Introduction to Quantum Field Theory

Matthew Schwartz
Harvard University

Fall 2008

Table of contents

Chapter 1
The Microscopic Theory of Radiation

1.1 Blackbody Radiation

Quantum Mechanics began on October 19, 2000 with Max Planck's explanation of the black-body radiation spectrum. Although his result led to the development of quantum mechanics, that is, the quantum mechanics of electrons, Planck's original observation was about the quantum nature of light, which is a topic for quantum field theory. So, radiation is a great way to motivate the need for a quantum theory of fields.

In 1900 people were very confused about how to explain the spectrum of radiation from hot objects. The equipartition theorem says that a body in thermal equilibrium should have energy equally distributed among the various modes. For a hot gas, this gives the Maxwell-Boltzmann distribution of thermal velocities, which is beautifully confirmed by data. But when you try to do the same thing for light, you get a very strange result.

Take a blackbody, which we can think of as a hot box of light (also known as a Jeans cube) of size L. We know that there are standing electromagnetic waves inside with frequencies

$$\nu_n = 2\pi \frac{|\vec{n}\,|}{L} \tag{1.1}$$

for integer 3-vectors $\vec{n}$. Before 1900 people thought you can have as much or as little energy in each mode as you want. Then the equipartition theorem says that since there are an infinite number of modes excited, the blackbody should emit light equally in all frequencies. In particular, there should be a lot of UV light – the "ultraviolet catastrophe." Instead, the distribution looks like a Maxwell-Boltzmann distribution.

This led Planck to postulate that the energy of each electromagnetic mode in the cavity is quantized in units of frequency

$$E_n = h\nu_n = \frac{2\pi}{L} h|\vec{n}\,| = |\vec{k}_n| \tag{1.2}$$

where h is Planck's constant. Einstein then interpreted this as saying light is made up of particles, called photons. Note that if the excitations are particles, then they are massless

$$m_n^2 = E_n^2 - |\vec{k}_n|^2 = 0 \tag{1.3}$$

Thus Planck and Einstein said that light is massless photons. Light having energy leads to the photoelectric effect; light having momentum leads to Compton scattering.

With Planck's energy hypothesis, the thermal distribution is easy to compute. Each mode of frequency ν_n can be excited j times, giving energy $E_n = j\nu_n$ in that mode. The probability of finding that much energy in the mode is the same as the probability of finding a certain amount of energy in *anything*, $P(E) = \exp(-E/k_B T)$. Thus the expectation value of the the energy in each mode is

$$\langle E_n \rangle = \frac{\sum_{j=0}^{\infty} (jE_n) e^{-jE_n\beta}}{\sum_{j=0}^{\infty} e^{-jE_n\beta}} = \frac{-\frac{d}{d\beta}\frac{1}{1-e^{-h\nu_n\beta}}}{\frac{1}{1-e^{h\nu_n\beta}}} = \frac{h\nu_n}{e^{h\nu_n\beta}-1} \tag{1.4}$$

where $\beta = 1/k_B T$. (This derivation is due to Debye. Nowadays we do the calculation using a canonical ensemble.).

Now let's take the continuum limit, $L \to \infty$. The distribution of energy in a blackbody becomes

$$E = \int d^3n \frac{h\nu_n}{e^{h\nu_n\beta} - 1} = 4\pi h \int d|\vec{n}| \frac{|\vec{n}|^2 \nu_n}{e^{h\nu_n\beta} - 1} \tag{1.5}$$

$$= 4\pi h \frac{L^3}{8\pi^3} \int d\nu \frac{\nu^3}{e^{h\nu\beta} - 1} \tag{1.6}$$

Thus,

$$I(\nu) = \frac{1}{V}\frac{dE}{d\nu} = \frac{h}{\pi^2} \frac{\nu^3}{e^{h\nu\beta} - 1} \tag{1.7}$$

Which is Plank's showed in 1900 correctly matches experiment. (NB: I threw in a factor of 2 for the 2 polarizations of light, which you shouldn't worry about at all yet.).

What does this have to do with Quantum Field Theory?

Well, in order for any of this to make sense, light has to be able to equilibrate. If we shine monochromatic light into the box, eventually it reaches all frequencies. But if different frequencies are different particles, these particles have to get created by something. Quantum field theory tells us how radiation works at the microscopic level.

1.2 Einstein Coefficients

The easiest handle on the creation of light comes from the coefficient of spontaneous emission. This is the rate for an excited atom to emit light. This phenomenon had been observed in chemical reactions, and as a form of radioactivity. But it was only understood statistically. In fact, Einstein in 1906 developed a beautiful proof of the relation between emission and absorption based on the existence of thermal equilibrium.

Einstein said: Suppose we have a cavity full of atoms with energy levels E_1 and E_2. Let $\omega = E_2 - E_1$. In equilibrium the number densities are determined by Boltzmann distributions.

$$n_1 = Ne^{-\beta E_1} \quad n_2 = Ne^{-\beta E_2} \tag{1.8}$$

The probability for an E_2 atom to emit a photon of energy ω and transition to state E_1 is called the coefficient for spontaneous emission A. The probability for a photon of energy ω to induce a transition from 2 to 1 is proportional to the coefficient of stimulated emission B_2. The induced rate is also proportional to the number of photons of energy ω. Then

$$dn_2 = [A + B_2 I(\omega)]\, n_2 \tag{1.9}$$

The probability for a photon to induce a transition from 1 to 2 is called the coefficient of stimulated absorption B_1. Then

$$dn_1 = B_1 I(\omega) n_1 \tag{1.10}$$

Now if we assume the system is in equilibrium then the rate going up must be the same as the rate going down: $dn_2 = dn_1$. Then

$$[A + B_2 I(\omega)]\, n_2 = B_1 I(\omega) n_1 \tag{1.11}$$

$$\left[B_1 e^{-\beta E_1} - B_2 e^{-\beta E_2}\right] I(\omega) = A e^{-\beta E_2} \tag{1.12}$$

$$I(\omega) = \frac{A}{B_1 e^{\beta\omega} - B_2} \tag{1.13}$$

But we already know that

$$I(\omega) = \frac{h}{2\pi^2} \frac{\omega^3}{e^{\beta\omega} - 1} \tag{1.14}$$

from above. Since equilibrium must be satisfied at any temperature,

$$B_2 = B_1 = B \tag{1.15}$$

and

$$\frac{A}{B} = \frac{\hbar}{2\pi^2}\omega^3 \tag{1.16}$$

This is a beautiful proof. If light has energy ω with a Planck distribution, then equilibrium relates A and B. However, why should you need equilibrium to say this? What causes an atom to emit a photon the first place?

The B coefficients, for stimulated emission/absorption can be calculated with a background electromagnetic field and quantum mechanics. Dirac did this in 1932. However, how do you calculated the coefficient of spontaneous emission B? It took 20 years to be able to calculate A/B from first principles. It took the invention of Quantum Field Theory.

1.3 Quantum Field Theory

The basic idea behind the calculation of the spontaneous emission coefficient was to treat the photon excitations like a quantum mechanics problem. Start by just looking at a single photon mode at a time, say of energy Δ. This mode can be excited j times. Each excitation adds energy Δ to the system.

There is a quantum mechanical system with this property – one you may remember from your quantum mechanics course: the simple harmonic oscillator. We will review it for homework. For now, let me just remind you of how it works. The easiest way to study the system is with creation and annihilation operators. These satisfy:

$$[a, a^\dagger] = 1 \tag{1.17}$$

There is also the number operator which counts modes

$$N = a^\dagger a \tag{1.18}$$

$$N|n\rangle = n|n\rangle \tag{1.19}$$

Then,

$$Na^\dagger|n\rangle = a^\dagger a a^\dagger|n\rangle = a^\dagger|n\rangle + a^\dagger a^\dagger a|n\rangle = (n+a)a^\dagger|n\rangle \tag{1.20}$$

and so

$$a^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle \tag{1.21}$$

$$a|n\rangle = \sqrt{n}\,|n-1\rangle \tag{1.22}$$

Where the normalization is set so that

$$\langle n|n\rangle = 1 \tag{1.23}$$

Now for the photon, we want to be able to arbitrarily excite any mode in the cavity. Thus

$$N = \sum_{\{k\}} a_k^\dagger a_k \tag{1.24}$$

This seems like absolutely nothing – just algebra. But it's not.

1.4 Spontaneous and stimulated emission

For example, let's look how a photon interacts with an atom. We want to use Fermi's golden rule to calculate the transition rate between states. Recall that the rule states that the rate is proportional to the matrix element squared

$$\Gamma \sim |M|^2 \delta(E_f - E_i) \tag{1.25}$$

where the δ-function serves to enforce energy conservations. Using perturbation theory, we know we can write interactions of the form

$$|M|^2 = |\langle f|H_{\rm int}|i\rangle|^2 \tag{1.26}$$

We don't know exactly what the interaction Hamiltonian $H_{\rm int}$ is, but it's got to have something to connect the initial and final atomic states and some creation operator or annihilation operator to create the photon. So it must look like

$$H_{\rm int} = H_0 a_k^\dagger + H_0 a_k + \cdots \tag{1.27}$$

For the $2 \to 1$ transition, the initial state is an excited atom with n photons of momenta k:

$$|i\rangle = |{\rm atom}'; n_k\rangle \tag{1.28}$$

The final state is a lower energy atom with $n+1$ photons of momentum k

$$\langle f| = \langle {\rm atom}; n_k + 1| \tag{1.29}$$

So,

$$M_{2\to 1} = \Big\langle {\rm atom}; n_k+1|H_0 a_k^\dagger + H_0 a_k|{\rm atom}'; n_k\Big\rangle \tag{1.30}$$

$$= \langle {\rm atom}|H_0|{\rm atom}'\rangle \Big\langle n_k+1|a_k^\dagger|n_k\Big\rangle \tag{1.31}$$

$$= M_0 \langle n_k+1|n_k+1\rangle \sqrt{n_k+1} \tag{1.32}$$

$$= M_0 \sqrt{n_k+1} \tag{1.33}$$

where

$$M_0 = \langle {\rm atom}|H_0|{\rm atom}'\rangle \tag{1.34}$$

Thus

$$|M_{2\to 1}|^2 = |M_0|^2 (n_k+1) \tag{1.35}$$

But if we are exciting an atom, then the initial state is an unexcited atom and n_Δ photons:

$$|i\rangle = |{\rm atom}; n_k\rangle \tag{1.36}$$

and the final state we have lost a photon and have an excited atom

$$\langle f| = \langle {\rm atom}'; n_k - 1| \tag{1.37}$$

Then

$$M_{1\to 2} = \Big\langle {\rm atom}'; n_k-1|H_0 a_k^\dagger + H_0 a_k|{\rm atom}; n_k\Big\rangle \tag{1.38}$$

$$= \langle {\rm atom}'|H_0|{\rm atom}\rangle \langle n_k - 1|a_k|n_k\rangle \tag{1.39}$$

$$= M_0 \sqrt{n_k} \tag{1.40}$$

Thus

$$dn_2 = |M_0|^2 n_2 (n_k+1) \tag{1.41}$$

$$dn_1 = |M_0|^2 \, n_2(n_k) \tag{1.42}$$

Remember Einstein's equations:

$$dn_2 = (A + B_2 I(\omega))\, n_2 \tag{1.43}$$

$$dn_1 = B_1 I(\omega) n_1 \tag{1.44}$$

Now,

$$E = \int \frac{d^3k}{(2\pi)^3}(hk)n_k = \int dk \frac{hk^3}{2\pi^2} n_k \tag{1.45}$$

$$I(k) = \frac{dE}{dk} = \frac{hk^3}{2\pi^2} n_k \tag{1.46}$$

$$\Rightarrow \quad n_k = \frac{\pi^2}{h} I(k) k^3 \tag{1.47}$$

Note that this doesn't depend on what n_k is, this is just a phase space integral. There is no mention of temperature or of equilibrium – it is true for any values of n_k.

Thus

$$dn_2 = Cn_2 \left[\frac{\pi^2}{\hbar k^3} I(k) + 1 \right] \tag{1.48}$$

$$dn_1 = Cn_1 \frac{\pi^2}{\hbar k^3} I(k) \tag{1.49}$$

Thus

$$B_1 = B_2 \tag{1.50}$$

$$\frac{A}{B} = \frac{2\pi^2}{\hbar k^3} \tag{1.51}$$

Which are Einstein's relations.

Beautiful!

This derivation is due to a paper of Dirac's from 1927.

Note that we never needed to talk about statistical mechanics, ensembles, or temperature.

Chapter 2
Lorentz Invariance and Conventions

2.1 Introduction

This lecture is meant to establish the conventions used in the course, and to review Lorentz invariance and notation. It is taken partly from Mike Luke's QFT lectures

$$\text{http://www2.physics.utoronto.ca/}\tilde{}\text{luke/PHY2403/References\_files/lecturenotes.pdf} \tag{2.1}$$

In particular section 1B. You might want to look at those lectures for more, and somewhat different, details.

2.2 Conventions

2.2.1 dimensional analysis

We will set

$$\hbar = \frac{h}{2\pi} = 1 \tag{2.2}$$

$$c = 1 \tag{2.3}$$

throughout.

This makes all quantities have dimensions of mass to some power. Sometimes we write the mass dimension of a quantity with brackets:

$$[M^d] = d \tag{2.4}$$

Really, everything is either a length (eg position x) or a mass (eg momentum p). With $\hbar = c = 1$ $[p] = 1/[x] = 1$. So

$$[x] = -1 \tag{2.5}$$

$$[t] = -1 \tag{2.6}$$

$$[p] = 1 \tag{2.7}$$

$$[\text{veolcity}] = \frac{[x]}{[t]} = 0 \tag{2.8}$$

Thus

$$[d^4x] = -4 \tag{2.9}$$

Action is dimensionless

$$[S] = [\int d^4x\, \mathcal{L}] = 0 \tag{2.10}$$

So Lagrangians have dimension 4

$$[\mathcal{L}] = 4 \tag{2.11}$$

For example, a free scalar field has Lagrangian

$$\mathcal{L} = \phi d^2 \phi \tag{2.12}$$

thus

$$[\phi] = 1 \tag{2.13}$$

and so on.

2.2.2 2π's

Keeping the factors of 2π straight can be a pain. They come from going back and forth between position and momentum space. For example,

$$\boxed{\int dk e^{ikx} = 2\pi\delta(x)} \tag{2.14}$$

In fact, you should internalize this relation as you will use it a lot.

The convention we will use is that momentum space integrals have 2πs and a $+$ sign in the exponent

$$f(x) = \int \frac{d^4k}{(2\pi)^4} \tilde{f}(k) e^{ikx} \tag{2.15}$$

The position space integrals have no $2\pi's$ and a minus sign in the exponent

$$\tilde{f}(k) = \int d^4x f(x) e^{-ikx} \tag{2.16}$$

2.3 Lorentz Invariance

Lorentz invariance implies a symmetry under rotations and boosts.

2.3.1 rotations

Rotations you should be extremely familiar with, and are more intuitive. For example, under rotations, a vector (x, y) transforms as

$$x \to x\cos\theta + y\sin\theta \tag{2.17}$$

$$y \to -x\sin\theta + y\cos\theta \tag{2.18}$$

We can write this as

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta + y\sin\theta \\ -x\sin\theta + y\cos\theta \end{pmatrix} \tag{2.19}$$

Or as

$$x_i \to R_{ij} x_j, \quad x_i = \begin{pmatrix} x \\ y \end{pmatrix}, \quad i = 1, 2 \tag{2.20}$$

We separate upper and lower indices to refer to column and row vectors respectively. So we can also write

$$x^i = (x, y) \to (x, y) \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = x^i R^T_{ij} \tag{2.21}$$

Note that $R^T = R^{-1}$. That is

$$R^T_{ij} R_{jk} = \delta_{jk} = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \tag{2.22}$$

or equivalently

$$R^T R = \mathbb{1} \tag{2.23}$$

This is actually enough to characterize it as a rotation, and we don't need the explicit form of the rotation matrix as a function of θ anymore.

We can contract upper and lower indices

$$x^i x_i = (x, y)\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2 \tag{2.24}$$

This is just the norm of the vector x_i and is rotation invariant. That $x^i x_i$ means sum over i is sometimes called the Einstein summation convention. In terms of matrices

$$x^i x_i \to x^i R^T R x_k = x^i x_i \tag{2.25}$$

since $R^T = R^{-1}$. In fact, we can formally describe the rotation group as preserving the inner product $x^i x_i = \delta_{ij} x^i x^j$

$$R_{ik} R_{jl} \delta_{kl} = [(R^T)(\delta)(R)]_{ij} = (R^T R)_{ij} = \delta_{ij} \tag{2.26}$$

which we can check explicitly

2.3.2 Lorentz transformations

Lorentz transformations work exactly like rotations, except we use vectors with 4 components with a Greek index V_μ, and the matrices satisfy

$$\Lambda^T \eta \Lambda = \eta = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{2.27}$$

$\eta_{\mu\nu}$ is known as the Minkoswki metric. So Lorentz transformations preserve the Minkowskian inner product

$$x^i x_j = x_i \eta_{ij} x_j = x_0^2 - x_1^2 - x_2^2 - x_3^2 \tag{2.28}$$

The Lorentz group is the most general set of transformations with this property.

A general Lorentz transformation can be written as a combination of 3 rotations and 3 boosts:

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_{xy} & -\sin\theta_{xy} & \\ 0 & \sin\theta_{xy} & \cos\theta_{xy} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & \cos\theta_{xz} & & -\sin\theta_{xz} \\ & & 1 & \\ & \sin\theta_{xz} & & \cos\theta_{xz} \end{pmatrix} \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \cos\theta_{xz} & -\sin\theta_{xz} \\ & & \sin\theta_{xz} & \cos\theta_{xz} \end{pmatrix} \tag{2.29}$$

$$\times \begin{pmatrix} \cosh\beta_x & \sinh\beta_x & & \\ \sinh\beta_x & \cosh\beta_x & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \begin{pmatrix} \cosh\beta_y & & \sinh\beta_y & \\ & 1 & & \\ \sinh\beta_y & & \cosh\beta_y & \\ & & & 1 \end{pmatrix} \begin{pmatrix} \cosh\beta_z & & & \sinh\beta_z \\ & 1 & & \\ & & 1 & \\ \sinh\beta_z & & & \cosh\beta_z \end{pmatrix} \tag{2.30}$$

Note that these matrices do not commute, so the order we do the rotations and boosts is important. We will only very rarely need an actual matrix representation of the group elements like this, but it is helpful to see it.

So now vectors transform as

$$V_\mu \to \Lambda^\nu_\mu V_\nu \tag{2.31}$$

$$W^\mu \to W^\nu (\Lambda^T)^\mu_\nu \tag{2.32}$$

That is, these vectors transform covariantly or contravariantly. The contraction of a covariant and contravariant vector is Lorentz *invariant*.

$$W^\mu V_\mu \to W^\nu (\Lambda^T)^\mu_\nu \Lambda^\alpha_\mu V_\alpha = W^\mu V_\mu \tag{2.33}$$

Only when you are using explicit representations of the Lorentz group elements to the index positions really matter. Otherwise *as long as the indices are contracted, I don't care whether they are up or down*:

$$V_\mu W^\mu = V^\mu W_\mu = V_\mu W_\mu \tag{2.34}$$

$$= V_0 W_0 - V_1 W_1 - V_2 W_2 - V_3 W_3 \tag{2.35}$$

This is a bad habit and I probably shouldn't be encouraging it, but as I said, this is meant to be a practical course.

Conventionally, position vectors have upper indices

$$x^\mu = (t, x, y, z) \tag{2.36}$$

derivatives have lower indices

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = (\partial_t, \partial_x, \partial_y, \partial_z) \tag{2.37}$$

momenta are like derivatives and have lower indices

$$p_\mu = (E, p_x, p_y, p_z) \tag{2.38}$$

Note that sometimes you will see (Mike Luke does this) $p_\mu = (E, -p_x, -p_y, -p_z)$. It doesn't matter if you put these signs in the individual vector elements as long as you always contract with the Minkowski metric $\eta = \mathrm{diag}(1, -1, -1, -1)$. That is, the signs only really matter when you write out the explicit contraction of a Lorentz invariant expression.

Scalar fields are Lorentz invariant

$$\phi(x) \to \phi(x) \tag{2.39}$$

Note that our definitions of x may be changing $x \to \Lambda x$, but the space-time point x is fixed, just our labels for it change. For example, think of temperature as a scalar. If I change frames, the labels for the points changes, but the temperature at that point stays the same.

Vectors transform as

$$V_\mu(x) \to \Lambda_{\mu\nu} V_\nu(x) \tag{2.40}$$

I'm not bothering with up and down indices, but the proper thing to do is write $V_\mu = \Lambda_\mu{}^\nu V_\nu$. Again, x transforms too, but passively. The difference with a scalar is that the components of a vector field at the point x rotate into each other as well.

Tensors transform as

$$T_{\mu\nu} \to \Lambda_{\mu\alpha}\Lambda_{\nu\beta} T_{\alpha\beta} \tag{2.41}$$

and so on.

We also write

$$\Box = \partial_\mu^2 = \partial^\mu \partial_\mu = \partial_t^2 - \partial_x^2 - \partial_y^2 - \partial_z^2 \tag{2.42}$$

$$\triangle = \nabla^2 = \partial_x^2 + \partial_y^2 + \partial_z^2 \tag{2.43}$$

Note that you should never have anything like

$$V_\mu W_\mu X_\mu \tag{2.44}$$

with 3 (or more) of the same indices. Be very careful to relabel things. For example, don't write

$$(V^2)(W^2) = V_\mu V_\mu W_\mu W_\mu \tag{2.45}$$

instead write

$$(V^2)(W^2) = V_\mu^2 W_\nu^2 = V_\mu V_\mu W_\nu W_\nu = \eta^{\mu\alpha}\eta^{\nu\beta} V_\mu V_\alpha W_\nu W_\beta \tag{2.46}$$

You will quickly get the hang of all this contracting.

In summary, we say

$$V^2 = V^\mu V_\mu, \quad \phi, \quad 1, \quad \partial_\mu V_\mu \tag{2.47}$$

are Lorentz *invariant*. But

$$V_\mu, F_{\mu\nu}, \partial_\mu, x^\mu \tag{2.48}$$

are Lorentz *covariant* (or contravariant, but only mathematicians say that).

2.4 Discrete transformations

Lorentz transformations are those which preserve the Minkowski metric

$$\Lambda^T \eta \Lambda = \eta \tag{2.49}$$

Equivalently, they are those which leave

$$x_\mu x^\mu = t^2 - x^2 - y^2 - z^2 \tag{2.50}$$

invariant.

So the transformations

$$P\colon (t, x, y, z) \to (t, -x, -y, -z) \quad \text{(parity)} \tag{2.51}$$

and

$$T\colon (t, x, y, z) \to (-t, x, y, z) \quad \text{(time reversal)} \tag{2.52}$$

are technically Lorentz transformations too. They can be written as

$$T = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \tag{2.53}$$

$$P = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{2.54}$$

We say that the Lorentz transformations which are continuously connected to the identity transformation are Proper Othonchronous Lorentz transformations. They can be characterized by

$$\det \Lambda = 1 \tag{2.55}$$

$$\Lambda_{00} \geqslant 0 \tag{2.56}$$

With T and P we get all Lorentz transformations. Note $\det \mathrm{P} = \det \mathrm{T} = -1$ and $T_{00} < 0$.

Also, we say a vector is *timelike* when

$$V_\mu^2 > 0, \quad \text{(timelike)} \tag{2.57}$$

and *spacelike* when

$$V_\mu^2 < 0, \quad \text{(spacelike)} \tag{2.58}$$

Obviously time $= (t, 0, 0, 0)$ is timelike and $(0, x, 0, 0)$ is spacelike. Whether something is timelike or spacelike is preserved under Lorentz transformations since the norm is preserved. If it's norm is zero we say it is *lightlike*

$$V_\mu^2 = 0, \quad \text{(lightlike)} \tag{2.59}$$

If V_μ is a 4-momentum, this says that it is massless. Photons are massless, which is the origin of the term lightlike.

Chapter 3
Second Quantization

3.1 Introduction

Last time we saw that if we treat each mode of the photon as a separate particle, and give it multiple excitations like it's a harmonic oscillator, we can derive the relation between the coefficients of induced and spontaneous emission without resorting to statistical mechanics. This was our first QED calculation. Today we will make this a little more precise, give it a little more justification, and pave the way for more detailed computations in Quantum Electrodynamics.

I should mention here that there are two ways of quantizing a field theory. The first is the called canonical quantization. It is historically how quantum field theory was understood, and closely follows what you learned in quantum mechanics. It's also easy to see the relation between the formalism and physical problems. The second way is called the Feynman Path Integral. It is more beautiful in a way, and more general, but it is hard to understand physically what you are calculating. In any case, both are necessary for a proper education in quantum field theory, so we will start with the traditional canonical quantization.

From now on I will set $\hbar = c = 1$. I will also be cavalier about signs and small numerical factors. This is mostly because I am too lazy to make sure they are right. But it is also because I want to keep talking about electromagnetic fields, which are physical, and to keep treating them like scalars, which are easy to study. It turns out that if you worry about the signs and treat the photon like a scalar, things just don't work (after all, that's how we know the photon is not a scalar!). Eventually, we will to QED properly and then I will be careful about signs, but not yet.

3.2 Simple Harmonic Oscillator

Recall the simple harmonic oscillator. Anything with a linear restoring potential (the simplest thing possible), like a spring, or a string with tension, or a wave, is a harmonic oscillator. For example, a spring has

$$m\frac{d^2x}{dt^2} - kx = 0 \quad \Rightarrow \quad x(t) = \cos\sqrt{\frac{k}{m}}\,t \tag{3.1}$$

I.e. it oscillates with frequency

$$\omega = \sqrt{\frac{k}{m}} \tag{3.2}$$

$$x(t) = c_1 e^{i\omega t} + c_2 e^{-i\omega t} \tag{3.3}$$

The classical Hamiltonian for this system is

$$H = \frac{1}{2}\frac{p^2}{m} - \frac{1}{2}m\omega^2 x^2 \tag{3.4}$$

To quantize it, we say

$$[x, p] = i\,h \tag{3.5}$$

A very very important trick is to write this as

$$a = \sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{ip}{m\omega}\right), \quad a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\left(x - \frac{ip}{m\omega}\right) \tag{3.6}$$

with

$$[a, a^\dagger] = 1 \tag{3.7}$$

Then

$$H = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right) \tag{3.8}$$

Thus energy eigenstates are eigenstates of the number operator

$$N = a^\dagger a \tag{3.9}$$

which is Hermetian. Also, all the stuff we saw yesterday is true.

$$N|n\rangle = n|n\rangle \tag{3.10}$$

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \tag{3.11}$$

$$a|n\rangle = \sqrt{n}|n-1\rangle \tag{3.12}$$

One more useful fact is how these operators evolve in time. Recall that in the Heisenberg picture, states don't evolve, operators do. Then

$$i\hbar\frac{d}{dt}a = [a, H] = \left[a, \hbar\omega\left(a^\dagger a + \frac{1}{2}\right)\right] = \hbar\omega(a a^\dagger a - a^\dagger a a) = -\hbar\omega[a^\dagger, a]a = \hbar\omega a \tag{3.13}$$

$$a(t) = e^{-i\omega t}a(0) \tag{3.14}$$

3.3 Plane waves as oscillators

The excitations of a plane electromagnetic wave in a cavity are very similar to the excitations of a harmonic oscillator. Recall that there's a nice Lorentz-covariant treatment of electromagnetism using

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & B_z & 0 & B_y \\ -E_z & -B_y & B_x & 0 \end{pmatrix} \tag{3.15}$$

F is antisymmetric. This makes Maxwell's equations have the simple form (in empty space)

$$\partial_\mu F_{\mu\nu} = 0 \tag{3.16}$$

The nice thing about using $F_{\mu\nu}$ is that it concisely encodes how E and B rotate into each other under boosts.

We can always write

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{3.17}$$

with

$$\partial_\mu A_\mu = 0 \tag{3.18}$$

Requiring $\partial_\mu A_\mu = 0$ is not necessary. It is a gauge choice (Lorentz gauge). We will discuss gauge invariance in great detail in a few weeks. For now, it is enough to know that the physical E and B fields can be combined in this way.

Then,

$$\partial_\mu F_{\mu\nu} = \Box A_\nu - \partial_\nu(\partial_\mu A_\mu) = \Box A_\nu = 0 \tag{3.19}$$

From here on, I will drop the subscript on A_ν, writing it as A. You can keep the ν if you want, but none of the qualitative conclusions that follow today will depend on it, and it makes things less cluttered.

So,

$$(\partial_t^2 - \partial_x^2)A = 0 \tag{3.20}$$

The solutions are plane waves. For example, one solution is

$$A(x) = a_k(t)e^{i\vec{k}\vec{x}} \tag{3.21}$$

where

$$(\partial_t^2 + \omega_k^2)a_k(t) = 0, \quad \omega_k = |\vec{k}\,| \tag{3.22}$$

This is exactly the equation of motion of a harmonic oscillator. Thus electromagnetic fields are oscillators.

We can decompose a general solution to Maxwell's equations as

$$A(x,t) = \int \frac{d^3k}{(2\pi)^3}\left[a_k(t)e^{ikx} + a_k^\star(t)e^{-ikx}\right] \tag{3.23}$$

3.4 Second Quantization

Since the modes of an electromagnetic field have the same classical equations as a simple harmonic oscillator, we can quantize them the same way

$$H = \int \frac{d^3k}{(2\pi)^3}\,\omega_k(a_k^\dagger a_k + \frac{1}{2}) \tag{3.24}$$

This is known as second quantization. First quantizing is saying that all the modes have energy proportional to their frequency. Second quantizing is saying that there are integer numbers of excitations of each of these modes. But this is somewhat misleading – the fact that there are discrete modes is a classical phenomenon. The two steps really are 1) interpret these modes as having momenta $k = h\nu$ and 2) quantize each mode as a harmonic oscillator. In that sense we are only quantizing once. But these are only words.

There are two new features in second quantization

1. We have many quantum mechanical systems – one for each k – all at the same time.
2. We interpret the n^{th} excitation of the k harmonic oscillator as having n photons.

Let's take a moment to appreciate this second point. Recall the old simple harmonic oscillator – the electron in a quadratic potential. We would never interpret the $|n\rangle$ states of this system as having n electrons, because that's completely wrong. (Actually, there is an interpretation of these n states as n particles called *phonons*. Phonons are a special topic which may be covered in semester 2.) The fact that a point-like electron in a quadratic potential has analogous equations of motion to a Fourier component of the electromagnetic field is just a coincidence. Don't let it confuse you. Both are just the simplest possible dynamical systems, with linear restoring forces.

To set up a proper analogy we need to first treat the electron as a classical field (we don't know how to do that yet), and find a set of solutions (like the discrete frequencies of the EM waves). Then we would quantize each of those solutions, allowing $|n\rangle$ excitations. However, if we did this, electrons would have Bose-Einstein statistics. Instead, they must have Fermi-Dirac statistics, so restrict $n = 0, 1$. There's a lot to learn before we can do electrons, so let's stick with photons for now.

3.4.1 more details

Now let's get a little more precise about what this Hamiltonian means. The natural generalization of

$$[a, a^\dagger] = 1 \tag{3.25}$$

is

$$[a_k, a_p^\dagger] = (2\pi)^3 \delta^3(p-k) \tag{3.26}$$

These $a_k^\dagger$ operators create particles with momentum k

$$a_k^\dagger |0\rangle = \frac{1}{\sqrt{2\omega_k}} |k\rangle \tag{3.27}$$

This factor of $\sqrt{2\omega_k}$ is just a convention, but it will make some calculations easier. You will show for homework that it has nice Lorentz transformation properties.

Also, we know how the momentum states $|k\rangle$ evolve from the way $a_k^\dagger$ evolves:

$$|k,t\rangle = a_k^\dagger(t)|0\rangle = e^{i\omega_k t} a_k^\dagger(0)|0\rangle = e^{i\omega_k t}|k\rangle \tag{3.28}$$

For normalization, we start with

$$\langle 0|0\rangle = 1 \tag{3.29}$$

Then

$$\langle p|k\rangle = 2\sqrt{\omega_p \omega_k}\langle 0|a_p a_k^\dagger|0\rangle = 2\omega_p (2\pi)^3 \delta^3(p-k) \tag{3.30}$$

And a complete set is

$$1 = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_p}|p\rangle\langle p| \tag{3.31}$$

So that

$$|k\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_p}|p\rangle\langle p|k\rangle = |p\rangle \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_p} 2\omega_p (2\pi)^3\delta^3(p-k) = |k\rangle \tag{3.32}$$

Finally, define

$$A(x) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2\omega_k}}\Big[a_k e^{ikx} + a_k^\dagger e^{-ikx}\Big] \tag{3.33}$$

This is the same as the free solutions to Maxwell's equations, equation (3.23) but instead of a_k being a function, it is now the creation/annihilation operator for that mode. The factor of $\sqrt{2\omega_k}$ has been added for later convenience, and in fact guarantees the $A(x)$ is Lorentz invariant. The connection with Eq. (3.23) should be considered only suggestive, and this taken as the definition of an operator $A(x)$ constructed from the creation and annihilation operators a_k and $a_k^\dagger$.

Note that:

$$\langle p|A(x)|0\rangle = \langle 0|\sqrt{2\omega_p}a_p \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2\omega_k}}\Big[a_k e^{ikx} + a_k^\dagger e^{-ikx}\Big]|0\rangle \tag{3.34}$$

$$= \int \frac{d^3k}{(2\pi)^3}\sqrt{\frac{\omega_p}{\omega_k}}\Big[e^{ikx}\langle 0|a_p a_k|0\rangle + e^{-ikx}\langle 0|a_p a_k^\dagger|0\rangle\Big] \tag{3.35}$$

$$= e^{-ikx} \tag{3.36}$$

So

$$A(x)|0\rangle = |x\rangle \tag{3.37}$$

$$\langle p|x\rangle = e^{-ikx} \tag{3.38}$$

That is $A(x)$ creates a photon at position x.

3.5 Normalization and Lorentz Invariance

The important formulas are

$$[a_k, a_p^\dagger] = (2\pi)^3 \delta^3(p-k) \tag{3.39}$$

$$a_k^\dagger |0\rangle = \frac{1}{\sqrt{2\omega_k}} |k\rangle \tag{3.40}$$

$$A(x) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_k}} \Big[a_k e^{ikx} + a_k^\dagger e^{-ikx} \Big] \tag{3.41}$$

Where

$$\omega_k \equiv \sqrt{\vec{k}^2 + m^2} \tag{3.42}$$

All these $2\omega_k$ factors are a conventional normalization, designed to make things easily Lorentz invariant. They come from

$$\int dk^0 \delta(k^2 - m^2)\theta(k_0) = \frac{1}{2\omega_k} \tag{3.43}$$

which implies

$$\int \frac{d^3k}{2\omega_k} \tag{3.44}$$

is Lorentz invariant. (You will check these relations in Problem Set 1). It also implies

$$2\omega_k \delta^3(p-k) \tag{3.45}$$

is Lorentz invariant.

Then,

$$\langle p|k\rangle = 2\sqrt{\omega_k \omega_p} \langle 0|a_p a_k^\dagger|0\rangle = 2\omega_k (2\pi)^3 \delta^3(p-k) \tag{3.46}$$

is Lorentz invariant, which is the whole point. Also,

$$A(x)|0\rangle = \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_p} e^{-i\vec{p}\cdot\vec{x}} |p\rangle \tag{3.47}$$

So we this same Lorentz invariant combination $\frac{d^3k}{2\omega_p}$ appears again.

3.6 Implications of Second Quantization

All we have done to quantize the electromagnetic field is to treat it as an infinite set of simple harmonic oscillators, one for each wavenumber k. So Quantum Field Theory is just quantum mechanics with an infinite number of fields.

Even if we put the system in a box of size L, so that $k = \frac{2\pi}{L} n$, we can still have $n \to \infty$. This will become a tremendous headache for us, since these $\infty's$ will creep into practically every calculation we do. A big chunk of this course will be devote to understanding them, and at some level they are still not understood.

In quantum mechanics we were used to treating a single electron in a background potential $V(x)$. Now that background, the electromagnetic system, is dynamical, so all kinds of funny things can happen. We already saw one, Dirac in his calculation of the Einstein coefficients. We can be a little bit more explicit about how we got this now.

We write the Hamiltonian as

$$H = H_0 + H_{\text{int}} \tag{3.48}$$

Where H_0 is some state that we can solve exactly. In this case it is the atom + the free photons

$$H_0 = H_{\text{atom}} + H_{\text{photon}} \tag{3.49}$$

The $H_{\rm int}$ is some modification that is small enough to let us use perturbation theory. Then all we need is Fermi's golden rule, the same one from quantum mechanics (which you might want to review). It says the rate for transitions between two states is proportional to the square of the matrix element of the interaction between the two states.

$$\Gamma = |\langle f|H_{\rm int}|i\rangle|^2 \delta\left(E_f - E_i\right) \tag{3.50}$$

In our case, $H_{\rm atom}$ gives the states $|\psi_n\rangle$ of the hydrogen atom, with energies E_n. $H_{\rm photon}$ is the Hamiltonian above,

$$H_{\rm photon} = \int \frac{d^3k}{(2\pi)^3}\omega_k(a_k^\dagger a_k + \frac{1}{2}) \tag{3.51}$$

and we can treat the interaction semi-classically:

$$H_{\rm int} = e\,A\,J_{\rm atom} \tag{3.52}$$

Here $J_{\rm atom}$ is shorthand for the atom treated as a source for photons. Essentially, we are treating the atom as a fixed background – that is, we won't need to consider how the atom gets deformed. The important point is that this interaction has an A field in it.

Then, according to Fermi's golden rule, the transition probability is proportional to the matrix element of the interaction squared

$$M_{1\to 2} = \langle {\rm atom}'; n_k - 1|H_{\rm int}|{\rm atom}; n_k\rangle = \langle {\rm atom}'|J_{\rm atom}|{\rm atom}\rangle\sqrt{n_k} \tag{3.53}$$

$$M_{2\to 1} = \langle {\rm atom}; n_k + 1|H_{\rm int}|{\rm atom}'; n_k\rangle = \langle {\rm atom}|J_{\rm atom}|{\rm atom}'\rangle\sqrt{n_k+1} \tag{3.54}$$

where we have used

$$\langle n_k - 1|A|n_k\rangle = \int \frac{d^3p}{(2\pi)^3}\langle n_k - 1|\,a_p|n_k\rangle = \sqrt{n_k} \tag{3.55}$$

$$\langle n_k + 1|A|n_k\rangle = \int \frac{d^3p}{(2\pi)^3}\left\langle n_k + 1|\,a_p^\dagger|n_k\right\rangle = \sqrt{n_k+1} \tag{3.56}$$

So,

$$dn_1 \propto |M_{1\to 2}|^2 = \langle {\rm atom}'|J_{\rm atom}|{\rm atom}\rangle^2 n_k \tag{3.57}$$

$$dn_2 = |M_{2\to 1}|^2 = \langle {\rm atom}'|J_{\rm atom}|{\rm atom}\rangle^2 (n_k+1) \tag{3.58}$$

This was derived by Dirac in his calculation of the Einstein coefficients.

Note that we only used one photon mode, of momentum k, so this was really just quantum mechanics. All QFT did was give us a delta-function from the d^3p integration.

