial \mathcal{L}}{\partial \dot{A}_0} = -\partial_\mu A^\mu$$
$$\pi^i = \frac{\partial \mathcal{L}}{\partial \dot{A}_i} = \partial^i A^0 - \dot{A}^i \tag{6.38}$$

Turning these classical fields into operators, we can simply impose the usual commutation relations,

$$[A_\mu(\vec{x}), A_\nu(\vec{y})] = [\pi^\mu(\vec{x}), \pi^\nu(\vec{y})] = 0$$
$$[A_\mu(\vec{x}), \pi_\nu(\vec{y})] = i\eta_{\mu\nu}\,\delta^{(3)}(\vec{x} - \vec{y}) \tag{6.39}$$

and we can make the usual expansion in terms of creation and annihilation operators and 4 polarization vectors $(\epsilon_\mu)^\lambda$, with $\lambda = 0, 1, 2, 3$.

$$A_\mu(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2|\vec{p}|}} \sum_{\lambda=0}^{3} \epsilon_\mu^\lambda(\vec{p}) \left[a_{\vec{p}}^\lambda e^{i\vec{p}\cdot\vec{x}} + a_{\vec{p}}^{\lambda\dagger} e^{-i\vec{p}\cdot\vec{x}} \right]$$

$$\pi^\mu(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} \sqrt{\frac{|\vec{p}|}{2}} (+i) \sum_{\lambda=0}^{3} (\epsilon^\mu)^\lambda(\vec{p}) \left[a_{\vec{p}}^\lambda e^{i\vec{p}\cdot\vec{x}} - a_{\vec{p}}^{\lambda\dagger} e^{-i\vec{p}\cdot\vec{x}} \right] \quad (6.40)$$

Note that the momentum π^μ comes with a factor of $(+i)$, rather than the familiar $(-i)$ that we've seen so far. This can be traced to the fact that the momentum (6.38) for the classical fields takes the form $\pi^\mu = -\dot{A}^\mu + \dots$ In the Heisenberg picture, it becomes clear that this descends to $(+i)$ in the definition of momentum.

There are now four polarization 4-vectors $\epsilon^\lambda(\vec{p})$, instead of the two polarization 3-vectors that we met in the Coulomb gauge. Of these four 4-vectors, we pick ϵ^0 to be timelike, while $\epsilon^{1,2,3}$ are spacelike. We pick the normalization

$$\epsilon^\lambda \cdot \epsilon^{\lambda'} = \eta^{\lambda\lambda'} \quad (6.41)$$

which also means that

$$(\epsilon_\mu)^\lambda (\epsilon_\nu)^{\lambda'} \eta_{\lambda\lambda'} = \eta_{\mu\nu} \quad (6.42)$$

The polarization vectors depend on the photon 4-momentum $p = (|\vec{p}|, \vec{p})$. Of the two spacelike polarizations, we will choose ϵ^1 and ϵ^2 to lie transverse to the momentum:

$$\epsilon^1 \cdot p = \epsilon^2 \cdot p = 0 \quad (6.43)$$

The third vector ϵ^3 is the longitudinal polarization. For example, if the momentum lies along the x^3 direction, so $p \sim (1, 0, 0, 1)$, then

$$\epsilon^0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad , \quad \epsilon^1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad , \quad \epsilon^2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad , \quad \epsilon^3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad (6.44)$$

For other 4-momenta, the polarization vectors are the appropriate Lorentz transformations of these vectors, since (6.43) are Lorentz invariant.

We do our usual trick, and translate the field commutation relations (6.39) into those for creation and annihilation operators. We find $[a_{\vec{p}}^\lambda, a_{\vec{q}}^{\lambda'}] = [a_{\vec{p}}^{\lambda\dagger}, a_{\vec{q}}^{\lambda'\dagger}] = 0$ and

$$[a_{\vec{p}}^\lambda, a_{\vec{q}}^{\lambda'\dagger}] = -\eta^{\lambda\lambda'} (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q}) \quad (6.45)$$

The minus signs here are odd to say the least! For spacelike $\lambda = 1, 2, 3$, everything looks fine,

$$[a_{\vec{p}}^{\lambda}, a_{\vec{q}}^{\lambda'\,\dagger}] = \delta^{\lambda\lambda'}\,(2\pi)^3\,\delta^{(3)}(\vec{p} - \vec{q}) \qquad \lambda, \lambda' = 1, 2, 3 \tag{6.46}$$

But for the timelike annihilation and creation operators, we have

$$[a_{\vec{p}}^0, a_{\vec{q}}^{0\,\dagger}] = -(2\pi)^3\,\delta^{(3)}(\vec{p} - \vec{q}) \tag{6.47}$$

This is very odd! To see just how strange this is, we take the Lorentz invariant vacuum $|0\rangle$ defined by

$$a_{\vec{p}}^{\lambda}\,|0\rangle = 0 \tag{6.48}$$

Then we can create one-particle states in the usual way,

$$|\vec{p}, \lambda\rangle = a_{\vec{p}}^{\lambda\,\dagger}\,|0\rangle \tag{6.49}$$

For spacelike polarization states, $\lambda = 1, 2, 3$, all seems well. But for the timelike polarization $\lambda = 0$, the state $|\vec{p}, 0\rangle$ has negative norm,

$$\langle\vec{p}, 0 |\, \vec{q}, 0\rangle = \langle 0|\, a_{\vec{p}}^0\, a_{\vec{q}}^{0\,\dagger}\,|0\rangle = -(2\pi)^3\,\delta^{(3)}(\vec{p} - \vec{q}) \tag{6.50}$$

Wtf? That's very very strange. A Hilbert space with negative norm means negative probabilities which makes no sense at all. We can trace this negative norm back to the wrong sign of the kinetic term for A_0 in our original Lagrangian: $\mathcal{L} = +\frac{1}{2}\dot{\vec{A}}^2 - \frac{1}{2}\dot{A}_0^2 + \ldots$.