Chapter 4
Classical Field Theory

4.1 Introduction

We have now seen how Quantum Field Theory is defined as quantum mechanics with an infinite number of oscillators. QFT can do some remarkable things, such as explain spontaneous emission without statistical mechanics. But it also seems to lead to absurdities, such as the infinite shift in the energy levels of the Hydrogen atom. To show that QFT is not absurd, but extremely predictive, we will have be very careful about how we do calculations. We'll begin by going through carefully some of the predictions which QFT gets right without infinities. These are the tree-level processes, which means they are leading order in perturbation theory. We will start to study them today.

4.2 Classical Field Theory

Before we get in to more QFT calculations, let's review some elements of classical field theory. The Hamiltonian is the sum of the kinetic and potential energies of a system

$$\mathcal{H} = K + V \tag{4.1}$$

So the Hamiltonian gives the total energy. The Hamiltonian has all the information about the time evolution of a system. This is true both classically, where time-translation is determined by a Poisson bracket, and quantum mechanically, where time-translation is determined by a commutator.

The problem with Hamiltonians is that they are not Lorentz invariant. The Hamiltonian picks out energy, which is not a Lorentz scalar, but the 0 component of a Lorentz vector: $V_\mu = (\mathcal{H}, \vec{P})$. It is great for non-relativistic systems, but for relativistic systems we will almost exclusively use Largrangians.

The Lagrangian (for fields, the Lagrangian density) is the difference between kinetic energy and potential energy

$$\mathcal{L} = K - V \tag{4.2}$$

So while the Hamiltonian corresponds to a conserved quantity, energy, the Lagrangian does not. Dynamics from a Lagrangian system is determined by trying to minimize the action, which is the integral over space and time of the Lagrangian density

$$\mathcal{S} = \int d^4x \, \mathcal{L}(x) \tag{4.3}$$

We don't usually talk about kinetic and potential energy in quantum field theory. Instead we talk about kinetic *terms* and *interactions*. Kinetic terms are *bilinear* in the fields. Bilinear means there are two fields. So kinetic terms look like

$$\mathcal{L}_K \supset \frac{1}{2}\phi\Box\phi, \quad \bar{\psi}\not{\partial}\psi, \quad \frac{1}{4}F_{\mu\nu}^2, \quad \frac{1}{2}m^2\phi^2, \quad \frac{1}{2}\phi_1\Box\phi_2, \quad \phi_1\partial_\mu A_\mu, \quad \cdots \tag{4.4}$$

where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu \mathrm{A}_\mu \tag{4.5}$$

So anything with two fields of the same or different type is a kinetic term. The kinetic terms tell you about the free fields. Fields with kinetic terms are said to be *dynamical* or *propagating.* If we want to be really particular, a field should actually have time-derivatives in its kinetic term to be dynamical (for example A_0 above does not), but we will be loose about this terminology.

I should also point out at this point that it is standard to use the letters ϕ or π for scalar fields, ψ, ξ, χ for fermions, A_μ, J_μ, V_μ for vectors and $h_{\mu\nu}$ for tensors.

Interactions have 3 or more fields

$$\mathcal{L}_I \supset \lambda\phi^3, \quad g\bar{\psi}\not{A}\psi, \quad g\partial_\mu\phi A_\mu\phi^\star, \quad g^2A_\mu^2A_\nu^2, \quad \frac{1}{M_P}\partial_\mu h_{\mu\nu}\partial_\nu h_{\alpha\beta}h_{\alpha\beta}, \quad \cdots \tag{4.6}$$

Since the interactions are the potential energy terms from the Hamiltonian, we also sometimes write H_{int} or V for $\mathcal{L}_I$. It's all just notation. Generally, we assume that the coefficients of the interaction terms are small in some sense, so that the fields are *weakly interacting* and we can do perturbation theory. If the interactions are not small, sometimes we can still solve the system, but it is usually very difficult. All we will talk about this semester is situations in which the interactions are weak.

4.3 Equations of Motion

To calculate the equations of motion from a Lagrangian, we can use the Euler Lagrange equations

$$\frac{\delta\mathcal{L}}{\delta\phi} - \partial_\alpha\frac{\delta\mathcal{L}}{\delta(\partial_\alpha\phi)} = 0 \tag{4.7}$$

For example, if our action is

$$\mathcal{S} = \int d^4x\Big[-\frac{1}{2}(\partial_\mu\phi)(\partial_\mu\phi) + V(\phi)\Big] \tag{4.8}$$

Then the equations are

$$V'(\phi) + \partial_\alpha[\partial_\alpha\phi] = 0 \tag{4.9}$$

$$\Box\phi + V'(\phi) = 0 \tag{4.10}$$

Where do the Euler-Lagrange equations come from? Basically, just from demanding that $\frac{\delta S}{\delta\phi} = 0$ when you assume $\mathcal{L} = \mathcal{L}(\phi,\ d\phi)$ (see Peskin and Schroeder section 2.2). This is the form that all "classical" Lagrangians had. Nowadays, we allow for much more general Lagrangians, for example, we can have terms like $d^3\phi$ in them.

It's actually pretty easy to derive the equations of motion for a general Lagrangian with any number of fields and derivatives. The way I like to do it is to integrate by parts, so at least there's one ϕ without derivatives. In this case,

$$\mathcal{S} = \int d^4x\Big[\frac{1}{2}\phi\Box\phi + V(\phi)\Big] \tag{4.11}$$

Then we vary with respect to ϕ.

$$\Box\phi + V'(\phi) = 0 \tag{4.12}$$

We get a factor of 2 because ϕ appears quadratically and symmetrically. That is, if we integrated by parts so that the other ϕ had no derivatives, we would have gotten the same term, and the 2 is the sum of both.

For example, a term like

$$L = (\Box\partial_\mu\phi)A^3\phi^3\partial_\mu\phi \tag{4.13}$$

would give

$$L = -\phi\Box\partial_\mu[A^3\phi^3\partial_\mu\phi] = \phi^3A^3\partial_\mu\phi(\Box\partial_\mu\phi) = -\phi\partial_\mu\big[(\Box\partial_\mu\phi)A^3\phi^3\big] \tag{4.14}$$

where I have isolated the 3 places ϕ appears. So then the equations of motion are

$$-\Box\partial_\mu[A^3\phi^3\partial_\mu\phi] + 3\phi^2A^3\partial_\mu\phi(\Box\partial_\mu\phi) - \partial_\mu\big[(\Box\partial_\mu\phi)A^3\phi^3\big] = 0 \tag{4.15}$$

Usually these things simplify, but this is the safest way to calculate equations of motion in general.

4.4 Currents

An extremely useful trick for doing calculations in field theory is to use *currents*. Currents are used in many ways

1. Currents can refer to something which is really a current, like the electrons moving in a wire. These are *external* currents. For example, the current for a single charge at the origin is

$$J_\mu(x): \quad \begin{cases} J_0(x) = \rho(x) \\ J_i(x) = v_i(x) \end{cases} \tag{4.16}$$

2. The second way currents are used is a *source* for fields. This is a term in the Lagrangian like

$$\mathcal{L} \supset A_\mu J_\mu \tag{4.17}$$

This current can either be an explicit external current, such as the charge current above, or just a formal place holder. The current is never a dynamical field, that is, it never has it's own kinetic terms. We may include dynamics of the source, for example, time dependence of $J_\mu(x,\, t)$, but we won't generally try to solve for the dynamics of J_μ at the same time as solving for the dynamics of real propagating fields like A_μ.

3. Currents can be place-holders for certain terms in a Lagrangian. For example, if our Lagrangian were

$$\mathcal{L} = \frac{1}{4}F_{\mu\nu}^2 + \phi^\star \Box \phi + A_\mu \phi^\star \partial_\mu \phi \tag{4.18}$$

We could write this as

$$\mathcal{L} = \frac{1}{4}F_{\mu\nu}^2 + \phi^\star \Box \phi + A_\mu J_\mu \tag{4.19}$$

with $J_\mu = \phi^\star \partial_\mu \phi$. The nice thing about this is that if we had a different kind of field interacting with A_μ, for example, $\bar{\psi}\gamma^\mu A_\mu \psi$, we could write this as $A_\mu J_\mu$ as well but now with $J_\mu = \bar{\psi}\gamma^\mu\psi$. Sometimes we don't care about whether we have a scalar ϕ or a fermion ψ, but only about the field A_μ, then this notation is particularly useful.

Using currents helps separate the problem into how the field ϕ or ψ produces the field A_μ and then how A_μ affects other fields. Sometimes we don't care about where J_μ comes from, like if it's a background magnetic field in the lab. Sometimes we care very much, like if we want to see how an atom spontaneously emits a photon.

4.5 Coulombs law – Classical

Let's recall how Coulomb's law comes out of classical field theory. Let's say we have a charge at the origin. This can be represented with an external current

$$J_\mu(x): \quad \begin{cases} J_0(x) = \rho(x) = -e\delta(x) \\ J_i(x) = 0 \end{cases} \tag{4.20}$$

The Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + H_{\text{int}} \tag{4.21}$$

The interaction of the current with the electromagnetic field is

$$H_{\rm int} = -A_\mu J^\mu = -A_\mu(x) J_\mu(x) \tag{4.22}$$

Let's calculate the equations of motion

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 - A_\mu J_\mu \tag{4.23}$$

$$= -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)F_{\mu\nu} - A_\mu J_\mu \tag{4.24}$$

$$= \frac{1}{4}A_\nu \partial_\mu F_{\mu\nu} - \frac{1}{4}A_\mu \partial_\nu F_{\mu\nu} - A_\mu J_\mu \tag{4.25}$$

$$= \frac{1}{2}A_\nu \partial_\mu F_{\mu\nu} - A_\nu J_\nu \tag{4.26}$$

Thus

$$\partial_\mu F_{\mu\nu} = J_\nu \tag{4.27}$$

Which are Maxwell's equations in the presence of a source.

Expanding

$$\partial_\mu(\partial_\mu A_\nu - \partial_\nu A_\mu) = J_\nu \tag{4.28}$$

$$\Box A_\nu - \partial_\nu(\partial_\mu A_\mu) = J_\nu \tag{4.29}$$

In Lorentz gauge, $\partial_\mu A_\mu = 0$, since $J_i = 0$, this is just

$$\Box A_0(x) = J_0(x) \tag{4.30}$$

So

$$A_0(x) = \frac{1}{\Box}J_0(x) = -\frac{e}{\Box}\delta(x) \tag{4.31}$$

This type of expression comes about in almost every calculation in quantum field theory. It says that the A field is determined by the source J after it propagates with the propagator

$$\Pi_A = \frac{1}{\Box} \tag{4.32}$$

We will understand these propagators in great detail, both today and in the rest of the course.

4.5.1 Fourier transform interlude

Continuing with the Coulomb calculation, we can take the the Fourier transform. Using

$$\delta(x) = \int \frac{d^3k}{(2\pi)^3}e^{ikx} \quad \Rightarrow \quad \tilde{\delta}(k) = 1 \tag{4.33}$$

$$\Box^n\delta(x) = \int \frac{d^3k}{(2\pi)^3}\Box^n e^{ikx} = \int \frac{d^3k}{(2\pi)^3}(-k^2)^n e^{ikx} \tag{4.34}$$

$$\Rightarrow \widetilde{\Box^n\delta}(k) = (-k^2)^n \tag{4.35}$$

More generally

$$\Box^n f(x) = \int \frac{d^3k}{(2\pi)^3}\Box^n \tilde{f}(k)e^{ikx} = \int \frac{d^3k}{(2\pi)^3}(-k^2)^n \tilde{f}(k)e^{ikx} \tag{4.36}$$

So

$$[\widetilde{\Box^n f}](k) = (-k^2)^n \tilde{f}(k) \tag{4.37}$$

Thus in general

$$\Box \leftrightarrow -k^2 \tag{4.38}$$

We will use this implicitly all the time. When a field theorist thinks "box" he thinks "$-k^2$".

4.5.2 Coulomb potential again

So,

$$A_0(x) = -\frac{e}{\Box}\delta(x) \tag{4.39}$$

Implies

$$A_0(x) = \int \frac{d^3k}{(2\pi)^3}\frac{e}{k^2}e^{ikx} \tag{4.40}$$

$$= \frac{e}{(2\pi)^3}\int_0^\infty k^2 dk \int_{-1}^1 d\cos\theta \int_0^{2\pi} d\phi \frac{1}{k^2}e^{ikr\cos\theta} \tag{4.41}$$

$$= \frac{e}{(2\pi)^2}\int_0^\infty dk \frac{e^{ikr} - e^{-ikr}}{ikr} \tag{4.42}$$

$$= -\frac{e}{8\pi^2}\frac{1}{ir}\int_{-\infty}^\infty dk \frac{e^{ikr} - e^{-ikr}}{k} \tag{4.43}$$

Note that the integrand does not blow up as $k \to 0$. So we can regulate with

$$\int_{-\infty}^\infty dk \frac{e^{ikr} - e^{-ikr}}{k} = \lim_{\varepsilon\to 0}\left[\int_{-\infty}^\infty dk \frac{e^{ikr} - e^{-ikr}}{k + i\varepsilon}\right] \tag{4.44}$$

Now assume $\varepsilon > 0$, then we pick up the pole at $k = -i\varepsilon$. For e^{-ikr} we close the contour up and miss the pole, so this term gives zero. For e^{ikr} we close the contour down and get

$$\int_{-\infty}^\infty dk \frac{e^{ikr}}{k + i\varepsilon} = (2\pi i)e^{\varepsilon r} \tag{4.45}$$

Thus

$$A_0(x) = -\frac{e}{4\pi}\frac{1}{r} \tag{4.46}$$

If you are not familiar with contour integration, please see Hitoshi Murayama's notes on Contour Integrals, which I have made available. We can also reproduce this result by taking the limit $m \to 0$ of a massive vector boson.

4.6 Green's functions

The important point is that we solved this by using

$$A_\mu = \frac{1}{\Box}J_\mu \tag{4.47}$$

Even if J_μ were much more complicated, producing all kinds of crazy looking electromagnetic fields, we could still use this equation.

For example, suppose we wanted to model the thing that generates the Coulombs law. For example, we might have a Lagrangian like

$$\mathcal{L} = \phi\Box\phi + F_{\mu\nu}^2 + \phi A_\mu \partial_\mu \phi^\star \tag{4.48}$$

where ψ represents a charged object which radiates the A field. Now A's equation of motion is (in Lorentz gauge)

$$\Box A_\mu = \phi\partial_\mu\phi^\star \tag{4.49}$$

So this is just what we had before but with $J_\mu = \phi\partial_\mu\phi^\star$. And again we will have $A_\mu = \frac{1}{\Box}J_\mu$

This is actually a really useful way to solve equations, and quite general. For example, let's suppose our electromagnetic field had a self-interaction. The electromagnetic field doesn't – it's equations of motion are linear. But there are plenty of self-interacting fields out there. The gluon is one. Another is the graviton. The Lagrangian for the graviton looks heuristically like

$$\mathcal{L} = \frac{1}{2}h\Box h + \frac{1}{3}\lambda h^3 + Jh \tag{4.50}$$

h represents the gravitational potential, like A_0 is the coulomb potential. (Don't get scared – we're just doing simple scalar field theory, I'm just calling it gravity). The h^3 term represents a graviton self interaction, which is present in general relativity. And we can think of $\lambda \sim 1/M_{\rm pl}$. The equations of motion are

$$\Box h + \lambda h^2 + J = 0 \tag{4.51}$$

Now we solve perturbatively in λ. For $\lambda = 0$,

$$h_0 = -\frac{1}{\Box}J \tag{4.52}$$

This is what we had before. Then we plug in

$$h = h_0 + h_1 \tag{4.53}$$

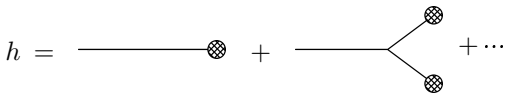
$$-\Box\frac{1}{\Box}J + \Box h_1 + \lambda h_0^2 + \cdots + J = 0 \tag{4.54}$$

$$\Rightarrow h_1 = -\lambda h_0\frac{1}{\Box}h_0 \tag{4.55}$$

Thus

$$h = -\frac{1}{\Box}J - \lambda\left(\frac{1}{\Box}J\right)\frac{1}{\Box}\left(\frac{1}{\Box}J\right) \tag{4.56}$$

We can keep this up, resulting in a nice expansion for h. In fact, it's not too hard to compute the perihelion shift of mercury using this kind of first order solution. There's a nice pictorial representation of this solution.



The line is a $\frac{1}{\Box}$ and the hatched circles represent the external currents. The vertex in the second graph is the interaction, proportional to λ. You can use this picture to guess the next term.

This is known as the Green's function method. The object

$$\Pi = \frac{1}{\Box} = -\frac{1}{k^2} \tag{4.57}$$

is also known as a propagator. Propagators are integral parts of quantum field theory. Here we see that classically, they tell us how the field propagates through space when it is sourced by a current $J_\mu(x)$. But note that the propagator has nothing to do with the source.

Chapter 5
Old Fashioned Perturbation Theory

5.1 Perturbation Theory in Quantum Field Theory

The simplest way to study perturbation theory is with Feynman path integrals. But while the path integral is easier to handle in many ways, it obscures the connection to quantum mechanics and perturbation theory. So today we will just plow ahead with the canonical formalism, using creation and annihilation operators. Since we have seen that Quantum Field Theory is just quantum mechanics with lots of fields, the rules for perturbation theory should not have changed. So we can begin by review them, and applying them to our second quantized photon.

5.2 Some early infinities

The first confusion about the second-quantized photon field was that the Hamiltonian

$$H_{\rm photon} = \int \frac{d^3k}{(2\pi)^3}\omega_k(a_k^\dagger a_k + \frac{1}{2}) \tag{5.1}$$

Seemed to say that the vacuum has infinite energy

$$E_0 = \langle 0|H|0\rangle = \int \frac{d^3k}{(2\pi)^3}\,\frac{1}{2} = \infty \tag{5.2}$$

This was initially quite confusing. However, there is an easy way out – how do you measure the energy of the vacuum? You don't! Only energy differences are measurable, and in these the zero-point energy, the energy of the ground state, drops out. Although we are getting way ahead of ourselves here, this is the basic idea behind renormalization – infinities can appear in intermediate calculations, but they must drop out of physical observables. (Actually, there is one way the absolute energy of the vacuum is observable: through gravity. Then the infinity is a really big problem, known as the cosmological constant problem).

In 1930, Oppenheimer thought to use perturbation theory to compute the shift of the energy of the Hydrogen atom due to the photons. (R. Oppenheimer, *Phys Rev*. **35** 461, 1930). He got infinity and concluded that quantum electrodynamics was wrong. In fact the result is not, infinite, but a finite calculable quantity known as the Lamb shift which agrees perfectly with data. However, it is instructive to understand Oppenheimer's argument.

Let's recall how perturbation theory works in quantum mechanics. We break the Hamiltonian into a part whose energy eigenstates we know, and a small perturbation:

$$H = H_0 + H_{\rm int} \tag{5.3}$$

With

$$H_0|\,\psi_n^0\rangle = E_n|\,\psi_n^0\rangle \tag{5.4}$$

Then,

$$\Delta|\psi_n\rangle = \sum_{m\neq n} \frac{\langle\psi_m|H_{\rm int}|\psi_n\rangle}{E_n - E_m}|\psi_m\rangle + \cdots \tag{5.5}$$

$$\Delta E_n = \langle\psi_n|H_{\rm int}|\psi_n\rangle + \sum_{m\neq n} \frac{|\langle\psi_n|H_{\rm int}|\psi_m\rangle|^2}{E_m - E_n} \tag{5.6}$$

This should be familiar from quantum mechanics.

Now, we can use this formula to compute the shift in energy levels of the hydrogen atom.

5.2.1 Stark Effect

First, take a fixed non-dynamical background field. Say an electric field in the z direction (this is the Stark effect). Then the potential energy is proportional to the electric field

$$H_{\rm int} = \vec{E}\cdot\vec{x} = e|E|z \tag{5.7}$$

Now an atom has no electric dipole moment, so the first order correction is zero

$$\langle\psi_n|H_{\rm int}|\psi_n\rangle = 0 \tag{5.8}$$

But to second order

$$\Delta E_0 = \sum_{m>0} \frac{|\langle\psi_0|H_{\rm int}|\psi_m\rangle|^2}{E_m - E_0} \leqslant \frac{1}{E_0}\sum_{m>0} \langle\psi_0|H_{\rm int}|\psi_m\rangle\langle\psi_m|H_{\rm int}|\psi_0\rangle = \frac{1}{E_0}\big\langle\psi_0|H_{\rm int}^2|\psi_0\big\rangle \tag{5.9}$$

where we have used the completeness relation

$$\sum_m |\psi_m\rangle\langle\psi_m| = 0 \tag{5.10}$$

We can (in principle) evaluate this matrix element, and it's some finite number, proportional to

$$\big\langle\psi_0|H_{\rm int}^2|\psi_0\big\rangle = e^2|E|^2\big\langle\psi_0|z^2|\psi_0\big\rangle \sim e^2|E|^2a_0^2 \tag{5.11}$$

So

$$\Delta E_0 \lesssim \frac{e^2|E|^2a_0^2}{E_0} \tag{5.12}$$

The only point I wanted to make is that this is not infinite.

5.2.2 Lamb shift a la Oppenheimer 1930

Now, instead, forget about the external electric field, but use our second quantized photons. Then

$$H_0 = H_0^{\rm atom} + H_0^{\rm photon} \tag{5.13}$$

The energy states of H_0 are

$$H_0|\psi_n;\{n_k\}\big\rangle = (E_n + n_k\omega_k)|\psi_n;\{n_k\}\big\rangle \tag{5.14}$$

Where we allow for any number of excitations of the photons of any momenta k.

The interaction is

$$H_{\rm int} = J(x)\,A(x) = J\int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2\omega_k}}\Big(a_k^\dagger e^{ikx} + a_k e^{-ikx}\Big) \tag{5.15}$$

$J(x)$ is a "current" (we will discuss currents next time) and contains the physics of how the photon interacts with the atom. It doesn't matter to the point we are trying to make, so don't worry about what exactly $J(x)$ is.

So now

$$\Delta E_0 = \int d^3k \sum_{m,n_k} \frac{|\langle\psi_0|H_{\rm int}|\psi_m;\{n_k\}\rangle|^2}{E_m + n_k\omega_k - E_0} \tag{5.16}$$

We can simplify the matrix elements as

$$|\langle\psi_0|H_{\rm int}|\psi_m;\{n_k\}\rangle|^2 = |\langle\psi_0|J(x)|\psi_m\rangle\langle 0|A(x)|n_k\rangle|^2 \tag{5.17}$$

Now $\langle n_k|A^\dagger|0\rangle = 0$ unless $n_k = 1$. So,

$$\langle n_k|A(x)|0\rangle = \langle k|A(x)|0\rangle = e^{ikx} \tag{5.18}$$

And so

$$\Delta E_0 = \sum_m \int d^3k \frac{|\langle\psi_0|J(x)e^{ikx}|\psi_m\rangle|^2}{E_m + \omega_k - E_0} \tag{5.19}$$

Now, to evaluate the numerator, let's just say $J = 1$ for simplicity (this would be the case in a Yukawa theory, for example). We'll also take m to be really excited, so it's effectively just a free electron with momentum m

$$\psi_m(x) \sim \int d^3m e^{im x} \tag{5.20}$$

Then

$$|\langle\psi_0|J(x)e^{ikx}|\psi_m\rangle|^2 = \int d^3x\, \psi_0(x)J(x)e^{i(k+m)x} \tag{5.21}$$

Then the sum over m goes into an integral, which forces $m = k$ and $E_m = \omega_k = |k|$, Then

$$\Delta E_0 = \int d^3m \int d^3k \int d^3x \frac{\psi_0(x)J(x)e^{i(k+m)x}}{|m| + |k| - E_0} \tag{5.22}$$

The phase factor forces, roughly, $k + m \sim 0$. This means that we started with small 4-momentum in the atom (~ 0) and the momenta gets split into a large k and a large m going in the opposite direction, but overall momentum is still conserved. We're not going to do the integral exactly, I just want to observe that it is divergent. For $p \sim k + m$ it goes roughly like

$$\Delta E_0 = \int \frac{d^3p}{p - E_0} \sim \int p\,dp = \infty \tag{5.23}$$

This means that there should be a quadratically divergent shift in the energy levels of the Hydrogen atom! Oppenheimer also showed that if you take the difference between two levels, the result is also divergent at large k, but only logarithmically so. He conclude that quantum field theory was a failure and must be abandoned.

What happened here? When it was just an atom and an external field, the shift was finite. But when we allow for the creation of photons, the shift diverges. There's more phase space to create these photons than the denominator in the energy sum suppresses them. There is nothing wrong with Oppenheimer's calculation, it's just that there are other infinities which he missed which eventually cancel this infinity (for example, the electron mass is infinite). We will understand how to handle these infinities in a controlled way later in the course, this was just a preview to demonstrate the complexities we'll be up against. The main point is that this way of doing perturbation theory can be very misleading because it is non-relativistic.

5.3 Old-Fashioned Perturbation Theory

We start with a bare Hamiltonian and a perturbation

$$H = H_0 + V \tag{5.24}$$

We will allow there to be a continuous range of energies. For example, in the case of the Hydrogen atom coupled to an EM field, the photon energies k are continuous.

Say we have a particular eigenstate of H_0, $|\phi\rangle$:

$$H_0|\phi\rangle = E|\phi\rangle \tag{5.25}$$

Then the eigenstate in the presence of the full H is

$$H|\psi\rangle = E|\psi\rangle \tag{5.26}$$

$$\Rightarrow |\psi\rangle = |\phi\rangle + \frac{1}{E-H_0}V|\psi\rangle \tag{5.27}$$

The inverted object appearing here is a kind of Green's function known as the Lippmann-Schwinger kernel. It satisfies

$$\Pi_{\rm LS} = \frac{1}{E-H_0+i\varepsilon} \tag{5.28}$$

we have added an $i\varepsilon$ to the denominator to make sure there aren't any unwanted singularities. It is not important for now. In practice, this $E-H_0$ means take the difference between the energy of the excited state and the energy of the ground state.

Then

$$\Delta E = \langle\psi|H|\psi\rangle - \langle\psi|H_0|\psi\rangle \tag{5.29}$$

$$= \langle\psi|V|\psi\rangle \tag{5.30}$$

$$= \left(\langle\phi| + \langle\psi|V\frac{1}{E-H_0}\right)V\left(|\phi\rangle + \frac{1}{E-H_0}V|\psi\rangle\right) \tag{5.31}$$

$$= \langle\phi|V|\phi\rangle + \langle\phi|V\frac{1}{E-H_0}V|\phi\rangle + \langle\phi|V\frac{1}{E-H_0}V\frac{1}{E-H_0}V|\phi\rangle + \cdots \tag{5.32}$$

$$= \langle\phi|V|\phi\rangle + \langle\phi|V\Pi_{\rm LS}V|\phi\rangle + \langle\phi|V\Pi_{\rm LS}V\Pi_{\rm LS}V|\phi\rangle + \cdots \tag{5.33}$$

$$= V_{ii} + V_{in}\Pi_{\rm LS}V_{ni} + V_{in}\Pi_{\rm LS}(n)V_{nm}\Pi_{\rm LS}(m)V_{mf} + \cdots \tag{5.34}$$

This is just good-old time-independent perturbation theory written in a funny language of propagators. So you see that we have been using Green's functions and propagators to do perturbation theory all along. We call this old-fashioned perturbation theory.

The equation says that to calculate the energy shift in perturbation theory there is a sum of terms. In each term the potential creates in intermediate state (n) which propagates with the propagator $\Pi_{\rm LS}(n)$ until it hits another potential, where it creates a new field (m) which then propagates and so on, until they hit the final potential factor, which transitions to the final state. There is a nice diagrammatic way of drawing this series, called Feynman graphs.

The rules are

- All states are *physical*, that is, they are on-shell at all times.
- Matrix elements V_{nm} will vanish unless momentum is conserved at each vertex.
- Therefore, energy is *not* conserved at each vertex.

By the way, on-shell means that the state satisfies its free-field equations of motion. For example, a scalar field satisfying $(\Box - m^2)\phi$ would have $p^2 = m^2$. We are in general allowed to consider off-shell states (and they will be critical to doing modern perturbation theory), but in quantum mechanics you don't use them. Thinking "off-shell" is one of the great conceptual advances in quantum field theory.

5.4 Examples

Let's do some illustrative examples to see how old-fashioned perturbation theory works in practice.

5.4.1 stark effect

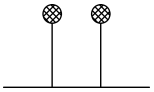
In the Stark effect calculation we did, the intermediate states were excited atoms. So we use

$$\Pi_n = \sum_{n>0} \frac{|\psi_n\rangle\langle\psi_n|}{H_{\text{atom}} - E_0} = \frac{\mathbb{1}}{H_{\text{atom}} - E_0} \tag{5.35}$$

Then,

$$\Delta E_0 = \sum_{m>0} \frac{\langle\psi_0|V|\psi_m\rangle\langle\psi_m|V|\psi_0\rangle}{E_m - E_0} = \langle\psi_0|V\Pi_{\text{atom}}V|\psi_0\rangle \tag{5.36}$$

This propagator represents an electron moving in a coulomb potential. We can draw this shift like

$\Delta E_0 \;=$ 

5.4.2 Lamb shift a la Oppenheimer 1930

Oppenheimer's failed attempt to calculate the Lamb shift works the same way. In this case, the intermediate state includes a propagating photon. So now what we were using as an external field is a separate propagator based on the free Hamiltonian

$$H_{\text{phot}} = \int d^3k\,\omega_k(a_k^\dagger a_k + \frac{1}{2}) \tag{5.37}$$

Now we the sum over excited atomic states and photon momenta:

$$\sum_{n>0} \to \int d^3p \sum_{n>0} \tag{5.38}$$

So the vertex splits the atom into an excited atom and a photon. Both propagate. The combined propagator is

$$\Pi_{n,p} = \int d^3p \sum_{n>0} \frac{|\psi_n;\,p\rangle\langle\psi_n;\,p|}{H_{\text{atom}} + H_{\text{photon}} - E_0} \tag{5.39}$$

Then the other vertex splits the same photon back into the atom and we have

$$\Delta E = \langle\psi_0|V\Pi_{n,p}V|\psi_0\rangle \tag{5.40}$$

Note that momentum is conserved, since the numerator is $|\psi_n;\,p\rangle\langle\psi_n;\,p|$, the photon and excited atom coming out of both vertices is the same state, so we can connect the lines and this diagram becomes a loop!

$\Delta E_0 \;=$ 

Keep in mind that this picture is just a shorthand for the perturbation expansion. But you can see it's really useful to remember what we're doing. The loop means that there is an unknown momentum, p, which we have to integrate over. We're only forcing momentum to be conserved overall, but it can be split between the atom and photon however it likes.

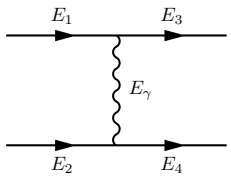
By the way, to see that momentum must be conserved is a little tricky with atomic states, since they are not momentum eigenstates. Nevertheless it's true, and we saw evidence of it in Oppenheimer's calculation when the phase factor $\exp(-i(k+m)x)$ forced $k + m \sim 0$. The amount by which $k + m \neq 0$ is the momentum of the original atomic state.

If you find this all a little vague, don't worry, we will be much more precise when we do perturbation theory properly in quantum field theory. It's also helpful to learn to live with a little vagueness if you have sound physical intuition.

5.5 Coulomb's law from OFPT

Let's try another example: an electron scattering off another electron. In this case, the states are free electrons, so they are momentum eigenstates and we can now be exact about the amplitudes.

The diagram for this process is



Say the intermediate photon has energy E_γ, and the incoming electrons have energies E_1 and E_2 and the outgoing electrons have energies E_3 and E_4. We need to calculate the transition matrix elements

$$T_{if} = V_{in}\frac{1}{E_i - E_n}V_{nf} \tag{5.41}$$

Where $E_i = E_f$ are the initial/final energies, and E_n is the energy of the intermediate state, which includes the energy of the photon.

Ignoring the spin of the electron and the photon, the interaction of the electron with the electromagnetic field is

$$V = e\psi_e A\psi_e = eJA, \quad J = \psi_e^\star\psi_e \tag{5.42}$$

So the matrix elements just give a factor of e. And we have

$$T_{if} = \frac{e^2}{E_i - E_n} \tag{5.43}$$

What is the intermediate state energy? In OFPT, where everything is on-the-mass-shell ($p^2 = m^2$), energy is not conserved in the interactions. That is, the energy of the intermediate state E_n is different from E_i and E_f. However, momentum is conserved. There are two possibilities:

Case 1) the first electron emits the photon and we look at the state before the photon hits the final state. Then the first electron is already in its final state, with energy E_3 and so the total intermediate state energy is

$$E_n^{(1)} = E_3 + E_2 + E_\gamma \tag{5.44}$$

$$\frac{1}{E_i - E_n^{(1)}} = \frac{1}{(E_1 + E_2) - (E_3 + E_2 + E_\gamma)} = \frac{1}{(E_1 - E_3) - E_\gamma} \tag{5.45}$$

Case 2) has the second electron emit the photon. Then the energy is

$$E_n = E_4 + E_1 + E_\gamma \tag{5.46}$$

$$\frac{1}{E_i - E_n^{(2)}} = \frac{1}{(E_1 + E_2) - (E_4 + E_1 + E_\gamma)} = \frac{1}{(E_2 - E_4) - E_\gamma} \tag{5.47}$$

Now overall energy conservation says $E_1 + E_2 = E_3 + E_4$, so $E_1 - E_3 = E_4 - E_2 \equiv \Delta E$. Adding these two cases gives

$$T_{if} = \frac{e^2}{E_i - E_n^{(1)}} + \frac{e^2}{E_i - E_n^{(2)}} = \frac{e^2}{\Delta E - E_\gamma} + \frac{e^2}{-\Delta E - E_\gamma} = \frac{2e^2 E_\gamma}{(E_\gamma)^2 - (\Delta E)^2} \tag{5.48}$$

Also, 3-momentum is conserved in OFPT, so $\vec{p}_\gamma = \vec{p}_3 - \vec{p}_1$. We also know that the intermediate photon is massless and on-shell (by definition in OFPT), so $E_\gamma = |\vec{p_\gamma}|$. Thus, let us *define* a 4-vector k^μ with

$$k^\mu = p_3^\mu - p_1^\mu = (\Delta E, \vec{p}_\gamma) \tag{5.49}$$

Then

$$k^2 = (\Delta E)^2 - (E_\gamma)^2 \tag{5.50}$$

So our transition matrix element is

$$T_{if} = -2E_\gamma \left(\frac{e^2}{k^2} \right) \tag{5.51}$$

Note that the Lorentz invariant propagator re-emerged! (The factor of $2E_\gamma$ just maintains the correct normalization).

So our diagram is really

$$T_{fi} = J \frac{1}{k^2} J \tag{5.52}$$

Where $J = \psi_e \psi_e$ is the electron source current. We have seen that if this source is treated as a static pointlike charge, we get the Coulomb potential. So now we have reproduced this observation from Old-Fashioned Perturbation Theory.

Note also that k^μ is not on-shell, since $k^2 \neq 0$ however now both energy and momentum are conserved. So it seems that there should be a formalism for perturbation theory in which 4-momentum is conserved at vertices and so propagators can be Lorentz invariant. This is time-dependent perturbation theory.

Chapter 6
Cross Sections and Decay Rates

6.1 Cross sections

The things we know how to calculate in quantum mechanics, or in quantum field theory (which is just quantum mechanics with lots of fields) are matrix elements

$$\langle f|i\rangle \tag{6.1}$$

where $|f\rangle$ is the final state and $|i\rangle$ the initial state. Matrix elements are the things which square to probabilities. This is the Schrodinger picture representation, where the states evolve in time. In the Heisenberg picture, we leave the states alone and put all the time evolution into an operator. In the special case where we scatter from $t=-\infty$ to $t=+\infty$, we give this operator a special name: the S-matrix. The S matrix is defined as

$$\langle f(t=-\infty)|i(t=+\infty)\rangle_{\text{schrodinger}} = \langle f|S|i\rangle \tag{6.2}$$

The S matrix has all the information about how the initial and final states evolve in time. Quantum Field Theory will tell us how to calculate S-matrix elements. Technically, the S-matrix is defined only for asymptotic ($t=\pm\infty$) states.

We also usually calculate the S-matrix perturbatively. In a free theory, it is simply the identity matrix $\mathbb{1}$. So we write

$$S = \mathbb{1} + iT \tag{6.3}$$

In general, the S matrix will conserve momentum. So it is helpful to factor an overall momentum conserving δ function to start with

$$T = (2\pi)^4\delta^4(\Sigma k_i)\,\mathcal{M} \tag{6.4}$$

Thus we can focus on computing the non-trivial part of the S-matrix, $\mathcal{M}$.

In a scattering process, we measure the probability by colliding beams of particles together and producing particular final states. Say 2 incoming particles with momenta p_1 and p_2 and a bunch of outgoing (final) particles with momenta p_j:

$$p_1 + p_2 \longrightarrow p_j \tag{6.5}$$

We would like to isolate, and eventually calculate the S-matrix element going into this process. Today we will discuss how to isolate it.

In reality, we collide beams with lots of particles. Since each collision is separate, we can write this as a sum of collisions in which there are only two particles colliding in the entire volume V and over the entire time T. To get something covariant, we have to also divide by the incoming flux. This gives us a cross section

$$d\sigma = \frac{1}{T}\frac{1}{\Phi}dP \tag{6.6}$$

where Φ is the flux and P is the probability.

The flux factor compensates for the boost when we do the calculation in a different frame. In the rest frame of one particle, the flux is just the magnitude of the velocity of the incoming particle divided by the total volume: $\Phi = |\vec{v}\,|/V$. If you like you can write the velocity as $\vec{v} = \vec{p}\,/p_0$. In a different frame, like the center-of-mass frame (COM), we have flux from both sides, so we have to add their velocities. Then $\Phi = |\vec{v}_1 - \vec{v}_2|/V$. This should be familiar from non-relativistic scattering. So we have

$$d\sigma = \frac{V}{T}\frac{1}{|\vec{v}_1 - \vec{v}_2|}dP \tag{6.7}$$

Next, we also know from quantum mechanics that probability is given by amplitudes squared. Since quantum field theory is just quantum mechanics a lot of fields, we have

$$dP = \frac{|\langle f|S|i\rangle|^2}{\langle f\,|f\rangle\langle i|i\rangle}d\Pi \tag{6.8}$$

Here $d\Pi$ is the region of final state momenta we are looking at. It is proportional to the product of the differential momentum d^3p_j of each final state and must integrate to 1. So

$$d\Pi = \prod_j \frac{V}{(2\pi)^3}d^3p_j \tag{6.9}$$

This has $\int d\Pi = 1$.

The $\langle f\,|f\rangle$ and $\langle i|i\rangle$ in the denominator come from the fact that the one particle states are defined at fixed time, and so are not normalized to 1 (this would not be Lorentz invariant). Instead we had

$$a_k^\dagger|0\rangle = \frac{1}{\sqrt{2\omega_k}}|k\rangle \quad [a_p, a_q^\dagger] = (2\pi)^3\delta^3(p-q) \tag{6.10}$$

So that

$$\langle p|p\rangle = (2\pi)^3(2\omega_p)\delta^3(0) \tag{6.11}$$

This $\delta^3(0)$ can be understood by using the relation

$$\delta(p) = = \frac{1}{2\pi}\int dx\, e^{ipx} \tag{6.12}$$

So,

$$\delta^3(0) = \frac{1}{(2\pi)^3}\int d^3x = \frac{V}{(2\pi)^3} \tag{6.13}$$

Thus,

$$\langle p|p\rangle = 2\omega_p V = 2E_p V \tag{6.14}$$

So,

$$\langle i|i\rangle = (2E_1V)(2E_2V), \quad \langle f\,|f\rangle = \prod_j (2E_jV) \tag{6.15}$$

Now we defined the matrix $\mathcal{M}$ so that

$$\langle f|S|i\rangle = \delta^4(\Sigma p)[1 + i\langle f\,|\,\mathcal{M}\,|i\rangle\,] \tag{6.16}$$

then

$$|\langle f|S|i\rangle|^2 = \delta^4(0)\delta^4(\Sigma p)|\mathcal{T}|^2 = TV\delta^4(\Sigma p)\,|\mathcal{T}|^2 \tag{6.17}$$

So

$$dP = \frac{\delta^{(4)}(\Sigma p)TV}{(2E_1V)(2E_2V)}\frac{1}{\prod_j (2E_jV)}|\mathcal{M}|^2\prod_j \frac{V}{(2\pi)^3}d^3p_j \tag{6.18}$$

$$dP = \frac{1}{V}\frac{1}{(2E_1)(2E_2)}|\mathcal{M}|^2 d\Pi_{\text{LIPS}} \tag{6.19}$$

where Lorentz Invariant Phase Space (LIPS) is defined as

$$d\Pi_{\text{LIPS}} = (2\pi)^4\delta^4(\Sigma p)\prod_{\text{final states } j}\frac{d^3p_j}{(2\pi)^3}\frac{1}{2E_{p_j}} \tag{6.20}$$

Putting everything together, we have

$$\boxed{d\sigma = \frac{1}{(2E_1)(2E_2)|\vec{v}_1 - \vec{v}_2|}|\mathcal{M}|^2 d\Pi_{\text{LIPS}}} \tag{6.21}$$

All the V's and $T's$ have dropped out, as they must, and it is helpful to recall that $\vec{v} = \frac{\vec{p}}{E}$.

6.2 Decay Rates

While we're at it, we may as well also calculate the formula for the decay rate. This is a 1 to many decay

$$p_1 \to p_j \tag{6.22}$$

You may be concerned that it is impossible for the incoming particle to be an asymptotic state at $-\infty$ if it is to decay, and so we shouldn't be able to use an S matrix at all. You are absolutely right. But it turns out that we can use it anyway. The basic reason is that we calculate the decay rate in perturbation theory, so it really is just like a $1 \to 2$ scattering process. If particle 1 decayed before our interaction, then it would be higher order in perturbation theory. Eventually, we will treat unstable particles properly, but that involves renormalization. So for now, let's just press on assuming it is not a decay but a $1 \to 2$ scattering.

Then everything is as before but we only have one $2E_iV$ in the denominator and there is no flux factor $\Phi = |\vec{v}\,|/V$. Then

$$\boxed{d\Gamma = \frac{1}{2E_1}|\mathcal{M}|^2 d\Pi_{\text{LIPS}}} \tag{6.23}$$

That was easy!

6.3 Special cases

We will often be calculating differential cross sections and decay rates for a certain special cases, like $2 \to 2$ scattering in the COM frame. Our formulas simplify in these cases.

For $2 \to 2$ scattering in the COM frame we have

$$p_1 + p_2 \longrightarrow p_3 + p_4 \tag{6.24}$$

with $\vec{p}_1 = -\vec{p}_2$ and $\vec{p}_3 = -\vec{p}_4$ and $E_1 + E_2 = E_3 + E_4 = E_{\text{cm}}$. Then

$$d\Pi_{\text{LIPS}} = (2\pi)^4 \delta^4(\Sigma p) \frac{d^3p_3}{(2\pi)^3} \frac{1}{2E_3} \frac{d^3p_4}{(2\pi)^3} \frac{1}{2E_4} \tag{6.25}$$

We don't need this entire differential distribution, so we will integrate over $\vec{p}_4$ using the delta function

$$d\Pi_{\text{LIPS}} = \frac{1}{16\pi^2} d\Omega \int dp_f \, \frac{p_f^2}{E_3} \frac{1}{E_4} \delta(E_3 + E_4 - E_{\text{cm}}) \tag{6.26}$$

where $p_f = |\vec{p}_3| = |\vec{p}_4|$. Now we want to integrate over p_f, but we can't forget about the mass shell conditions. These are $E_3 = \sqrt{m_3^2 + p_f^2}$ and also $E_4 = \sqrt{m_4^2 + p_f^2}$. So we have to carefully change variables in the delta function. Let $x = E_3 + E_4 - E_{\text{cm}} = x(p_f)$. Then

$$\frac{dx}{dp} = \frac{d}{dp}(E_3 + E_4 - E_{\text{cm}}) = \frac{p_f}{E_3} + \frac{p_f}{E_4} = p_f \left(\frac{E_{\text{cm}}}{E_3 E_4} \right) \tag{6.27}$$

Thus

$$d\Pi_{\text{LIPS}} = \frac{1}{16\pi^2} \int d\Omega \int_0^\infty dx \, \frac{p_f(x)}{E_{\text{cm}}} \delta(x) = \frac{1}{16\pi^2} \int d\Omega \, \frac{|p_f|}{E_{\text{cm}}} \tag{6.28}$$

Thus,

$$\left(\frac{d\sigma}{d\Omega}\right)_{\rm cm} = \frac{1}{(2E_1)(2E_2)|\bar{v}_1 - \vec{v}_2|}\frac{1}{16\pi^2}\frac{|p_f|}{E_{\rm cm}}|\mathcal{M}|^2 \tag{6.29}$$

Also

$$|\bar{v}_1 - \vec{v}_2| = \left|\frac{|p_i|}{E_1} + \frac{|p_i|}{E_2}\right| = |p_i|\frac{E_{\rm cm}}{E_1 E_2} \tag{6.30}$$

So we end up with the fairly simple formula

$$\boxed{\left(\frac{d\sigma}{d\Omega}\right)_{\rm cm} = \frac{1}{64\pi^2 E_{\rm cm}^2}\frac{|p_f|}{|p_i|}|\mathcal{M}|^2} \tag{6.31}$$

If all the masses are equal than $|p_f| = |p_i|$ and this formula simplifies further:

$$\left(\frac{d\sigma}{d\Omega}\right)_{\rm cm} = \frac{1}{64\pi^2 E_{\rm cm}^2}|\mathcal{M}|^2 \quad \text{(masses equal)} \tag{6.32}$$

6.4 Non-relativistic limit

To make sure we are not fooling ourselves, let's take the non-relativistic limit and show that we get the regular result from quantum mechanics. We'll consider the case where an electron ϕ_e of mass m_e scatters off a proton ϕ_p of mass m_p. From non-relativistic quantum mechanics, we know the cross section should be given by the Born approximation:

$$\left(\frac{d\sigma}{d\Omega}\right)_{\rm born} = \frac{1}{4\pi^2}m_e^2|\tilde{V}(k)]^2 \tag{6.33}$$

where the Fourier transform of the potential is given by

$$\tilde{V}(k) = \int d^3x\, e^{-ikx}V(x) \tag{6.34}$$

and k is the difference in the electron momentum before and after scattering (the momentum *transfer*). For example, if this is a Coulomb potential, $\tilde{V}(k) = \frac{e^2}{k^2}$ so

$$\left(\frac{d\sigma}{d\Omega}\right)_{\rm born} = \frac{1}{4\pi^2}m_e^2(\frac{e^2}{k^2})^2 \tag{6.35}$$

Now, for the field theory version, in the center-of-mass frame is the proton rest frame and $E_{\rm cm} = m_p$. Also the scattering is elastic so $|p_i| = |p_f|$. Then the prediction is

$$\left(\frac{d\sigma}{d\Omega}\right)_{\rm cm} = \frac{1}{64\pi^2 m_p^2}|\mathcal{M}|^2 \tag{6.36}$$

First, let's check the dimensions in these formulas: $[V] = [\frac{1}{r}] = M^1$, thus $[\tilde{V}] = M^{-2}$ and so $[(d\sigma)_{\rm born}] = M^{-2}$, which is the correct dimension for a cross section. For the relativistic formula, to get a cross section, we must have $[\mathcal{M}] = 1$.

Since we are taking the non-relativistic limit, we don't care about the electron and proton spin. Treating them as scalars, the minimal Lagrangian which couples a scalar to the electromagnetic field is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + \frac{1}{2}|\partial_\mu + eA_\mu\phi|^2 - \frac{1}{2}m^2\phi^2 \tag{6.37}$$

$$= \phi(\Box - m^2)\phi + eA_\mu\phi\partial_\mu\phi + \cdots \tag{6.38}$$

Now in the non-relativistic limit the momentum $p_\mu = (E, \vec{p})$ is close to being at rest $(m, 0)$, so $E \sim m$, that is $\partial_t \phi \sim m\phi$ and $|\vec{p}\,| \ll m$. Thus this becomes roughly

$$\mathcal{L} = \phi(\vec{\partial})^2\phi + e m A_0 \phi^2 \tag{6.39}$$

where we have one ϕ for the proton and one ϕ for the electron. Then the matrix element has a piece proportional to $e m_p$ from the interaction where the p^+ creates a photon, a factor of the propagator $\frac{1}{k^2}$ from the photon moving, and a piece proportional to $-e m_p$ from where the photon hits the electron. Thus

$$\mathcal{M} \sim e\, m_p \frac{1}{k^2} e\, m_e \tag{6.40}$$

Thus,

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{cm}} = \frac{e^4 m_e^2}{64\pi^2}\frac{1}{k^4} \tag{6.41}$$

which is the same thing up to factors of 2π which I lost somewhere along the way. Thus the non-relativistic limit works.

Note that the factor of m in the electron-photon interaction is very unconventional. It is more standard to rescale $\phi \to \frac{1}{\sqrt{2m}}\phi$ so that

$$\mathcal{L}_{\text{nr}} = \frac{1}{2m_e}\phi_e(\vec{\partial})^2\phi_e + \frac{1}{2m_p}\phi_p(\vec{\partial})^2\phi_p + e A^0 \phi_e^2 - e A^0 \phi_p^2 \tag{6.42}$$

which has the usual $\frac{p^2}{2m}$ for the kinetic term and an interaction with just a charge e and no mass m in the coupling. But of course it amounts to the same thing in the end. By the way, since in the non-relativistic limit $\frac{1}{\sqrt{2m}} = \frac{1}{\sqrt{2E}}$, this rescaling is closely related to our normalization factors $\frac{1}{\sqrt{2\omega_p}}$.

6.5 $e^+e^- \to \mu^+\mu^-$

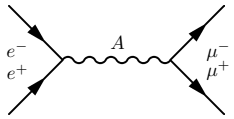
Let's do a simple example of calculating a matrix element. We will consider the process of electron-positron annihilation into muon-antimuon (taken from chapter 1 of Peskin and Schroeder). We want to calculate

$$\mathcal{M} = \langle e^+e^-|\mu^+\mu^-\rangle \tag{6.43}$$

Now we know the electron doesn't interact with the muon directly, only through the electromagnetism. So we'll have to sum over intermediate states

$$\mathcal{M} = \langle e^+e^-|\mu^+\mu^-\rangle = \sum_A \langle e^+e^-|A\rangle\langle A|\mu^+\mu^-\rangle \tag{6.44}$$

For future reference, the Feynman diagram looks like



(6.45)

This diagram has a precise meaning, as we will see soon, but for now just think of it as a pictorial drawing of the process: the e^+e^- annihilate into the photon, which propagates along, then decays into a $\mu^+\mu^-$ pair.

Let's try to guess the form of these matrix elements by using angular momentum. We will go to the center-of-mass frame, in which the e^- has momentum $p^- = (E, 0, 0, E)$ and the e^+ has momentum $p^+ = (E, 0, 0, -E)$. We know the electron is spin $1/2$, so we have 4 possibilities for the spin of the incoming state.

$$|i\rangle = |\uparrow\uparrow\rangle, \quad |i\rangle = |\downarrow\downarrow\rangle, \quad |i\rangle = |\uparrow\downarrow\rangle, \quad |i\rangle = |\downarrow\uparrow\rangle \tag{6.46}$$

Since the e^+ and e^- are coming from opposite directions (in the CM frame) there is no orbital angular momentum.

Now, the photon has spin 1, so only the first 2 initial states could possibly annihilate into a photon. Since the electron spin is perpendicular to it's momentum, the photon spin will be either in the x or y direction. The two possible resulting photon polarizations are

$$\epsilon^1 = (0,1,0,0), \quad \epsilon^2 = (0,0,1,0) \tag{6.47}$$

Both of these polarizations are produced by the e^+e^- annihilation. We can just pick a convention, say $|\uparrow\uparrow\rangle$ produces ϵ^1 and $|\downarrow\downarrow\rangle$ produces ϵ^2.

Next, the muon is also spin 1/2 so it has 4 possible spin states too. Similarly only 2 of them can annihilate into a photon. However, the μ^+ and μ^- are not going in the z direction. Their momenta can be written as

$$p_{\mu^-} = E(1,0,\sin\theta,\cos\theta), \quad p_{\mu^+} = E(1,0,-\sin\theta,-\cos\theta) \tag{6.48}$$

Where θ is the angle to the e^+e^- axis. There is also an azimuthal angle ϕ about the z axis which we have set to 0 for simplicity (the problem has cylindrical symmetry). So in this case the two possible directions for the photon polarization

$$\bar{\epsilon}^{\,1} = (0,1,0,0), \quad \bar{\epsilon}^{\,2} = (1,0,\cos\theta,-\sin\theta) \tag{6.49}$$

You can check that these are orthogonal to p_{μ^-} and p_{μ^+}.

Now the matrix element is given by summing over all possible intermediate polarizations for a particular initial state. So we have, summing over the two possible final states

$$\mathcal{M}_1 = \mathcal{M}(\,|\uparrow\uparrow\rangle \to |f\rangle) = \epsilon^1\bar{\epsilon}^{\,1} + \epsilon^1\bar{\epsilon}^{\,2} = 1 \tag{6.50}$$

$$\mathcal{M}_2 = \mathcal{M}(|\downarrow\downarrow\rangle \to |f\rangle) = \epsilon^2\bar{\epsilon}^{\,1} + \epsilon^2\bar{\epsilon}^{\,2} = \cos\theta \tag{6.51}$$

Thus, the final cross section, summing over initial states is

$$|\mathcal{M}|^2 = |\mathcal{M}_1|^2 + |\mathcal{M}_2|^2 = 1 + \cos^2\theta \tag{6.52}$$

This leads to

$$\frac{d\sigma}{d\Omega} = \frac{e^4}{64\pi^2 E_{\rm cm}^2}\big(1+\cos^2\theta\big) \tag{6.53}$$

Which gives the correct angular distribution in $e^+e^- \to \mu^+\mu^-$.

Peskin and Schroeder in their derivation neglect to point out that there is a factor of the photon propagator $\frac{1}{k^2}$ which they ignore. Since k is the virtual photon momentum, $k = (E_{\rm cm},\, 0,\, 0,\, 0)$ so $k^2 = E_{\rm cm}^2$. There is another factor of $2E = E_{\rm cm}$ from the e^+e^- side and a factor of $2E = E_{\rm cm}$ from the $\mu^+\mu^-$ side, which come from the relativistic normalization described in the previous section. Thus all these $E_{\rm cm}$'s cancel.

Chapter 7
The LSZ reduction formula and the Feynman propagator

7.1 LSZ reduction theorem

Most of the calculations we will be interested in in Quantum Field Theory are scattering cross sections. That is, we have a collider where two particles at $t = -\infty$ are collided with each other and we are interested in what happens at $t = +\infty$. We would like to do the calculation in momentum space, because generally we know the incoming momenta of the particles.

Let us add time dependence explicitly to our creation operators

$$a_p^\dagger(t) = a_p^\dagger e^{i\omega_p t} \tag{7.1}$$

Since we know how a_p and $a_p^\dagger$ depend on time, we know how the free quantum field depends on time

$$\phi(x,t) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2\omega_k}}\Big(a_k^\dagger e^{ikx} + a_k e^{-ikx}\Big) \tag{7.2}$$

where here $kx = \omega_k t - \vec{k}\vec{x}$. Note that there is no separate dynamical energy E, the energy is determined exactly by the momentum by $\omega_k \equiv \sqrt{\vec{k}^{\,2} + m^2}$ – these are on-shell external states.

So we want to set up the initial state as

$$|i\rangle = |p_1 p_2\rangle = \sqrt{2\omega_{p_1}}\sqrt{2\omega_{p_2}}a_{p_1}^\dagger(-\infty)a_{p_2}^\dagger(-\infty)|0\rangle \tag{7.3}$$

$$|f\rangle = \sqrt{2\omega_{p_3}}\cdots\sqrt{2\omega_{p_n}}a_{p_3}^\dagger(\infty)\cdots a_{p_n}^\dagger(\infty)|0\rangle \tag{7.4}$$

So,

$$\langle f|S|i\rangle = 4\sqrt{\omega_1\omega_2\omega_3\cdots\omega_n}\Big\langle 0|a_{p_3}(\infty)\cdots a_{p_n}(\infty)a_{p_1}^\dagger(-\infty)a_{p_2}^\dagger(-\infty)|0\Big\rangle \tag{7.5}$$

This expression is not terribly useful as is. We would like to relate it to something we can compute with our Lorentz invariant quantum fields $\phi(x,\, t)$.The LSZ (Lehmann-Symanzik-Zimmermann) reduction theorem relates $\langle f|i\rangle$ to an expression involving ϕ's. The following proof is an extended version Srednicki's. You can find various different versions in various textbooks.

To begin let's try to isolate the creation and annihilation operators. We can observe that

$$(i\partial_t + \omega_k)\Big(a_k^\dagger e^{ikx} + a_k e^{-ikx}\Big) = 2\omega_k a_k e^{ikx} \tag{7.6}$$

To get a_k out of $\phi(x,t)$ we simply have to Fourier transform:

$$\int d^3x\, e^{-i\vec{p}\vec{x}}(i\partial_t + \omega_p)\phi(x,t) = \int d^3x\, e^{-i\vec{p}\vec{x}}(i\partial_t + \omega_p)\int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2\omega_k}}\Big(a_k^\dagger e^{ikx} + a_k e^{-ikx}\Big) \tag{7.7}$$

$$= \int \frac{d^3k}{(2\pi)^3}\int d^3x\left[\left(\frac{-\omega_k + \omega_p}{\sqrt{2\omega_k}}\right)a_k^\dagger e^{ikx}e^{-i\vec{p}\vec{x}} + \left(\frac{\omega_k + \omega_p}{\sqrt{2\omega_k}}\right)a_k e^{-ikx}e^{-i\vec{p}\vec{x}}\right] \tag{7.8}$$

The x integral gives a $\delta^{(3)}(p+k)$ in the first term and a $\delta^{(3)}(p-k)$ in the second term. Either way, it forces $\omega_k = \omega_p$ and so we get

$$\int d^3x\, e^{-i\,\vec{p}\,\vec{x}}(i\partial_t + \omega_p)\phi(x,t) = \sqrt{2\omega_p}\, a_p e^{i\omega_p t} \tag{7.9}$$

To derive LSZ we will compute something very close to this:

$$\int d^4x\, e^{ipx}(\Box + m^2)\phi(x,t) \tag{7.10}$$

This integral is over 4 x's and p_μ is an arbitrary 4 vector, so that $p_0 \neq \sqrt{\vec{p}^{\,2} + m^2}$. We will have to be careful about boundary conditions at $t = \pm\infty$ but we will still assume that everything vanishes smoothly at $x = \pm\infty$ and so we can integrate the ∂_i^2 term by parts, which brings down a $-\vec{p}^{\,2}$.

Now,

$$\int d^4x\, e^{ipx}(\Box + m^2)\phi(x,t) = \int d^{\,4}x\, e^{ipx}[\partial_t^2 - \partial_i^2 + m^2]\phi(x,t) \tag{7.11}$$

$$= \int d^{\,4}x\, e^{ipx}[\partial_t^2 + \vec{p}^{\,2} + m^2]\phi(x,t) = \int d^{\,4}x\, e^{ipx}[\partial_t^2 + \omega_p^2]\phi(x,t) \tag{7.12}$$

And also

$$-i\partial_t\big[\, e^{ipx}(i\partial_t + \omega_p)\phi(x,t)\,\big] = \big[\, p_0 e^{ipx}(i\partial_t + \omega_p) + e^{ipx}(\partial_t^2 - i\omega_p\partial_t)\,\big]\phi(x,t) \tag{7.13}$$

$$= e^{ipx}(\partial_t^2 + \omega_p p_0 + i\,(p_0 - \omega_p)\partial_t)\phi(x,t) \tag{7.14}$$

So, in the limit $p_0 \to \omega_p$ this becomes

$$= e^{ipx}(\partial_t^2 + \omega_p^2)\phi(x,t) \tag{7.15}$$

which is what we had above. So we get

$$\int d^{\,4}x\, e^{ipx}[\partial_t^2 + \omega_p^2]\phi(x,t) = -\,i\int d^{\,4}x\,\partial_t\big[\, e^{ipx}(-i\partial_t + \omega_p)\,\phi(x,t)\,\big] \tag{7.16}$$

$$= -\,i\int dt\,\partial_t\big[\,\sqrt{2\omega_p}\, a_p e^{i\omega_p t}\,\big] \tag{7.17}$$

Now ordinarily we would drop something like this, but we have to be careful because we are precisely interested in the boundaries where $t = \pm\infty$. Thus putting everything together, we find

$$i\int d^4x\, e^{ipx}(\Box + m^2)\phi(x,t) = \sqrt{2\omega_p}\int dt\,\partial_t[a_p(t)] = \sqrt{2\omega_p}[a_p(\infty) - a_p(-\infty)] \tag{7.18}$$

That is,

$$\sqrt{2\omega_p}[a_p(\infty) - a_p(-\infty)] = -\,i\int d^4x\, e^{-ipx}(\Box + m^2)\phi(x,t) \tag{7.19}$$

Similarly (by taking the Hermetian conjugate)

$$\sqrt{2\omega_p}\Big[\, a_p^\dagger(\infty) - a_p^\dagger(-\infty)\,\Big] = i\int d^4x\, e^{-ipx}(\Box + m^2)\phi(x,t) \tag{7.20}$$

We're almost done. We wanted to compute

$$\langle f\,|S|\,i\rangle = 4\sqrt{\omega_1\omega_2\omega_3\omega_4}\Big\langle 0|a_{p_3}(\infty)\cdots a_n(\infty)a_{p_1}^\dagger(-\infty)a_{p_2}^\dagger(-\infty)|0\Big\rangle \tag{7.21}$$

and we have an expression for $a_p(\infty) - a_p(-\infty)$. Note that all the initial states have $a_p^\dagger$'s and $-\infty$'s and the final states have $a_p's$ and $+\infty$'s. This leads to a very important trick. We can rewrite the S-matrix element as

$$\langle f\,|S|\,i\rangle = 4\sqrt{\omega_1\omega_2\omega_3\omega_4}\,\big\langle 0\big|T\{[a_{p_3}(\infty) - a_{p_3}(-\infty)][a_{p_4}(\infty) - a_{p_4}(-\infty)] \tag{7.22}$$

$$\times\,[a_{p_1}^\dagger(\infty) - a_{p_1}^\dagger(-\infty)][a_{p_2}^\dagger(\infty) - a_{p_2}^\dagger(-\infty)]\Big|0\Big\rangle \tag{7.23}$$

where the time-ordering operator $T\{\}$ indicates that all the operators should be ordered so that those at later times are always to the left of those at earlier times. This just migrates all the unwanted $a^\dagger(\infty)$'s associated with the initial states to the left, where they annihilate on $\langle f|$ and all the unwanted $a(-\infty)'s$ to the right, where they annihilate $|i\rangle$. So all the stuff we don't want drops out of this expression. Note that we get a -1 for all the initial states, which turns the $-i's$ into $i's$.

Thus, we get

$$\langle p_3 p_4|p_1 p_2\rangle = \left[i\int d^4x_1\, e^{ip_1 x_1}(\Box_1 + m^2)\right]\cdots\left[i\int d^4x_2\, e^{-ip_4 x_4}(\Box_4 + m^2)\right] \tag{7.24}$$

$$\times\, \langle 0|T\{\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\}|0\rangle \tag{7.25}$$

where $\Box_i = (\frac{\partial}{\partial x_i^\mu})^2$. This is the LSZ reduction formula.

It says that to calculate a scattering matrix element, put some extra $\Box + m^2$ factors in front. Remember that we had to take $p_0 \to \omega_p$ at some point in the proof. This puts the external states in $|i\rangle$ and $|f\rangle$ on shell. It also causes their propagators, which look like $\frac{1}{\Box + m^2}$ to blow up. However, the the factors of $\Box + m^2$ will precisely cancel these divergences. They will also kill anything which does not have a divergence. That's the whole point of the LSZ formula: it isolates the asymptotic states by looking for a pole.

Now, we did this derivation for free fields. If you want ϕ to be an interacting field, the result is the same, but the derivation is significantly more complicated. We need to make sure the asymptotic states at $-\infty$ and $+\infty$ are one-particle, not multi-particle states. To do this, the proof proceeds by constructing wave packets for the initial and final states and shows that the multi-particle components of the wave-packets disperse differently allowing the one-particle components to be isolated. To really prove the theorem, we also need elements of renormalization. Basically, we have to require that

$$\langle 0|\phi(x)|0\rangle = 0 \tag{7.26}$$

and

$$\big\langle k|\phi(x)|0\rangle = e^{-ikx} \tag{7.27}$$

We showed these hold in the free theory, but unless we are careful, these will not hold when ϕ can interact. The bottom line is that we *demand* that these relations hold, by redefining, or renormalizing, at each order in perturbation theory. We will come back to this in a few weeks.

The physical reason that LSZ works for interacting fields is that $\Box + m^2$ is just a funny way of writing zero for an asymptotic state, whether it's free or interacting. But only a 1-particle state will have a $\frac{1}{\Box + m^2}$ pole in the amplitude to cancel the zero. Thus the $\Box + m^2$ terms project out the 1-particle states. In any case, I don't know of anyone who has ever needed to use the details of LSZ for anything. So I encourage you to skim the proof on your own, but don't stress out about it if you can't master it on first pass. You can study it again in the future if you want to – it's not going anywhere.

It's easy to think that LSZ is totally trivial, but it isn't! It's the only thing that tells us what the initial states are (the things created from the vacuum at $t = -\infty$) and what the final states are (the things that annihilate into the vacuum at $t = +\infty$). The $\pm i$ in the phase factors in the $\Box + m^2$ terms project these out. The time ordering is totally physical: all the creation of the initial states happens before the annihilation of the final states. In fact, because this is true not just for free fields, all the crazy stuff which happens at intermediate times in an interacting theory must be time-ordered too. But the great thing is that we don't need to know which are the initial states and which are the final states anymore when we do the hard part of the computation. We just have to calculate time ordered products, and the LSZ theorem sorts out what is being scattered for us.

7.2 Feynman propagator

Let's warm up by calculating this time-ordered product in the free theory. We start with the free field operator

$$\phi(x,t) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2\omega_k}}\Big[a_k^\dagger e^{ikx} + a_k e^{-ikx}\Big] \tag{7.28}$$

where $k_0 = \omega_k = \sqrt{m^2+k^2}$. Then

$$\langle 0|\phi(x_1)\phi(x_2)|0\rangle = \int \frac{d^3k_1}{(2\pi)^3}\int \frac{d^3k_2}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{k_1}}}\frac{1}{\sqrt{2\omega_{k_2}}}\Big\langle 0|a_{k_1}a_{k_2}^\dagger|0\Big\rangle e^{i(k_2x_2-k_1x_1)}$$

The $\Big\langle 0|a_{k_1}a_{k_2}^\dagger|0\Big\rangle$ gives a $(2\pi)^3\delta^3(k_1-k_2)$ so that

$$\langle 0|\phi(x_1)\phi(x_2)|0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{ik(x_2-x_1)} \tag{7.29}$$

Now, we are interested in $\langle 0|T\{\phi(x_1)\phi(x_2)\}|0\rangle$. This is (remember latest to the left)

$$\langle 0|T\{\phi(x_1)\phi(x_2)\}|0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}\Big[e^{ik(x_2-x_1)}\,\theta(t_1-t_2) + e^{ik(x_1-x_2)}\,\theta(t_2-t_1)\Big] \tag{7.30}$$

$$= \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}\Big[e^{i\,\vec{k}\,(\vec{x}_1-\vec{x}_2)}\,e^{-i\omega_k\tau}\theta(\tau) + e^{-i\,\vec{k}\,(\vec{x}_1-\vec{x}_2)}\,e^{i\omega_k\tau}\,\theta(-\tau)\Big] \tag{7.31}$$

where $\tau = t_1 - t_2$. Taking $k \to -k$ leaves the volume integral $\int d^3k$ invariant and gives

$$\langle 0|T\{\phi(x_1)\phi(x_2)\}|0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{-i\,\vec{k}\,(\vec{x}_1-\vec{x}_2)}\big[e^{i\omega_k\tau}\theta(-\tau) + e^{-i\omega_k\tau}\theta(\tau)\big] \tag{7.32}$$

Now it turns out that

$$e^{-i\omega_k\tau}\theta(\tau) + e^{i\omega_k\tau}\theta(-\tau) = \lim_{\varepsilon\to 0}\frac{-2\omega_k}{2\pi i}\int_{-\infty}^{\infty}\frac{d\omega}{\omega^2-\omega_k^2+i\varepsilon}e^{i\omega\tau} = \tag{7.33}$$

To see this, we first separate out the pole

$$\frac{1}{\omega^2-\omega_k^2+i\varepsilon} = \frac{1}{[\omega-(\omega_k-i\varepsilon)][\omega-(-\omega_k+i\epsilon)]} = \frac{1}{2\omega_k}\left[\frac{1}{\omega-(\omega_k-i\epsilon)} - \frac{1}{\omega-(-\omega_k+i\varepsilon)}\right] \tag{7.34}$$

where we have dropped terms of order ε^2 and write $\varepsilon\omega_k = \varepsilon$, which is ok since we will take $\varepsilon \to 0$ in the end. Now,

$$\int_{-\infty}^{\infty}\frac{1}{\omega-(-\omega_k+i\epsilon)}e^{i\omega\tau} = -2\pi i e^{-i\omega_k\tau}\theta(\tau) + O(\varepsilon) \tag{7.35}$$

since we only pick up the pole if we close the contour up, which we would only do for $\tau > 0$, and the minus sign comes from doing the contour counterclockwise. Similarly,

$$\int_{-\infty}^{\infty}\frac{1}{\omega-(\omega_k-i\epsilon)}e^{i\omega\tau} = 2\pi i e^{i\omega_k\tau}\theta(-\tau) + O(\varepsilon) \tag{7.36}$$

Thus,

$$\lim_{\varepsilon\to 0}\int_{-\infty}^{\infty}\frac{d\omega}{\omega^2-\omega_k^2+i\varepsilon}e^{i\omega\tau} = -\frac{2\pi i}{2\omega_k}\big[e^{i\omega_k\tau}\theta(-\tau) + e^{-i\omega_k\tau}\theta(\tau)\big] \tag{7.37}$$

Then

$$\langle 0|T\{\phi(x_1)\phi(x_2)\}|0\rangle = \lim_{\varepsilon\to 0}\int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{-i\,\vec{k}\,(\vec{x}_1-\vec{x}_2)}\int d\omega\frac{-2\omega_k}{2\pi i}\frac{1}{\omega^2-\omega_k^2+i\varepsilon}e^{i\omega\tau} \tag{7.38}$$

Letting the limit be implicit, this is:

$$\boxed{\langle 0|T\{\phi(x_1)\phi(x_2)\}|0\rangle = \int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2-m^2+i\varepsilon}e^{ik(x_1-x_2)}} \tag{7.39}$$

This beautiful Lorentz invariant object is called the Feynman propagator.

Keep in mind:

- $k_0 \neq \sqrt{\vec{k}^2 + m^2}$ anymore. It is a separate integration variable. The propagating field can be off-shell!
- The i comes from a contour integral. We will always get factors of i in two-point functions of real fields.
- The ε is just a trick for representing the time ordering a simple way. We will always take $\varepsilon \to 0$ at the end, and often leave it implicit. You always need a $+i\varepsilon$ for time ordered products, but is really just short hand for a pole-prescription in the contour integral, which is exactly equivalent to adding various $\theta(t)$ factors.
- For $\varepsilon = 0$ this looks just like the classical Green's function for the Klein Gordon equation $\Box\phi = J$. That's because it *is* the same thing. We are just computing classical propagation in a really complicated way.

Chapter 8
Feynman Rules

8.1 Time-Dependent Perturbation Theory

From the LSZ theorem, we need objects like

$$\langle f|i\rangle \sim \langle 0|T\{\phi(x_1)\cdots\phi(x_n)\}|0\rangle \tag{8.1}$$

where these $\phi's$ are not free but the fields which evolve due to the full Hamiltonian.

First, we separate, as usual, the Hamiltonian into H_0 and the perturbation V.

$$H = H_0 + V \tag{8.2}$$

We already know how the free fields evolve, due to H_0

$$\phi_I(x,t) = e^{iH_0t}\phi_0(x)e^{-iH_0t} = \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2\omega_k}}\Big[a_k e^{ikx} + a_k^\dagger e^{-ikx}\Big] \tag{8.3}$$

The subscript I stands for "interaction", since we are now in the interaction picture. Note that the free fields are defined at a reference time $t = t_0 = 0$ and $\omega_k = \sqrt{\vec{k}^2 + m^2}$. So going to the interaction picture is easy in QFT, and particularly nice because the phase factors become Lorentz Invariant.

So really what we computed the Feynman propagator last time, we were computing

$$\langle 0|T\{\phi_I(x)\phi_I(y)\}|0\rangle = \int \frac{d^4k}{(2\pi)^4}\frac{1}{k^2 - m^2 + i\varepsilon}e^{ik(x-y)} \tag{8.4}$$

Now let's define an operator U which expresses the full field $\phi(x,t)$ in terms of the evolving free field $\phi_I(x,t)$ where 0 is the reference time at which the free fields are equal to the interacting fields. We are relating ϕ to ϕ_I at the *same time* t, but the evolution has started from time t_0.

We know

$$\phi(x,t) = e^{iHt}\phi_0(x,0)e^{-iHt} = e^{iHt}e^{-iH_0t}(e^{iH_0t}\phi_0(x,0)e^{-iH_0t})e^{iH_0t}e^{-iHt} \tag{8.5}$$

Then

$$U(t,0) = e^{iH_0t}e^{-iHt} \tag{8.6}$$

It follows that

$$i\partial_t U(t,0) = e^{iH_0t}(-H_0 + H)e^{-iHt} \tag{8.7}$$

$$= \big(e^{iH_0t}Ve^{-iH_0t}\big)e^{iH_0t}e^{-iHt} \tag{8.8}$$

So

$$i\partial_t U(t,0) = V_I(t)U(t,0) \tag{8.9}$$

where

$$V_I(t) = e^{iH_0t}Ve^{-iH_0t} \tag{8.10}$$

All we've done is conjugate everything with e^{iH_0t} and we are left with the Schrodinger equation with bunch of I's.

Now we need to solve

$$i\partial_t U(t,t_0) = V_I(t)U(t,t_0) \tag{8.11}$$

If everything commuted, the solution would just be $U_I(t,t_0) = \exp(-i\int_{t_0}^{t} V_I(t')dt')$. But $V_I(t_1)$ does not necessarily commute with $V_I(t_2)$, so this is not the right answer. It turns out the right answer is very similar

$$U(t,t_0) = T\left\{\exp[-i\int_{t_0}^{t} dt'\, V_I(t')]\right\} \tag{8.12}$$

where T is the time ordering operator. This solution works because time ordering effectively makes everything inside commute:

$$T\{AB\} = T\{BA\} \tag{8.13}$$

Which is why we get the exponential solution.

Time ordering of an exponential is defined in the obvious way through its expansion

$$U(t,0) = 1 - i\int_0^t dt' V_I(t') - \frac{1}{2}\int_0^t dt' \int_0^{t'} dt'' T\{V_I(t')V_I(t'')\} + \cdots \tag{8.14}$$

8.1.1 perturbative solution for Dyson's series

If you don't find the previous argument compelling, we can solve for $U(t)$ the hard way, using perturbation theory.

$$i\partial_t U(t) = V_I(t)U(t) \tag{8.15}$$

At zeroth order

$$U(t) = 1 \tag{8.16}$$

To first order in V_I

$$U(t) = 1 + \int_0^t dt' V_I(t')U(t') = 1 + \int_0^t dt' V_I(t') \tag{8.17}$$

To second order

$$U(t) = 1 - \int_0^{t'} dt' V_I(t')\left[1 - \int_0^{t'} dt'' V_I(t'')U(t'')\right] \tag{8.18}$$

$$= 1 - \int_0^t dt' V_I(t') - \int_0^t dt' \int_0^{t'} dt'' V_I(t')(t'') \tag{8.19}$$

Note that $t' > t''$ so we can write this as

$$U(t) = 1 - \int_0^t dt'(t') - \int_0^t dt' \int_0^{t'} dt'' T\{V_I(t')V_I(t'')\} \tag{8.20}$$

where T means the things in brackets are time ordered. So this is

$$U(t) = 1 - \int_0^t dt' V_I(t') - \frac{1}{2}\int_0^t dt' \int_0^{t'} dt'' T\{V_I(t')V_I(t'')\} + \cdots \tag{8.21}$$

And in fact

$$U(t) = T\left\{\exp[-i\int_0^t dt' V_I(t')]\right\} \tag{8.22}$$

More generally

$$U(t,t_0) = T\left\{\exp[-i\int_{t_0}^t dt' V_I(t')]\right\} \tag{8.23}$$

8.1.2 U relations

It is convenient to abbreviate U with

$$U_{21} \equiv U(t_2, t_1) = T\left\{ \exp[-i \int_{t_1}^{t_2} dt' V_I(t')] \right\} \tag{8.24}$$

Remember that in QFT we always have later times on the LEFT. It follows that

$$U_{21} U_{12} = 1 \quad \Rightarrow \quad U_{21}^{-1} = U_{21}^\dagger = U_{12} \tag{8.25}$$

and for $t_1 < t_2 < t_3$

$$U_{32} U_{21} = U_{31} \tag{8.26}$$

So,

$$\phi(x,t) = U^\dagger(t,0) \phi_I(x,t) U(t,0) \tag{8.27}$$

$$\phi(x,t_1) = U^\dagger(t_1,0) \phi_I(x,t_1) U(t_1,0) = U_{10}^\dagger \phi(x,t_1) U_{10} \tag{8.28}$$

$$= U_{01} \phi(x,t_1) U_{10} \tag{8.29}$$

8.1.3 Ground State

The last thing we need to know is what the ground state is in the interacting theory. Call it $|\Omega\rangle$. In the free theory

$$|0\rangle_I = e^{iH_0 t} |0\rangle = |0\rangle \tag{8.30}$$

Then

$$|\Omega(t)\rangle = U(t, -\infty) |0\rangle = U_{t-\infty} |0\rangle \tag{8.31}$$

That is, the reference time, for which the $|\Omega\rangle = |0\rangle$ is $t = -\infty$.