At this point we should remember our constraint equation, $\partial_\mu A^\mu = 0$, which, until now, we've not imposed on our theory. This is going to come to our rescue. We will see that it will remove the timelike, negative norm states, and cut the physical polarizations down to two. We work in the Heisenberg picture, so that

$$\partial_\mu A^\mu = 0 \tag{6.51}$$

makes sense as an operator equation. Then we could try implementing the constraint in the quantum theory in a number of different ways. Let's look at a number of increasingly weak ways to do this

- We could ask that $\partial_\mu A^\mu = 0$ is imposed as an equation on operators. But this can't possibly work because the commutation relations (6.39) won't be obeyed for $\pi^0 = -\partial_\mu A^\mu$. We need some weaker condition.

- We could try to impose the condition on the Hilbert space instead of directly on the operators. After all, that's where the trouble lies! We could imagine that there's some way to split the Hilbert space up into good states $|\Psi\rangle$ and bad states that somehow decouple from the system. With luck, our bad states will include the weird negative norm states that we're so disgusted by. But how can we define the good states? One idea is to impose

$$\partial_\mu A^\mu |\Psi\rangle = 0 \tag{6.52}$$

on all good, physical states $|\Psi\rangle$. But this can't work either! Again, the condition is too strong. For example, suppose we decompose $A_\mu(x) = A_\mu^+(x) + A_\mu^-(x)$ with

$$A_\mu^+(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2|\vec{p}|}} \sum_{\lambda=0}^{3} \epsilon_\mu^\lambda a_{\vec{p}}^\lambda e^{-ip\cdot x}$$

$$A_\mu^-(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2|\vec{p}|}} \sum_{\lambda=0}^{3} \epsilon_\mu^\lambda a_{\vec{p}}^{\lambda\dagger} e^{+ip\cdot x} \tag{6.53}$$

Then, on the vacuum $A_\mu^+ |0\rangle = 0$ automatically, but $\partial^\mu A_\mu^- |0\rangle \neq 0$. So not even the vacuum is a physical state if we use (6.52) as our constraint

- Our final attempt will be the correct one. In order to keep the vacuum as a good physical state, we can ask that physical states $|\Psi\rangle$ are defined by

$$\partial^\mu A_\mu^+ |\Psi\rangle = 0 \tag{6.54}$$

This ensures that

$$\langle \Psi' | \partial_\mu A^\mu |\Psi\rangle = 0 \tag{6.55}$$

so that the operator $\partial_\mu A^\mu$ has vanishing matrix elements between physical states. Equation (6.54) is known as the *Gupta-Bleuler* condition. The linearity of the constraint means that the physical states $|\Psi\rangle$ span a physical Hilbert space $\mathcal{H}_{\text{phys}}$.

So what does the physical Hilbert space $\mathcal{H}_{\text{phys}}$ look like? And, in particular, have we rid ourselves of those nasty negative norm states so that $\mathcal{H}_{\text{phys}}$ has a positive definite inner product defined on it? The answer is actually no, but almost!

Let's consider a basis of states for the Fock space. We can decompose any element of this basis as $|\Psi\rangle = |\psi_T\rangle |\phi\rangle$, where $|\psi_T\rangle$ contains only transverse photons, created by

$a_{\vec{p}}^{1,2\,\dagger}$, while $|\phi\rangle$ contains the timelike photons created by $a_{\vec{p}}^{0\,\dagger}$ and longitudinal photons created by $a_{\vec{p}}^{3\,\dagger}$. The Gupta-Bleuler condition (6.54) requires

$$(a_{\vec{p}}^3 - a_{\vec{p}}^0)\,|\phi\rangle = 0 \tag{6.56}$$

This means that the physical states must contain combinations of timelike and longitudinal photons. Whenever the state contains a timelike photon of momentum $\vec{p}$, it must also contain a longitudinal photon with the same momentum. In general $|\phi\rangle$ will be a linear combination of states $|\phi_n\rangle$ containing n pairs of timelike and longitudinal photons, which we can write as

$$|\phi\rangle = \sum_{n=0}^{\infty} C_n\,|\phi_n\rangle \tag{6.57}$$

where $|\phi_0\rangle = |0\rangle$ is simply the vacuum. It's not hard to show that although the condition (6.56) does indeed decouple the negative norm states, all the remaining states involving timelike and longitudinal photons have zero norm

$$\langle \phi_m |\, \phi_n \rangle = \delta_{n0}\delta_{m0} \tag{6.58}$$

This means that the inner product on $\mathcal{H}_{\text{phys}}$ is positive semi-definite. Which is an improvement. But we still need to deal with all these zero norm states.

The way we cope with the zero norm states is to treat them as gauge equivalent to the vacuum. Two states that differ only in their timelike and longitudinal photon content, $|\phi_n\rangle$ with $n \geq 1$ are said to be physically equivalent. We can think of the gauge symmetry of the classical theory as descending to the Hilbert space of the quantum theory. Of course, we can't just stipulate that two states are physically identical unless they give the same expectation value for all physical observables. We can check that this is true for the Hamiltonian, which can be easily computed to be

$$H = \int \frac{d^3p}{(2\pi)^3}\,|\vec{p}| \left(\sum_{i=1}^{3} a_{\vec{p}}^{i\,\dagger}a_{\vec{p}}^i - a_{\vec{p}}^{0\,\dagger}a_{\vec{p}}^0 \right) \tag{6.59}$$

But the condition (6.56) ensures that $\langle\Psi|\,a_{\vec{p}}^{3\,\dagger}a_{\vec{p}}^3\,|\Psi\rangle = \langle\Psi|\,a_{\vec{p}}^{0\,\dagger}a_{\vec{p}}^0\,|\Psi\rangle$ so that the contributions from the timelike and longitudinal photons cancel amongst themselves in the Hamiltonian. This also renders the Hamiltonian positive definite, leaving us just with the contribution from the transverse photons as we would expect.

In general, one can show that the expectation values of all gauge invariant operators evaluated on physical states are independent of the coefficients C_n in (6.57).