Thus we can evolve everything to a common time t_0. Recalling

$$\phi(x,t_1) = U_{01} \phi(x,t_1) U_{10} \tag{8.32}$$

we find

$$\langle \Omega | \phi(x_1) \phi(x_2) | \Omega \rangle = \langle 0 | U_{\infty 0} U_{01} \phi_I(x_1) U_{10} U_{02} \phi_I(x_2) U_{20} U_{0-\infty} | 0 \rangle \tag{8.33}$$

And time-ordered products have a particularly nice structure, because we can commute things around and combine the $U's$:

$$\langle \Omega | T\{\phi(x_1)\phi(x_2)\} | \Omega \rangle = \langle 0 | T\{\phi_I(x_1)\phi_I(x_2) U_{\infty 0} U_{01} U_{10} U_{02} U_{20} U_{0-\infty}\} | 0 \rangle \tag{8.34}$$

$$= \langle 0 | T\{\phi_I(x_1)\phi_I(x_2) U_{\infty -\infty}\} | 0 \rangle \tag{8.35}$$

$$= \left\langle 0 \middle| T\left\{ \phi_I(x_1)\phi_I(x_2) \exp[-i \int_{-\infty}^{\infty} dt V_I(t)] \right\} \middle| 0 \right\rangle \tag{8.36}$$

Which is a remarkably simply final result.

8.1.4 meaning of I subscript

We saw that ϕ_I is just the free field ϕ_0 with explicit time-dependence in the phase factors

$$\phi_I(x,t) = e^{iH_0 t} \phi_0(x) e^{-iH_0 t} = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_k}} \Big[a_k e^{ikx} + a_k^\dagger e^{-ikx} \Big] \tag{8.37}$$

Similarly the interaction potential is just the potential written in terms of interacting $\phi's$. That is, suppose we know the interaction in terms of the fields at some time t_0:

$$V = \int d^3x \big[\lambda\phi^3\big] = \int d^3x \big[\lambda\phi_0^3\big] \tag{8.38}$$

We use ϕ_0 instead of the interacting field ϕ not because the fields are not interacting, but because our basis of fields at the time $t = t_0$ is the basis of free fields. Then

$$V_I = e^{iH_0t} \int d^3x \big[\lambda\phi_0^3\big] e^{-iH_0t} = \int d^3x \big[\lambda\phi_I^3\big] \tag{8.39}$$

We will usually just drop the I subscripts on V and ϕ, because going to the interaction picture just means adding $it_0\omega_k$ to the phase factors, which is unambiguous. For this lecture, however, we will keep I's in to make an explicit distinction between free and interacting fields.

8.1.5 Summary

In summary, the I subscript refers to the free theory evolving with the free Hamiltonian. No subscript means the full interacting fields. Matrix elements are related by

$$\langle\Omega|\phi(x_1)\phi(x_2)|\Omega\rangle = \langle 0|U_{\infty 0}U_{01}\phi_I(x_1)U_{10}U_{02}\phi_I(x_2)U_{20}U_{0-\infty}|0\rangle \tag{8.40}$$

where

$$U_{ij} = T\left\{\exp[-i\int_{t_j}^{t_i} dt'\, V_I(t')]\right\} \tag{8.41}$$

and $|\Omega\rangle$ is the ground state in the interacting theory.

For the special case of time ordered products, such as what we need for S-matrix elements, this simplifies to

$$\boxed{\langle\Omega|T\{\phi(x_1)\phi(x_2)\}|\Omega\rangle = \left\langle 0\middle|T\left\{\phi_I(x_1)\phi_I(x_2)\exp[-i\int_{-\infty}^{\infty} dt V_I(t)]\right\}\middle|0\right\rangle} \tag{8.42}$$

8.2 Time ordered products and contractions

Let's use this formula to calculate something. Let's take our interaction to be of the form ϕ^3. Then

$$V_I = \lambda \int d^3x \phi(x)^3 \tag{8.43}$$

To next order, we use the formula

$$\langle\Omega|T\{\phi(x_1)\phi(x_2)\}|\Omega\rangle = \left\langle 0\middle|T\left\{\phi_I(x_1)\phi_I(x_2)\exp\left[-i\int_{-\infty}^{\infty} dt V_I(t)\right]\right\}\middle|0\right\rangle \tag{8.44}$$

$$= \langle 0|T\{\phi_I(x_1)\phi_I(x_2)|0\rangle - i\lambda\int d^4x \big\langle 0\big|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(x)^3\big|0\big\rangle \tag{8.45}$$

$$-\lambda^2\int d^4x \int d^4y \big\langle 0\big|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(x)^3\phi_I(y)^3\big|0\big\rangle + \cdots \tag{8.46}$$

Notice how the $\int_{-\infty}^{\infty} dt$ combined with the $\int d^3x$ in V to give a Lorentz invariant integral. So now we need to be able to compute expressions like

$$\big\langle 0\big|T\big\{\phi_I(x_1)\phi_I(x_2)\phi_I(x)^3\phi_I(y)^3\big\}\big|0\big\rangle \tag{8.47}$$

To begin, recall that each $\phi_I(x)$ has a creation and an annihilation operator and some exponentials

$$\phi_I(x) = \int \frac{d^4p}{(2\pi)^4}\frac{1}{\sqrt{2\omega_p}}\Big(a_p^\dagger e^{ipx} + a_p e^{-ipx}\Big) \tag{8.48}$$

So the expression becomes a bunch of $a_p's$ and $a_p^{\dagger\prime}s$ and exponentials. In our example, we get something like

$$(a_{p_1}^\dagger e^{ip_1x_1} + h.c)(a_{p_2}^\dagger e^{ip_2x_2} + h.c)(a_{p_3}^\dagger e^{ip_3x} + h.c.)(a_{p_4}^\dagger e^{ip_3x} + h.c.)\cdots(a_{p_8}^\dagger e^{ip_8x} + h.c.) \tag{8.49}$$

integrated over x, y and the 8 $p's$.

For this to give a non-zero matrix element on the vacuum, all the a_{p_i}'s will have to be paired off with all the $a_{p_j}^\dagger$'s. This is really easy to do – for each field, from right to left, either create a field from the vacuum, or annihilate one that's already been created. e just sum over all ways of doing this. Then each pairing will give

$$\langle 0|T\{\phi_I(x)\phi_I(y)\}|0\rangle = \int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2 - m^2 + i\varepsilon}e^{ik(x-y)} \equiv D_F(x,y) \tag{8.50}$$

If there's more $a_p's$ than $a_p^{\dagger\prime}s$ the answer will be zero. If two $p's$ are the same, then there could be some 2-particle states, or more generally n-particle states, which might give a factors of n. So we'll have to be careful to divide out by symmetry factors for identical particles.

For our example interaction

$$\big\langle 0\big|T\big\{\,\phi_I(x_1)\phi_I(x_2)\phi_I(x)^3\phi_I(y)^3\big\}\big|0\big\rangle \tag{8.51}$$

There are 3 possible contractions

1. $D_F(x_1,x_2)D_F(x,y)^3$
2. $D_F(x_1,x)D_F(x_2,y)D_F(x,y)D_F(x,y)$
3. $D_F(x_1,x)D_F(x_2,x)D_F(x,y)D_F(y,y)$

So the result is the sum of these, plus the phases, with appropriate symmetry factors.

8.2.1 Wick's theorem

This procedure we described follow from a more general theorem called Wick's theorem. I don't really know why you need the more general form, but we can't really cover field theory without mentioning it.

Recall the Feynman propagator

$$\langle 0|T\{\phi_I(x)\phi_I(y)\}|0\rangle = \int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2 - m^2 + i\varepsilon}e^{ik(x-y)} \equiv D_F(x,y) \tag{8.52}$$

This came from

$$\Big\langle 0|(a_{p_1}^\dagger e^{ip_1x} + a_{p_1}e^{-ip_1x})(a_{p_2}^\dagger e^{ip_2y} + a_{p_2}e^{-ip_2y}|0\Big\rangle \tag{8.53}$$

We only used the $a_{p_1}^\dagger a_{p_2}$ term in the expansion. In fact really, what we used was $[a_{p_1}^\dagger, a_{p_2}]$, since this is a delta function. If we expand the operator we find in addition to the $[a_{p_1}, a_{p_2}^\dagger]$ term used for the Feynman propagator, terms like

$$a_{p_1}^\dagger a_{p_2} + a_{p_1}^\dagger a_{p_2}^\dagger + a_{p_1}a_{p_2} \tag{8.54}$$

All of these have $a_p^\dagger$'s on the left and a_p's on the right, so they have vanishing matrix elements in the vacuum. We call such terms *normal ordered*, and indicate it with a $::$.

- normal ordering: all the $a_p^\dagger$'s are on the right, and all the $a_p's$ are on the left. Represented with colons.

Thus for the two point function:

$$T\{\phi_I(x)\phi_I(y)\} = D_F(x,y) + :\phi_I(x)\phi_I(y): \tag{8.55}$$

Wick's theorem says that

$$T\{\phi_I(x_1)\cdots\phi_I(x_n)\} = :\phi_I(x_1)\cdots\phi_I(x_n) + \text{all possible contractions}: \tag{8.56}$$

where a contraction means take two fields $\phi(x_i)$ and $\phi(x_j)$ from anywhere in the series and replace them with a factor of $D_F(x_i, x_j)$ for each pair of fields. All possible contractions includes one contraction, two contractions, etc, involving any of the fields. But each field can only be contracted once. Since normal ordered products vanish unless all the fields are contracted this implies that

The proof is by induction. For one field,

$$T\{\phi_I(x)\} = :\phi_I(x):\tag{8.57}$$

Now suppose the theorem holds for n fields. Since we can commute the fields within a time ordering symbol, let's just put them all in time order, and reliable the $x's$ appropriately (we could have called them $y's$ if you like). Then

$$T\{\phi_I(x_1)\phi_I(x_2)\cdots\phi_I(x_n)\} = \phi_I(x_1)T\{\phi_I(x_2)\cdots\phi_I(x_n)\} \tag{8.58}$$

$$= (a^\dagger_{p_1}e^{ip_1x} + a_{p_1}e^{-ip_1x})[:\phi_I(x_2)\cdots\phi_I(x_n) + \text{contractions of}\{x_2,\cdots,x_n\}:] \tag{8.59}$$

Now we need to get the a_{p_1} and $a^\dagger_{p_1}$ inside the normal ordering. For a_{p_1} it's easy since it is already on the left, where it needs to be. The term with $a^\dagger_{p_1}$ has to be commuted passed everybody. Each time we get a

$$a^\dagger_{p_1}\phi(x_j) = \cdots a^\dagger_{p_1}(a^\dagger_{p_j} + a_{p_j})\cdots = \cdots\phi(x_j)a^\dagger_{p_1}\cdots + \cdots\Big[a^\dagger_{p_1}, a_{p_j}\Big]\cdots \tag{8.60}$$

But this commutator is just a delta function, which is the same thing we used to calculate the Feynman propagator. So this gives a factor of $D_F(x_1, x_j)$. As we move $a^\dagger_p$ along, we pick up all the possible contractions, which is all there is to the proof.

The result of Wick's theorem is that time-ordered products are given by a bunch of contractions plus normal ordered products. Since the normal-ordered products vanish in vacuum matrix elements, we get the special case we used above. The bottom line is that all of the S-matrix elements we will compute can be reduced to a bunch of $D_F s$ and a bunch of phase factors.

8.3 Feynman diagrams

There's a very important way to rewrite the sum of contractions called Feynman diagrams. It says draw a point for each x_i or x or y in the expression. Then each $D_F(x, y)$ is a line between x and y. Then there are as many lines coming out of the $x's$ as there are insertions. The final amplitude will be a sum of all the ways of contracting everything together.

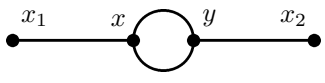
Let's work through our example. We started from

$$V_I = \lambda\phi(x)^3 \tag{8.61}$$

and were lead to

$$\big\langle 0\big|T\big\{\phi(x_1)\phi(x_2)\phi(x)^3\phi(y)^3\big\}\big|0\big\rangle = \lambda^2\int d^4x\int d^4y D_F(x_1,x)D_F(x_2,y)D_F(x,y)D_F(x,y) + \cdots$$

This term corresponds to one particular Feynman diagram:



To evaluate this diagram, first write every D_F in momentum space

$$D_F(x,y) = \int\frac{d^4p}{(2\pi)^4}\frac{i}{p^2 - m^2 + i\varepsilon}e^{ip(x-y)} \tag{8.62}$$

Thus all the $x's$ appear only in exponentials. There were 8 $p's$ to begin with, and we paired them off, so there are 4 p's left from the Fourier transforms of the 4 D_F's. We have

$$M = \lambda^2 \int d^4x \int d^4y \Big[\int \frac{d^4p}{(2\pi)^4} \Big]^4 e^{ip_1(x_1-x)} e^{ip_2(x_2-y)} e^{ip_3(x-y)} e^{ip_4(x-y)} \frac{i}{p_1^2+i\varepsilon} \frac{i}{p_2^2+i\varepsilon} \frac{i}{p_3^2+i\varepsilon} \frac{i}{p_4^2+i\varepsilon}$$

Now we can do the x and y integrals, which produce $\delta^{(4)}(p_1+p_3+p_4)$ and $\delta^{(4)}(p_2+p_3+p_4)$. This says that momentum is conserved at each vertex in the Feynman diagram. If we integrate over p_3 using the first δ-function, $p_3 = -p_1 - p_4$ the second δ function becomes $\delta^{(4)}(p_1+p_2)$. Then we have, relabelling $p_4 = k$:

$$M = -\lambda^2 \int \frac{d^4k}{(2\pi)^4} \int \frac{d^4p_1}{(2\pi)^4} \int \frac{d^4p_2}{(2\pi)^4} e^{ip_1x_1} e^{ip_2x_2} \frac{i}{p_1^2+i\varepsilon} \frac{i}{p_2^2+i\varepsilon} \frac{i}{(p_1+k)^2+i\varepsilon} \frac{i}{k^2+i\varepsilon} \delta^{(4)}(p_1+p_2)$$

Next, remember that we are interested in an S-matrix element. From the LSZ theorem,

$$\langle f|S|i\rangle = \Big[i \int d^4x_1 \, e^{-ip_ix_1} (\Box_1^2 - m^2) \Big] \Big[i \int d^4x_2 e^{ip_fx_2} (\Box_2^2 - m^2) \Big] \langle \Omega | T\{\phi(x_1)\phi(x_2)\} | \Omega \rangle \tag{8.63}$$

$$= \lambda^2 \int d^4x_1 \, e^{-ip_ix_1} (-p_i)^2 \int d^4x_2 e^{ip_fx_2} (-p_f^2) \Big] M \tag{8.64}$$

Now we note that the x_1 integral gives $\delta^4(p_1 - p_i)$ and the x_2 integral gives a $\delta^4(p_2 + p_f)$. So we can now do the p_1 and p_2 integrals, giving

$$\langle f|S|i\rangle = \lambda^2 \int d^4k \frac{i}{(p_i+k)^2+i\varepsilon} \frac{i}{k^2+i\varepsilon} \delta^{(4)}(p_i - p_f) \tag{8.65}$$

Note how the two propagator factors in the beginning get canceled. This always happens for external legs – remember the point of LSZ was to force the external lines to be on-shell.

This integral is obviously infinite. We will spend a large part of the remainder of the course trying to make sense out of these things. Finally, the $\delta^4(p_1 - p_2)$ term in the answer forces overall momentum conservation, and will always be present in any calculation. But we will always factor it out, like we did when we related differential scattering amplitudes to S-matrix elements.

We can summarize this procedure with the Feynman rules

- Internal lines get propagators $\frac{i}{p^2 - m^2 + i\varepsilon}$
- Vertices come from interactions in the Lagrangian. They get factors of the coupling constant times i from the $\exp(iV)$.
- External lines don't get propagators, because these are canceled by terms from the LSZ reduction formula.
- Momentum is conserved at each vertex
- Integrate over undetermined momenta.

8.3.1 symmetry factors

We haven't kept track of overall normalization. There are lots of factors of $2, 3, 4, \cdots$ that show up, which have to be carefully kept track of. In practice, we would draw each unique diagram, then try to figure out the appropriate symmetry factor. In the end, the rules are quite simple.

- There is a factor of $\frac{1}{m!}$ from the expansion of $\exp(iV) = \sum \frac{1}{m!}(iV)^m$. But if we expand to order m there will be m identical vertices in the same diagram. We can also swap these vertices around, leaving the diagram looking the same. So if we calculate the diagram once, the $m!$ from permuting the diagrams will cancel the $\frac{1}{m!}$ from the exponential. Thus, we can forget about this part.

- Consider a $\phi^n(x)$ interaction. It leads to a vertex at the point x with n lines sticking out. When we connect those lines up to the rest of the diagram, we get $D_F(x,\, y_i)^n$. The $y_i's$ can also be x's, but first lets suppose they're not. If we swap around which line connects to which y_i, we will get the same expression, corresponding to a different contraction with the same Feynman diagram. So we can calculate the Feynman diagram once and just multiply by factor of $n!$ for every vertex with n lines sticking out. Instead of doing this, it is conventional to rescale the vertex by a factor of $n!$

$$\frac{\lambda}{4!}\phi^4, \quad \frac{g}{3!}\phi^3, \quad ... \tag{8.66}$$

- Now suppose some of the $y_i's$ are $x's$. For example, if we have $\phi(x)^2 \to D_F(x,x)^2$. This looks like a bubble starting at x and ending at x. In this case, there is only one contraction, so there are fewer diagrams than we thought and we have to divide by 2! to compensate. Similarly, if we had $\phi(x)^2\phi(y)^2$, the diagram with two lines connecting x and y corresponds to only one contraction, $D_F(x,\, y)^2$, so we need to put back a 2! here too. In general, we need to put back a factor corresponding to the number of geometric symmetries of the diagram. See chapter 4 of Peskin and Schroeder for some more examples.

- Identical particles. Identical particles are already taken care of in Wick's theorem. Moving around the a_p's and $a_p^\dagger$'s has the algebra of identical particles in them. The only time it matters is when the identical particles are external particles, that is, when we can't distinguish the particles we are scattering. And then it's only for final states, since we distinguish our initial states by the setup of the experiment. Thus when n of the same particles are produced, we have to divide the *cross section* by $n!$.

8.3.2 signs of momenta

There is a standard convention about how to choose the direction that the momenta are going. For external momenta it makes sense to assign them their physical values, which should have positive energy. Then momentum conservation becomes

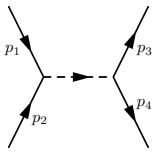
$$\sum p_i = \sum p_f \tag{8.67}$$

which appears in a δ-function as

$$\delta^{(4)}(p_i - p_f) \tag{8.68}$$

This is not the same as $\delta^4(\Sigma p)$ which we sometimes write as shorthand, but would obviously require some of the momenta to have negative energy

For internal lines, we integrate over the momenta, so the result is symmetry to $k_\mu \to -\, k_\mu$. But it is important to keep track of which way the momentum is going so that all the δ-functions at the vertices are also $\Sigma(p_{\rm in} - p_{\rm out})$. So we draw arrows on the lines to indicate this



Actually, the arrows will eventually come to be associated with the direction the charge is flowing, and then we will sometimes add additional arrows to refer to just the momentum. But for now, let's stick to the arrows representing the direction momentum flow.

Also, sometimes Feynman diagrams are drawn with time going upwards. I prefer to have time going to the right, but again, this is just a matter of taste. You could have time going diagonally northeast and the diagram would have the same meaning.

8.4 Vacuum diagrams

A lot of the contractions will result in diagrams where the external lines just connect to each other, and all the non-trivial stuff appears in bubbles. There are also diagrams representing the time ordered product of nothing $\langle\Omega|\Omega\rangle$ which are just the bubbles. It is not hard to relate these two, then we can factor out the vacuum fluctuations $\langle\Omega|\Omega\rangle$ and just concentrate on the interesting physics.

Call the internal points x_i and the external points y_i. The bubbles are just functions of the x_i, for example, some bubbles would give

$$f(x_i) = D_F(x_i, x_j)\cdots D_F(x_k, x_l) \tag{8.69}$$

These factor out completely from any Feynman diagram

$$\mathcal{M} = f_n(x_j)g(y_i) \tag{8.70}$$

So if we sum over all possible bubbles for a fixed way of connecting the external lines we get

$$\sum \mathcal{M} = [\sum_n f_n(x_i)]g(y_i) \tag{8.71}$$

But this sum over bubbles is exactly $\langle\Omega|\Omega\rangle$. So the thing we really want to calculate is

$$\frac{\langle\Omega|\phi(y_1)\cdots\phi(y_i)|\Omega\rangle}{\langle\Omega|\Omega\rangle} = \langle 0|T\{\phi(y_1)\cdots\phi_I(y_n)\}|0\rangle_{\text{no bubbles}} \tag{8.72}$$

This gets rid of all the bubbles.

Now remember how we defined the transfer matrix by

$$\mathcal{S} = \mathbb{1} + i\mathcal{T} \tag{8.73}$$

with the M-matrix, which is the thing we're ultimately trying to calculate, is defined by

$$\mathcal{T} = \delta^4(\Sigma p)\mathcal{M} \tag{8.74}$$

What contributes to the $\mathbb{1}$? Well, anything in which the final states and initial states are identical will contribute to the $\mathbb{1}$. This includes not only the tree level graph, at 0th order in perturbation theory, where the initial states just move along into final states. But also modifications of these graphs where a bubble comes out of one of these trivial lines.

The $\mathcal{M}$ matrix only describes the interesting physics, where some scattering actually happens. Thus when we calculate $\mathcal{M}$ we only have to evaluate diagrams in which all the initial states and final states are connected together, where we can get from any of the external $y_i's$ to any of the others by moving along lines in the graph. Thus

$$\mathcal{M} = \langle 0|T\{\phi(y_1)\cdots\phi_I(y_n)\}|0\rangle_{\text{fully connected}} \tag{8.75}$$

These are the only graphs that have interesting physics in them. Everything else is factored out and normalized away as "nothing happens".

8.5 Tree-level ϕ scattering

Now let's do an example in ϕ^3 theory. Take the Lagrangian to be

$$\mathcal{L} = -\frac{1}{2}\phi\Box\phi - \frac{1}{2}m^2\phi^2 + \frac{g}{3!}\phi^3 \tag{8.76}$$

We want to calculate

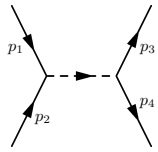
$$\frac{d\sigma}{d\Omega}(\phi\phi \to \phi\phi) = \frac{1}{64\pi^2 E_{\text{cm}}^2}|\mathcal{M}|^2 \tag{8.77}$$

where remember

$$\mathcal{S} = \mathbb{1} + (2\pi)^4 \delta^4(\Sigma p_i) i\mathcal{M} \tag{8.78}$$

If the incoming momenta are p_1 and p_2 and the outgoing momenta are p_3 and p_4. There are 3 diagrams.

The first

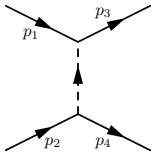


gives

$$i\mathcal{M}_1 = (ig)\Pi(p_1+p_2)(ig) = (ig)\frac{i}{(p_1+p_2)^2 - m^2 + i\varepsilon}(ig) = \frac{ig^2}{s - m^2 + i\varepsilon} \tag{8.79}$$

where $s = (p_1+p_2)^2$.

The second

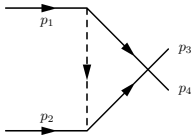


turns into

$$i\mathcal{M}_2 = (ig)\frac{i}{(p_1+p_3)^2 - m^2 + i\varepsilon}(ig) = \frac{ig^2}{t - m^2 + i\varepsilon} \tag{8.80}$$

where $t = (p_1 - p_3)^2$.

The final diagram is



This evaluates to

$$i\mathcal{M}_3 = (ig)\frac{i}{(p_1+p_3)^2 - m^2 + i\varepsilon}(ig) = \frac{ig^2}{u - m^2 + i\varepsilon} \tag{8.81}$$

where $u = (p_1 - p_4)^2$. So the sum is

$$\frac{d\sigma}{d\Omega}(\phi\phi \to \phi\phi) = \frac{g^4}{64\pi^2 E_{\text{cm}}^2}\left[\frac{1}{s - m^2} + \frac{1}{t - m^2} + \frac{1}{u - m^2}\right]^2 \tag{8.82}$$

We have dropped the $i\varepsilon$, which is fine as long as s, t, u are not equal to m^2. For that to happen, the intermediate scalar would have to go on-shell in one of the diagrams, which is a degenerate situation, usually contributing only to $\mathbb{1}$ in the S-matrix. The $i\varepsilon$'s will be necessary for loops, but in tree-level diagrams you can pretty much ignore them.

8.5.1 Mandelstam variables

These variables s, t, u are called Mandelstam variables. They are a great shorthand, used almost exclusively in $2 \to 2$ scattering and in $1 \to 3$ decays. For $2 \to 2$ scattering, with initial momenta p_1 and p_2 and final momenta p_3 and p_4, they are defined by

$$s = (p_1+p_2)^2 = (p_3+p_4)^2 \tag{8.83}$$

$$t = (p_1-p_3)^2 = (p_2-p_4)^2 \tag{8.84}$$

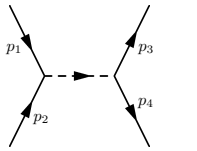
$$u = (p_1-p_4)^2 = (p_2-p_3)^2 \tag{8.85}$$

These satisfy

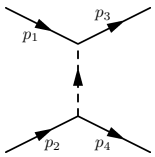
$$s+t+u=\sum m_i^2 \tag{8.86}$$

where m_i are the invariant masses of the particles.

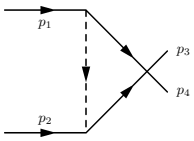
As we say in the previous example, s,t and u correspond to particular diagrams where the propagator has invariant mass s, t or u. So we say s-channel for annihilation diagrams. In these the intermediate state has $p_\mu^2=s>0$. The t- and u- channels to scattering diagrams



$s-$channel



$t-$channel



$u-$channel

s,t and u are great because they are Lorentz invariant. So we compute $\mathcal{M}^2(s,t,u)$ in the CM frame, and then we can easily find out what it is in any other frame, for example the frame of the lab in which we are doing the experiment. So you see s,t,u can be very useful and we will use them a lot!

8.6 With derivative couplings

Suppose we have an interaction with derivatives in it, like

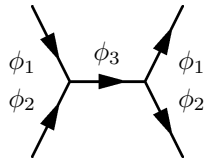
$$\mathcal{L}=\lambda\phi_1(\partial_\mu\phi_2)(\partial_\mu\phi_3) \tag{8.87}$$

where I have included 3 different scalar fields for clarity. In momentum space, these ∂_μ's give factors of momenta. But now remember that

$$\phi(x)=\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\Big[a_p^\dagger e^{ipx}+a_p e^{-ipx}\Big] \tag{8.88}$$

So if the particle is being created, that is emerging from a vertex, it gets a factor of ip_μ, and if it's being destroyed, that is entering from a vertex, it gets a factor if $-ip_\mu$. So a $-$ for incoming momentum and a $+$ for outgoing momentum. In this case, it's quite important to keep track of whether momentum is flowing into or out of the vertex.

For example, take the diagram



Label the initial momenta p_1^μ and p_2^μ and the final momenta are $p_1^{\mu\prime}$ and $p_2^{\mu\prime}$. The exchange momenta is $k^\mu = p_1^\mu + p_2^\mu = p_1' + p_2'$. The diagram gives

$$i\mathcal{M} = (-i\lambda)^2(-ip_2^\mu)(ik^\mu)\frac{i}{k^2}\big(ip_2^{\nu\prime}\big)(-ik^\nu) \tag{8.89}$$

$$= -i\lambda^2\frac{[p_2p_1 + (p_2)^2][p_2'p_1' + (p_2')^2]}{(p_1+p_2)^2} \tag{8.90}$$

As a cross check, suppose we integrated this term by parts

$$\mathcal{L} = -\lambda\phi_3[(\partial_\mu\phi_1)(\partial_\mu\phi_2) + \phi_1\Box\phi_2] \tag{8.91}$$

Now our 1 diagram becomes 4 diagrams, from the 2 types of vertices on the two sides. But we can just add up the contributions to the vertices before multiplying, which gives

$$\mathcal{M} = (i\lambda)^2\Big[(-ip_2^\mu)(-ip_1^\mu) + (-ip_2)^2\Big]\frac{i}{k^2}\Big[\big(ip_2^{\nu\prime}\big)\big(ip_1'^{\nu}\big) + \big(ip_2'\big)^2\Big] \tag{8.92}$$

$$= -i\lambda^2\frac{[p_2p_1 + (p_2)^2][p_2'p_1' + (p_2')^2]}{(p_1+p_2)^2} \tag{8.93}$$

which is exactly what we had above! So integrating by parts doesn't effect the matrix elements.

To see that in general, we just have to show that total derivatives don't contribute to matrix elements. Suppose we have a term

$$\mathcal{L} = \partial_\mu(\phi_1\cdots\phi_n) \tag{8.94}$$

where there are any number of fields in this term. This would give a contribution from the derivative acting on each field, which gives a contribution that fields momenta. So if the vertex would have given V without the derivative, adding the derivative makes it

$$\Big(\sum_{\text{incoming}} p_\mu^i - \sum_{\text{outgoing}} p_\mu^j\Big)V \tag{8.95}$$

but the sum of incoming momenta is equal to the sum of outgoing momenta, because momentum is conserved at the vertex. Therefore, total derivatives don't have an effect.

To be precise, total derivatives don't have an effect in perturbation theory. It turns out a term like

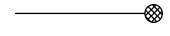
$$\tilde{F}F \equiv \varepsilon_{\mu\nu\alpha\beta}F_{\mu\nu}F_{\alpha\beta} = \partial_\mu(\varepsilon_{\mu\nu\alpha\beta}A_\alpha\partial_\beta A_\nu) \tag{8.96}$$

is a total derivative. Thus if we add a term $\theta\tilde{F}F$ to the Lagrangian, nothing happens. That is, nothing happens in perturbation theory. It turns out that there are effects of this term that will never show up in Feynman diagrams, but are perfectly real. They have physical consequences. For example, this term leads to the violation of time-reversal invariance (the strong CP problem), and a term like it may be responsible for the existence of matter over antimatter in the universe. These kinds of effects are related to pseudo-particles called instantons and sphalerons and are a purely non-perturbative phenomenon.

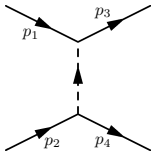
8.7 Summary

We saw that in classical field theory, there is way to solve the equations of motion using greens functions. For example, in the presence of a classical source, this led to Feynman diagrams of the form for the Coulomb potential

$$\phi(r) \sim A_0 \sim \frac{e^2}{r} \sim -\frac{e^2}{k^2}J_0 \sim$$



Now we have seen the same amplitude reproduced with the "real" Feynman diagrams



$$= \frac{e^2}{(p_3 - p_1)^2} = \frac{e^2}{k^2}$$

If fact, if you really think about it, what we showed was that

$$d\sigma \sim \mathcal{M}^2 \sim \frac{e^4}{k^4} \tag{8.97}$$

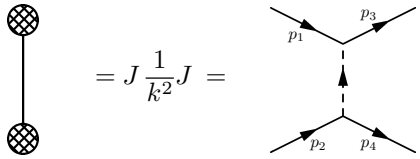
and then we used the Born approximation from non-relativistic quantum mechanics

$$d\sigma \sim \tilde{V}(k)^2 \tag{8.98}$$

to extract $V(k) = \frac{e^2}{k^2}$. But it's the same thing – Coulomb's law is coming from a Lagrangian

$$\mathcal{L} = -\frac{1}{2} A \Box A + A J \tag{8.99}$$

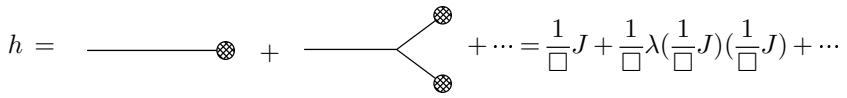
For the classical case, we just took $J_\mu = \delta_{\mu 0}\delta^{(3)}(x)$, to simulate a point charge. For the Feynman diagram case, we replaced J by some interaction $J = e\phi A\phi$, so that the point charge is now thought of like a scalar field, which is the non-relativistic version of an electron. Diagrammatically



I think it is helpful also to see more directly why the Feynman rules and the classical perturbation theory are the same. Starting with a Lagrangian with a cubic interaction

$$\mathcal{L} = h \Box h + \lambda h^3 + h J \tag{8.100}$$

In classical field theory, we found a perturbation expansion that looked like



where $\frac{1}{\Box} \equiv G(x)$ is shorthand for the Green's function which solves the classical equation

$$\Box G(x) = \delta(x) \tag{8.101}$$

It is also sometimes called a Kernel. The Feynman rules are the same thing, but the Green's function is replaced by the Feynman propagator

$$D_F(x,0) = \langle 0 | T\{\phi(x)\phi(0)\} | 0 \rangle = \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 + i\varepsilon} e^{ikx} \tag{8.102}$$

But this is also a Green's function, satisfying

$$\Box D_F(x,0) = \delta(x) \tag{8.103}$$

But the time ordering makes it a Green's function for scattering processes. Both Green's functions are essentially just $\frac{1}{k^2}$ – don't let the $i\varepsilon$ distract you from this essential point.

What are these Green's functions? In the classical case, solving $\Box G(x) = \delta(x)$ gives you $G(x)$ which is the size of a field if you add a source it at $x = 0$. If the source is moving around, the movement propagates at the speed of light (not faster!), since $\Box$ is Lorentz invariant. So the Green's function tells you how this movement propagates. In the quantum case, we saw that the field $\phi(x)$ creates a particle at position x

$$\langle \phi(x)|p\rangle = e^{ipx} \tag{8.104}$$

Thus $\langle 0|\phi(x)\phi(0)|0\rangle$ tells you the amplitude for a particle created at 0 to get to x. That's the same thing as what the Green's function is telling you. So both of these things are propagators. Obviously the quantum one can interfere with itself and pair produce things and do all kinds of fun quantum mechanical stuff that the classical field can't do. But in the classical limit, they are the same, and we can see that immediately in the equations of motion.

We also saw that the same amplitude comes from old-fashioned perturbation theory (OFPT), where the sum over intermediate states includes different time orderings

$$T_{if} = \frac{e^2}{E_i - E_n^{(1)}} + \frac{e^2}{E_i - E_n^{(2)}} = \frac{2e^2 E_\gamma}{k^2} \tag{8.105}$$

Now $2E_\gamma$ goes into the normalization of the fields, in the $\frac{1}{\sqrt{2\omega_p}}$ factor, so this transition matrix element, T_{if} is the same thing as our Lorentz-invariant matrix element $\mathcal{M}$, but calculated a third way, using time-independent perturbation theory from quantum mechanics. So there's lots of ways to do the same calculation.

The Feynman rules have the advantages over OFPT that

- They are manifestly Lorentz invariant – useful in relativistic settings
- 4-momentum is conserved at each vertex, as opposed to just 3-momentum. This forces us to sacrifice having intermediate particles be on-shell. But that's ok, since intermediate particles are not observable – thinking off-shell is now thought of as a great conceptual advance.
- They are systematic for higher order corrections, in which virtual particles are pair created. As we saw with Oppenheimer's attempt at the Lamb shift. As far as I know, nobody has ever been able to sort out loop corrections (renormalization) in OFPT

The Feynman rules have the advantage over classical field theory that they are quantum mechanical!

Finally, let's break down the steps that led to the Feynman rules. We started with wanting to calculate cross sections for scattering processes. These were given by some kinematic factors times S-matrix elements

$$d\sigma \sim \text{kinematics} \times \langle f|S|i\rangle^2 \tag{8.106}$$

We factored out the uninteresting part of S-matrix

$$1 - S = i\delta^4(\Sigma p)\mathcal{M} \tag{8.107}$$

$$\Rightarrow \quad d\sigma \sim \text{kinematics} \times \mathcal{M}^2 \tag{8.108}$$

Then we derived the LSZ theorem which said to add some phases and project out the poles of the external lines and to write the matrix element as a time ordered product

$$\mathcal{M} = p_i^2 e^{ip_i x_1} \cdots p_f^2 e^{-ip_f^2} \times \langle 0|T\{\phi \cdots \phi\}|0\rangle \tag{8.109}$$

Then the time ordered product could be expressed in perturbation theory in terms of interactions and free particle propagators

$$\langle 0|T\{\phi \cdots \phi\}|0\rangle = \frac{1}{p_i^2} \cdots \frac{1}{p_f^2} \left(\lambda^n \frac{1}{k_1^2} \lambda^m \frac{1}{k_2^2} \cdots \right) \tag{8.110}$$

The propagators from the external legs canceled the p^2 from LSZ and the phases give delta functions, enforcing momentum conservation. In the end only the interesting stuff contributes to the cross section

$$d\sigma \sim \text{kinematics} \times \left(\lambda^n \frac{1}{k_1^2} \lambda^m \frac{1}{k_2^2} \cdots \right) \tag{8.111}$$

Keep in mind that all the messy intermediate steps in getting to the Feynman rules took 30 years, from 1930-1960 to be worked out. So although the final seems simple, and you might even have guessed it from thinking about classical theory, historically there was a *lot* that went into it. The Feynman rules are *fantastic* because they are not only simple and intuitive (if you know classical field theory), but also exact and systematic.

Chapter 9
Gauge Invariance

9.1 Introduction

Up until now, we have dealt with general features of quantum field theories. For example, we have seen how to calculate scattering amplitudes starting from a general Lagrangian. Now we will begin to explore what the Lagrangian of the real world could possibly be. Then, of course, we will discuss what it actually is, or at least what we have figured out about it so far.

A good way to start understanding the consistency requirements of the physical universe is with a discussion of spin. Spin should be familiar from quantum mechanics. Spin is a quantum number, labelled by J, which refers to representations of the rotation group SO(3). Fermions have half integer spin $J = \frac{1}{2}, \frac{3}{2}, \cdots$ and bosons have integer spin $J = 0, 1, 2, 3, \cdots$.

Spin 1 particles are critical to quantum electrodynamics, because the photon is spin 1. But they are also essential to understanding quantum field theory in general. I cannot emphasize enough how important spin 1 fields are – it has hard to do any calculation or study any theoretical question in quantum field theory without having to deal with spin 1. In order to understand spin 1, we first have to understand some of the requirements of a consistent quantum field theory.

9.2 Unitary representations of the Poincare group

In quantum field theory, we are interested particles, which transform covariantly under translations and Lorentz transformations. That is, they should form representations of the Poincare group. This means, there is some way to write the Poincare transformation so that

$$|\psi\rangle \to \mathcal{P}|\psi\rangle \tag{9.1}$$

More explicitly, there is some kind of basis we can choose for our states $|\psi\rangle$, call it $|\psi_i\rangle$ so that

$$|\psi_i\rangle \to \mathcal{P}_{ij}|\psi_j\rangle \tag{9.2}$$

So when we say *representation* we mean a way to write the transformation so that when it acts the basis closes in to itself.

In addition, we want *unitary* representations. The reason for this is that the things we compute in field theory are matrix elements

$$\mathcal{M} = \langle\psi_1|\psi_2\rangle \tag{9.3}$$

which should be Poincare *invariant*. If $\mathcal{M}$ is Poincare invariant, and $|\psi_1\rangle$ and $|\psi_2\rangle$ transform covariantly under a Poincare transformation $\mathcal{P}$, we find

$$\mathcal{M} = \langle\psi_1|\mathcal{P}^\dagger\mathcal{P}|\psi_2\rangle \tag{9.4}$$

So we need $\mathcal{P}^\dagger\mathcal{P} = 1$, which is the definition of unitarity. There are many more representations of the Poincare group than the unitary ones, but the unitary ones are the only ones we'll be able to compute Poincare invariant matrix elements from.

As an aside, it is worth pointing out here that there is an even stronger requirement on physical theories: the S matrix must be unitary. Recall that the S-matrix is defined by

$$|f\rangle = S|i\rangle \tag{9.5}$$

Say we start with some normalized state $\langle i|i\rangle = 1$, meaning that the probability for finding anything at time $t = -\infty$ is $P = \langle i|i\rangle^2 = 1$. Then $\langle f|f\rangle = \langle i|S^\dagger S|i\rangle$. So we better have $S^\dagger S = 1$, or else we end up at $t = +\infty$ with less than we started! Thus unitarity of the S-matrix is equivalent to conservation of probability, which seems to be a property of our universe as far as anyone can tell. We'll see consequences of unitary of S later on, but for now let's stick to the discussion of spin.

The unitary representations of the Poincare group were classified for the first time by Eugene Wigner in 1939. As you might imagine, before that people didn't really know what the rules were for constructing physical theories, and by trial and error they were coming across all kinds of problems. Wigner showed that unitary representations are uniquely classified by mass m and spin j. m can take any value and spin is spin of 3D rotations. I'm not going to go through the proof here, but it is done in great detail in Weinberg's book.

The rules are that for a general mass $m \neq 0$, a field of spin j has $2j+1$ polarizations, labeled by $\sigma = -j, -j+1, \cdots, j$ (you can think of $\sigma = j_z$ if you like). So a massive field of spin 0 has one state, spin $\frac{1}{2}$ has two, spin 1 has three, and so on. For massless states, all spins have only two polarizations: $\sigma = \pm j$ (massless spin 0 has only one polarization since $j = -j$). It's not hard to prove this using group theory, but we will content ourselves with understanding it using fields and Lagrangians. Again, see Weinberg's book for details.

The whole trick in constructing a local Lorentz invariant field theory is to embed these spin states into objects with spacetime indices. That is we want to squeeze states of spin $0, \frac{1}{2}, 1, \frac{3}{2}, 2$ etc into scalars, vectors, tensors, and spinors. That way we can write down simple looking Lagrangians and develop general methods for doing calculations. Then we end up with a complication: tensors have 0, 4, 16, 64, $\cdots$, 4^n elements (forgetting half integer spin for now), but spin states have 0, 3, 5, 7, $\cdots$, $2j+1$ physical degrees of freedom. The embedding of the $2j+1$ states in the 4^n dimensional tensors causes all the trouble. But in the end, nobody has figured out a better way to do things (although recent excitement about twistors indicates that there may be an alternative).

9.2.1 a simple example

We don't need all that fancy language to see the problem. Say we were naive and guessed that a spin 1 state should have a Lagrangian like

$$\mathcal{L} = V_\mu \Box V_\mu \tag{9.6}$$

The free field solutions are simply plane waves

$$V_\mu = \varepsilon_\mu e^{ikx} \tag{9.7}$$

where for polarization ε_μ, any 4-component vector is allowed. We can just declare that ε_μ transforms like a vector, then $|V_\mu|^2 = \varepsilon_\mu^2$ is Lorentz invariant. However, if we try to make particles this way, we find their norms are

$$\langle V_\mu|V_\mu\rangle = \varepsilon_\mu^2 = \varepsilon_0^2 - \varepsilon_1^2 - \varepsilon_2^2 - \varepsilon_3^2 \tag{9.8}$$

The norm is Lorentz invariant, but not positive definite!

We can't get anywhere with a theory without a positive definite inner product (norm). For example, if I have a particle with $\varepsilon = \varepsilon_0 = (1, 0, 0, 0)$, the total probability is just $|\varepsilon^2|^2 = 1$. If I instead took a state with polarization $\varepsilon = \varepsilon_1 = (0, 1, 0, 0)$ the total probability would still be $|\varepsilon^2|^2 = 1$. However if I consider them at the same time I would have a state $|\psi\rangle = |\varepsilon_0\rangle \oplus |\varepsilon_1\rangle$ (a two particle state), whose norm is

$$\langle\psi|\psi\rangle = \langle\varepsilon_0|\varepsilon_0\rangle + \langle\varepsilon_1|\varepsilon_0\rangle + \langle\varepsilon_0|\varepsilon_1\rangle + \langle\varepsilon_1|\varepsilon_1\rangle = 1 - 1 = 0 \tag{9.9}$$

So the probability of finding anything is 0. That doesn't make any sense, since I was suppose to have two particles! I can even separate the particles far away in the universe. Still, I find my answer depends on whether I consider the particles together or separately. Weinberg uses the term cluster decomposition to describe these kinds of considerations. So having the Lagrangian $V_\mu \Box V_\mu$ violates cluster decomposition.

What went wrong in this example? The vector V_μ has 4 degrees of freedom, which is three for spin 1 and one for spin 0 (it is a *reducible* representation). We found that when we try to put both of these spins in the same field V_μ negative norm states pop out. So here's where the fancy language helps. Under unitary transformations, the norm can never change. So if we can make sure each field only describes particles of a single spin (a *irreducible* representation), we can give the particle positive norm states and we will never have a problem. Thus the challenge is to embed irreducible unitary representations of the Poincare group (characterized by mass and spin) into fields which transform under general representations of the Lorentz group, like $V_\mu(x)$.

9.3 Embedding particles into fields

Let's now explore how we can construct Lagrangians for fields which contain only particles of single spins.

9.3.1 spin 0

For spin-0, the embedding is easy, we just put the 1 degree of freedom into a scalar ϕ.

9.3.2 massive spin 1

For spin 1, there are three degrees of freedom if it's massive or two if it's massless (we'll derive this shortly). So the smallest tensor field we could possible embed these in is a vector field A_μ which has 4-components. The challenge is to make it so that not all of those degrees of freedom are active.

Let's start with massive spin one. If we had the Lagrangian

$$\mathcal{L} = \frac{1}{2} A_\mu \Box A_\mu + \frac{1}{2} m^2 A_\mu^2 \tag{9.10}$$

then the equations of motion are

$$(\Box + m^2) A_\mu = 0 \tag{9.11}$$

So we have four propagating modes. This is the three we want, for massive spin-1, plus another, which turns out to be spin 0. To see that, note that some of these $A'_\mu s$ can be written as

$$A_\mu = \partial_\mu \phi \tag{9.12}$$

So that the Lorentz transformation properties simply emerge from those of ∂_μ, but the actual degree of freedom is a spin-0 scalar ϕ. So we would like to isolate the *other* degrees of freedom.

There is a nice way to do that. If we take the Lagrangian to be

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}^2 + m^2 A_\mu^2 \tag{9.13}$$

We find the equations of motion are

$$(\Box + m^2) A_\mu - \partial_\mu (\partial_\nu A_\nu) = 0 \tag{9.14}$$

Acting on this with ∂_μ we find that A_μ satisfies a constraint

$$m^2 (\partial_\mu A_\mu) = 0 \tag{9.15}$$

So that one degree of freedom is projected out. Thus there only 3 degrees of freedom in a vector field starting from this Lagrangian. These are the 3 degrees of freedom of a spin-1 field. The equations of motion also simplify to

$$(\Box + m^2) A_\mu = 0 \tag{9.16}$$

In Fourier space, we can write

$$A_\mu(x) = \int \frac{d^4k}{(2\pi)^4}\varepsilon_\mu(k)e^{ikx} \tag{9.17}$$

where ε_μ is called a polarization. By the equations of motion, $k^2 = m^2$ and $k_\mu\varepsilon_\mu = 0$. These polarization vectors ε_μ are conventionally normalized so that $\varepsilon_\mu^2 = 1$. We can also choose a canonical basis. Take the momentum of the field to be in the z-direction

$$k_\mu = (E, 0, 0, p), \quad E^2 - p^2 = m^2 \tag{9.18}$$

then two obvious vectors satisfying $k_\mu\varepsilon_\mu = 0$ and $\varepsilon_\mu^2 = 1$ are

$$\epsilon_\mu^+ = (0, 1, 0, 0), \quad \epsilon_\mu^- = (0, 0, 1, 0) \tag{9.19}$$

the other one is

$$\epsilon_\mu^L = (\frac{p}{m}, 0, 0, \frac{E}{m}) \tag{9.20}$$

It's easy to check that $(\epsilon_\mu^L)^2 = 1$ and $k_\mu\epsilon_\mu^L = 0$.

By the way, massive spin 1 fields are not a purely theoretical concept: they exist! There is one called the ρ-meson, that is lighter than the proton, but unstable, so we don't often see it. More importantly, there are really heavy ones, the W and Z bosons, which mediate the weak force and radioactivity. You will study them in great detail in a standard model class. But there's an important feature of these heavy bosons that is easy to see already. Note that at high energy $E \gg m$, the longitudinal polarization becomes

$$\epsilon_\mu^L \sim \frac{E}{m}(1, 0, 0, 1) \tag{9.21}$$

So if we scatter these modes, we might have a cross section that goes like $d\sigma \sim \varepsilon_\mu^2 \sim g^2\frac{E^2}{m^2}$, where g is the coupling constant. So, no matter how small g is, if we go to high enough energies, this cross section blows up. However, cross sections can't be arbitrarily big. After all, they are probabilities, which are bounded by 1. So at some scale, what we are calculating becomes not a very good representation of what is really going on. In other words, our perturbation theory is breaking down. We can see already that this happens at $E \sim \frac{m}{g}$. If $m \sim 100$ GeV and $g \sim 0.1$, we find $E \sim 1$ TeV. That's why we expect that our theory, with the W or Z bosons of $m_W \sim 100\,\text{GeV}$, should act funny at the LHC. We don't know exactly how to calculate things at that scale. We have some ideas, but can't prove any of them yet, which is why we are building the machine.

By the way, the fact that there is a spin-1 particle in this Lagrangian follows completely from the Lagrangian itself. We never have to impose any spin-oney conditions in addition. We didn't even have to talk about spin, or Lorentz invariance at all – all the properties associated with that spin would just have fallen out when we tried to calculate physical quantities. That's the beauty of symmetries: they work even if you don't know about them! It would be fine to think of A_μ as 4 scalar fields that happen to conspire so that when you compute something in one frame certain ones contribute and when you compute in a different frame, other ones contribute, but the final answer is frame independent. Obviously it's a lot easier to do the calculation if we know this ahead of time, so we can choose a nice frame, but in no way is it required.

9.3.3 massless spin 1

The easiest way to come up with a theory of massless spin 1 is to simply take the $m \to 0$ limit of the massive spin 1 theory. Then the Lagrangian becomes

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 \tag{9.22}$$

which is the Lagrangian for electrodynamics. So this is the right answer.

But we also know that massless spin 1 should have only 2 polarizations. What happened to the third? Well, first of all, note that as $m \to 0$ the longitudinal polarizations blows up:

$$\epsilon_\mu^L = (\frac{p}{m}, 0, 0, \frac{E}{m}) \to \infty \tag{9.23}$$

$$k_\mu \to (E, 0, 0, E) \tag{9.24}$$

Partly, this is due to normalization. Note that we also have $p \to E$ so $\epsilon_\mu^L \to k_\mu$ up to the divergent normalization. We'll come back to this. The other problem is that the constraint equation $m^2(\partial_\mu A_\mu) = 0$ is automatically satisfied for $m = 0$, so we no longer automatically have $\partial_\mu A_\mu = 0$. Thus taking the massless limit isn't so easy.

Instead, let's just postulate the Lagrangian and start over with analyzing the degrees of freedom. So we start with

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}^2, \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{9.25}$$

This has an important property that the massive Lagrangian didn't: gauge invariance. It is invariant under the transformation

$$A_\mu(x) \to A_\mu(x) - \partial_\mu \alpha(x) \tag{9.26}$$

For any function $\alpha(x)$. Thus two fields A_μ which differ by the derivative of a scalar are physically equivalent.

The equations of motion following from the Lagrangian are

$$\Box A_\mu - \partial_\mu(\partial_\nu A_\nu) = 0 \tag{9.27}$$

This is really 4 equations and it is helpful to separate out the 0 and i components:

$$-\partial_j^2 A_0 - \partial_j A_j = 0 \tag{9.28}$$

$$\Box A_i + \partial_i(\partial_t A_0 - \partial_j A_j) = 0 \tag{9.29}$$

To count the physical degrees of freedom, let's choose a gauge. Instead of Lorentz gauge, we will use Coulomb gauge $\partial_j A_j = 0$. Then the A_0 equation of motion becomes

$$\partial_j^2 A_0 = 0 \tag{9.30}$$

which has no time derivative. Now under gauge transformations $\partial_i A_i \to \partial_i A_i + \partial_i^2 \alpha$, so Coulomb gauge is preserved under $A_\mu \to A_\mu + \partial_\mu \alpha$ for any α satisfying $\partial_i^2 \alpha = 0$. Since $A_0 \to A_0 + \partial_t \alpha$ and A_0 also satisfies $\partial_i^2 A_0 = 0$ we have exactly the residual symmetry we need to set $A_0 = 0$. Thus we have eliminated one degree of freedom from A_μ completely, and we are down to three. One more to go!

The other equations are

$$\Box A_i = 0 \tag{9.31}$$

which seem to propagate 3 modes. But don't forget that A_i is constrained by $\partial_i A_i = 0$. In Fourier space

$$A_\mu(x) = \int \frac{d^4k}{(2\pi)^4} \epsilon_\mu(k) e^{ikx} \tag{9.32}$$

And we have $k^2 = 0$ (equations of motion), $k_i \epsilon_i$ (gauge choice), and $\epsilon_0 = 0$ (gauge choice). Choosing a frame, we can write the momentum as $k_\mu = (E, 0, 0, E)$. Then these equations are solved by

$$\epsilon_\mu^1 = (0, 1, 0, 0), \quad \epsilon_\mu^2 = (0, 0, 1, 0) \tag{9.33}$$

Often people use a different basis

$$\epsilon_\mu^L = (0, 1, i, 0), \quad \epsilon_\mu^R = (0, 1, -i, 0) \tag{9.34}$$

which are called helicity eigenstates, and correspond to circularly polarized light. But I have always found it easier to use the ϵ^1 and ϵ^2 basis. In any case we have found that the massless photon has two polarizations.

We could also have used Lorentz gauge, in which case we would have found that three vectors satisfy $k_\mu \epsilon_\mu = 0$

$$\epsilon^1_\mu = (0,1,0,0), \quad \epsilon^2_\mu = (0,0,1,0), \quad \epsilon^f_\mu = (1,0,0,1) \tag{9.35}$$

However, the forward polarized photon ϵ^f_μ is not normalizable, so we have to throw it out. There is actually residual gauge symmetry in Lorentz gauge which can remove this polarization too. If $A_\mu = \partial_\mu \phi$ for some ϕ, then we can shift with $\alpha = \phi$ to remove this state. Again, this is the residual scalar degree of freedom. In the same way, if we had not imposed the second Coulomb gauge condition $\epsilon_0 = 0$, we would have found the other polarization satisfying $k_i \varepsilon_i = 0$ is $\epsilon^0 = (1,0,0,0)$. But this also cannot be normalized so that $\epsilon^i_\mu \epsilon^j_\mu = -\delta^{ij}$, since ϵ^0 is timelike.

9.3.4 summary

To summarize, for massive spin 1, we needed to have the kinetic term $F^2_{\mu\nu} + m^2 A^2_\mu$ in order to enforce $\partial_\mu A_\mu = 0$ which eliminated one degree of freedom from A_μ, leaving the 3 for massive spin 1. The Lagrangian for a massive spin 1 particle does not have gauge invariance, but we still need $F^2_{\mu\nu}$.

For the massless case, having $F^2_{\mu\nu}$ gives us gauge invariance. This allows us to remove an additional polarization, leaving 2 which is the correct number for a massless spin-1 representation of the Poincare group. Sometimes people say that 1 mode is eliminated because A_0 doesn't propagate, and the other mode is eliminated by a gauge transformation. Some people say the two modes are both eliminated by gauge invariance. I don't know if these words make a difference to anything, or even what exactly they mean. The only concrete statement is that the gauge invariant Lagrangian for a vector field propagates two massless spin 1 degrees of freedom.

9.4 Covariant derivatives

In order not to screw up our counting of degrees of freedom, the interactions in the Lagrangian should respect gauge invariance as well. For example, you might try to couple

$$\mathcal{L} = \cdots + A_\mu \phi \partial_\mu \phi \tag{9.36}$$

But this is not invariant.

$$A_\mu \phi \partial_\mu \phi \to A_\mu \phi \partial_\mu \phi + (\partial_\mu \alpha) \phi \partial_\mu \phi \tag{9.37}$$

In fact, it's impossible to couple A_μ to a field with only one degree of freedom, like the scalar field ϕ. We must be able to make ϕ transform to compensate for the gauge transformation of A_μ, to cancel the $\partial_\mu \alpha$ term. But if there is only one field ϕ, it has nothing to mix with so it can't transform.

Thus we need at least two $\phi's$, ϕ_1 and ϕ_2. It's easiest to deal with such a doublet by putting them together into a complex field $\phi = \phi_1 + i\phi_2$, and then to work with ϕ and $\phi^\star$. Under a gauge transformation, ϕ can transform like

$$\phi \to e^{i\alpha(x)} \phi \tag{9.38}$$

Which makes

$$m^2 \phi^\star \phi \tag{9.39}$$

gauge invariant. But what about the derivatives? $|\partial_\mu \phi|^2$ is not invariant.

We can in fact make the kinetic term gauge invariant using covariant derivatives. If a field transforms as

$$\phi \to e^{i\alpha(x)} \phi \tag{9.40}$$

Then

$$(\partial_\mu + i g A_\mu)\phi \to (\partial_\mu + i g A_\mu - i\partial_\mu \alpha) e^{i\alpha(x)}\phi = e^{i\alpha(x)}(\partial_\mu + i g A_\mu)\phi \tag{9.41}$$

So we define the *covariant* derivative by

$$D_\mu \phi \equiv (\partial_\mu + i g A_\mu)\phi \to e^{i\alpha(x)} D_\mu \phi \tag{9.42}$$

which transforms just like the field does. Thus

$$\mathcal{L} = (D_\mu \phi)^\star (D_\mu \phi) + m^2 \phi^\star \phi \tag{9.43}$$

is gauge invariant.

9.5 Conserved Currents

We can expand out the Lagrangian we just found (dropping the mass for simplicity):

$$\mathcal{L} = (D_\mu \phi)^\star (D_\mu \phi) = \partial_\mu \phi^\star \partial_\mu \phi + i g A_\mu \phi^\star \partial_\mu \phi - i g A_\mu (\partial_\mu \phi^\star)\phi + g^2 A_\mu^2 \phi^\star \phi \tag{9.44}$$

$$= \partial_\mu \phi^\star \partial_\mu \phi + A_\mu J_\mu + \mathcal{O}(g^2) \tag{9.45}$$

$$J_\mu = i g[(\partial_\mu \phi^\star)\phi - \phi^\star(\partial_\mu \phi)] \tag{9.46}$$

The equations of motion for the free field ϕ are simply $\Box\phi = 0$, which directly implies $\partial_\mu J_\mu = 0$. So by just demanding this symmetry hold for the interactions, we were automatically led to a conserved current. It turns out this follows from a general result.

Let's just write the thing that the gauge field couples to as J_μ:

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}^2 + A_\mu J_\mu \tag{9.47}$$

then under a gauge transformation the action transforms as

$$\int d^4x\, A_\mu J_\mu \to \int d^4x (A_\mu + \partial_\mu \alpha(x)) J_\mu \tag{9.48}$$

$$= \int d^4x [A_\mu J_\mu - \alpha(x)\partial_\mu J_\mu] \tag{9.49}$$

where we have integrated by parts in the last step. For the action to be invariant for any function $\alpha(x)$ we must have $\partial_\mu J_\mu = 0$, that is the current must be conserved.

We say "conserved" because the total charge

$$Q = \int d^3x\, J_0 \tag{9.50}$$

satisfies

$$\partial_t Q = \int d^3x \partial_t J_0 = \int d^3x \vec{\nabla}\vec{J} \tag{9.51}$$

which is Gauss's Law. If $\vec{J}$ vanishes at the boundary, meaning that no charge is leaving our experiment, then $\partial_t Q = 0$. Thus the total charge doesn't change with time, and is conserved.

9.5.1 Noether's theorem

The existence of conserved currents is quite a general consequence of symmetries in a Lagrangian. The direct connection is embodied in Noether's theorem. It's a beautiful theorem, and like many great theorems, is almost trivial to prove.

Say we have a Lagrangian $\mathcal{L}(\phi_i)$ which is invariant when a bunch of fields ϕ_i transform some way under a symmetry parameter α. We don't even need α to be a function of spacetime for this theorem, it can be a number. But it has to be a continuous symmetry (*i.e.* not $\phi \to -\phi$), so that we can take $\delta\phi_i$ arbitrarily small. Then

$$0 = \delta\mathcal{L} = \frac{\delta\mathcal{L}}{\delta\phi_i}\frac{\delta\phi_i}{\delta\alpha} + \frac{\delta\mathcal{L}}{\delta\partial_\mu\phi_i}\frac{\delta(\partial_\mu\phi_i)}{\delta\alpha} \tag{9.52}$$

Now if we use the equations of motion (the Euler Lagrange equations)

$$\frac{\delta\mathcal{L}}{\delta\phi_i} - \partial_\mu\left(\frac{\delta\mathcal{L}}{\delta\partial_\mu\phi_i}\right) = 0 \tag{9.53}$$

So that

$$0 = \partial_\mu\left(\frac{\delta\mathcal{L}}{\delta\partial_\mu\phi_i}\right)\frac{\delta\phi_i}{\delta\alpha} + \frac{\delta\mathcal{L}}{\delta\partial_\mu\phi_i}\delta\frac{\partial_\mu\phi_i}{\delta\alpha} \tag{9.54}$$

$$= \partial_\mu\left(\frac{\delta\mathcal{L}}{\delta\partial_\mu\phi_i}\frac{\delta\phi_i}{\delta\alpha}\right) \tag{9.55}$$

Thus our conserved current is simply

$$J_\mu = \frac{\delta\mathcal{L}}{\delta\partial_\mu\phi_i}\frac{\delta\phi_i}{\delta\alpha} \tag{9.56}$$

You can make this more general, by allowing for $\mathcal{L}$ not just to depend on ϕ and $\partial_\mu\phi$ but also on higher derivatives, and also by allowing that the symmetry only leaves the Lagrangian invariant up to a total derivative. But those cases are unenlightening so we'll stop here.

In summary

- Noether's Theorem: If the action has a continuous symmetry, there exists a current associated with that symmetry which is conserved on the equations of motion.

To emphasize

- the symmetry must be continuous
- it works for *global* symmetries, parametrized by numbers α not just *local* (gauge) symmetries parametrized by functions $\alpha(x)$
- the current is conserved *on-shell*, that is, when the equations of motion are satisfied
- the theorem itself has nothing to do with gauge fields. but gauge invariance require continuous symmetries, so we often use the theorem in the context of gauge fields.

9.5.2 example

The easiest way to understand this therm is with an example. Take a scalar field, with Lagrangian

$$\mathcal{L} = |\partial_\mu\phi|^2 - m^2|\phi|^2 \tag{9.57}$$

The Lagrangian is invariant under $\phi \to e^{i\alpha}\phi$, whose infinitesimal form is

$$\phi \to \phi + i\alpha\phi, \quad \phi^\star \to \phi^\star - i\alpha\phi^\star \tag{9.58}$$

That is

$$\frac{\delta\phi}{\delta\alpha} = i\phi, \quad \frac{\delta\phi^\star}{\delta\alpha} = -i\phi^\star \tag{9.59}$$

Then,

$$J_\mu = \frac{\delta\mathcal{L}}{\delta\partial_\mu\phi_i}\frac{\delta\phi_i}{\delta\alpha} = i[(\partial_\mu\phi^\star)\phi - \phi^\star\partial_\mu\phi] \tag{9.60}$$

This is the same thing we saw was conserved before, because it's the same thing that couples to the gauge field when the symmetry is gauged. However, the current is conserved even if there is no gauge field. All we need is the *global* symmetry to have a conserved current.

We could derived J_μ. We could have done this before discussing gauge fields. Then we would have written down $A_\mu J_\mu$ by hand to generate the interactions, which would be interpretable as covariant derivatives. This approach is not so important for this symmetry, but is a nice way to think about gravity, described clearly in Feynman's lectures on gravity, which I strongly recommend.

9.5.3 summary

To summarize, we saw that to have a massless unitary representation of spin 1 we needed gauge invariance. This required a conserved current, which in turn required that charge be conserved. To couple the photon to matter, we needed more than one degree of freedom so we were led to ϕ and $\phi^\star$, which are related by symmetry transformations. So there are a lot of implications of demanding unitary representations!

9.6 Quantizing spin 1

To quantize a spin one field, we just allow for the polarizations to be separate degrees of freedom. So as far as the free fields are concerned, it is just 2 or 3 scalar fields. But we have to be a little careful about the Lorentz transformation properties of the polarization vectors.

9.6.1 massive spin 1

$$A_\mu(x,t) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2\omega_k}}\sum_{i=1}^{3}\left(\epsilon_\mu^i(k)a_{k,i}^\dagger e^{ikx} + \epsilon_\mu^{i\star}(k)a_{k,i}e^{-ikx}\right) \tag{9.61}$$

This just means we have separate creation and annihilation operators for each of the polarizations, and we sum over them. $\varepsilon_\mu^i(k)$ represents a canonical set of basis vectors in the frame where $p_\mu = k_\mu$, which we will now explain. $\epsilon_\mu^{i\,\star}$ is the complex conjugate polarization, in case we choose a complex basis.

Let's think a little more about how the polarization vectors change with Lorentz transformations. Obviously any of the 4 degrees of freedom in $\varepsilon_\mu = (\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3)$ will mix if we transform them as

$$\varepsilon_\mu \to \Lambda_{\mu\nu}\varepsilon_\nu \tag{9.62}$$

So what we mean by a representation is that we also transform the momentum vector p_μ, and we have to re-expand our polarization vectors in terms of the rotated basis in the new frame. In that sense, they won't mix.

For the massive case, our previous basis had momentum in the z direction

$$p_\mu = (E, 0, 0, p) \tag{9.63}$$

and so we can write our basis as

$$\epsilon_\mu^+(p) = (0,1,0,0), \quad \epsilon_\mu^-(p) = (\,0,0,1,0), \quad \epsilon_\mu^L(p) = (\frac{p}{m}, 0, 0, \frac{E}{m}) \tag{9.64}$$

where these all satisfy $\epsilon_\mu^2 = 1$ and $\epsilon_\mu p_\mu = 0$.

The fourth degree of freedom in $\varepsilon_\mu^S(p) = \frac{1}{m}p_\mu = (\frac{E}{m}, 0, 0, \frac{p}{m})$ which in position space is $\epsilon_\mu^S = \partial_\mu\alpha(x)$ for some function $\alpha(x)$, which transforms as a scalar. To see that the polarization based on the scalar doesn't mix with the other three is trivial: if something is the divergence of a function $\alpha(x)$, under a Lorentz transformation it will still be the divergence of the same function, just in a different frame! So if we are doing things right, the vectors in the spin-1 representation (the $\epsilon_\mu^i{}'s$) shouldn't mix with the vector in the spin-0 representation ϵ_μ^S.

Now, if our basis vectors are changing as we boost, why haven't we made everything entirely trivial. The answer is that there are Lorentz transformations which leave k_μ alone, and therefore leave our basis alone. For those Lorentz transformation, our basis vectors must rotate into each other and form a complete representation.

For example, suppose we go to the frame

$$q_\mu = (m, 0, 0, 0) \tag{9.65}$$

Then we can choose our polarization basis vectors as

$$\epsilon^1_\mu(q) = (0, 1, 0, 0), \quad \epsilon^2_\mu(q) = (0, 0, 1, 0), \quad \epsilon^3_\mu(q) = (0, 0, 0, 1) \tag{9.66}$$

and $\epsilon^S_\mu = (1, 0, 0, 0)$. Now the group which preserves q_μ is simply the 3D rotation group SO(3). Clearly, the 3 ε^i_μ polarizations will rotate among each other under 3D rotations and ϵ^S_μ stays the same. If we boost it looks like the ϵ^i_μ will mix with ϵ^S_μ. But all bets are off because the basis vectors will also change, for example to $\varepsilon^+_\mu, \varepsilon^-_\mu$ and ε^L_μ above. So it's enough just to fix the 4-momentum k_μ when trying to decide if we have a representation of not.

This method of studying representations of the Lorentz group is called the method of induced representations. The representations are *induced* by representations of the *little group*, which is defined as the group preserving the 4-momentum k_μ.

So now you know what $\epsilon^i_\mu(k)$ means.

9.6.2 massless spin 1

We quantize massless spin one exactly like massive spin 1, but summing over 2 polarizations instead of 3.

$$A_\mu(x,t) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2\omega_k}}\sum_{i=1}^{2}(\epsilon^i_\mu(k)a^\dagger_{k,i}e^{ikx} + \epsilon^i_\mu(k)a_{k,i}e^{-ikx}) \tag{9.67}$$

A sample basis is

$$p_\mu = (E, 0, 0, E) \tag{9.68}$$

$$\epsilon^1_\mu(p) = (0, 1, 0, 0), \quad \epsilon^2_\mu(p) = (0, 0, 1, 0) \tag{9.69}$$

these satisfy $\epsilon^2_\mu = 1$ and $\epsilon_\mu p_\mu = 0$. The bad polarizations are

$$\epsilon^f_\mu(p) = (1, 0, 0, 1), \quad \epsilon^b_\mu(p) = (1, 0, 0, -1) \tag{9.70}$$

where f and b stand for forward and backward.

But now we have a problem. From the massive case, we know that ϵ^1_μ and ϵ^2_μ rotate into each other under Lorentz transformations and also into the longitudinal mode $\epsilon^L_\mu(p) = (\frac{p}{m}, 0, 0, \frac{E}{m})$. As $m \to 0$, this longitudinal mode becomes the same as our forward-polarized photon, up to normalization

$$\lim_{m\to 0}\epsilon^L_\mu(p) = \epsilon^f_\mu(p) = p_\mu \tag{9.71}$$

In fact, you can show that this really happens. For some Lorentz transformations

$$\epsilon^1_\mu(p) \to a_1(\Lambda)\epsilon^1_\mu(p) + a_2(\Lambda)\epsilon^2_\mu(p) + a_3(\Lambda)p_\mu \tag{9.72}$$

$$\epsilon^2_\mu(p) \to b_1(\Lambda)\epsilon^1_\mu(p) + b_2(\Lambda)\epsilon^2_\mu(p) + b_3(\Lambda)p_\mu \tag{9.73}$$

where the $a's$ and $b's$ are numbers depending on the particular boost and rotation angles specified by Λ.

This is EXTREMELY IMPORTANT. We have found that under Lorentz transformations the massless polarizations transform as

$$\epsilon_\mu \to \epsilon_\mu + k_\mu \tag{9.74}$$

When we calculate something in QED we will get matrix elements

$$|\mathcal{M}|^2 = \varepsilon_\mu M_\mu \tag{9.75}$$

for some M_μ transforming like a Lorentz vector. Since we only sum over the 2 physical polarizations, if we Lorentz transform this expression we get

$$|\mathcal{M}|^2 = \varepsilon_\mu M_\mu + k_\mu M_\mu \tag{9.76}$$

Thus the matrix elements can only be Lorentz invariant if

$$\boxed{k_\mu M_\mu = 0} \tag{9.77}$$

This is known as a Ward identity. It must hold by LORENTZ INVARIANCE and the fact that UNITARY representations for MASSLESS PHOTONS have TWO POLARIZATIONS.

That it actually does hold in QED is complicated to show, but we will do it eventually.

It is closely related to gauge invariance. Since the Lagrangian is invariant under $A_\mu \to A_\mu + \partial_\mu \alpha$, in momentum space, this is $A_\mu \to A_\mu + k_\mu$. Since $A_\mu = \varepsilon_\mu e^{ikx}$, this should directly imply that $\varepsilon_\mu \to \varepsilon_\mu + k_\mu$ is a symmetry of the theory, which is the ward identity.

9.7 Photon propagator

In order to calculate anything with a photon, we are going to need to know it's propagator

$$\Pi_A = \langle 0 | T\{A_\mu A_\nu\} | 0 \rangle \tag{9.78}$$

The easiest way to derive it is to just look for the classical Green's function. Then turning it into a time ordered product is exactly the same as for a scalar.

Let us first try to calculate the photon propagator by using the equations of motion, without choosing a gauge. In the presence of a current, the equations of motion are

$$\partial_\mu F_{\mu\nu} = J_\nu \tag{9.79}$$

$$\Rightarrow \partial_\mu \partial_\mu A_\nu - \partial_\mu \partial_\nu A_\mu = J_\nu \tag{9.80}$$

$$\Rightarrow (k^2 \eta_{\mu\nu} - k_\nu k_\mu) A_\mu = J_\nu \tag{9.81}$$

We would like to write $A_\mu = \Pi_{\mu\nu} J_\nu$, so that $(k^2 \eta_{\mu\nu} - k_\mu k_\nu)\Pi_{\nu\alpha} = \delta_{\mu\alpha}$. That is, we want to invert the kinetic term. The problem is that

$$\det(k^2 \eta_{\mu\nu} - k_\alpha k_\mu) = 0 \tag{9.82}$$

We can see this because $k^2 \eta_{\mu\nu} - k_\alpha k_\mu$ has a zero eigenvalue, with eigenvector k_μ. This simply represents the gauge invariance: A_μ is not uniquely determined by J_μ; different gauges will give different values for A_μ.

So what do we do? We could try to just choose a gauge, for example $\partial_\mu A_\mu = 0$. This would reduce the Lagrangian to

$$-\frac{1}{4} F_{\mu\nu} \to \frac{1}{2} A_\mu \Box A_\mu \tag{9.83}$$

However, now it looks like there are 4 propagating degrees of freedom in A_μ instead of 2. In fact, you can do this, but you have to keep track of the gauge constraint $\partial_\mu A_\mu = 0$ all along. This is known as a second-class constraint, and it's a real pain to work with constrained systems. We would need more information than what is contained in the Lagrangian.

The solution is to add a new auxiliary (non-propagating) field which acts like a Lagrange multiplier to enforce the constraint through the equations of motion

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}^2 + \frac{1}{\xi} (\partial_\mu A_\mu)^2 + J_\mu A_\mu \tag{9.84}$$

(Writing $\frac{1}{\xi}$ instead of ξ is just a convention). The equations of motion of ξ are just $\partial_\mu A_\mu = 0$ which was the Lorentz gauge constraint. In fact, we could just treat ξ as a number, and for very small ξ there is a tremendous pressure on the Lagrangian to have $\partial_\mu A_\mu = 0$ to stay near the minimum.

Now, the equations of motion for A_μ are

$$\left[k^2 \eta_{\mu\nu} - (1 - \frac{1}{\xi}) k_\mu k_\nu \right] A_\nu = J_\mu \tag{9.85}$$

Which can now be inverted to

$$\Pi^{\text{classical}}_{\mu\nu} = \frac{\eta_{\mu\nu} - (1-\xi)\frac{k_\mu k_\nu}{k^2}}{k^2} \tag{9.86}$$

(classical refers to this being the classical Green's function, as opposed to the Feynman propagator, which is the time-ordered two-point function). We can check:

$$(k^2\eta_{\mu\alpha} - (1-\frac{1}{\xi})k_\mu k_\alpha)\Pi^{\text{classical}}_{\alpha\nu} = \Big[(k^2\eta_{\mu\alpha} - (1-\frac{1}{\xi})k_\mu k_\alpha) \Big]\big[k^2\eta_{\alpha\nu} - (1-\xi)k_\alpha k_\nu \big]\frac{1}{k^4} \tag{9.87}$$

$$= \eta_{\mu\nu} + \Big[-(1-\frac{1}{\xi}) - (1-\xi) + (1-\frac{1}{\xi})(1-\xi) \Big]\frac{k_\mu k_\nu}{k^2} \tag{9.88}$$

$$= \eta_{\mu\nu} \tag{9.89}$$

To work out the Feynman propagator, the steps are just like the scalar field, and the answer is

$$\langle 0|T\{A_\mu(x)A_\nu(y)\}|0\rangle = \int \frac{d^4k}{(2\pi)^4}\Pi_{\mu\nu} \tag{9.90}$$

$$\Pi_{\mu\nu} = \frac{-i}{k^2 + i\varepsilon}\Big[\eta_{\mu\nu} - (1-\xi)\frac{k_\mu k_\nu}{k^2} \Big] \tag{9.91}$$

Where the $i\varepsilon$ for time ordering, the i comes from the contour integral going from d^3k in A_μ to d^4k and the -1 indicates that it is the spacial A_i components which propagate like scalars, since

$$-i\eta_{\mu\nu} = \begin{pmatrix} -i & & & \\ & i & & \\ & & i & \\ & & & i \end{pmatrix} \tag{9.92}$$

and the scalar propagator was $\Pi_S = \frac{i}{k^2 + i\varepsilon}$.

9.7.1 gauges

For any value of ξ we get a different Lorentz-invariant gauge. Some useful gauges are:

- Feynman gauge $\xi = 1$

$$\Pi_{\mu\nu} = -i\frac{\eta_{\mu\nu}}{k^2 + i\varepsilon} \tag{9.93}$$

  This is the gauge we will use for most calculations

- Lorentz Gauge $\xi \to 0$

$$\Pi_{\mu\nu} = -i\frac{\eta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}}{k^2 + i\varepsilon} \tag{9.94}$$

  We saw that $\xi \to 0$ forces $\partial_\mu A_\mu = 0$. Note that we couldn't set $\xi = 0$ and then invert the kinetic term, but we can invert and then set $\xi = 0$. This gauge isn't the most useful for calculations.

- Unitary gauge $\xi = \infty$. This gauge is useless for QED, since the propagator blows up. But it is extremely useful for the gauge theory of the weak interactions.

9.7.2 gauge invariance

The final answer for any Lorentz invariant quantity had better be gauge invariant. Since

$$\Pi_{\mu\nu} = \frac{-i}{k^2 + i\varepsilon}\Big[\eta_{\mu\nu} - (1-\xi)\frac{k_\mu k_\nu}{k^2} \Big] \tag{9.95}$$

This means that whatever we contract $\Pi_{\mu\nu}$ with had better give 0 if $\Pi_{\mu\nu} = k_\mu k_\nu$. This is very similar to the requirement of the Ward identities, which says that the matrix elements vanish if the physical external polarization is replaced by $\epsilon_\mu \to k_\mu$. It turns out the proof of the two requirements are very similar. In fact, if we can prove gauge invariance (ξ-independence) it will imply the Ward identity, since if the k_μ part of the propagator always gives zero it will also give zero on an external state.

Now, I don't know of a general diagrammatic proof of gauge invariance (again, I just mean ξ-independence of Feynman diagrams by this – it's easy to prove that a Lagrangian is gauge invariant). We will do the proof for the special case of scalar QED (where the electron has no spin) and real QED (where the electron is a spinor). But if you have all kinds of crazy interactions like $(D_\mu\phi)^{10}$, I think you would have to do the proof all over again. That's a little unsatisfying. However, there is a way to show in general that if we start with 2 polarizations for the photon, we will never produce that unwanted third one (or the fourth), which was the whole point of gauge invariance to begin with. That proof uses something called BRST invariance, which we will get to eventually, and is completely non-perturbative. But we need a lot more sophistication to understand it, in particular we need the path integrals, so we'll have to wait.