Propagators

Finally, it's a simple matter to compute the propagator in Lorentz gauge. It is given by

$$\langle 0| \, T \, A_\mu(x) A_\nu(y) \, |0\rangle = \int \frac{d^4 p}{(2\pi)^4} \frac{-i\eta_{\mu\nu}}{p^2 + i\epsilon} \, e^{-ip\cdot(x-y)} \tag{6.60}$$

This is a lot nicer than the propagator we found in Coulomb gauge: in particular, it's Lorentz invariant. We could also return to the Lagrangian (6.37). Had we pushed through the calculation with arbitrary coefficient α, we would find the propagator,

$$\langle 0| \, T \, A_\mu(x) A_\nu(y) \, |0\rangle = \int \frac{d^4 p}{(2\pi)^4} \frac{-i}{p^2 + i\epsilon} \left(\eta_{\mu\nu} + (\alpha - 1)\frac{p_\mu p_\nu}{p^2} \right) e^{-ip\cdot(x-y)} \tag{6.61}$$

6.3 Coupling to Matter

Let's now build an interacting theory of light and matter. We want to write down a Lagrangian which couples A_μ to some matter fields, either scalars or spinors. For example, we could write something like

$$\mathcal{L} = -\tfrac{1}{4} F_{\mu\nu} F^{\mu\nu} - j^\mu A_\mu \tag{6.62}$$

where j^μ is some function of the matter fields. The equations of motion read

$$\partial_\mu F^{\mu\nu} = j^\nu \tag{6.63}$$

so, for consistency, we require

$$\partial_\mu j^\mu = 0 \tag{6.64}$$

In other words, j^μ must be a conserved current. But we've got lots of those! Let's look at how we can couple two of them to electromagnetism.

6.3.1 Coupling to Fermions

The Dirac Lagrangian

$$\mathcal{L} = \bar{\psi}(i \, \partial\!\!\!/ - m)\psi \tag{6.65}$$

has an internal symmetry $\psi \to e^{-i\alpha}\psi$ and $\bar\psi \to e^{+i\alpha}\bar\psi$, with $\alpha \in \mathbf{R}$. This gives rise to the conserved current $j_V^\mu = \bar\psi\gamma^\mu\psi$. So we could look at the theory of electromagnetism coupled to fermions, with the Lagrangian,

$$\mathcal{L} = -\tfrac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}(i \, \partial\!\!\!/ - m)\psi - e\bar\psi\gamma^\mu A_\mu\psi \tag{6.66}$$

where we've introduced a coupling constant e. For the free Maxwell theory, we have seen that the existence of a gauge symmetry was crucial in order to cut down the physical degrees of freedom to the requisite 2. Does our interacting theory above still have a gauge symmetry? The answer is yes. To see this, let's rewrite the Lagrangian as

$$\mathcal{L} = -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\not{D} - m)\psi \tag{6.67}$$

where $D_\mu\psi = \partial_\mu\psi + ieA_\mu\psi$ is called the *covariant derivative*. This Lagrangian is invariant under gauge transformations which act as

$$A_\mu \to A_\mu + \partial_\mu\lambda \quad \text{and} \quad \psi \to e^{-ie\lambda}\psi \tag{6.68}$$

for an arbitrary function $\lambda(x)$. The tricky term is the derivative acting on ψ, since this will also hit the $e^{-ie\lambda}$ piece after the transformation. To see that all is well, let's look at how the covariant derivative transforms. We have

$$\begin{aligned} D_\mu\psi &= \partial_\mu\psi + ieA_\mu\psi \\ &\to \partial_\mu(e^{-ie\lambda}\psi) + ie(A_\mu + \partial_\mu\lambda)(e^{-ie\lambda}\psi) \\ &= e^{-ie\lambda}D_\mu\psi \end{aligned} \tag{6.69}$$

so the covariant derivative has the nice property that it merely picks up a phase under the gauge transformation, with the derivative of $e^{-ie\lambda}$ cancelling the transformation of the gauge field. This ensures that the whole Lagrangian is invariant, since $\bar{\psi} \to e^{+ie\lambda(x)}\bar{\psi}$.

Electric Charge

The coupling e has the interpretation of the electric charge of the ψ particle. This follows from the equations of motion of classical electromagnetism $\partial_\mu F^{\mu\nu} = j^\nu$: we know that the j^0 component is the charge density. We therefore have the total charge Q given by

$$Q = e\int d^3x\, \bar{\psi}(\vec{x})\gamma^0\psi(\vec{x}) \tag{6.70}$$

After treating this as a quantum equation, we have

$$Q = e\int \frac{d^3p}{(2\pi)^3}\sum_{s=1}^{2}(b_{\vec{p}}^{s\,\dagger}b_{\vec{p}}^{s} - c_{\vec{p}}^{s\,\dagger}c_{\vec{p}}^{s}) \tag{6.71}$$

which is the number of particles, minus the number of antiparticles. Note that the particle and the anti-particle are required by the formalism to have opposite electric

charge. For QED, the theory of light interacting with electrons, the electric charge is usually written in terms of the dimensionless ratio α, known as the fine structure constant

$$\alpha = \frac{e^2}{4\pi\hbar c} \approx \frac{1}{137} \tag{6.72}$$

Setting $\hbar = c = 1$, we have $e = \sqrt{4\pi\alpha} \approx 0.3$.

There's a small subtlety here that's worth elaborating on. I stressed that there's a radical difference between the interpretation of a global symmetry and a gauge symmetry. The former takes you from one physical state to another with the same properties and results in a conserved current through Noether's theorem. The latter is a redundancy in our description of the system. Yet in electromagnetism, the gauge symmetry $\psi \to e^{+ie\lambda(x)}\psi$ seems to lead to a conservation law, namely the conservation of electric charge. This is because among the infinite number of gauge symmetries parameterized by a function $\lambda(x)$, there is also a single global symmetry: that with $\lambda(x) = $ constant. This is a true symmetry of the system, meaning that it takes us to another physical state. More generally, the subset of global symmetries from among the gauge symmetries are those for which $\lambda(x) \to \alpha = $ constant as $x \to \infty$. These take us from one physical state to another.