9.8 Why gauge invariance?

The symmetry $A_\mu \to A_\mu + \partial_\mu\alpha(x)$ is known as a gauge symmetry, because it depends on a function of spacetime $\alpha(x)$. (The word *gauge* means size, and the original symmetry of this type was conceived by Hermann Weyl as an invariance under scale transformations, now known as Weyl invariance). Another word for *gauge* symmetry is *local* symmetry. The alternative is global symmetry, for which α is just a number. Any theory which has a gauge symmetry also has a global symmetry, since we can just take $\alpha(x) =$ constant. And we saw that the global symmetry is all that's necessary to have a conserved current.

Note that under a global symmetry, the gauge field A_μ is fixed. So if we set $A_\mu = 0$, the rest of the Lagrangian will still have a global symmetry, since A_μ doesn't transform. Moreover terms like

$$m^2 A_\mu^2 \tag{9.96}$$

are invariant under the global symmetry, even though they violate gauge invariance.

Also note that even if we choose a gauge, such as $\partial_\mu A_\mu = 0$, the global symmetry is still present, but the gauge symmetry is broken (the gauge fixing $\partial_\mu A_\mu = 0$ is not invariant). But we know that we can always choose a gauge, and the physics is going to be exactly the same. So what does this mean? What's the point of having a local symmetry if we can just choose a gauge (in fact, we *have to* choose a gauge to do any computations) and the physics is the same?

There are two answers to this question. First, it is fair to say that gauge symmetries are a total fake. They are just redundancies of description and really do have no physically observable consequences. In contrast, global symmetries are real features of nature with observable consequences. For example, global symmetries imply the existence of conserved charges, which we can test. So the first answer is that we technically don't need gauge symmetries at all.

The second answer is that local symmetries make it much easier to do computations. You might wonder why we even bother introducing this field A_μ which has this huge redundancy to it. Instead, why not just quantize the electric and magnetic fields, that is $F_{\mu\nu}$, itself? Well you could do that, but it turns out to be more of a pain than using A_μ. To see that, first note that $F_{\mu\nu}$ as a field does not propagate with the Lagrangian $\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2$. All the dynamics will be moved to the interactions. Moreover, if we include interactions, either with a simple current $A_\mu J_\mu$ or with a scalar field $\phi^\star A_\mu \partial_\mu \phi$ or with a fermion $\bar\psi\gamma_\mu A_\mu\psi$, we see that they naturally involve A_μ. If we want to write these in terms of $F_{\mu\nu}$ we have to solve for A_μ in terms of $F_{\mu\nu}$ and we will get some crazy non-local thing like $A_\nu = \frac{\partial_\nu}{\Box}F_{\mu\nu}$. Then we'd have to spend all our time showing that the theory is actually local and causal. It turns out to be much easier to deal with a little redundancy so that we don't have to check locality all the time.

Another reason is that all of the physics of the electromagnetic field is not entirely contained in $F_{\mu\nu}$. In fact there are global properties of A_μ that are not contained in $F_{\mu\nu}$ but that can be measured. This is the Aharanov-Bohm effect, that you might remember from quantum mechanics. Thus we are going to accept that using the field A_μ instead of $F_{\mu\nu}$ is a necessary complication.

So there's no physics in gauge invariance but it makes it a lot easier to do field theory. The physical content is what we saw in the previous section with the Lorentz transformation properties of spin 1 fields

Massless spin 1 fields have 2 polarizations

One implication of this is the Ward identity, which is related to gauge invariance. But I'm sorry to tell you gauge invariance doesn't seem to be necessary. Here are two examples to illustrate why gauge invariance isn't as powerful as it seems.

First, we saw that we could calculate the photon propagator from the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + \frac{1}{\xi}(\partial_\mu A_\mu)^2 \tag{9.97}$$

This Lagrangian breaks gauge invariance, but the photon still has only 2 polarizations.

Second, consider the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + m^2(A_\mu + \partial_\mu \pi) \tag{9.98}$$

This has a gauge invariance under which $A_\mu\,(x) \to A_\mu\,(x) + \partial_\mu \alpha(x)$ and $\pi(x) \to \pi(x) - \alpha(x)$. However, I can use that symmetry to set $\pi = 0$ everywhere. Then the Lagrangian reduces to that of a massive gauge boson. So the physics is that of 3 polarizations of a spin 1 field, but there is an exact gauge invariance in the Lagrangian. I could do something even more crazy with this Lagrangian: integrate out π. By setting π equal to it's equations of motion and substituting back into the Lagrangian, it becomes

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + F_{\mu\nu}\frac{m^2}{\Box}F_{\mu\nu} \tag{9.99}$$

This Lagrangian is also manifestly gauge invariant, but it is non-local. It looks like a massless spin 1 field, but we know that it's exactly the same as the massive spin-1 Lagrangian. It is not hard to see that any Lagrangian involving A_μ can be made gauge invariant by adding π in this way tracking information can provide resolution. without changing the physical content of the theory. So gauge invariance it's tricky business indeed.

Now that I've convinced you that gauge invariance is misleading, let me unconvince you. It turns out that the gauge fixing in Eq. (9.97) is very special. We can only get away with it because the gauge symmetry of QED we are using is particularly simple (technically, because it is Abelian). When we do path integrals, and study non-Abelian gauge invariance, you will see that you can't just drop a gauge fixing term in the Lagrangian, but have to add it in a very controlled way. Doing it properly leaves a residual symmetry in the quantum theory called BRST invariance. In the classical theory you are fine, but in the quantum theory you have to be careful, or you violate unitarity (probabilities don't add up to 1). Ok, but then you can argue that it is not gauge invariance that's fundamental but BRST invariance. However, it turns out that sometimes you need to break BRST invariance. However, this has to be controlled, leading to something called the Bitalin-Vilkovisky formalism. Then you get Slavnov-Taylor identities, which are a type of generalized Ward identities. So it keeps getting more and more complicated, and I don't think anybody really understands what's going on.

The bottom line, and the point of all this, is that gauge invariance (and BRST invariance), is a feature of the way we compute things in quantum field theory. It is not a feature of our universe. The observable feature of our universe is that *massless spin 1 fields have 2 physical polarizations*.

Chapter 10
Scalar QED

10.1 Introduction

Now that we have Feynman rules and we know how to quantize the photon, we are very close to quantum electrodynamics. All we need is the electron, which is a spinor. Before we get into spinors, however, it is useful to explore a theory which is an approximation to QED in which the spin of the electron can be neglected. This is called scalar QED. The Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 - |D_\mu\phi|^2 - m^2|\phi|^2 \tag{10.1}$$

$$D_\mu\phi = \partial_\mu\phi + ieA_\mu\phi \tag{10.2}$$

There actually are charged scalar fields in the world which do have exactly this Lagrangian: the charged pions. But that's not really the point. Spinors are messy, so starting with this simplified Lagrangian will let us understand some elements of field theory without the mess.

10.2 Quantizing Complex Scalar fields

We saw that for a scalar field to couple to A_μ it has to be complex. This is because the charge is associated with a global symmetry under which

$$\phi \to e^{i\alpha}\phi \tag{10.3}$$

this can only happen for a complex field – we need the global symmetry so that there can be a local, gauge symmetry, which A_μ will represent. The first thing to notice is that the classical equations of motion for ϕ and $\phi^\star$ are

$$(\Box + m^2)\phi - ieA_\mu\partial_\mu\phi = 0 \tag{10.4}$$

$$(\Box + m^2)\phi^\star + ieA_\mu\partial_\mu\phi^\star = 0 \tag{10.5}$$

So that ϕ and $\phi^\star$ couple to the electromagnetic field with opposite charge but have the same mass. Thus in a relativistic theory with a gauge field, we *automatically* get anti-particles!

For a real scalar we had

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\Big(a_p^\dagger e^{ipx} + a_p e^{-ipx}\Big) \tag{10.6}$$

For complex scalar the field is no longer Hermetian. We can write in general

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\Big(a_p^\dagger e^{ipx} + b_p e^{-ipx}\Big) \tag{10.7}$$

Then, by complex conjugation

$$\phi^\star(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\Big(b_p^\dagger e^{ipx} + a_p e^{-ipx}\Big) \tag{10.8}$$

Thus we can conclude that $b_p^\dagger$ creates particles of the opposite charge and same mass to what $a_p^\dagger$ creates, we call these antiparticles. Note that in both cases $\omega_p = \sqrt{p^2+m^2} > 0$.

All we used was the fact that the field was complex. Clearly $a_p^\dagger \neq b_p^\dagger$ as these operators create particles of opposite charge. So charge implies complex fields, which implies antiparticles. That is

- Matter coupled to massless spin 1 particles automatically implies the existence of particles of the same mass and opposite charge.

That is a very profound observation.

10.2.1 Dirac's sea

Historically, it was the Dirac equation which led to antiparticles. In fact, in 1931 Dirac predicted there should be a particle like the electron but with opposite charge, a year before the positron was discovered by Anderson in 1932.

Actually, Dirac had an interpretation of ϕ which sounds really funny in retrospect, but was much more logical to him for historical reasons. Suppose we had written

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\left(a_p^\dagger e^{ipx} + c_p^\dagger e^{-ipx}\right) \tag{10.9}$$

where both $a_p^\dagger$ and $c_p^\dagger$ are creation operators. Then $c_p^\dagger$ seems to be creating states of negative frequency, or equivalently negative energy. Dirac interpreted it as removing something of positive energy, and creating an energy hole. But an energy hole in what? His answer was that the universe is a sea full of positive energy states. Then $b_p^\dagger$ creates a hole in this sea, which moves around like an independent excitation.

Then why does the sea stay full, and not collapse to the lower energy configuration? Dirac's explanation for this was to invoke the Fermi exclusion principle. The sea is like the orbitals of an atom, We can only see particle or hole excitations on that sea, like when our atom is ionized. So this is really confusing, since it should work only for Fermions, not our scalar field which is a boson. It's much easier to write $c_p^\dagger = b_p$, which cleans everything up immediately.

You can use the sea language if you want, and then the proper answer for why it stays full is that there are no interactions in the theory to collapse it – since ϕ creates particles and holes, you can never lower the energy. In any case, these days Physicists accept Dirac's negative frequency excursion and the Dirac sea as historical curiosity, like the aether or phlogiston, and don't give it much thought, although it does come up from time to time.

10.3 Feynman rules for scalar QED

Expanding out the scalar QED Lagrangian we find

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 - \phi^\star(\Box + m^2)\phi - ieA_\mu[\phi^\star(\partial_\mu\phi) - (\partial_\mu\phi^\star)\phi] + e^2A_\mu^2|\phi|^2 \tag{10.10}$$

This thing that A_μ couples to is just the conserved current for a complex scalar field

$$J_\mu = i[\phi^\star(\partial_\mu\phi) - (\partial_\mu\phi^\star)\phi] \tag{10.11}$$

On the (free) equations of motion for ϕ

$$(\Box + m^2)\phi = 0 \tag{10.12}$$

J_μ is conserved. We also understood J_μ as the Noether current for charge conservation.

We can read off the Feynman rules from the Lagrangian. The scalar propagator is

$$\langle 0|T\{\phi^\star(x)\phi(y)\}|0\rangle = \int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2 - m^2 + i\varepsilon}e^{ik(x-y)} \tag{10.13}$$

This propagates both ϕ and $\phi^\star$, that is both particles and antiparticles at the same time – they cannot be disentangled! The vertex

$$ie A_\mu[\phi^\star(\partial_\mu\phi) - (\partial_\mu\phi^\star)\phi] \tag{10.14}$$

connects A_μ with ϕ and $\phi^\star$. To see which momentum factors we get, look back at the quantized fields:

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\Big(a_p^\dagger e^{ipx} + b_p e^{-ipx}\Big) \tag{10.15}$$

$$\phi^\star(x) = \int \frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\Big(b_p^\dagger e^{ipx} + a_p e^{-ipx}\Big) \tag{10.16}$$

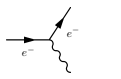
we see $\phi(x)$ in the interaction implies the creation of a particle with momentum p_μ or the annihilation of a antiparticle with momentum $-p_\mu$. $\phi^\star$ implies the creation of an antiparticle with momentum p_μ or the annihilation of an antiparticle with momentum $-p_\mu$.

Since the vertex has the form

$$J_\mu = ie[\phi^\star(\partial_\mu\phi) - (\partial_\mu\phi^\star)\phi] \tag{10.17}$$

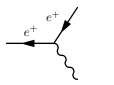
it always has one ϕ and one $\phi^\star$. Also, if the derivative acts on the antiparticle, it gets a minus sign. In addition, each p comes with an i, and there is another i from the expansion of $\exp(iV)$, so we always get an overall minus sign. There are 4 possibilities

- annihilate e^- and create e^- – particle scattering (p_1 is on the left)



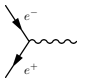
$$= \quad -ie(p_\mu^1 + p_\mu^2)$$

- annihilate e^+ and create e^+ – antiparticle scattering (p_1 is on the left)



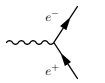
$$= \quad -ie(-p_\mu^1 - p_\mu^2)$$

- annihilate a e^- and annihilate an e^+ – pair annihilation (p_1 is the e^-)



$$= \quad -ie(p_\mu^1 - p_\mu^2)$$

- create a e^- and create e^+ – pair creation (p_1 is the e^-)



$$= \quad -ie(p_\mu^1 - p_\mu^2)$$

First of all, we see that there are only 4 types of vertices. It is impossible for a vertex to create two e^-'s, annihilate two e^-'s, create two $e^{+\prime}s$ or annihilate two $e^{+\prime}s$. So the Feynman rules guarantee that charge is conserved.

Secondly, we done something that may seem bizarre – we reversed the arrows on the e^+ lines. This is actually really helpful to remind us that charge is conserved. Forget about the arrows for a moment and just think of physical momentum flowing from left to right. Let's check the signs. For the first diagram, the term $\phi^\star(\partial_\mu\phi)$ gives a $+\,p_\mu^2$ because the e^- is created by ϕ and the $-\,(\partial_\mu\phi^\star)\phi$ gives a $-\,(-\,p_\mu^1)$ because an e^- is being destroyed by $\phi^\star$. For the second diagram, we find $\phi^\star(\partial_\mu\phi)$ annihilates the e^+ giving $-\,p_\mu^2$ and $-\,(\partial_\mu\phi^\star)\phi$ creates an e^+ giving $-\,(+\,p_\mu^1)$. The second two you can figure out on your own.

What's the pattern? If the arrow is going to the right, we get a $+\,iep_\mu$, if the arrow is going to the left, we get a $-\,iep_\mu$. So the arrow represents charge flow, not momentum flow. Then the rules are simple

- Arrows are charge flow = charge $\times$ momentum

- Arrow to the right $= iep^\mu$; arrow to the left $-iep^\mu$. Sum over the two lines in vertex.

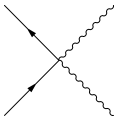
One way to interpret these rules is that particle number flowing forward in time is the same as anti-particle number flowing backward in time. This is known as the Feynman-Stuckleberg interpretation, and is related to Dirac's sea and the negative frequency hole states. It's kind of a cute observation and gives a funny picture of the universe with electrons flying around, bouncing off photons and going back in time, etc. You can have fun thinking about this. But to my knowledge. nothing practical has ever come out this language.

You also see that the arrows on the lines are connected. This is particularly handy in loops. But keep in mind that for physical external anti-particles, the momentum is actually flowing backwards to the direction of the arrow. That's not hard to remember – if your diagram represents a physical process, particle arrows should always point right and antiparticle lines should always point left.

We can't forget that there is another 4-point vertex in scalar QED

$$\mathcal{L}_{\text{int}} = e^2 A_\mu^2 |\phi|^2 \tag{10.18}$$

Which looks like



$$= 2\,i e^2 g_{\mu\nu} \tag{10.19}$$

Remember, this vertex comes from $|D_\mu \phi|^2$, so it is forced by gauge invariance. The 2 comes from the symmetry factor for the 2 $A's$. There would not have been a 2 if we had written $\frac{1}{2}e^2 A_\mu^2 |\phi|^2$, but this is not what the Lagrangian gives us. The i comes from the expansion of $\exp(iV)$ which we always have for Feynman rules.

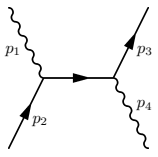
10.4 External states

Now we know the propagators for the photon and the complex scalar field. The only thing left is external states. For a scalar field, this is easy – we just get a factor 1. That's because a complex scalar field is just two real scalar fields, so we just take the real scalar field result.

For external lines, we really should rederive the LSZ theorem. However, it's easier just to observe that the photon field is

$$A_\mu(x,t) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2\omega_k}} \sum_{i=1}^{2} \left(\epsilon_\mu^i(k) a_{k,i}^\dagger e^{ikx} + \epsilon_\mu^i(k) a_{k,i} e^{-ikx}\right) \tag{10.20}$$

So as far as free states are concerned, which is all we need for LSZ, the photon is just a bunch of scalar fields integrated against some polarizations $\epsilon_\mu^i(k)$. Since we know k of the external states, by definition of the scattering amplitude, we get a factor of the photon polarization. So a diagram like this

$i\mathcal{M} =$ 

$$= (-ie)\varepsilon_\mu^1 (p_2^\mu + k^\mu)\frac{i}{k^2 - m^2 + i\varepsilon}(-ie)(p_2^\nu + k^\nu)\epsilon_\nu^4$$

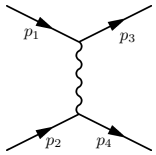
$$k_\mu = p_1 + p_2 \tag{10.21}$$

The first polarization (ϵ_μ^1) is the polarization of the photon labeled with p_μ^1. It gets contracted with the momenta $p_2^\mu + k^\mu$ which comes from the $ieA_\mu[\phi^\star(\partial_\mu\phi) - (\partial_\mu\phi^\star)\phi]$ vertex, we just substitute the free field A_μ instead of a propagating field (its propagator is canceled by LSZ). The second polarization (ϵ_μ^4) is the polarization of the photon labeled with p_μ^4.

10.5 Scattering in scalar QED

Now we have a real theory with photons and antiparticles and we can calculate something for real. Let's try $e^-e^- \to e^-e^-$ scattering. There are two diagrams. In the t-channel diagram gives

$$i\mathcal{M}_t = \quad = (-ie)(p_1^\mu + p_3^\mu)\frac{-i[g_{\mu\nu} + (1-\xi)\frac{k_\mu k_\nu}{k^2}]}{k^2}(-ie)(p_2^\nu + p_4^\nu)$$



with

$$k^\mu = p_3^\mu - p_1^\mu \tag{10.22}$$

But note that

$$k^\mu(p_1^\mu + p_3^\mu) = (p_3^\mu - p_1^\mu)(p_3^\mu + p_1^\mu) = p_3^2 - p_1^2 = m^2 - m^2 = 0 \tag{10.23}$$

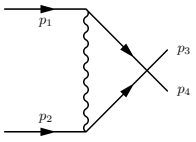
So this simplifies to

$$\mathcal{M}_t = -e^2\frac{(p_1^\mu + p_3^\mu)(p_2^\mu + p_4^\mu)}{t}$$

This of course had to happen, by gauge invariance. But also, we are just calculating exactly $k_\mu J_\mu = 0$. In any case, it's important to do this kind of check to make sure our algebra is right.

The u-channel gives

$$i\mathcal{M}_u = \quad = (-ie)(p_1^\mu + p_4^\mu)\frac{-i[g_{\mu\nu} + (1-\xi)\frac{k_\mu k_\nu}{k^2}]}{k^2}(-ie)(p_2^\nu + p_3^\nu),$$



where

$$k^\mu = p_4^\mu - p_1^\mu \tag{10.24}$$

For this

$$k^\mu(p_1^\mu + p_4^\mu) = p_1^2 - p_4^2 = 0 \tag{10.25}$$

So

$$\mathcal{M}_u = e^2\frac{(p_1^\mu + p_4^\mu)(p_2^\mu + p_3^\mu)}{u}$$

Thus the cross section for scalar Moller scattering is

$$\frac{d\sigma(e^-e^- \to e^-e^-)}{d\Omega} = \frac{e^4}{64\pi^2 E_{\rm cm}^2}\left[\frac{(p_1^\mu + p_3^\mu)(p_2^\mu + p_4^\mu)}{t} + \frac{(p_1^\mu + p_4^\mu)(p_2^\mu + p_3^\mu)}{u}\right]^2 \tag{10.26}$$

You can simplify this further, if you want. For example, in the massless limit this becomes

$$\frac{d\sigma(e^-e^- \to e^-e^-)}{d\Omega} = \alpha_e^2\left(\frac{1}{t} + \frac{1}{u}\right)^2 \tag{10.27}$$

where $\alpha_e = \frac{e^2}{4\pi}$ is the fine-structure constant.

10.6 Gauge invariance in scalar QED

We saw in the previous example that the matrix elements for a particular amplitude in scalar QED were independent of the gauge parameter ξ. Let's see if we can prove this in more generality.

The photon propagator is

$$\Pi_{\mu\nu} = \frac{-i\left[g_{\mu\nu} + (1-\xi)\frac{k_\mu k_\nu}{k^2}\right]}{k^2} \tag{10.28}$$

So we have to show that the matrix elements vanish if we take

$$\Pi_{\mu\nu} = \xi k_\mu k_\nu \tag{10.29}$$

In the previous example, we used that $p_1^2 = p_3^2$ which says that the mass of the incoming electron was the same as the mass of the outgoing electron. That *must* be true for particles, or for antiparticle. But suppose the photon connected not to an external line, which is on-shell, but to an off-shell internal line. These lines do not have to have $p^2 = m^2$, nor is the p^2 of the line entering the vertex the same as the p^2 of the line leaving the vertex. This does not mean the amplitudes with internal lines are not gauge invariant, because we have to sum over all the relevant diagrams to get the gauge invariant result.

Imagine a general diagram with all kinds of ϕ lines sticking out, and all kinds of other things sticking out too, and lots of loops. We can even have lots of types of charged particles ϕ_i with different masses m_i. Now pick a particular ϕ which attaches to some photon. Now we can do something very convenient, we can trace the ϕ line around the diagram. That's because charge is conserved, as we can see in our Feynman rules. Now we know that the final amplitude will be the some of lots of diagrams. So we can swap our photon with a different photon in the diagram and we get another contribution to the same amplitude. We will have to sum over all possible contributions.

Now pick a particular photon in that diagram, whose propagator has a $k^\mu k^\nu$ factor. Call the particle that the photon connects to ϕ. Now we can do something very convenient, we can trace the ϕ line around the diagram. That's because charge is conserved, as we can see in our Feynman rules. Now we know that the final amplitude will be the some of lots of diagrams. So we can swap our photon with a different photon in the diagram and we get another contribution to the same amplitude. We will have to sum over all possible contributions.

In each contribution, the momenta along the ϕ line segments are the same, so the ϕ propagator contributions will be the same. All that we care about is the factor from $\Pi_{\mu\nu}$ contracted with the vertex, as we move around to different vertices. Actually, the other end of the photon line also changes the diagram in a messy way, so I'm going to just cheat and claim that all that stuff doesn't matter. This is easy to see in the case where the other end of the photon connects to the line of a really heavy particle, much heavier than all the photon momentum. Then my relatively small momentum insertion makes little difference to either the vertices or the propagators. Gauge invariance in scalar QED is true in general, but if I take this simplification, I can show the essential elements of the proof without the extra complication.

The possible vertices are either from e^-e^+ annihilation or from $e^- \to e^-$ scattering. If the interaction is of the annihilation type,

$$= \quad -ie(p_\mu^1 - p_\mu^2), \qquad k_\mu = p_\mu^{i+1} + p_\mu^i.$$

we get

$$\mathcal{M}_i = (p_\mu^i - p_\mu^{i+1})\Pi_{\mu\nu} M_\nu \tag{10.30}$$

$$= \xi(p_\mu^i - p_\mu^{i+1}) k_\mu k_\nu M_\nu \tag{10.31}$$

$$= \xi[(p_\mu^i)^2 - (p_\mu^{i+1})^2] k_\nu M_\nu \tag{10.32}$$

where M_ν is the rest of the diagram, which doesn't change when I move the photon around. k_ν is fixed too – it is the momentum of the photon line, but it's helpful to substitute the explicit expression in the first term (the k_μ term) to simplify things.

If the interaction is of the scattering type, we get

$$= \quad -ie(p_\mu^1 + p_\mu^2), \quad k_\mu = p_\mu^{i+1} - p_\mu^i$$

$$\mathcal{M}_i = \xi(p_\mu^i + p_\mu^{i+1}) k_\mu k_\nu M_\nu \tag{10.33}$$

$$= \xi[(p_\mu^i)^2 - (p_\mu^{i+1})^2] k_\nu M_\nu \tag{10.34}$$

So it's the same thing.

Now we have to add up all the contributions.

$$M_{\rm tot} = (\xi k_\nu M_\nu) \sum_{i=1}^{n} (p_{i+1}^2 - p_i^2) \tag{10.35}$$

First, take the case that the connected charge current line which we are inserting the photon on to ends in external particles/antiparticles p_1 and p_n. Then

$$M_{\rm tot} = (\xi k_\nu M_\nu) \sum_{i=1}^{n} (p_{i+1}^2 - p_i^2) = (\xi k_{\ \nu} M_\nu)(p_n^2 - p_1^2) \tag{10.36}$$

Now, since p_1 and p_n are the external states which are on-shell and have the same mass ($m_{e^-} = m_{e^+}$), we see that the result is always zero.

Next, suppose the photon comes out of a loop. Then everything is the same, but we have to sum over the whole loop. The last insertion connects the beginning to the end of the loop so we get

$$(\xi k_\nu M_\nu) \left[p_1^2 - p_n^2 + \sum_{i=1}^{n} (p_i^2 - p_{i+1}^2) \right] = 0 \tag{10.37}$$

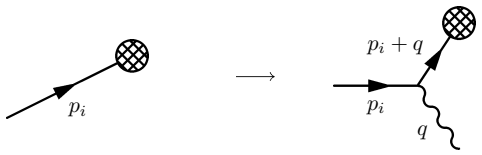
So it works no matter what. This is the way you prove gauge invariance diagrammatically for scalar QED.

10.7 Lorentz invariance and charge conservation

There is a beautiful and direct connection between Lorentz invariance and charge conjugation that bypasses gauge invariance completely. This is due to Weinberg, and more details can be found in section 13.1 of his Quantum Field Theory text.

The calculation is similar to what we used for gauge invariance, but we'll put back in the scalar propagators and take the limit that the photon momentum is soft (i.e. goes to zero). So imagine we have some diagram with lots of external legs and loops and things. Say the matrix element for this process is M_0.

Now I want to add tack one more incoming photon of momentum q_μ onto an external leg. Let me stick it into leg i, which I take to be an incoming e^-.



Then it modifies amplitude by adding a factor

$$(-ie)\frac{p_\mu^i + (p_\mu^i + q_\mu)}{(p_i+q)^2 - m^2}\epsilon_\mu \tag{10.38}$$

The numerator is just the photon scalar vertex, and the denominator is the propagator for the new leg. Now, by soft we mean at least that $q \ll p_i$ for all the external momenta p_i. So this term reduces to

$$-ie\frac{p_\mu^i + (p_\mu^i + q_\mu)}{(p_i+q)^2 - m^2}\epsilon_\mu \approx -ie\frac{2p_\mu^i}{p_i^2 + 2p_i q + q^2 - m^2}\epsilon_\mu \approx (-ie)\frac{p_\mu^i \varepsilon_\mu}{p_i q} \tag{10.39}$$

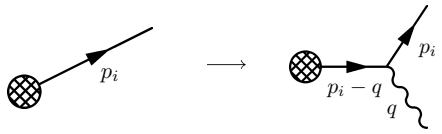
Of course, this modification also affects other parts of the diagram (inside the blob) because it contributes momentum to lots of lines. But if these lines are not close to being on-shell $p^2 \neq m^2$ then $(p + q)^2 - m^2 \approx p^2 - m^2$, because the photon is soft. There's certainly nothing diverging as $q \to 0$ like the medication we just found. So the dominant effect (and the only effect in the limit $q \to 0$) from the addition of a soft photon is from modifying the external legs, because they are on-shell.

If the leg is an incoming e^+, we would get

$$(ie)\frac{p^i_\mu \epsilon_\mu}{p_i q} \tag{10.40}$$

where the sign flip comes from the charge of the e^+.

If the leg is an outgoing, it's a little different. The photon is still incoming, so we have



So the amplitude is modified to

$$(-ie)\frac{p^i_\mu + (p^i_\mu - q_\mu)}{(p_i - q)^2 - m^2}\epsilon_\mu \approx (ie)\frac{p^i_\mu \epsilon_\mu}{p_i q} \tag{10.41}$$

Which is the same with an opposite sign. Similarly for an outgoing positron, we would get another sign flip and

$$(-ie)\frac{p^i_\mu \epsilon_\mu}{p_i q} \tag{10.42}$$

In fact if we have a more general crazy theory with particles of any charge, c_i, we will get similar formulas.

Summing over all the particles we get

$$\mathcal{M} = M_0 \sum_{\text{incoming}} (\text{charge})_i \frac{p^i_\mu \epsilon_\mu}{p_i q} - M_0 \sum_{\text{outgoing}} (\text{charge})_j \frac{p^j_\mu \epsilon_\mu}{p_j q} \tag{10.43}$$

where $(\text{charge})_i$ is the charge of particle i.

Here comes the punchline. We can now use the fact that under a Lorentz transformations

$$\epsilon_\mu \to \epsilon_\mu + q_\mu \tag{10.44}$$

which is the root of all the trouble with massless spin 1 particles. For the result to be gauge invariant we then must have

$$\sum_{\text{incoming}} (\text{charge})_i = \sum_{\text{outgoing}} (\text{charge})_j \tag{10.45}$$

Which says that charge is conserved

Massless spin 1 particles imply conservation of charge

What's the big deal?, you say, we knew that already. But in the previous derivation, we had to use gauge invariance, gauge fix, work out the photon propagator, etc. Those are all artifacts of trying to write down nice simple Lagrangian. For this result we didn't need to talk about the photon propagator at all, or choose a gauge or anything. We just needed to use that a massless particle of spin 1 has 2 polarizations and take the soft limit. Note that you can't take a soft limit if the photon is not massless! So everything is interconnected.

It's hard to emphasize enough how fantastic this result is. It's one of the coolest results in field theory, in my opinion. If you don't buy it, let me show you one more example.

10.7.1 Lorentz invariance for spin 2

There's another wonderful application of the same calculation, this time to a massless spin 2 field. A massless spin 2 field has 2 polarizations $\epsilon^i_{\mu\nu}$ which rotate into each other under Lorentz transformations, but also into $q_\mu q_\nu$:

$$\epsilon^i_{\mu\nu} \to c_{ij}(\Lambda)\epsilon^j_{\mu\nu} + q_\mu \Lambda_\nu + \Lambda_\nu q_\nu + f(\Lambda) q_\mu q_\nu \tag{10.46}$$

where these $\Lambda's$ have to do with the explicit way the Lorentz group acts, which we don't care about so much. Thus any theory involving a massless spin 2 field should satisfy a Ward identity: if we replace even one index of the polarization tensor by k_μ the matrix elements much vanish.

What do the interactions look like? Well, we have to couple something to a tensor field $h_{\mu\nu}$. We will be emitting one h so it should be linear in $h_{\mu\nu}$. By Lorentz invariance, we need two derivatives to fill up the $h_{\mu\nu}$ indices. We need 2 more fields to have an interaction, so we can have

$$\mathcal{L} = (\text{charge}) h_{\mu\nu} \partial_\mu \phi_1 \partial_\nu \phi_2 \tag{10.47}$$

or with the two derivatives acting elsewhere. We could also have more derivatives, but then the charge would have a different dimension and we can consider those interactions separately. If there are any derivatives acting on $h_{\mu\nu}$ I can move the over to the $\phi's$ by integrating by parts. This is totally general, and works for any kinds of fields

Going through the same argument as before, adding up diagrams for soft graviton emission, we find

$$\mathcal{M} = M_0 \left[\sum_{\text{incoming}} (\text{charge})_i \frac{p_\mu^i}{p_i q} \epsilon_{\mu\nu} p_\nu^i - \sum_{\text{outgoing}} (\text{charge})_j \frac{p_\mu^j}{p_j q} \epsilon_{\mu\nu} p_\nu^j \right] \tag{10.48}$$

where these charges are the gravitational charges of the things. The extra factor of p_ν^i has to be there by Lorentz invariance. The only other thing it could have been is q_ν, which could come from a derivative acting on the graviton, but we already moved those away by integration by parts.

By Lorentz invariance, requiring a massless spin 2 field, this should vanish if $\epsilon_{\mu\nu} = q_\mu \Lambda_\nu$ for any Λ_ν. So

$$M_0 \Lambda_\nu \left[\sum_{\text{incoming}} (\text{charge})_i p_\nu^i - \sum_{\text{outgoing}} (\text{charge})_j p_\nu^j \right] = 0 \tag{10.49}$$

$$\sum_{\text{incoming}} (\text{charge})_i p_\nu^i = \sum_{\text{outgoing}} (\text{charge})_j p_\nu^j \tag{10.50}$$

This says that the sum of $(\text{charge})_i p_i^\mu$ is conserved. But we already know that the sum of p_i^μ is conserved – momentum conservation. So, for example, we can solve for p_1^μ in terms of the others. If we and another constraint on the p_i^μ then there would be a different solution for p_1^μ, which is impossible unless all the p_i^μ are zero. The only way we can have non-trivial scattering is for all the charges to be the same

$$(\text{charge})_i = (\text{charge})_j \tag{10.51}$$

But that's exactly what gravity does! All particles gravitate with the same strength $\frac{1}{M_P}$. Gravity is universal

Massless spin 2 particles imply gravity is universal

We can keep going. For massless spin 3 we would need

$$\sum_{\text{incoming}} (\text{charge})_i p_\nu^i p_\mu^i = \sum_{\text{outgoing}} (\text{charge})_j p_\nu^j p_\mu^j \tag{10.52}$$

For example, the $\mu = \nu = 0$ component of this says

$$\sum_{\text{incoming}} (\text{charge})_i E_i^2 = \sum_{\text{outgoing}} (\text{charge})_j E_j^2 \tag{10.53}$$

that is the sum of the squares of the energies times some charges are conserved. That's way too constraining. The only way out is if all the charges are 0, which is a boring, non-interacting theory of free massless spin 3 field. So,

There are no interacting theories of masssless particles with spin>2

And in fact, no massless particles with spin > 2 have ever been seen.

Now you can see why Weinberg's theorem is so fantastic.q If we can find a nice way to ignore them, then perhaps we will get something finite.

Chapter 11
Spinors and the Dirac Equation

11.1 Introduction

From non-relativistic quantum mechanics, we already know that the electron has spin $\frac{1}{2}$. We usually write it as a doublet

$$|\psi\rangle = \left|\begin{pmatrix} \uparrow \\ \downarrow \end{pmatrix}\right\rangle \tag{11.1}$$

You might also remember the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{11.2}$$

which simplify the treatment of spin.

For example, the interaction of the electron with a magnetic field appears in what is sometimes called the Schrodinger-Pauli equation:

$$i\partial_t\psi = H\psi = \left[\left(\frac{p^2}{2m} + V(r) - \mu_B\vec{B}\cdot\vec{L}\right)\mathbb{1}_{2\times 2} - 2\mu_B\vec{B}\cdot\vec{\sigma}\right]\psi \tag{11.3}$$

where $\mu_B = \frac{e}{2m_e}$ is the "Bohr magneton" (the size of the electron's orbital magnetic moment) and $\vec{L} = \vec{x} \times \vec{p}$ is the angular momentum operator. We have also written $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ to call attention to the fact that these $\sigma_i's$ transform as the components of a vector, just like the magnetic field B_i. Thus $(\vec{\sigma}\cdot\vec{B})\psi$ is rotationally invariant. This is non-trivial, and only works because

$$[\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k \tag{11.4}$$

which are the same algebraic relations satisfied by infinitesimal rotations (we will review this shortly).

Keep in mind that σ_i do not change under rotations – they are always given by Eq. (2) in any frame. ψ is changing and B_i is changing, and these changes cancel when we write $(\vec{\sigma}\cdot\vec{B})\psi$.

We could also have written down a rotationally invariant equation of motion for ψ

$$\mathbb{1}\cdot\partial_t\psi - \partial_i\sigma_i\psi = 0 \tag{11.5}$$

Since ∂_i transforms like a 3-vector and so does $\sigma_i\psi$, this equation is rotationally invariant. It turns out it is Lorentz invariant too. In fact, this is just the Dirac equation!

If we write

$$\sigma_\mu = (\mathbb{1}_{2\times 2}, \sigma_1, \sigma_2, \sigma_3) \tag{11.6}$$

Then it is

$$\sigma^\mu\partial_\mu\psi = 0 \tag{11.7}$$

which is nice and simple looking. Actually, this is the Dirac equation for a Weyl spinor, which is not exactly the same as the equation commonly called the Dirac equation.

By the way, it does not follow that this equation is Lorentz invariant just because we've written it as $\sigma^\mu\partial_\mu$. For example,

$$(\sigma^\mu\partial_\mu + m)\psi = 0 \tag{11.8}$$

is *not* Lorentz invariant. To understand these enigmatic transformation properties, we have to know how to construct the Lorentz group out of the σ_μ's and see how those act on ψ. But instead of just guessing the answer (or looking it up), we're going to do something a little more general, and motivate what kinds of transformation properties are possible.

11.2 Representations of the Lorentz Group

We have already argued that particles in our world should transform under unitary representations of the Poincare group. These are characterized by mass m and spin j. We motivated this in our study of the spin 1 fields. Mass is Lorentz invariant, so it is an obvious quantum number. Momentum is also conserved, but it is Lorentz *covariant.* If choose a frame in which the momentum has some canonical form, for example $p^\mu = (m,\, 0,\, 0,\, 0)$ for $m > 0$, then the particles are characterized by the group that holds this momentum fixed. That is the little group, in this case the group of 3D rotations, SO(3). The group of 3D rotations provides the second quantum number, j. The way the polarizations transform under the full Lorentz group is then induced by the transformations under SO(3) and the way the momentum transforms under boosts.

We also saw that to do field theory, we have to write down Largrangians involving fields. These are things like V_μ or ϕ or $T_{\mu\nu}$ which don't have to involve particles *per se*, although obviously in a physical theory they will. As we saw for spin 1, there's a lot a trouble that comes from having to embed particles of fixed mass and spin into fields like V_μ. For example, V_μ has 4 degrees of freedom which describes spin 0 and spin 1, so the Lagrangian has to be carefully chosen to make sure the physical theory never excites the spin 0 component. In addition, when we want the particle to have $m = 0$ and spin 1, we needed to go further and make sure the longitudinal polarization is never produced. This led directly to charge conservation. The next logical step to make these embeddings a bit more systematic is to see what kinds of Lorentz-invariant fields we can write down at all. This will reveal the existence of the spin $\frac{1}{2}$ states, and help us characterize their embeddings into fields.

A *group* is a set of elements $\{g_i\}$ and a rule $g_i \otimes g_j \to g_k$ which tells how each pair of elements is multiplied to get a third. The rule defines the group, independent of any particular way to write the group elements down as matrices. A *representation* is particular embedding of these $g_i's$ into matrices. Often we talk about the vectors on which the matrices act as being the representation, but technically the matrix embedding is the representation. Any group has the trivial representations r: $g_i \to \mathbb{1}$. But more generally we care about embeddings which are *faithful*, for which each element gets it's own matrix.

Recall that the Lorentz group is the set of rotations and boosts which preserve the Minkoswki metric

$$\Lambda^T \eta \Lambda = \eta \tag{11.9}$$

Lorentz transformation can act on space-time 4-vectors as

$$V_\mu \to \Lambda_{\mu\nu} V_\nu \tag{11.10}$$

Where Λ is a combination of rotations and boosts. We saw that we could write a Lorentz transformation as the product of 3 rotations and 3 boosts:

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_{xy} & -\sin\theta_{xy} & \\ 0 & \sin\theta_{xy} & \cos\theta_{xy} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & \cos\theta_{xz} & & -\sin\theta_{xz} \\ & & 1 & \\ & \sin\theta_{xz} & & \cos\theta_{xz} \end{pmatrix} \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \cos\theta_{xz} & -\sin\theta_{xz} \\ & & \sin\theta_{xz} & \cos\theta_{xz} \end{pmatrix} \tag{11.11}$$

$$\times \begin{pmatrix} \cosh\beta_x & \sinh\beta_x & & \\ \sinh\beta_x & \cosh\beta_x & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \begin{pmatrix} \cosh\beta_y & & \sinh\beta_y & \\ & 1 & & \\ \sinh\beta_y & & \cosh\beta_y & \\ & & & 1 \end{pmatrix} \begin{pmatrix} \cosh\beta_z & & & \sinh\beta_z \\ & 1 & & \\ & & 1 & \\ \sinh\beta_z & & & \cosh\beta_z \end{pmatrix} \tag{11.12}$$

This is a particular representation. of the Lorentz group. That is, it is one embedding of the group into a set of matrices.

The group itself is a mathematical object independent of any particular representation. To extract the group away from its representations, it is easiest to look at infinitesimal transformations. We can always write these group elements as

$$\Lambda = \exp(i\theta_i \lambda_i) = 1 + i\theta_i \lambda_i + \cdots \tag{11.13}$$

where θ_i are 6 numbers, corresponding to the 6 angles $\theta_{xy}, \theta_{xz}, \theta_{yz}, \beta_{xy}, \beta_{xz}, \beta_{yz}$ and Λ_i are called *generators*. For the regular, vector, representation above, the generators are

$$\lambda_1 = V_{12} = i\begin{pmatrix} 0 & & & \\ & 0 & -1 & \\ & 1 & 0 & \\ & & & 0 \end{pmatrix}, \quad \lambda_2 = V_{13} = i\begin{pmatrix} 0 & & & \\ & 0 & & -1 \\ & & 0 & \\ & 1 & & 0 \end{pmatrix}, \quad \lambda_3 = V_{23} = i\begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & -1 \\ & & 1 & 0 \end{pmatrix}$$

$$\lambda_4 = V_{01} = i\begin{pmatrix} 0 & 1 & & \\ 1 & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}, \quad \lambda_5 = V_{02} = i\begin{pmatrix} 0 & & 1 & \\ & 0 & & \\ 1 & & 0 & \\ & & & 0 \end{pmatrix}, \quad \lambda_6 = V_{03} = i\begin{pmatrix} 0 & & & 1 \\ & 0 & & \\ & & 0 & \\ 1 & & & 0 \end{pmatrix} \tag{11.14}$$

The $\lambda's$ and the $V's$ are just different notations for the same objects. The generators satisfy:

$$[V_{\mu\nu}, V_{\rho\sigma}] = i(\eta_{\nu\rho} V_{\mu\sigma} - \eta_{\mu\rho} V_{\nu\sigma} - \eta_{\nu\sigma} V_{\mu\rho} + \eta_{\mu\sigma} V_{\nu\rho}) \tag{11.15}$$

which means

$$[V_{01}, V_{12}] = i\, V_{02}, \quad [V_{12}, V_{23}] = i\, V_{13}, \quad \ldots \tag{11.16}$$

So now we can define the Lorentz group as the set of transformations generated by these generators.

Another representation or the Lorentz generators is given

$$J_{\mu\nu} = i(x_\mu \partial_\nu - x_\nu \partial_\mu) \tag{11.17}$$

These are the classical generators of angular momentum generalized to include time. You can check that $J_{\mu\nu}$ satisfy the commutation relations of the Lorentz algebra.

Technically, we say the generators make up the Lorentz Algebra *so(1,3)* which generates the Lorentz Group SO(1, 3). A common convention is to use lowercase letters for the names of algebras and uppercase letters for groups. As another technicality, we note that it is possible for two different groups to have the same algebra. For example, the Proper Orthonchronous Lorentz group and the Lorentz Group have the same algebra, but the Lorentz group has in addition the discrete symmetries time reversal and parity reversal. It's a small difference, but it's a difference and the groups are not identical.

11.3 General representations

There is a very nice way to study the representations of the Lorentz group. Start with the rotation generators J_i and the boost generators K_j. For example, you could take $J_1 = V_{23}$, $K_1 = V_{01}$, etc. as a particular representation of this basis. These satisfy

$$[J_i, J_j] = i\epsilon_{ijk} J_k \tag{11.18}$$

$$[J_i, K_j] = i\epsilon_{ijk} K_k \tag{11.19}$$

$$[K_i, K_j] = -i\epsilon_{ijk} J_k \tag{11.20}$$

where ϵ_{ijk} is defined by $\varepsilon_{123} = 1$ and being antisymmetric in all of its indices (it is called the *totally antisymmetric tensor*.) As is probably clear to you already, $[J_i, J_j] = i\epsilon_{ijk} J_k$ is the algebra for rotations, so(3) so the J_i generate the subgroup of rotations in 3D.

Now take the linear combinations

$$J_i^+ = \frac{1}{2}(J_i + iK_i) \tag{11.21}$$

$$J_i^- = \frac{1}{2}(J_i - iK_i) \tag{11.22}$$

which satisfy

$$[J_i^+, J_j^+] = i\epsilon_{ijk}J_k^+ \tag{11.23}$$

$$[J_i^-, J_j^-] = i\epsilon_{ijk}J_k^- \tag{11.24}$$

$$[J_i^+, J_j^-] = 0 \tag{11.25}$$

So we have found 2 commuting subgroups of the Lorentz group. The algebra of the $J's$ is the 3D rotation algebra, so(3), more commonly called so(2). So we have shown that

$$\mathrm{so}(1,3) = \mathrm{su}(2) \times \mathrm{su}(2) \tag{11.26}$$

(technically, it should really be $\mathrm{sl}(2,\mathbb{R}) = \mathrm{so}(1,1)$ not su(2), but these algebras are the same up to some i's, and physicists usually just say su(2).).

The names so(n) and su(n) stand for special orthogonal algebra and special unitary algebra. Orthogonal means it preserves a norm defined with transpose: V^TV is invariant under SO(n). Unitary means it preserves a norm defined with adjoint $V^\dagger V$ is invariant under SU(n). Special means the determinant is 1, which just normalizes the elements and gets rid of phases.

The decomposition $\mathrm{so}(3,1) = \mathrm{su}(2) \times \mathrm{su}(2)$ makes studying the representations very easy. We already know what the representations are of su(2), since this is the algebra of Pauli matrices, which generates the 3D rotation group SO(3) ($\mathrm{so}(3) = \mathrm{su}(2)$). The representations are characterized by a quantum number j, and have $2j+1$ elements, labeled by $m = j_z$, So representations of the Lorentz group are characterized by two numbers A and B. The (A,B) representation has $(2A+1)(2B+1)$ degrees of freedom.

The regular rotation generators are $\vec{J} = \vec{J}^+ + \vec{J}^-$, where I now use the vector superscript to call attention the fact that the spins must be added vectorially (remember Clebsch-Gordon coefficients?). The bottom line is that each (A,B) representation contains particles of spin $J = A+B, A+B-1, \cdots, |A-B|$. For example,

rep of $\mathrm{su}(2)\times\mathrm{su}(2)$	rep of $3D$ rotation group
$(A,B) = (0,0)$	$J = 0$
$(A,B) = (\frac{1}{2},0)$	$J = \frac{1}{2}$
$(A,B) = (0,\frac{1}{2})$	$J = \frac{1}{2}$
$(A,B) = (\frac{1}{2},\frac{1}{2})$	$J = 1,0$
$(A,B) = (1,0)$	$J = 1$
$(A,B) = (1,1)$	$J = 2,1,0$

(11.27)

So we see that the regular 4D vector representation A_μ contains spins 1,0 so it corresponds to the $(\frac{1}{2},\frac{1}{2})$ representation of the Lorentz group.

11.3.1 unitary representations

Representations of the Lorentz group are not always *unitary*. Unitarity means $\Lambda^\dagger\Lambda = 1$, which is necessary to have Lorentz invariant matrix elements

$$\langle\psi|\psi\rangle \to \langle\psi|\Lambda^\dagger\Lambda|\psi\rangle \tag{11.28}$$

Since the group element is the exponential of the generator $\Lambda = e^{i\lambda}$, unitarity requires that $\lambda^\dagger = \lambda$, that is, that λ be Hermetian. Looking back at our explicit 4D representation for the Lorentz generators, we see that the rotations

$$J_1 = i\begin{pmatrix} 0 & & & \\ & 0 & -1 & \\ & 1 & 0 & \\ & & & 0 \end{pmatrix}, \quad J_2 = i\begin{pmatrix} 0 & & & \\ & 0 & & -1 \\ & & 0 & \\ & 1 & & 0 \end{pmatrix}, \quad J_3 = i\begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & -1 \\ & & 1 & 0 \end{pmatrix}$$

are Hermetian. While the boosts

$$K_1 = i\begin{pmatrix} 0 & 1 & & \\ 1 & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}, \quad K_2 = i\begin{pmatrix} 0 & & 1 & \\ & 0 & & \\ 1 & & 0 & \\ & & & 0 \end{pmatrix}, \quad K_3 = i\begin{pmatrix} 0 & & & 1 \\ & 0 & & \\ & & 0 & \\ 1 & & & 0 \end{pmatrix} \tag{11.29}$$

Are not. (They are anti-Hermetian $\lambda^\dagger = -\lambda$.)

Non-unitary representations of the Lorentz group are ok for fields, but not for particles. Particles *must* transform under unitary representations of the Poincare group, so that we can construct physical amplitudes with them. As Wigner showed, and we discussed when we introduced gauge invariance, unitary representations and are characterized by mass and spin. This spin indicates a representation of the 3D rotation group, which is unitary.

Now, the generators for the su(2) × su(2) decomposition $\vec{J}_\pm = \vec{J} \pm i\,\vec{K}$ are Hermetian. This is because su(2) is the special unitary group, so of course it has unitary representations. But this doesn't mean that the corresponding representations of the Lorentz group are unitary. You have to be very careful about the factors of i and how these groups are defined. We said that a Lorentz group element is

$$\Lambda = \exp(i\theta_i J_i + i\beta_i K_i) \tag{11.30}$$

where the θ_i are the rotation angles and β_i the boosts "angles". These are *real* numbers. If we took the basis $\lambda_i = J^\pm$, then the corresponding $\theta_i's$ would have to be complex. So the representations of the Lorentz group are determined by the algebra su(2) × su(2), but since the angles θ are real, the corresponding representations of the Lorentz group are not unitary.

11.3.2 summary

In summary, the irreducible representations of the Lorentz group are characterized by two spins (A, B). Physical *particles* are characterized by mass, m and spin j. Fields generally describe particles of many spins, and one has to make sure that a physical theory only has causally propagating particles of the spins we want.

11.4 Spin $\frac{1}{2}$ representation

Now that we know that the irreducible representations of the Lorentz group are characterized by two spins, (A, B), how do we find the explicit representations? In particular, what do the $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ representations look like? We need to find 2x2 matrices that satisfy

$$[J_i^+, J_j^+] = i\epsilon_{ijk}J_k^+ \tag{11.31}$$

$$[J_i^-, J_j^-] = i\epsilon_{ijk}J_k^- \tag{11.32}$$

$$[J_i^+, J_j^-] = 0 \tag{11.33}$$

But we already know such matrices: the Pauli matrices

$$[\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k \tag{11.34}$$

rescaling we find

$$[\frac{\sigma_i}{2}, \frac{\sigma_j}{2}] = i\varepsilon_{ijk}\frac{\sigma_k}{2} \tag{11.35}$$

Another useful fact is that

$$\{\sigma_i, \sigma_j\} = \sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij} \tag{11.36}$$

Thus we can set $J_i^+ = \frac{\sigma i}{2}$. This is the $(\frac{1}{2},)$. What about J_i^-? This should be the $(, 0)$, so the obvious thing to do is just take the trivial representation $J_i^- = 0$. So the $(\frac{1}{2}, 0)$ representation is

$$(\frac{1}{2}, 0): \quad J_i^+ = \frac{\sigma_i}{2}, \quad J_i^- = 0 \tag{11.37}$$

Similarly, the $(0, \frac{1}{2})$ representation is

$$(0, \frac{1}{2}): \quad J_i^+ = 0, \quad J_i^- = \frac{\sigma_i}{2} \tag{11.38}$$

What does this mean for actual Lorentz transformations? Well, the rotations are $\vec{J} = \vec{J}_+ + \vec{J}_-$ and the boosts are $\vec{K} = i(J_- - J_+)$ so

$$(\frac{1}{2}, 0): \quad J_i = \frac{\sigma_i}{2}, \quad K_i = -i\frac{\sigma_i}{2} \tag{11.39}$$

$$(0, \frac{1}{2}): \quad J_i = \frac{\sigma_i}{2}, \quad K_i = i\frac{\sigma_i}{2} \tag{11.40}$$

Since the Pauli matrices are Hermetian $\sigma_i^\dagger = \sigma_i$ we see that the rotations are Hermetian but the boosts are anti-Hermetian. This is the same as what we found for the vector representation. Also notice that these two representations are complex conjugates of each other. In contrast, the vector representation was *real*.

Explicitly, if ψ_L is a $(\frac{1}{2}, 0)$ spinor, known also as a left-handed Weyl spinor, then under rotations angles θ_i and boost angles β_i

$$\psi_L \to e^{\frac{1}{2}(i\theta_i\sigma_i + \beta_i\sigma_i)}\psi_L = (1 + \frac{1}{2}i\theta_i\sigma_i + \frac{1}{2}\beta_i\sigma_i + \cdots)\psi_L \tag{11.41}$$

Similarly,

$$\psi_R \to e^{i\theta_i\sigma_i - \beta_i\sigma_i}\psi_R = (1 + \frac{1}{2}i\theta_i\sigma_i - \frac{1}{2}\beta_i\sigma_i + \cdots)\psi_R \tag{11.42}$$

Infinitesimally,

$$\delta\psi_L = \frac{1}{2}(i\theta_i + \beta_i)\sigma_i\psi_L \tag{11.43}$$

$$\delta\psi_R = \frac{1}{2}(i\theta_i - \beta_i)\sigma_i\psi_R \tag{11.44}$$

Note again the angles θ_i and β_i are real numbers. Although we mapped J_i^- or J_i^+ to 0, we still have non-trivial action of all the Lorentz generators. So these are *faithful* irreducible representations of the Lorentz group. Similarly

$$\delta\psi_L^\dagger = \frac{1}{2}(-i\theta_i + \beta_i)\psi_L^\dagger\sigma_i \tag{11.45}$$

$$\delta\psi_R^\dagger = \frac{1}{2}(-i\theta_i - \beta_i)\psi_L^\dagger\sigma_i \tag{11.46}$$

11.4.1 Lorentz invariants

Now the simplest thing to do would be to write down a Lagrangian with terms like

$$(\psi_L)^\dagger\Box\psi_L + m^2(\psi_L)^\dagger\psi_L \tag{11.47}$$

However, we can see from the above that this is not Lorentz invariant:

$$\delta\psi_L^\dagger\psi_L = \frac{1}{2}(\psi_L)^\dagger[(i\theta_i+\beta_i)\sigma_i\psi_L] + \frac{1}{2}[(\psi_L)^\dagger(-i\theta_i+\beta_i)\sigma_i]\psi_L \tag{11.48}$$

$$= \frac{\beta_i}{2}\psi_L^\dagger\sigma_i\psi_L \neq 0 \tag{11.49}$$

This is just the manifestation of the fact that the representation is not unitary because the boost generators are anti-Hermetian. If we allow ourselves two fields, ψ_L and ψ_R, we can write down terms like $\psi_R^\dagger\psi_L$. Under infinitesimal Lorentz transformations,

$$\delta(\psi_R^\dagger\psi_L) = \psi_R^\dagger\left[\frac{1}{2}(-i\theta_i-\beta_i)\sigma_i^\dagger\right]\psi_L + \psi_R^\dagger\left[\frac{1}{2}(i\theta_i+\beta_i)\sigma_i\right]\psi_L = 0 \tag{11.50}$$

Which is great. However, this term is not real. But then we can just add the Hermetian conjugate, so

$$m\Big(\psi_R^\dagger\psi_L + \psi_L^\dagger\psi_R\Big) \tag{11.51}$$

is ok.

What about kinetic terms? We could try

$$\psi_R^\dagger\Box\psi_L + \psi_L^\dagger\Box\psi_R \tag{11.52}$$

which is both Lorentz invariant and real. But this is actually not a very interesting Lagrangian. We can always split up our field into components $\psi_L = \left(\begin{smallmatrix}\psi_1\\ \psi_2\end{smallmatrix}\right)$, where ψ_1 and ψ_2 are just regular fields. Then we see that this is just the Lagrangian for a couple of scalars. So it's not enough to declare the Lorentz transformation properties of something, the Lagrangian has to *force* those transformation properties. In the same way, a vector field is just 4 scalars until we contract it with ∂_μ in the Lagrangian.

To proceed, let's look at

$$\psi_L^\dagger\sigma_j\psi_L \tag{11.53}$$

This transforms as

$$\delta\psi_L^\dagger\sigma_j\psi_L = \frac{1}{2}\psi_L^\dagger\sigma_j[(i\theta_i+\beta_i)\sigma_i\psi_L] + \frac{1}{2}[\psi_L^\dagger(-i\theta_i+\beta_i)\sigma_i]\sigma_j\psi_L \tag{11.54}$$

$$= \frac{i\theta_i}{2}\psi_L^\dagger(\sigma_j\sigma_i-\sigma_j\sigma_i)\psi_L + \frac{\beta_i}{2}\psi_L^\dagger(\sigma_i\sigma_j+\sigma_j\sigma_i)\psi_L \tag{11.55}$$

$$= \theta_i\varepsilon_{ijk}\psi_L^\dagger\sigma_k\psi_L + \beta_j\psi_L^\dagger\psi_L \tag{11.56}$$

Thus we have found that

$$\delta\psi_L^\dagger\psi_L = \frac{\beta_i}{2}\psi_L^\dagger\sigma_i\psi_L \tag{11.57}$$

$$\delta\psi_L^\dagger\sigma_j\psi_L = \theta_i\varepsilon_{ijk}\psi_L^\dagger\sigma_k\psi_L + \beta_j\psi_L^\dagger\psi_L \tag{11.58}$$

Similarly,

$$\delta\psi_R^\dagger\psi_R = -\frac{\beta_i}{2}\psi_L^\dagger\sigma_i\psi_L \tag{11.59}$$

$$\delta\psi_R^\dagger\sigma_j\psi_R = \theta_i\varepsilon_{ijk}\psi_L^\dagger\sigma_k\psi_L + -\beta_j\psi_L^\dagger\psi_L \tag{11.60}$$

Now, recall how a vector transforms

$$\delta V_0 = \beta_i V_0 \tag{11.61}$$

$$\delta V_j = \beta_j V_0 + \theta_i\varepsilon_{ijk}V_k \tag{11.62}$$

This is just the transformation of the vector $V_\mu^L = (\psi_L^\dagger\psi_L, \psi_L^\dagger\sigma_j\psi_L)$. Thus

$$\psi_L^\dagger\partial_t\psi_L + \psi_L^\dagger\partial_j\sigma_j\psi_L \tag{11.63}$$

is Lorentz invariant. Similarly $V_\mu^R = (\psi_R^\dagger \psi_R, \psi_R^\dagger \sigma_j \psi_R)$ transforms like a vector

Defining

$$\sigma_\mu = (\mathbb{1}, \vec{\sigma}), \quad \bar{\sigma}_\mu = (\mathbb{1}, -\vec{\sigma}) \tag{11.64}$$

We can write all the Lorentz-invariant terms we have found as

$$\mathcal{L} = i\psi_L^\dagger \sigma_\mu \partial_\mu \psi_L + i\psi_R^\dagger \bar{\sigma}_\mu \partial_\mu \psi_R + m\Big(\psi_R^\dagger \psi_L + \psi_L^\dagger \psi_R\Big) \tag{11.65}$$

I have added a factor of i on the kinetic terms for later convenience.

There's an even shorter-hand way to write this. If we write our spinor as a doublet

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{11.66}$$

Then, define

$$\bar{\psi} = \begin{pmatrix} \psi_R^\dagger & \psi_L^\dagger \end{pmatrix} \tag{11.67}$$

And using the 4x4 matrices

$$\gamma_\mu = \begin{pmatrix} & \sigma_\mu \\ \bar{\sigma}_\mu & \end{pmatrix} \tag{11.68}$$

Our Lagrangian becomes

$$\mathcal{L} = \bar{\psi}(\gamma^\mu \partial_\mu - m)\psi \tag{11.69}$$

Which is the conventional form of the Dirac equation.

11.5 Dirac matrices

Expanding them out, the Dirac matrices are

$$\gamma_0 = \begin{pmatrix} & \mathbb{1} \\ \mathbb{1} & \end{pmatrix} \quad \gamma_i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \tag{11.70}$$

Or,

$$\gamma_0 = \begin{pmatrix} 0 & & 1 & \\ & 0 & & 1 \\ 1 & & 0 & \\ & 1 & & 0 \end{pmatrix}, \quad \gamma_1 = \begin{pmatrix} 0 & & & 1 \\ & 0 & 1 & \\ & -1 & 0 & \\ -1 & & & 0 \end{pmatrix} \tag{11.71}$$

$$\gamma_2 = \begin{pmatrix} 0 & & & -i \\ & 0 & i & \\ & i & 0 & \\ -i & & & 0 \end{pmatrix}, \quad \gamma_3 = \begin{pmatrix} 0 & & 1 & \\ & 0 & & -1 \\ -1 & & 0 & \\ & 1 & & 0 \end{pmatrix} \tag{11.72}$$

They satisfy

$$\{\gamma_\mu, \gamma_\nu\} = 2\eta_{\mu\nu} \tag{11.73}$$

In the same way that the algebra of the Lorentz group is more fundamental than any particular representation, the algebra of the $\gamma's$ is more fundamental than any particular representation of them. We say the $\gamma's$ form the Dirac algebra, which is a special case of a Clifford algebra. This particular form of the Dirac matrices is known as the Weyl representation.

The Lorentz generators are

$$S_{\mu\nu} = \frac{i}{4}[\gamma_\mu, \gamma_\nu] \tag{11.74}$$

They satisfy the Lorentz algebra for any $\gamma's$ satisfying the Clifford algebra. That is, you can derive from $\{\gamma_\mu, \gamma_\nu\} = 2\eta_{\mu\nu}$ that

$$[S_{\mu\nu}, S_{\rho\sigma}] = i(\eta_{\nu\rho} S_{\mu\sigma} - \eta_{\mu\rho} S_{\nu\sigma} - \eta_{\nu\sigma} S_{\mu\rho} + \eta_{\mu\sigma} S_{\nu\rho}) \tag{11.75}$$

In the Weyl representation, the Lorentz generators are

$$S_{ij}=\frac{1}{2}\varepsilon_{ijk}\begin{pmatrix}\sigma_k & \\ & \sigma_k\end{pmatrix},\quad K_i=S_{0i}=-\frac{i}{2}\begin{pmatrix}\sigma_i & \\ & -\sigma_i\end{pmatrix} \tag{11.76}$$

Or, very explicitly

$$S_{12}=\frac{1}{2}\begin{pmatrix}1&&&\\&-1&&\\&&1&\\&&&-1\end{pmatrix},\quad S_{13}=\frac{i}{2}\begin{pmatrix}0&1&&\\-1&0&&\\&&0&1\\&&-1&0\end{pmatrix},\quad S_{23}=\frac{1}{2}\begin{pmatrix}0&1&&\\1&0&&\\&&0&1\\&&1&0\end{pmatrix} \tag{11.77}$$

$$S_{01}=\frac{i}{2}\begin{pmatrix}0&-1&&\\-1&0&&\\&&0&1\\&&1&0\end{pmatrix},\quad S_{02}=\frac{1}{2}\begin{pmatrix}0&-1&&\\1&0&&\\&&0&1\\&&-1&0\end{pmatrix},\quad S_{03}=\frac{i}{2}\begin{pmatrix}-1&&&\\&1&&\\&&1&\\&&&-1\end{pmatrix} \tag{11.78}$$

These are block diagonal. These are the same generators we used for the $(\frac{1}{2},0)$ and $(0,\frac{1}{2})$ representations above. It makes it clear that the Dirac representation is reducible, it is the sum of a left-handed and a right-handed spinor representation.

Another important representation is the Majorana representation

$$\gamma^0=\begin{pmatrix}0&\sigma^2\\\sigma^2&0\end{pmatrix},\quad \gamma^1=\begin{pmatrix}i\sigma^3&0\\0&i\sigma^3\end{pmatrix},\quad \gamma^2=\begin{pmatrix}0&-\sigma^2\\\sigma^2&0\end{pmatrix}\quad \gamma^3=\begin{pmatrix}-i\sigma^1&0\\0&-i\sigma^1\end{pmatrix} \tag{11.79}$$

In this basis the $\gamma's$ are purely imaginary.

The Weyl spinors, ψ_L and ψ_R are more fundamental, because they correspond to irreducible representations of the Lorentz group. But the electron is a Dirac spinor. Thus to do QED, it is easiest just to stick to the $\gamma's$ and to get used to manipulating them. Eventually, when you do supersymmetry, or study the weak interactions, you will need to use the L and R representations again.

11.6 Rotations

Now let's see what happens when we rotate by an angle θ in the xy plane. We use

$$\Lambda=\exp(i\theta_{xy}V_{12}) \tag{11.80}$$

How do we exponentiate a matrix? The standard trick is to first diagonalizing it with a unitary transformation, do the rotation, then transform back. This unitary transformation is like choosing a direction, except it is purely mathematical as the direction must be complex!

First, for the vector representation

$$V_{12}=i\begin{pmatrix}0&&&\\&0&1&\\&-1&0&\\&&&0\end{pmatrix}=U^{-1}\begin{pmatrix}0&&&\\&-1&&\\&&1&\\&&&0\end{pmatrix}U \tag{11.81}$$

So,

$$\Lambda(\theta_{xy})=\exp(i\theta_{xy}V_{12})=U^{-1}\begin{pmatrix}0&&&\\&\exp(-i\theta_{xy})&&\\&&\exp(i\theta_{xy})&\\&&&0\end{pmatrix}U \tag{11.82}$$

$$\Lambda(2\pi)=\mathbb{1} \tag{11.83}$$

That is, we rotate 360 degrees and we're back to where we started.

For the spinor representation

$$\Lambda_s = \exp(i\theta_{\mu\nu}S_{\mu\nu}) \tag{11.84}$$

the 12 rotation is already diagonal:

$$S_{12} = \frac{1}{2}\begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix} \tag{11.85}$$

So,

$$\Lambda(\theta_{xy}) = \exp(i\theta_{xy}S_{12}) = \begin{pmatrix} \exp(\frac{i}{2}\theta_{\mathrm{xy}}) & & & \\ & -\exp(\frac{i}{2}\theta_{\mathrm{xy}}) & & \\ & & \exp(\frac{i}{2}\theta_{\mathrm{xy}}) & \\ & & & -\exp(\frac{i}{2}\theta_{\mathrm{xy}}) \end{pmatrix} \tag{11.86}$$

$$\Lambda(2\pi) = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & -1 & \\ & & & 1 \end{pmatrix} \tag{11.87}$$

Thus a 2π rotation does not bring us back where we started! If we rotate by 4π it would. So we say spinors are spin $\frac{1}{2}$. What does this mean physically? I have no idea. There are plenty of physical consequences of spinors being spin $\frac{1}{2}$, but this business of rotating by 2π is not actually a physical thing you can do.

As an aside, note that the factor of 2 is determined by the normalization of the matrices, which is set by the Lie Algebra. For each representation by itself, this would be arbitrary, but it is important for expressions which combine the representations to be invariant.

11.7 Lorentz Invariants

The γ matrices themselves transform nicely under the Lorentz group.

$$\Lambda_s^{-1}\gamma_\mu\Lambda_s = (\Lambda_V)_{\mu\nu}\gamma_\nu \tag{11.88}$$

where the Λ_s are the Lorentz transformations acting on each γ_μ individually, as a matrix, and the Λ_V is the vector representation which mixes up the Lorentz indices. That is, writing out the matrix indices $\gamma_\mu^{\alpha\beta}$, this means

$$(\Lambda_s^{-1})_{\delta\alpha}\gamma_\mu^{\alpha\beta}(\Lambda_s)_{\beta\gamma} = (\Lambda_V)_{\mu\nu}\gamma_\nu^{\alpha\beta} \tag{11.89}$$

where μ refers to which γ matrix, and α and β index the elements of that matrix.

Then the equation

$$\{\gamma_\mu, \gamma_\nu\} = 2\eta_{\mu\nu} \tag{11.90}$$

really means

$$\gamma_\mu^{\alpha\gamma}\gamma_\nu^{\gamma\beta} + \gamma_\nu^{\alpha\gamma}\gamma_\mu^{\gamma\beta} = 2\eta_{\mu\nu}\delta^{\alpha\beta} \tag{11.91}$$

And the equation

$$S_{\mu\nu} = \frac{i}{4}[\gamma_\mu, \gamma_\nu] \tag{11.92}$$

Should really be written as

$$S_{\mu\nu}^{\alpha\beta} = \frac{i}{4}\left(\gamma_\mu^{\alpha\gamma}\gamma_\nu^{\gamma\beta} - \gamma_\nu^{\alpha\gamma}\gamma_\mu^{\gamma\beta}\right) \tag{11.93}$$

For an expression like

$$V_\mu V_\mu = V_\mu\eta_{\mu\nu}V_\nu = V_\mu\{\gamma_\mu, \gamma_\nu\}V_\nu \tag{11.94}$$

To be invariant, it must be that μ transforms in the vector representation.

Next consider

$$\psi^\dagger\psi \to \Big(\psi^\dagger\Lambda_s^\dagger\Big)(\Lambda_s\psi) \tag{11.95}$$

For this to be invariant, we would need $\Lambda_s^\dagger = \Lambda_s^{-1}$, that is, for the representation of the Lorentz group to be unitary. The spinor representation, like any other representation is *not* unitary, because the boost generators are anti-Hermetian.

It is useful to study the properties of the Lorentz generators from the Dirac algebra itself, without needing to choose a particular basis for the γ_μ. First note that

$$\{\gamma_\mu, \gamma_\nu\} = 2\eta_{\mu\nu} \quad \Rightarrow \quad \gamma_0^2 = 1, \quad \gamma_i^2 = -1 \tag{11.96}$$

So the eigenvalues of γ_0 are ± 1 and the eigenvalues of γ_i are $\pm i$. Thus if we diagonalize γ_0, we would find it Hermetian, and if we diagonalize γ_1 γ_2 or γ_3 we would find they are anti-Hermetian. This is true in general.

$$\gamma_0^\dagger = \gamma_0, \quad \gamma_i^\dagger = -\gamma_i \tag{11.97}$$

Then,

$$(S_{\mu\nu})^\dagger = \left(\frac{i}{4}[\gamma_\mu, \gamma_\nu]\right)^\dagger = -\frac{i}{4}\Big[\gamma_\nu^\dagger, \gamma_\mu^\dagger\Big] = \frac{i}{4}\Big[\gamma_\mu^\dagger, \gamma_\nu^\dagger\Big] \tag{11.98}$$

Which implies

$$S_{ij}^\dagger = S_{ij} \quad S_{0i}^\dagger = -S_{0i} \tag{11.99}$$

Again, we see that the rotations are unitary and the boosts are not. You can see this from the explicit representations above. But because we showed it algebraically, it is true in ANY representation.

Now, observe that *one* of the Dirac matrices is Hermetian, γ_0 (γ_0 is the only Hermetian Dirac matrix because the metric signature is (1,-1,-1-,1)). Moreover

$$\gamma_0\gamma_i\gamma_0 = -\gamma_i = \gamma_i^\dagger, \quad \gamma_0\gamma_0\gamma_0 = \gamma_0 = \gamma_0^\dagger \tag{11.100}$$

$$\Rightarrow \quad \gamma_\mu^\dagger = \gamma_0\gamma_\mu\gamma_0 \tag{11.101}$$

$$\Rightarrow \gamma_0 S_{\mu\nu}^\dagger\gamma_0 = \gamma_0\frac{i}{4}\Big[\gamma_\mu^\dagger, \gamma_\nu^\dagger\Big]\gamma_0 = \frac{i}{4}\Big[\gamma_0\gamma_\mu^\dagger\gamma_0, \gamma_0\gamma_\nu^\dagger\gamma_0\Big] = \frac{i}{4}[\gamma_\mu, \gamma_\nu] = S_{\mu\nu} \tag{11.102}$$

$$(\gamma_0\Lambda_s\gamma_0)^\dagger = \gamma_0\exp(i\theta_{\mu\nu}S_{\mu\nu})^\dagger\gamma_0 = \exp(-i\theta_{\mu\nu}\gamma_0 S_{\mu\nu}^\dagger\gamma_0) = \exp(-i\theta_{\mu\nu}S_{\mu\nu}) = \Lambda_s^{-1} \tag{11.103}$$

Then, finally,

$$\psi^\dagger\gamma_0\psi \to \Big(\psi^\dagger\Lambda_s^\dagger\Big)\gamma_0(\Lambda_s\psi) = \Big(\psi^\dagger\gamma_0\Lambda_s^{-1}\Lambda_s\psi\Big) = \psi^\dagger\gamma_0\psi \tag{11.104}$$

which is Lorentz invariant.

We have just been rederiving from the Dirac algebra point of view what we found by hand from the Weyl point of view. The conjugate of ψ is not $\psi^\dagger$ but

$$\bar\psi \equiv \psi^\dagger\gamma_0 \tag{11.105}$$

The point is that $\bar\psi$ transforms according to Λ_s^{-1}. Thus $\bar\psi\psi$ is Lorentz invariant.

We can also construct objects like like

$$\bar\psi\gamma_\mu\psi, \quad \bar\psi\gamma_\mu\gamma_\nu\psi, \quad \bar\psi\partial_\mu y \tag{11.106}$$

all transform nicely under the Lorentz group. Also

$$\mathcal{L} = \bar\psi\,(i\gamma^\mu\partial_\mu - m)\psi \tag{11.107}$$

is Lorentz invariant. We abbreviate this with

$$\mathcal{L} = \bar\psi\,(i\not\partial - m)\psi \tag{11.108}$$

This is the Dirac Lagrangian.

The Dirac equation follows from this Lagrangian by the equations of motion

$$(i\not{\partial} - m)\psi = 0 \tag{11.109}$$

To be explicit, this is shorthand for

$$(i\gamma_\mu^{\alpha\beta}\partial_\mu - m\delta^{\alpha\beta})\psi_\beta = 0 \tag{11.110}$$

By multiplying on the right, we find

$$0 = (i\not{\partial} + m)(i\not{\partial} - m)\psi = (-\frac{1}{2}\partial_\mu\partial_\nu\{\gamma_\mu, \gamma_\nu\} - \frac{1}{2}\partial_\mu\partial_\nu[\gamma_\mu, \gamma_\nu] - m^2)\psi \tag{11.111}$$

$$= -(\partial^2 + m^2)\psi \tag{11.112}$$

So ψ satisfies the Klein-Gordon equation

$$(\Box + m^2)\psi = 0 \tag{11.113}$$

It is in this sense that people sometimes say the Dirac equation is the square-root of the Klein-Gordon equation.

We can integrate the Lagrangian by parts to derive the equations of motion for $\bar{\psi}$:

$$\mathcal{L} = \bar{\psi}(i\not{\partial} - m)\psi = -i\partial_\mu\bar{\psi}(\gamma_\mu - m)\psi \tag{11.114}$$

So,

$$-i\partial_\mu\bar{\psi}\gamma^\mu - m\bar{\psi} = 0 \tag{11.115}$$

This γ_μ on the right is a little annoying, so we often hide it by writing

$$\bar{\psi}(-i\overleftarrow{\not{\partial}} - m) = 0 \tag{11.116}$$

where the derivative acts to the left. This makes the conjugate equation look more like the original Dirac equation.

11.8 Coupling to the photon

Under a gauge transform ψ transforms just like a scalar, since the gauge invariance has nothing to do with spin. So,

$$\psi \to e^{i\alpha}\psi \tag{11.117}$$

Then we use the same covariant derivative $\partial_\mu - ieA_\mu$ as for a scalar. So

$$D_\mu\psi = (\partial_\mu - ieA_\mu)\psi \tag{11.118}$$

Then the Dirac equation becomes

$$(i\not{\partial} + e\not{A} - m)\psi = 0 \tag{11.119}$$

Now we try to reproduce the Klein Gordon equation for a scalar field coupled to A_μ.

$$[(i\partial_\mu + eA_\mu)^2 - m^2]\phi \tag{11.120}$$

Following the same route as before

$$0 = (i\not{\partial} + e\not{A} + m)(i\not{\partial} + e\not{A} - m)\psi \tag{11.121}$$

$$= [(i\partial_\mu + eA_\mu)(i\partial_\nu + eA_\nu)\gamma^\mu\gamma^\nu - m^2]\psi \tag{11.122}$$

$$= \left(\frac{1}{4}\{i\partial_\mu + eA_\mu, i\partial_\nu + eA_\nu\}\{\gamma^\mu, \gamma^\nu\} + \frac{1}{4}[i\partial_\mu + eA_\mu, i\partial_\nu + eA_\nu][\gamma^\mu, \gamma^\nu] - m^2\right)\psi \tag{11.123}$$

Before the antisymmetric combination dropped out, but now we find

$$\frac{1}{4}[i\partial_\mu + e\,A_\mu, i\partial_\nu + e\,A_\nu] = e[i\partial_\mu A_\nu - i\partial_\nu A_\mu] = e\,i\,F_{\mu\nu} \tag{11.124}$$

So we get

$$\left((i\partial_\mu + e\,A_\mu)^2 + \frac{i\,e}{4}F_{\mu\nu}[\gamma^\mu, \gamma^\nu] - m^2\right)\psi \tag{11.125}$$

Which contains an extra term compared to the Klein-Gordon equation.

What is this term? Well, recall that $\frac{i}{4}[\gamma^\mu, \gamma^\nu] = S^{\mu\nu}$, our Lorentz generators. These have the form (in the Weyl representation)

$$S_{ij} = \frac{1}{2}\varepsilon_{ijk}\begin{pmatrix} \sigma_k & \\ & \sigma_k \end{pmatrix}, \quad S_{0i} = -\frac{i}{2}\begin{pmatrix} \sigma_i & \\ & -\sigma_i \end{pmatrix} \tag{11.126}$$

And since

$$F_{0i} = E_i, \quad F_{ij} = \varepsilon_{ijk}B_k \tag{11.127}$$

We get

$$\left\{(\partial_\mu - i\,e\,A_\mu)^2 + m^2 - \frac{e}{2}\begin{pmatrix} (\vec{B} + i\,\vec{E})\vec{\sigma} & \\ & (\vec{B} - i\vec{E})\vec{\sigma} \end{pmatrix}\right\}\psi = 0 \tag{11.128}$$

This corresponds to a magnetic dipole moment.