Finally, let's check that the 4×4 matrix C that we introduced in Section 4.5 really deserves the name "charge conjugation matrix". If we take the complex conjugation of the Dirac equation, we have

$$(i\gamma^\mu \partial_\mu - e\gamma^\mu A_\mu - m)\psi = 0 \quad \Rightarrow \quad (-i(\gamma^\mu)^\star \partial_\mu - e(\gamma^\mu)^\star A_\mu - m)\psi^\star = 0$$

Now using the defining equation $C^\dagger \gamma^\mu C = -(\gamma^\mu)^\star$, and the definition $\psi^{(c)} = C\psi^\star$, we see that the charge conjugate spinor $\psi^{(c)}$ satisfies

$$(i\gamma^\mu \partial_\mu + e\gamma^\mu A_\mu - m)\psi^{(c)} = 0 \tag{6.73}$$

So we see that the charge conjugate spinor $\psi^{(c)}$ satisfies the Dirac equation, but with charge $-e$ instead of $+e$.

6.3.2 Coupling to Scalars

For a real scalar field, we have no suitable conserved current. This means that we can't couple a real scalar field to a gauge field.

Let's now consider a complex scalar field φ. (For this section, I'll depart from our previous notation and call the scalar field φ to avoid confusing it with the spinor). We have a symmetry $\varphi \to e^{-i\alpha}\varphi$. We could try to couple the associated current to the gauge field,

$$\mathcal{L}_{\text{int}} = -i((\partial_\mu \varphi^\star)\varphi - \varphi^\star \partial_\mu \varphi)A^\mu \tag{6.74}$$

But this doesn't work because

- The theory is no longer gauge invariant

- The current j^μ that we coupled to A_μ depends on $\partial_\mu \varphi$. This means that if we try to compute the current associated to the symmetry, it will now pick up a contribution from the $j^\mu A_\mu$ term. So the whole procedure wasn't consistent.

We solve both of these problems simultaneously by remembering the covariant derivative. In this scalar theory, the combination

$$\mathcal{D}_\mu \varphi = \partial_\mu \varphi + ieA_\mu \varphi \tag{6.75}$$

again transforms as $\mathcal{D}_\mu \varphi \to e^{-ie\lambda}\mathcal{D}_\mu \varphi$ under a gauge transformation $A_\mu \to A_\mu + \partial_\mu \lambda$ and $\varphi \to e^{-ie\lambda}\varphi$. This means that we can construct a gauge invariant action for a charged scalar field coupled to a photon simply by promoting all derivatives to covariant derivatives

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \mathcal{D}_\mu \varphi^\star \mathcal{D}^\mu \varphi - m^2 |\varphi|^2 \tag{6.76}$$

In general, this trick works for any theory. If we have a $U(1)$ symmetry that we wish to couple to a gauge field, we may do so by replacing all derivatives by suitable covariant derivatives. This procedure is known as *minimal coupling*.

6.4 QED

Let's now work out the Feynman rules for the full theory of quantum electrodynamics (QED) – the theory of electrons interacting with light. The Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\not{D} - m)\psi \tag{6.77}$$

where $D_\mu = \partial_\mu + ieA_\mu$.

The route we take now depends on the gauge choice. If we worked in Lorentz gauge previously, then we can jump straight to Section 6.5 where the Feynman rules for QED are written down. If, however, we worked in Coulomb gauge, then we still have a bit of work in front of us in order to massage the photon propagator into something Lorentz invariant. We will now do that.

In Coulomb gauge $\nabla \cdot \vec{A} = 0$, the equation of motion arising from varying A_0 is now

$$-\nabla^2 A_0 = e\psi^\dagger \psi \equiv ej^0 \qquad (6.78)$$

which has the solution

$$A_0(\vec{x}, t) = e \int d^3x' \, \frac{j^0(\vec{x}', t)}{4\pi|\vec{x} - \vec{x}'|} \qquad (6.79)$$

In Coulomb gauge we can rewrite the Maxwell part of the Lagrangian as

$$L_{\text{Maxwell}} = \int d^3x \, \tfrac{1}{2}\vec{E}^2 - \tfrac{1}{2}\vec{B}^2$$

$$= \int d^3x \, \tfrac{1}{2}(\dot{\vec{A}} + \nabla A_0)^2 - \tfrac{1}{2}\vec{B}^2$$

$$= \int d^3x \, \tfrac{1}{2}\dot{\vec{A}}^2 + \tfrac{1}{2}(\nabla A_0)^2 - \tfrac{1}{2}\vec{B}^2 \qquad (6.80)$$

where the cross-term has vanished using $\nabla \cdot \vec{A} = 0$. After integrating the second term by parts and inserting the equation for A_0, we have

$$L_{\text{Maxwell}} = \int d^3x \, \left[\tfrac{1}{2}\dot{\vec{A}}^2 - \tfrac{1}{2}\vec{B}^2 + \frac{e^2}{2} \int d^3x' \, \frac{j_0(\vec{x})j_0(\vec{x}')}{4\pi|\vec{x} - \vec{x}'|} \right] \qquad (6.81)$$

We find ourselves with a nonlocal term in the action. This is exactly the type of interaction that we boasted in Section 1.1.4 never arises in Nature! It appears here as an artifact of working in Coulomb gauge: it does not mean that the theory of QED is nonlocal. For example, it wouldn't appear if we worked in Lorentz gauge.

We now compute the Hamiltonian. Changing notation slightly from previous chapters, we have the conjugate momenta,

$$\vec{\Pi} = \frac{\partial \mathcal{L}}{\partial \dot{\vec{A}}} = \dot{\vec{A}}$$

$$\pi_\psi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\psi^\dagger \qquad (6.82)$$
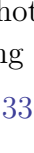
which gives us the Hamiltonian
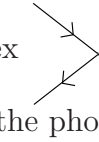
$$H = \int d^3x \, \left[\tfrac{1}{2}\dot{\vec{A}}^2 + \tfrac{1}{2}\vec{B}^2 + \bar{\psi}(-i\gamma^i\partial_i + m)\psi - e\vec{j} \cdot \vec{A} + \frac{e^2}{2} \int d^3x' \, \frac{j^0(\vec{x})j^0(\vec{x}')}{4\pi|\vec{x} - \vec{x}'|} \right]$$
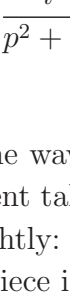
where $\vec{j} = \bar{\psi}\vec{\gamma}\psi$ and $j^0 = \bar{\psi}\gamma^0\psi$.

6.4.1 Naive Feynman Rules

We want to determine the Feynman rules for this theory. For fermions, the rules are the same as those given in Section 5. The new pieces are:

- We denote the photon by a wavy line. Each end of the line comes with an $i, j = 1, 2, 3$ index telling us the component of $\vec{A}$. We calculated the transverse photon propagator in (6.33): it is $\sim\!\!\sim\!\!\sim\!\!\sim$ and contributes $D_{ij}^{\mathrm{tr}} = \dfrac{i}{p^2 + i\epsilon}\left(\delta_{ij} - \dfrac{p_i p_j}{|\vec{p}|^2}\right)$

- The vertex $\nearrow\!\!\!\sim\!\!\!\sim$ contributes $-ie\gamma^i$. The index on γ^i contracts with the index on the photon line.

- The non-local interaction which, in position space, is given by $\nearrow\!\!\cdots\!\!\nwarrow$ contributes a factor of $\dfrac{i(e\gamma^0)^2 \delta(x^0 - y^0)}{4\pi|\vec{x} - \vec{y}|}$

These Feynman rules are rather messy. This is the price we've paid for working in Coulomb gauge. We'll now show that we can massage these expressions into something much more simple and Lorentz invariant. Let's start with the offending instantaneous interaction. Since it comes from the A_0 component of the gauge field, we could try to redefine the propagator to include a D_{00} piece which will capture this term. In fact, it fits quite nicely in this form: if we look in momentum space, we have

$$\frac{\delta(x^0 - y^0)}{4\pi|\vec{x} - \vec{y}|} = \int \frac{d^4p}{(2\pi)^4} \frac{e^{ip\cdot(x-y)}}{|\vec{p}|^2} \tag{6.83}$$

so we can combine the non-local interaction with the transverse photon propagator by defining a new photon propagator

$$D_{\mu\nu}(p) = \begin{cases} +\dfrac{i}{|\vec{p}|^2} & \mu, \nu = 0 \\[2mm] \dfrac{i}{p^2 + i\epsilon}\left(\delta_{ij} - \dfrac{p_i p_j}{|\vec{p}|^2}\right) & \mu = i \neq 0, \nu = j \neq 0 \\[2mm] 0 & \text{otherwise} \end{cases} \tag{6.84}$$
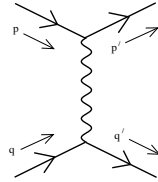
With this propagator, the wavy photon line now carries a $\mu, \nu = 0, 1, 2, 3$ index, with the extra $\mu = 0$ component taking care of the instantaneous interaction. We now need to change our vertex slightly: the $-ie\gamma^i$ above gets replaced by $-ie\gamma^\mu$ which correctly accounts for the $(e\gamma^0)^2$ piece in the instantaneous interaction.

The D_{00} piece of the propagator doesn't look a whole lot different from the transverse photon propagator. But wouldn't it be nice if they were both part of something more symmetric! In fact, they are. We have the following:

Claim: We can replace the propagator $D_{\mu\nu}(p)$ with the simpler, Lorentz invariant propagator

$$D_{\mu\nu}(p) = -i\frac{\eta_{\mu\nu}}{p^2} \tag{6.85}$$

Proof: There is a general proof using current conservation. Here we'll be more pedestrian and show that we can do this for certain Feynman diagrams. In particular, we focus on a particular tree-level diagram that contributes to $e^- e^- \to e^- e^-$ scattering,

 $\sim\ e^2 [\bar{u}(p')\gamma^\mu u(p)]\, D_{\mu\nu}(k)\, [\bar{u}(q')\gamma^\nu u(q)] \tag{6.86}$

where $k = p - p' = q' - q$. Recall that $u(\vec{p})$ satisfies the equation

$$(\not{p} - m)u(\vec{p}) = 0 \tag{6.87}$$
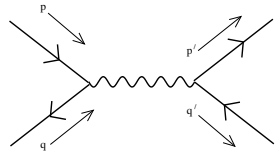
Let's define the spinor contractions $\alpha^\mu = \bar{u}(\vec{p}')\gamma^\mu u(\vec{p})$ and $\beta^\nu = \bar{u}(\vec{q}')\gamma^\nu u(\vec{q})$. Then since $k = p - p' = q' - q$, we have

$$k_\mu \alpha^\mu = \bar{u}(\vec{p}')(\not{p} - \not{p}')u(\vec{p}) = \bar{u}(\vec{p}')(m - m)u(\vec{p}) = 0 \tag{6.88}$$

and, similarly, $k_\nu \beta^\nu = 0$. Using this fact, the diagram can be written as

$$
\begin{aligned}
\alpha^\mu D_{\mu\nu}\beta^\nu &= i\left(\frac{\vec{\alpha}\cdot\vec{\beta}}{k^2} - \frac{(\vec{\alpha}\cdot\vec{k})(\vec{\beta}\cdot\vec{k})}{k^2|\vec{k}|^2} + \frac{\alpha^0 \beta^0}{|\vec{k}|^2}\right) \\
&= i\left(\frac{\vec{\alpha}\cdot\vec{\beta}}{k^2} - \frac{k_0^2 \alpha_0 \beta_0}{k^2|\vec{k}|^2} + \frac{\alpha_0 \beta_0}{|\vec{k}|^2}\right) \\
&= i\left(\frac{\vec{\alpha}\cdot\vec{\beta}}{k^2} - \frac{1}{k^2|\vec{k}|^2}\,(k_0^2 - k^2)\,\alpha_0 \beta_0\right) \\
&= -\frac{i}{k^2}\alpha\cdot\beta = \alpha^\mu \left(-\frac{i\eta_{\mu\nu}}{k^2}\right)\beta^\nu
\end{aligned} \tag{6.89}
$$

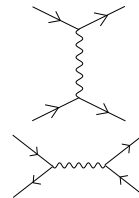which is the claimed result. You can similarly check that the same substitution is legal in the diagram

 $$\sim\ e^2[\bar{v}(\vec{q})\gamma^\mu u(\vec{p})]D_{\mu\nu}(k)[\bar{u}(\vec{p}')\gamma^\nu v(\vec{q}')] \tag{6.90}$$

In fact, although we won't show it here, it's a general fact that in every Feynman diagram we may use the very nice, Lorentz invariant propagator $D_{\mu\nu} = -i\eta_{\mu\nu}/p^2$. $\qquad\square$

Note: This is the propagator we found when quantizing in Lorentz gauge (using the Feynman gauge parameter). In general, quantizing the Lagrangian (6.37) in Lorentz gauge, we have the propagator

$$D_{\mu\nu} = -\frac{i}{p^2}\left(\eta_{\mu\nu} + (\alpha - 1)\frac{p_\mu p_\nu}{p^2}\right) \tag{6.91}$$
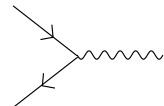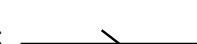
Using similar arguments to those given above, you can show that the $p_\mu p_\nu/p^2$ term cancels in all diagrams. For example, in the following diagrams the $p_\mu p_\nu$ piece of the propagator contributes as

 $$\sim \bar{u}(p')\gamma^\mu u(p)\,k_\mu = \bar{u}(p')(\not{p} - \not{p}')u(p) = 0$$

 $$\sim \bar{v}(p)\gamma^\mu u(q)\,k_\mu = \bar{u}(p)(\not{p} + \not{q}')u(q) = 0 \tag{6.92}$$

6.5 Feynman Rules

Finally, we have the Feynman rules for QED. For vertices and internal lines, we write

- Vertex: $-ie\gamma^\mu$

- Photon Propagator: $-\dfrac{i\eta_{\mu\nu}}{p^2 + i\epsilon}$

- Fermion Propagator: $\dfrac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon}$

For external lines in the diagram, we attach

- Photons: We add a polarization vector $\epsilon^\mu_{\text{in}}/\epsilon^\mu_{\text{out}}$ for incoming/outgoing photons. In Coulomb gauge, $\epsilon^0 = 0$ and $\vec{\epsilon}\cdot\vec{p} = 0$.

- Fermions: We add a spinor $u^r(\vec{p})/\bar{u}^r(\vec{p})$ for incoming/outgoing fermions. We add a spinor $\bar{v}^r(\vec{p})/v^r(\vec{p})$ for incoming/outgoing anti-fermions.

6.5.1 Charged Scalars

> "Pauli asked me to calculate the cross section for pair creation of scalar particles by photons. It was only a short time after Bethe and Heitler had solved the same problem for electrons and positrons. I met Bethe in Copenhagen at a conference and asked him to tell me how he did the calculations. I also inquired how long it would take to perform this task; he answered, "It would take me three days, but you will need about three weeks." He was right, as usual; furthermore, the published cross sections were wrong by a factor of four."
>
> *Viki Weisskopf*

The interaction terms in the Lagrangian for charged scalars come from the covariant derivative terms,

$$\mathcal{L} = \mathcal{D}_\mu \psi^\dagger \, \mathcal{D}^\mu \psi = \partial_\mu \psi^\dagger \partial^\mu \psi - ieA_\mu(\psi^\dagger \partial^\mu \psi - \psi \partial^\mu \psi^\dagger) + e^2 A_\mu A^\mu \psi^\dagger \psi \qquad (6.93)$$

This gives rise to two interaction vertices. But the cubic vertex is something we haven't seen before: it contains kinetic terms. How do these appear in the Feynman rules? After a Fourier transform, the derivative term means that the interaction is stronger for fermions with higher momentum, so we include a momentum factor in the Feynman rule. There is also a second, "seagull" graph. The two Feynman rules are

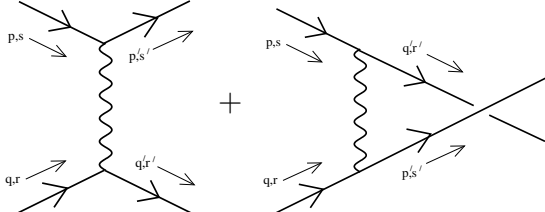$$- ie(p + q)_\mu \qquad \text{and} \qquad + 2ie^2\eta_{\mu\nu}$$

The factor of two in the seagull diagram arises because of the two identical particles appearing in the vertex. (It's the same reason that the 1/4! didn't appear in the Feynman rules for ϕ^4 theory).

6.6 Scattering in QED

Let's now calculate some amplitudes for various processes in quantum electrodynamics, with a photon coupled to a single fermion. We will consider the analogous set of processes that we saw in Section 3 and Section 5. We have

Electron Scattering

Electron scattering $e^- e^- \rightarrow e^- e^-$ is described by the two leading order Feynman diagrams, given by
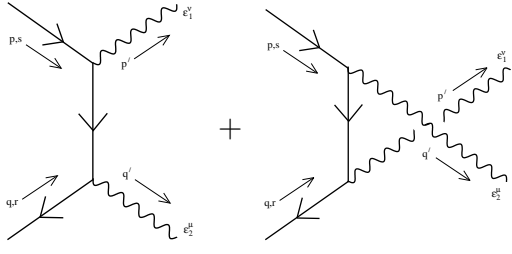

$$= -i(-ie)^2 \left(\frac{[\bar{u}^{s'}(\vec{p}')\gamma^\mu u^s(\vec{p})]\,[\bar{u}^{r'}(\vec{q}')\gamma_\mu u^r(\vec{q})]}{(p'-p)^2} \right.$$

$$\left. - \frac{[\bar{u}^{s'}(\vec{p}')\gamma^\mu u^r(\vec{q})]\,[\bar{u}^{r'}(\vec{q}')\gamma_\mu u^s(\vec{p})]}{(p-q')^2} \right)$$

The overall $-i$ comes from the $-i\eta_{\mu\nu}$ in the propagator, which contract the indices on the γ-matrices (remember that it's really positive for $\mu, \nu = 1, 2, 3$).

Electron Positron Annihilation
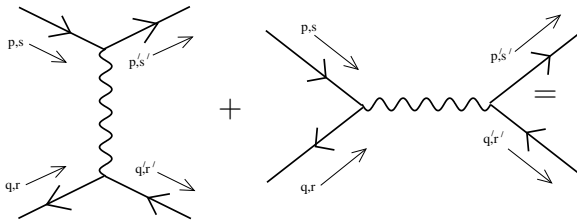
Let's now look at $e^- e^+ \rightarrow 2\gamma$, two gamma rays. The two lowest order Feynman diagrams are,

$$= i(-ie)^2\, \bar{v}^r(\vec{q}) \left(\frac{\gamma_\mu(\not{p} - \not{p}' + m)\gamma_\nu}{(p-p')^2 - m^2} \right.$$

$$\left. + \frac{\gamma_\nu(\not{p} - \not{q}' + m)\gamma_\mu}{(p-q')^2 - m^2} \right) u^s(\vec{p})\epsilon_1^\nu(\vec{p}')\epsilon_2^\mu(\vec{q}')$$

Electron Positron Scattering

For $e^- e^+ \rightarrow e^- e^+$ scattering (sometimes known as Bhabha scattering) the two lowest order Feynman diagrams are
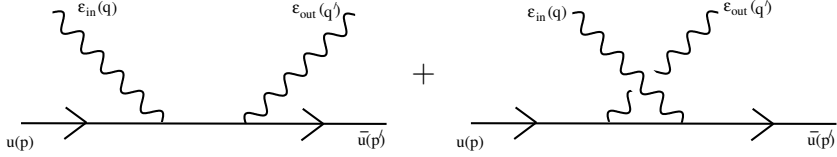

$$= -i(-ie)^2 \left(-\frac{[\bar{u}^{s'}(\vec{p}')\gamma^\mu u^s(\vec{p})]\,[\bar{v}^r(\vec{q})\gamma_\mu v^{r'}(\vec{q}')]}{(p-p')^2} \right.$$

$$\left. + \frac{[\bar{v}^r(\vec{q})\gamma^\mu u^s(\vec{p})]\,[\bar{u}^{s'}(\vec{p}')\gamma_\mu v^{r'}(\vec{q}')]}{(p+q)^2} \right)$$

Compton Scattering

The scattering of photons (in particular x-rays) off electrons $e^- \gamma \rightarrow e^- \gamma$ is known as Compton scattering. Historically, the change in wavelength of the photon in the

scattering process was one of the conclusive pieces of evidence that light could behave as a particle. The amplitude is given by,

$$= \; i(-ie)^2 \bar{u}^{r'}(\vec{p}') \left(\frac{\gamma_\mu (\not{p} + \not{q} + m)\gamma_\nu}{(p+q)^2 - m^2} + \frac{\gamma_\nu (\not{p} - \not{q}' + m)\gamma_\mu}{(p-q')^2 - m^2} \right) u^s(\vec{p}) \, \epsilon_{\text{in}}^\nu \, \epsilon_{\text{out}}^\mu$$

This amplitude vanishes for longitudinal photons. For example, suppose $\epsilon_{\text{in}} \sim q$. Then, using momentum conservation $p + q = p' + q'$, we may write the amplitude as

$$i\mathcal{A} = i(-ie)^2 \bar{u}^{r'}(\vec{p}') \left(\frac{\not{\epsilon}_{\text{out}}(\not{p} + \not{q} + m)\not{q}}{(p+q)^2 - m^2} + \frac{\not{q}(\not{p}' - \not{q} + m)\not{\epsilon}_{\text{out}}}{(p'-q)^2 - m^2} \right) u^s(\vec{p})$$

$$= i(-ie)^2 \bar{u}^{r'}(\vec{p}') \, \not{\epsilon}_{\text{out}} u^s(\vec{p}) \left(\frac{2p \cdot q}{(p+q)^2 - m^2} + \frac{2p' \cdot q}{(p'-q)^2 - m^2} \right) \qquad (6.94)$$

where, in going to the second line, we've performed some γ-matrix manipulations, and invoked the fact that q is null, so $\not{q}\not{q} = 0$, together with the spinor equations $(\not{p} - m)u(\vec{p})$ and $\bar{u}(\vec{p}')(\not{p}' - m) = 0$. We now recall the fact that q is a null vector, while $p^2 = (p')^2 = m^2$ since the external legs are on mass-shell. This means that the two denominators in the amplitude read $(p + q)^2 - m^2 = 2p \cdot q$ and $(p' - q)^2 - m^2 = -2p' \cdot q$. This ensures that the combined amplitude vanishes for longitudinal photons as promised. A similar result holds when $\not{\epsilon}_{\text{out}} \sim q'$.

Photon Scattering

In QED, photons no longer pass through each other unimpeded. At one-loop, there is a diagram which leads to photon scattering. Although naively logarithmically divergent, the diagram is actually rendered finite by gauge invariance.
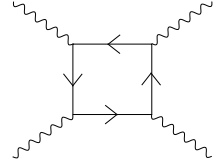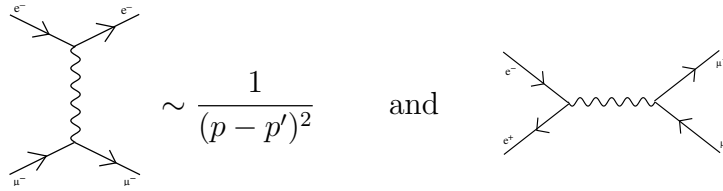

Figure 30:

Adding Muons

Adding a second fermion into the mix, which we could identify as a muon, new processes become possible. For example, we can now have processes such as $e^- \mu^- \to e^- \mu^-$ scattering, and $e^+ e^-$ annihilation into a muon anti-muon pair. Using our standard notation of p and q for incoming momenta, and p' and q' for outgoing

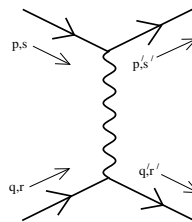momenta, we have the amplitudes given by

$$\sim \frac{1}{(p - p')^2} \qquad \text{and} \qquad \sim \frac{1}{(p + q)^2} \qquad (6.95)$$

6.6.1 The Coulomb Potential

We've come a long way. We've understood how to compute quantum amplitudes in a large array of field theories. To end this course, we use our newfound knowledge to rederive a result you learnt in kindergarten: Coulomb's law.

To do this, we repeat our calculation that led us to the Yukawa force in Sections 3.5.2 and 5.7.2. We start by looking at $e^- e^- \to e^- e^-$ scattering. We have
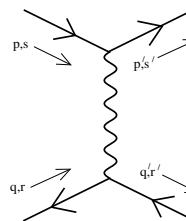
$$= -i(-ie)^2 \frac{[\bar{u}(\vec{p}')\gamma^\mu u(\vec{p})]\,[\bar{u}(\vec{q}')\gamma_\mu u(\vec{q})]}{(p' - p)^2} \qquad (6.96)$$

Following (5.49), the non-relativistic limit of the spinor is $u(p) \to \sqrt{m}\begin{pmatrix} \xi \\ \xi \end{pmatrix}$. This means that the γ^0 piece of the interaction gives a term $\bar{u}^s(\vec{p})\gamma^0 u^r(\vec{q}) \to 2m\delta^{rs}$, while the spatial γ^i, $i = 1, 2, 3$ pieces vanish in the non-relativistic limit: $\bar{u}^s(\vec{p})\gamma^i u^r(\vec{q}) \to 0$. Comparing the scattering amplitude in this limit to that of non-relativistic quantum mechanics, we have the effective potential between two electrons given by,

$$U(\vec{r}) = +e^2 \int \frac{d^3 p}{(2\pi)^3} \frac{e^{i\vec{p}\cdot\vec{r}}}{|\vec{p}|^2} = +\frac{e^2}{4\pi r} \qquad (6.97)$$

We find the familiar repulsive Coulomb potential. We can trace the minus sign that gives a repulsive potential to the fact that only the A_0 component of the intermediate propagator $\sim -i\eta_{\mu\nu}$ contributes in the non-relativistic limit.

For $e^- e^+ \to e^- e^+$ scattering, the amplitude is

$$= +i(-ie)^2 \frac{[\bar{u}(\vec{p}')\gamma^\mu u(\vec{p})]\,[\bar{v}(\vec{q})\gamma_\mu v(\vec{q}')]}{(p' - p)^2} \qquad (6.98)$$
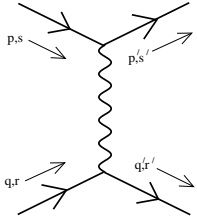
The overall + sign comes from treating the fermions correctly: we saw the same minus sign when studying scattering in Yukawa theory. The difference now comes from looking at the non-relativistic limit. We have $\bar{v}\gamma^0 v \to 2m$, giving us the potential between opposite charges,

$$U(\vec{r}) = -e^2 \int \frac{d^3 p}{(2\pi)^3} \frac{e^{i\vec{p}\cdot\vec{r}}}{|\vec{p}|^2} = -\frac{e^2}{4\pi r} \tag{6.99}$$

Reassuringly, we find an attractive force between an electron and positron. The difference from the calculation of the Yukawa force comes again from the zeroth component of the gauge field, this time in the guise of the γ^0 sandwiched between $\bar{v}\gamma^0 v \to 2m$, rather than the $\bar{v}v \to -2m$ that we saw in the Yukawa case.

The Coulomb Potential for Scalars

There are many minus signs in the above calculation which somewhat obscure the crucial one which gives rise to the repulsive force. A careful study reveals the offending sign to be that which sits in front of the A_0 piece of the photon propagator $-i\eta_{\mu\nu}/p^2$. Note that with our signature $(+---)$, the propagating A_i have the correct sign, while A_0 comes with the wrong sign. This is simpler to see in the case of scalar QED, where we don't have to worry about the gamma matrices. From the Feynman rules of Section 6.5.1, we have the non-relativistic limit of scalar $e^- e^-$ scattering,

$$= -i\eta_{\mu\nu}(-ie)^2 \frac{(p+p')^\mu (q+q')_\nu}{(p'-p)^2} \to -i(-ie)^2 \frac{(2m)^2}{-(\vec{p}-\vec{p}')^2}$$

where the non-relativistic limit in the numerator involves $(p+p')\cdot(q+q') \approx (p+p')^0(q+q')_0 \approx (2m)^2$ and is responsible for selecting the A_0 part of the photon propagator rather than the A_i piece. This shows that the Coulomb potential for spin 0 particles of the same charge is again repulsive, just as it is for fermions. For $e^- e^+$ scattering, the amplitude picks up an extra minus sign because the arrows on the legs of the Feynman rules in Section 6.5.1 are correlated with the momentum arrows. Flipping the arrows on one pair of legs in the amplitude introduces the relevant minus sign to ensure that the non-relativistic potential between $e^- e^+$ is attractive as expected.

6.7 Afterword

In this course, we have laid the foundational framework for quantum field theory. Most of the developments that we've seen were already in place by the middle of the 1930s, pioneered by people such as Jordan, Dirac, Heisenberg, Pauli and Weisskopf[5].

Yet by the end of the 1930s, physicists were ready to give up on quantum field theory. The difficulty lies in the next terms in perturbation theory. These are the terms that correspond to Feynamn diagrams with loops in them, which we have scrupulously avoided computing in this course. The reason we've avoided them is because they typically give infinity! And, after ten years of trying, and failing, to make sense of this, the general feeling was that one should do something else. This from Dirac in 1937,

> *Because of its extreme complexity, most physicists will be glad to see the end of QED*

But the leading figures of the day gave up too soon. It took a new generation of postwar physicists — people like Schwinger, Feynman, Tomonaga and Dyson — to return to quantum field theory and tame the infinities. The story of how they did that will be told in next term's course.

---

[5]For more details on the history of quantum field theory, see the excellent book "QED and the Men who Made it" by Sam Schweber.