This is pretty remarkable. For a free spinor, we reproduce the equation of motion of a scalar field. But when the spinor is coupled to the photon, we find an additional interaction corresponds to a magnetic dipole moment. We can read off that the electron has spin $\frac{1}{2}$. Note: the coupling to the electric field is not an electric dipole moment – that would not have an i, but is simply the effect of a magnetic moment in a boosted frame.

11.9 Weyl spinors – Helicity

Dirac spinors, what we have been using, are 4 component complex fields in the $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ representation of the Lorentz group. Let's return for a moment to thinking about the 2-component fields.

In the Weyl basis, the γ matrices have the form

$$\gamma_\mu = \begin{pmatrix} 0 & \sigma_\mu \\ \bar{\sigma}_\mu & 0 \end{pmatrix} \tag{11.129}$$

and Lorentz generators are block diagonal

$$\delta\psi = \begin{pmatrix} (i\theta_i + \beta_i)\sigma_i & \\ & (i\theta_i - \beta_i)\sigma_i \end{pmatrix}\psi \tag{11.130}$$

We can write our spinor as the left and right helicity parts

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{11.131}$$

The Dirac equation is

$$\begin{pmatrix} -m & i\sigma_\mu D_\mu \\ i\bar{\sigma}_\mu D_\mu & -m \end{pmatrix}\begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \tag{11.132}$$

meaning

$$(i\bar{\sigma}_\mu\partial_\mu + e\bar{\sigma}_\mu A_\mu)\psi_R = m\psi_L \tag{11.133}$$

$$(i\sigma_\mu\partial_\mu + e\sigma_\mu A_\mu)\psi_L = m\psi_R \tag{11.134}$$

So the electron mass mixes the left and right handed states.

In the absence of a mass, this implies

$$0 = i\sigma_\mu \partial_\mu \psi_R = (E + \vec{\sigma}\cdot\vec{p})\psi_R \tag{11.135}$$

$$0 = i\bar{\sigma}_\mu \partial_\mu \psi_L = (E - \vec{\sigma}\cdot\vec{p})\psi_L \tag{11.136}$$

So the left and right handed states are eigenstates of the operator $\vec{\sigma}\cdot\vec{p}$ with opposite eigenvalue. This operator projects the spin on the momentum direction. We call spin projected on the direction of motion the helicity, so the left and right handed states have opposite helicity. Since the mass mixes left and right handed states, massive fermions are not eigenstates of helicity. This makes ssene – in the electron's rest frame, there is no direction of motion so helicity is not well definedl in the massless limit, there is no rest frame, and a left-handed electron can never turn into a right-handed electron, and a right-handed electron can never turn into a left-handed electron, even in the presence of electromagnetic fields.

The fact that projection of spin on the direction of momentum is a good quantum number for massless particles works for massless particles of any spin. For any spin, we will always find $(E \pm s\vec{J}\vec{p})\Psi_s = 0$, where $\vec{J}$ are the rotation generators of spin s. For spin $1/2$, $\vec{J} = \frac{\vec{\sigma}}{2}$. For photons, the rotation generators are listed in section 2. For example, $J_z = V_{23}$ has eigenvalues ± 1 with eigenstates $(0, i, 1, 0)$ and $(0, -i, 1, 0)$. These are the states of circularly polarized light in the z direction. They are helicity eigenstates. So massless particles always have two helicity states. It is true for spin $1/2$ and spin 1, as we have seen, it is true for gravitons (spin 2), Rarita-Schwing fields (spin $3/2$) and spins $s > 2$ (although, as we have seen, it is impossible to have interacting theories with massless fields of spin $s > 2$).

We have seen that the ψ_L and ψ_R states

- do not mix under Lorentz Transformation
- ψ_L and ψ_R each have two components on which the $\sigma' s$ act. These are the two spin states of the electron – both left and right handed spinors have 2 spin states.
- ψ_L and ψ_R have opposite helicity.

Using

$$\bar{\psi} = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}^\dagger \gamma_0 = \begin{pmatrix} \psi_R^\dagger & \psi_L^\dagger \end{pmatrix} \tag{11.137}$$

The Lagrangian

$$\mathcal{L} = \bar{\psi}\,(i\gamma^\mu \partial_\mu + e\gamma^\mu A_\mu - m)\psi \tag{11.138}$$

becomes

$$\mathcal{L} = i\psi_L^\dagger \bar{\sigma}^\mu D_\mu \psi_L + i\psi_R^\dagger \sigma^\mu D_\mu \psi_R - m(\psi_L^\dagger \psi_R + \psi_R^\dagger \psi_L) \tag{11.139}$$

Which is what we derived in the beginning. Note that ψ_L and ψ_R by themselves must be massless . To right down a mass term, we need both a ψ_L and a ψ_R.

It is helpful to be able to project out the left or right handed Weyl spinors from a Dirac spinor. We can do that with the γ_5 matrix

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{11.140}$$

in the Weyl representation

$$\gamma^5 = \begin{pmatrix} -1 & \\ & 1 \end{pmatrix}, \tag{11.141}$$

So

$$P_R = \frac{1+\gamma^5}{2} = \begin{pmatrix} 0 & \\ & 1 \end{pmatrix}, \quad P_L = \frac{1-\gamma^5}{2} = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix} \tag{11.142}$$

These are projection operators since $P_R^2 = P_R$ and $P_L^2 = P_L$ and

$$P_R \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \begin{pmatrix} 0 \\ \psi_R \end{pmatrix}, \quad P_L \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \begin{pmatrix} \psi_L \\ 0 \end{pmatrix} \tag{11.143}$$

These projectors are useful because they are basis-independent.

11.10 Spin and statistics

Recall that we couldn't write down a mass term $\psi_L^\dagger\psi_L$ for just a left handed spinor because

$$\delta\psi_L = \frac{1}{2}(i\theta_i + \beta_i)\sigma_i\psi_L \tag{11.144}$$

$$\delta\psi_L^\dagger = \frac{1}{2}(-i\theta_i + \beta_i)\psi_L^\dagger\sigma_i \tag{11.145}$$

$$\delta\psi_L^\dagger\psi_L = \beta_i\psi_L^\dagger\psi_L \neq 0 \tag{11.146}$$

This just indicates that the boosts are not unitary transformations.

To get something new, recall that for the Pauli matrices, σ_1 and σ_3 are real, and σ_2 is imaginary.

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{11.147}$$

We can see that

$$\sigma_1^\star = \sigma_1, \quad \sigma_2^\star = -\sigma_2, \quad \sigma_3^\star = \sigma_3 \tag{11.148}$$

$$\sigma_1^T = \sigma_1, \quad \sigma_2^T = -\sigma_2, \quad \sigma_3^T = \sigma_3 \tag{11.149}$$

For the $i = 1$ and $i = 3$, $\sigma_i^T\sigma_2 = \sigma_i\sigma_2 = -\sigma_2\sigma_1$. For $i = 2$, $\sigma_i^T\sigma_2 = -\sigma_2\sigma_2 = -\sigma_2\sigma_i$. So

$$\sigma_i^T\sigma_2 = -\sigma_2\sigma_i \tag{11.150}$$

Then

$$\delta(\psi_L^T\sigma_2) = \frac{1}{2}(i\theta_i + \beta_i)\psi_L^T\sigma_i^T\sigma_2 = \frac{1}{2}(-i\theta_i - \beta_i)\big(\psi_L^T\sigma_2\big)\sigma_i \tag{11.151}$$

Which cancels the transformation property of ψ_L. Thus

$$\psi_L^T\sigma_2\psi_L \tag{11.152}$$

is Lorentz invariant.

The only problem is that $\sigma_2 = \begin{pmatrix} & -i \\ i & \end{pmatrix}$ so this is just

$$\psi_L^T\sigma_2\psi_L = \begin{pmatrix} \psi_1 & \psi_2 \end{pmatrix}\begin{pmatrix} & i \\ -i & \end{pmatrix}\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = i(\psi_1\psi_2 - \psi_2\psi_1) \tag{11.153}$$

This is true for any spinor components ψ_1 and ψ_2. So,

$$\boxed{\psi_1\psi_2 - \psi_2\psi_1 \text{ is Lorentz invariant}} \tag{11.154}$$

This is an example of the **spin-statistics theorem**. I think it's the simplest way to see the connection between anti-symmetrization and Lorentz invariance.

You know it from quantum mechanics. Fermions come in anti-symmetric states. It may look more familiar if we use arrows for the ψ_1 and ψ_2 states:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle) \tag{11.155}$$

We see that wavefunctions are antisymmetric to the exchange of particles. This is the Pauli exclusion principle. Here we see that it follows simply from Lorentz invariance.

Obviously if the fermion components commute this is just zero. So fermion components can't be regular numbers, they must be anticommuting numbers. Such things are called *Grassman* numbers and satisfy a Grassman algebra. We will see that it's easy to make the quantum fields anti-commute, by demanding that the creation and annihilation operators *anti-commute*, $a_p^\dagger a_q^\dagger = -a_q^\dagger a_p^\dagger$,. And we'll rederive spin-statistics from properties of the S-matrix. But here we are seeing that if we are to make any sense of spinors classically, or in any way at all, they *must* anticommute.

11.11 Majorana fermions

If we allow fermions to be Grassman numbers, then we can write down a Lagrangian for a single Weyl spinor with a mass term

$$\mathcal{L} = i\psi_L^\dagger \sigma_\mu \partial_\mu \psi_L + i\frac{m}{2}(\psi_L^\dagger \sigma_2 \psi_L^\star - \psi_L^T \sigma_2 \psi_L) \tag{11.156}$$

These kinds of mass terms are called Majorana masses. Note that this mass term breaks the symmetry under $\psi \to e^{i\alpha}\psi$, since

$$\psi_L^T \sigma_2 \psi_L \to \psi_L^T e^{i\alpha} \sigma_2 e^{i\alpha} \psi_L = e^{2i\alpha} \psi_L^T \sigma_2 \psi_L \tag{11.157}$$

So a particle with a Majorana mass cannot be charged.

The equation of motion for the Majorana fermion is

$$\sigma_\mu \partial_\mu \psi_L + m\sigma_2 \psi_L^\star = 0 \tag{11.158}$$

which follows from the Lagrangian above.

If we have a Majorana fermion, we can still use the Dirac algebra to describe it, but we have to put it into a 4-component spinor

$$\psi = \begin{pmatrix} \psi_L \\ -i\sigma_2 \psi_L^\star \end{pmatrix} \tag{11.159}$$

This transforms like a Dirac spinor because $\sigma_2 \psi_L^\star$ transforms like ψ_R.

Since (in the Weyl basis)

$$-i\gamma_2 \psi^\star = -i \begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix} \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}^\star = \begin{pmatrix} i\sigma_2 \psi_R^\star \\ -i\sigma_2 \psi_L^\star \end{pmatrix} \tag{11.160}$$

We can see that a Majorana fermion is one satisfying

$$\psi = \psi_c \equiv -i\gamma_2 \psi^\star \tag{11.161}$$

We call ψ_c the charge conjugate fermion. A Majorana fermion is it's own charge conjugate. Since it is real, it is als its own antiparticle.

Finally, note that in the Weyl basis γ_2 is imaginary and γ_0, γ_1 and γ_3 are real. Of course, we we could just as well have taken γ_3 imaginary and γ_2 real, but it's conventional to pick out γ_2. We can also define a new representation of the γ matrices by $\hat{\gamma}_\mu = \gamma_2 \gamma_\mu^\star \gamma_2$. This satisfies the Dirac Algebra because $\gamma_2^2 = -1$. Now define

$$\psi_c = -i\gamma_2 \psi^\star \quad \Leftrightarrow \quad \psi^\star = -i\gamma_2 \psi_c \tag{11.162}$$

If we take the Dirac equation

$$(i\not{\partial} + e\not{A} - m)\psi = 0 \tag{11.163}$$

and take the complex conjugate we get

$$(-i\gamma_\mu^\star \partial_\mu + e\gamma_\mu^\star A_\mu - m)\psi^\star = 0 \tag{11.164}$$

$$\Rightarrow \quad \gamma_2(-i\gamma_\mu^\star \partial_\mu + e\gamma_\mu^\star A_\mu - m)\gamma_2 \psi_c = 0 \tag{11.165}$$

$$\Rightarrow \quad (i\hat{\gamma}_\mu \partial_\mu - e\hat{\gamma}_\mu A_\mu - m)\psi_c = 0 \tag{11.166}$$

So ψ_c has the oppisite charge from ψ, which is another reason that Majorana fermions can't be charged.

11.11.1 summary

In summary,

We have seen three types of spinors

- Dirac Spinors: massive, left AND right handed
- Weyl Spinors: massless, left OR right handed

- Majorana spinors: real constrained Dirac spinors.

We are not sure if the neutrino is a Majorana or a Dirac spinor. Weyl spinors play a very important role in supersymmetry, and in the theory of the Weak interactions. But for QED, we can just stick with Dirac spinors.

11.12 Solving the Dirac equation

Let's take a break from the Dirac equation for a moment, and recall how we discovered antiparticles for complex scalar fields. The Lagrangian was

$$\mathcal{L} = \frac{1}{2}[(\partial_\mu + ieA_\mu)\phi^\star][(\partial_\mu - ieA_\mu)\phi] + m^2\phi^\star\phi \tag{11.167}$$

$$\mathcal{L} = \phi^\star(-\partial_\mu + ieA_\mu)(\partial_\mu - ieA_\mu)\phi] + m^2\phi^\star\phi \tag{11.168}$$

The equations of motion are

$$(\partial_\mu - ieA_\mu)^2\phi + m^2\phi = 0 \tag{11.169}$$

$$(\partial_\mu + ieA_\mu)^2\phi^\star + m^2\phi^\star = 0 \tag{11.170}$$

So we see that ϕ and $\phi^\star$ have opposite charge, and we interpret them as particle and antiparticle. Recall that when we quantized the field ϕ, it created a particle and destroyed and antiparticle, and vice-versa for $\phi^\star$. But at the classical level, we can just think of ϕ as particle and $\phi^\star$ as antiparticle.

How do we know if we have a particle or an antiparticle? There is an easy way to answer this. The free equations for ϕ and $\phi^\star$ are the same:

$$(\Box + m^2)\phi = (\Box + m^2)\phi^\star = 0 \tag{11.171}$$

These have plane wave solutions

$$\phi = \phi_p e^{ip_\mu x^\mu} \tag{11.172}$$

In the rest frame, $p_0^2 = m^2$, so $p_0 = \pm m$. The solution with $p_0 = -m$ is confusing. It is a legitimate solution to the equation of motion, but it says that these particles are going backward in time! But note that

$$\phi = \phi_p e^{ip_0 t} \quad \Leftrightarrow \quad \phi^\star = \phi_p^\star e^{-ip_0 t} \tag{11.173}$$

So we can just as easily interpret these solutions as anti-particles going forward in time. Obviously this interpretation is easier to swallow, but Feynman spent some time showing that there really is no physically distinguishable difference.

Now back to spinors. The Dirac equation is

$$(i\not{\partial} + e\not{A} - m)\psi = 0 \tag{11.174}$$

$$\bar{\psi}(-i\overleftarrow{\not{\partial}} - e\not{A} - m) = 0 \tag{11.175}$$

So $\bar{\psi}$ is a particle with mass m and opposite charge to ψ: the positron.

11.12.1 free field solutions

Let's expand the Dirac spinor in plane wave solutions

$$\psi_p(x) = u_\beta(p)e^{ipx} \tag{11.176}$$

u_β and v_β are the polarizations. With this Ansatz, the Dirac equation becomes

$$\begin{pmatrix} -m & p_\mu\sigma_\mu \\ p_\mu\bar{\sigma}_\mu & -m \end{pmatrix} u_\beta(p) = 0 \tag{11.177}$$

Since the spinor also satisfies $(\Box + m^2)\psi = 0$, it follows that $p^2 = m^2$. We already know that the negative frequency solutions $p_0 = -m$ are not a problem since the antiparticles

$$\bar{\psi} = \bar{u}e^{-ipx} \tag{11.178}$$

have positive frequency. Still, it is somewhat convenient to keep talking about $p_0 < 0$ for now. For $p_0 < 0$ we call the polarizations $v_\beta(p)$. Then the solutions are

$$\psi_{p,s}(x) = u_s(p)e^{ipx}, \quad p_0 > 0 \tag{11.179}$$

$$\psi_{p,s}(x) = v_s(p)e^{-ipx}, \quad p_0 < 0 \tag{11.180}$$

u_s and v_s are the polarizations and $s = 1, 2$ labels the spin.

In the rest frame, $p = (m, 0, 0, 0)$, the equations of motion reduce to

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} u_s = 0 \quad \text{and} \quad \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} v_s = 0 \tag{11.181}$$

So solutions are constants

$$u_s = \begin{pmatrix} \xi_s \\ \xi_s \end{pmatrix} \tag{11.182}$$

and

$$v_s = \begin{pmatrix} \eta_s \\ -\eta_s \end{pmatrix} \tag{11.183}$$

For any two-component spinors ξ and η. Thus for $s = 1$ we can take the basis spinor to be $\xi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $s = 2$ may be $\xi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Now let's boost in the z-direction. Peskin and Schroeder do the actual boost. But we'll just solve the equations again in the boosted frame and match the normalization. If $p = (E, 0, 0, p_z)$ then

$$p^\mu \sigma^\mu = \begin{pmatrix} E - p_z & 0 \\ 0 & E + p_z \end{pmatrix}, \quad p_\mu \bar{\sigma}_\mu = \begin{pmatrix} E + p_z & 0 \\ 0 & E - p_z \end{pmatrix} \tag{11.184}$$

Let $x = \sqrt{E - p_z}$ and $y = \sqrt{E + p_z}$, then $m^2 = (E - p_z)(E + p_z) = xy$.

Then the Dirac equation is

$$(i\not{\partial} - m)\psi = \begin{pmatrix} -xy & 0 & x^2 & 0 \\ 0 & -xy & 0 & y^2 \\ y^2 & 0 & -xy & 0 \\ 0 & x^2 & 0 & -xy \end{pmatrix} u_\beta(p) = 0 \tag{11.185}$$

The solutions are

$$u_s = u_\xi(p) = \frac{1}{\sqrt{m}} \begin{pmatrix} x\xi_1 \\ y\xi_2 \\ y\xi_1 \\ x\xi_2 \end{pmatrix} = \frac{1}{\sqrt{m}} \begin{pmatrix} \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} \xi \\ \begin{pmatrix} y & 0 \\ 0 & x \end{pmatrix} \xi \end{pmatrix} \tag{11.186}$$

Check

$$-xy + x^2 y^2 = 0, \quad \cdots \tag{11.187}$$

Note also that in the rest frame $p_z = 0$ in the rest frame, $x^2 = y^2 = m$ so we are back to the ξ states.

So the solutions in the p_z frame are

$$u_s(p) = \begin{pmatrix} \begin{pmatrix} \sqrt{E - p_z} & 0 \\ 0 & \sqrt{E + p_z} \end{pmatrix} \xi_s \\ \begin{pmatrix} \sqrt{E + p_z} & 0 \\ 0 & \sqrt{E - p_z} \end{pmatrix} \xi_s \end{pmatrix}, \quad v_s(p) = \begin{pmatrix} \begin{pmatrix} \sqrt{E - p_z} & 0 \\ 0 & \sqrt{E + p_z} \end{pmatrix} \eta_s \\ \begin{pmatrix} \sqrt{E + p_z} & 0 \\ 0 & \sqrt{E - p_z} \end{pmatrix} \eta_s \end{pmatrix} \tag{11.188}$$

Using

$$\sqrt{p\cdot\sigma}=\begin{pmatrix}\sqrt{E-p_z} & 0\\ 0 & \sqrt{E-p_z}\end{pmatrix},\quad \sqrt{p\cdot\bar{\sigma}}=\begin{pmatrix}\sqrt{E+p_z} & 0\\ 0 & \sqrt{E-p_z}\end{pmatrix} \tag{11.189}$$

We can write more generally

$$u_s(p)=\begin{pmatrix}\sqrt{p\cdot\sigma}\,\xi_s\\ \sqrt{p\cdot\bar{\sigma}}\,\xi_s\end{pmatrix},\quad v_s(p)=\begin{pmatrix}\sqrt{p\cdot\sigma}\,\eta_s\\ -\sqrt{p\cdot\bar{\sigma}}\,\eta_s\end{pmatrix} \tag{11.190}$$

where the square root of a matrix means diagonalize it and then take the square root. In practice, we will always pick p along the z axis, so we don't really need to know how to make sense of these formulas. But they are useful because they are explicitly Lorentz covariant so we can manipulate them cleanly.

11.12.2 normalization and spin sums

We have chosen a normalization here, which amounts to

$$\bar{\psi}\psi'=u_s^\dagger(p)\gamma_0 u_{s'}(p)=\begin{pmatrix}\sqrt{p\cdot\sigma}\,\xi_s\\ \sqrt{p\cdot\bar{\sigma}}\,\xi_s\end{pmatrix}^\dagger\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\sqrt{p\cdot\sigma}\,\xi_{s'}\\ \sqrt{p\cdot\bar{\sigma}}\,\xi_{s'}\end{pmatrix} \tag{11.191}$$

$$=2\sqrt{\begin{pmatrix}E-p_z & 0\\ 0 & E+p_z\end{pmatrix}\begin{pmatrix}E+p_z & 0\\ 0 & E-p_z\end{pmatrix}}\,\xi_s^\dagger\xi_{s'} \tag{11.192}$$

$$=2m\delta_{ss'} \tag{11.193}$$

We taken $\xi_s^\dagger\xi_{s'}=\delta_{ss'}$ by convention. This is the normalization for the *inner product*.

For $m=0$ this way of writing the condition is not useful. Instead we can just set the normalization in the frame of the particle;s direction by normalizing

$$u_s^\dagger u_s=\begin{pmatrix}\sqrt{p\cdot\sigma}\,\xi_s\\ \sqrt{p\cdot\bar{\sigma}}\,\xi_s\end{pmatrix}^\dagger\begin{pmatrix}\sqrt{p\cdot\sigma}\,\xi_{s'}\\ \sqrt{p\cdot\bar{\sigma}}\,\xi_{s'}\end{pmatrix}=2E\xi_s^\dagger\xi_{s'}=2E\delta_{ss'} \tag{11.194}$$

which has the same $\sqrt{2E}$ factors we've been using.

We can also compute the *outer product*

$$\sum_{s=1}^{2}u_s(p)\bar{u}_s(p)=\not{p}+m \tag{11.195}$$

where the sum is over $s=1,2$ corresponding to $\xi=\begin{pmatrix}1\\0\end{pmatrix}$ and $\xi=\begin{pmatrix}0\\1\end{pmatrix}$. For the antiparticles,

$$\sum_{s=1}^{2}v_s(p)\bar{v}_s(p)=\not{p}-m \tag{11.196}$$

It is also true that

$$\sum_{s=1}^{2}u_s\bar{v}_s=\sum_{s=1}^{2}\bar{u}_s v_s=0 \tag{11.197}$$

You should try to verify these relations on your own.

To keep straight the inner and outer products, it may be helpful to compare to spin 1 particles. In that case we had for massive spin 1

$$[\epsilon_\mu^i(p)]^\star\epsilon_\mu^j(p)=-\delta^{ij}\quad\leftrightarrow\quad \bar{u}_s(p)u_{s'}(p)=2m\delta_{ss'}\quad\text{(inner product)} \tag{11.198}$$

$$\sum_{i=1}^{3}[\epsilon_\mu^i(p)]^\star\epsilon_\nu^i(p)=\eta_{\mu\nu}-\frac{p^\mu p^\nu}{m^2}\quad\leftrightarrow\quad \sum_{s=1}^{2}u_s(p)\bar{u}_s(p)=\not{p}+m\quad\text{(outer product)} \tag{11.199}$$

So when we sum Lorentz indices or internal spinor indices, we use an inner product and get a number. When we sum over polarizations/spins, we get a matrix.

11.12.3 charge conjugation

In the rest frame,

$$u_s = \begin{pmatrix} \xi_s \\ \xi_s \end{pmatrix}, \quad v_s = \begin{pmatrix} \eta_s \\ -\eta_s \end{pmatrix} \tag{11.200}$$

where ξ and η are constants, for spin up and spin down.

Let us see how the spin states transform under charge conjugation

$$|\uparrow\rangle^c = \begin{pmatrix} 1 \\ 0 \end{pmatrix}^c = -i\sigma_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix}^c = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -i \end{pmatrix} = -i|\downarrow\rangle \tag{11.201}$$

$$|\downarrow\rangle^c = \begin{pmatrix} 0 \\ 1 \end{pmatrix}^c = -i\sigma_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}^c = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} i \\ 0 \end{pmatrix} = i|\uparrow\rangle \tag{11.202}$$

So charge congugation flips the spin. Thus

$$(u_p^s)^c = \begin{pmatrix} \xi_{-s} \\ -\xi_{-s} \end{pmatrix} \tag{11.203}$$

This will be relevant to studying charge-conjugation invariance.

Chapter 12
Quantum Electrodynamics

12.1 Introduction

We have found the Lagrangian for QED:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + i\bar{\psi}\not{D}\psi - m\bar{\psi}\psi \tag{12.1}$$

In order to compute anything with it in the quantum theory, we first have to quantize the spinors.

12.2 Quantizing Spinors

For complex scalar the field is we had

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\Big(a_p^\dagger e^{ipx} + b_p e^{-ipx}\Big) \tag{12.2}$$

$$\phi^\star(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\Big(b_p^\dagger e^{ipx} + a_p e^{-ipx}\Big) \tag{12.3}$$

which impels that $b_p^\dagger$ creates particles of the opposite charge and same mass, which we called antiparticles.

For the Dirac equation the Lagrangian is

$$\mathcal{L} = \bar{\psi}\,(i\not{D} - m)\,\psi \tag{12.4}$$

And the equations of motion are

$$(i\not{\partial} + e\not{A} - m)\,\psi = 0 \tag{12.5}$$

$$\bar{\psi}\,(-i\overleftarrow{\not{\partial}} - e\not{A} - m) = 0 \tag{12.6}$$

The free field solutions can be labeled by constant two-component spinors ξ and η. The actual solutions depend on momenta in a funny way. Labelling them u^s and v^s where $s = 1, 2$ is the spin, we have

$$u_p^1 = \begin{pmatrix} \sqrt{p\cdot\sigma} \\ 0 \\ \sqrt{p\cdot\bar{\sigma}} \\ 0 \end{pmatrix}, u_p^2 = \begin{pmatrix} 0 \\ \sqrt{p\cdot\sigma} \\ 0 \\ \sqrt{p\cdot\sigma} \end{pmatrix}, v_p^1 = \begin{pmatrix} \sqrt{p\cdot\sigma} \\ 0 \\ -\sqrt{p\cdot\bar{\sigma}} \\ 0 \end{pmatrix}, v_p^2 = \begin{pmatrix} 0 \\ \sqrt{p\cdot\sigma} \\ 0 \\ -\sqrt{p\cdot\sigma} \end{pmatrix} \tag{12.7}$$

The u^s are the positive frequency electrons, and the v^s are negative frequency electrons, or equivalently, positive frequency positrons.

Thus we take

$$\psi(x) = \sum_s \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\left(a_p^{s\,\dagger}u_p^s e^{ipx} + b_p^s v_p^s e^{-ipx}\right) \tag{12.8}$$

$$\bar{\psi}(x) = \sum_s \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\left(b_p^{s\,\dagger}\bar{v}_p^{\,s} e^{ipx} + a_p^s \bar{u}_p^{\,s} e^{-ipx}\right) \tag{12.9}$$

Where as always the energy is always positive and determined by the 3-momentum $\omega_p = \sqrt{\vec{p}^{\,2} + m^2}$.

Note the convention that $\psi(x)$ gets both u and v and $\bar{\psi}$ gets $\bar{u}$ and $\bar{v}$. It might have made sense to use u and $\bar{v}$ in ψ, but since when we found the free states, v were the negative frequency solution, this way of assigning u and v to ψ and $\bar{\psi}$ is consistent. This is just a convention.

12.3 Dirac Propagator

Next, let's calculate the propagator for a Dirac spinor.

$$\left\langle 0|T\{\bar{\psi}(x)\psi(y)\}|0\right\rangle \tag{12.10}$$

Let's first review what happens with a complex scalar we had

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\left(a_p^\dagger e^{ipx} + b_p e^{-ipx}\right) \tag{12.11}$$

$$\phi^\star(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\left(b_p^\dagger e^{ipx} + a_p e^{-ipx}\right) \tag{12.12}$$

Then

$$\langle 0|\phi^\star(x)\phi(0)|0\rangle = \int \frac{d^3p}{(2\pi)^3}\int \frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\frac{1}{\sqrt{2\omega_q}}\langle 0|\left(b_p^\dagger e^{ipx} + a_p e^{-ipx}\right)\left(a_p^{s\,\dagger} + b_p^s\right)|0\rangle$$

$$= \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{-ikx} = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{-i\omega_k t + i\vec{k}\vec{x}} \tag{12.13}$$

$$\langle 0|\phi(0)\phi^\star(x)|0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{ikx} = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{i\omega_k t - i\,\vec{k}\vec{x}} \tag{12.14}$$

Then,

$$\langle 0|T\{\phi^\star(x)\phi(0)\}|0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}\left[e^{i\vec{k}\vec{x} - i\omega_k t}\theta(t) + e^{-i\,\vec{k}\vec{x} + i\omega_k t}\theta(-t)\right] \tag{12.15}$$

$$= \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{i\,\vec{k}\vec{x}}e^{-i\omega_k|t|} \tag{12.16}$$

Then using

$$e^{-i\omega_k|t|} = (2\omega_k)\int_{-\infty}^{\infty}\frac{d\omega}{(2\pi)}\frac{i}{\omega^2 - \omega_k^2 + i\varepsilon}e^{i\omega t} \tag{12.17}$$

We find

$$\langle 0|T\{\phi^\star(x)\phi(0)\}|0\rangle = \int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2 - m^2 + i\varepsilon}e^{ikx} \tag{12.18}$$

Which is the regular scalar propagator.

For a fermion, we start the same way

$$\psi(x) = \sum_s \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\left(a_p^{s\,\dagger}u_p^s e^{ipx} + b_p^s v_p^s e^{-ipx}\right) \tag{12.19}$$

$$\bar{\psi}(x) = \sum_s \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\left(b_p^{s\,\dagger}\bar{v}_p^{\,s} e^{ipx} + a_p^s \bar{u}_p^{\,s} e^{-ipx}\right) \tag{12.20}$$

Then,

$$\langle 0|\psi(0)\bar{\psi}(x)|0\rangle = \int \frac{d^3p}{(2\pi)^3}\int \frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\frac{1}{\sqrt{2\omega_q}}\times$$

$$\sum_{s,s'}\Big\langle 0|\big(a_p^{s\dagger}u_p^s + b_p^s v_p^s\big)\Big(b_q^{s'\dagger}\bar{v}_q^{s'}e^{ipx} + a_q^{s'}\bar{u}_q^{s'}e^{-ipx}\Big)|0\Big\rangle \tag{12.21}$$

$$= \int \frac{d^3p}{(2\pi)^3}\int \frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\frac{1}{\sqrt{2\omega_q}}\sum_{s,s'} v_p^s\bar{v}_q^{s'}\Big\langle 0|b_p^s b_q^{s'\dagger}|0\Big\rangle e^{ipx}$$

$$= \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}\sum_s v_k^s\bar{v}_k^s e^{ikx}$$

Now we sum over polarizations using the outer product

$$\sum_{s=1}^{2} u_s(p)\bar{u}_s(p) = \not{p} + m, \quad \sum_{s=1}^{2} v_s(p)\bar{v}_s(p) = \not{p} - m \tag{12.22}$$

Giving

$$\langle 0|\psi(0)\bar{\psi}(x)|0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{ikx}(\not{k} + m)$$

Similarly

$$\langle 0|\bar{\psi}(x)\psi(0)|0\rangle = \int \frac{d^3p}{(2\pi)^3}\int \frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\frac{1}{\sqrt{2\omega_q}}\langle 0|\big(b_p^s\bar{v}_p^s e^{-ipx} + a_p^{s\dagger}\bar{u}_p^s e^{ipx}\big)\big(a_q^s u_q^s + b_q^{s\dagger}v_q^s\big)|0\rangle$$

$$= \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{-ikx}(\not{k} - m) \tag{12.23}$$

So,

$$\langle 0|\psi(0)\bar{\psi}(x)|0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{ikx}(\not{k} + m) = (-i\not{\partial} + m)\int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{-i\vec{k}\vec{x} + i\omega_k t} \tag{12.24}$$

$$\langle 0|\bar{\psi}(x)\psi(0)|0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{-ikx}(\not{k} - m) = (i\not{\partial} - m)\int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{i\vec{k}\vec{x} - i\omega_k t} \tag{12.25}$$

Then,

$$\langle 0|T\{\bar{\psi}(x)\psi(0)\}|0\rangle = (-i\not{\partial} + m)\int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}\Big[e^{i\vec{k}\vec{x} - i\omega_k t}\theta(t) - e^{-i\vec{k}\vec{x} + i\omega_k t}\theta(-t)\Big] \tag{12.26}$$

$$= (-i\not{\partial} + m)\int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{i\vec{k}\vec{x}}e^{-i\omega_k|t|}[\theta(t) - \theta(-t)] \tag{12.27}$$

$$= (-i\not{\partial} + m)\int \frac{d^4k}{(2\pi)^4}\frac{i}{\omega^2 - \omega_k^2 + i\varepsilon}e^{ikx}(-1)^{\frac{|t|}{t}} \tag{12.28}$$

Wait a minute – this should have simplified. What happened?

We really want that minus sign to be a plus. We can get that by defining the time ordered propagator as

$$T\{\bar{\psi}(x)\psi(0)\} = \bar{\psi}(x)\psi(0)\theta(t) - \psi(0)\bar{\psi}(x)\theta(-t) \tag{12.29}$$

Then,

$$\langle 0|T\{\bar{\psi}(x)\psi(0)\}|0\rangle = (-i\not{\partial} + m)\int \frac{d^4k}{(2\pi)^4}\frac{i}{\omega^2 - \omega_k^2 + i\varepsilon}e^{ikx} \tag{12.30}$$

$$\boxed{\langle 0|T\{\bar{\psi}(x)\psi(0)\}|0\rangle = \int \frac{d^4k}{(2\pi)^4}\frac{i(\not{k} + m)}{k^2 - m^2 + i\varepsilon}} \tag{12.31}$$

This is beautifully Lorentz invariant.

So we want to have the time-ordered product of spinors have an extra minus sign. But then for arbitrary spinors

$$T\{\psi(x)\psi(y)\} = \psi(x)\psi(y)\theta(x_0 - y_0) - \psi(y)\psi(x)\theta(y_0 - x_0) = -T\{\psi(y)\psi(x)\} \tag{12.32}$$

Thus unless we want this to vanish, we must assume that $\psi(y)$ and $\psi(x)$ anticommute.

$$\{\psi(x), \psi(y)\} \equiv \psi(x)\psi(y) + \psi(y)\psi(x) = 0 \tag{12.33}$$

Similarly, fields must anticommute with their anti-fields, up to a delta function

$$\{\psi(x), \bar{\psi}(y)\} = \delta^3(x - y) \tag{12.34}$$

This is equivalent to

$$\{a_p^r, a_q^{s\dagger}\} = \{b_p^r, b_q^{s\dagger}\} = (2\pi)^3\delta^{(3)}(p - q)\delta^{rs} \tag{12.35}$$

So we see that in order to have Lorentz invariant time-ordered products, fermions must anti-commute.

Now consider what happens if we try to construct a state with two of the same field in it:

$$a_p^\dagger a_p^\dagger|0\rangle = -a_p^\dagger a_p^\dagger|0\rangle = 0 \tag{12.36}$$

So we can't! This is the Fermi-exclusion principle, and it follows directly from the anticommutation relations.

This is an example of the spin-statistics theorem. Let's trace back to what happened. We found

$$\langle 0|\phi^\star(x)\phi(0)|0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{ikx} \tag{12.37}$$

$$\langle 0|\phi(0)\phi^\star(x)|0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{-ikx} \tag{12.38}$$

As compared to (for $m = 0$):

$$\langle 0|\bar{\psi}(x)\psi(0)|0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{\not{k}}{2\omega_k}e^{ikx} \tag{12.39}$$

$$\langle 0|\psi(0)\bar{\psi}(x)|0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{\not{k}}{2\omega_k}e^{-ikx} \tag{12.40}$$

Now we can see that the problem is that $\not{k}$ is odd under the rotation which takes $k \to -k$, so that an extra -1 is generated when we try to combine the time-ordered sum for the Fermion. Rotating $k \to -k$ is a rotation by π. We saw that this gives a factor of i in the fermion case. So here we have 2 fermions, and we get a -1. So it's directly related to the spin $\frac{1}{2}$. This will happen for any half integer spin, which gets an extra -1 in the rotation.

Another way to look at it is that they $\not{k}$ factor comes from the polarization sum, which in turn comes from the requirement that the free solutions satisfy $\not{k}u_s(k) = \not{k}v_s(k) = 0$. In fact, we can now see directly that the same problem will happen for any particle of half-integer spin. A particle of spin $n + \frac{1}{2}$ for integer n will have a field with n vector indices and a spinor index, $\psi_{\mu_1\cdots\mu_n}$. So polarization sum must have a factor of γ_μ and the only thing around to contract γ_μ with is it's momentum k_μ. Thus we always get a $\not{k}$, and the time-ordered product can never be Lorentz invariant unless the fields anticommute. These are fermions. They obey Fermi-Dirac statistics.

In contrast for integer spin there can only be an even number of k_μ^2 in the polarization sum (there had better be no $k'_\mu s$ since $k^2 = 0$!). So these fields must commute to have Lorentz invariant time ordered products. These are bosons. They obey Bose-Einstein statistics

12.4 Spin statistics

We saw that to have a Lorentz invariant time-ordered product for fermions, we need them to antisymmetric. That is

$$\psi(x)\psi(y) = -\psi(y)\psi(x) \tag{12.41}$$

or equivalently

$$\left\{a_p^r, a_q^{s\,\dagger}\right\} = (2\pi)^3\delta^{(3)}(p-q)\delta^{rs} \tag{12.42}$$

Acting on the vacuum, this has that multiparticle states are antisymmetric on the exchange of two particles

$$\psi_{\cdots s_1 p_1 \cdots s_2 p_2 \cdots} = \cdots a_{s_1}^\dagger(p_1)\cdots a_{s_2}^\dagger(p_2)\cdots|0\Big\rangle = -\,\psi_{\cdots s_2 p_2 \cdots s_1 p_2 \cdots} \tag{12.43}$$

In general, we could ask what kinds of statistics are possible. (This discussion is taken from section 4.1 of Weinberg).

Suppose we have a state with many identical particles. By identical, we mean that if we swap how the two particles are represented, we still have the same physical state. Only the normalization can change. So

$$\psi_{\cdots s_1 p_1 \cdots s_2 p_2 \cdots} = \alpha\psi_{\cdots s_2 p_2 \cdots s_1 p_2 \cdots} \tag{12.44}$$

where α is a number. This is the definition of *identical particles*.

If we chage back, we find that

$$\psi_{\cdots s_1 p_1 \cdots s_2 p_2 \cdots} = \alpha\psi_{\cdots s_2 p_2 \cdots s_1 p_2 \cdots} = \alpha^2\psi_{\cdots s_1 p_1 \cdots s_2 p_2 \cdots} \tag{12.45}$$

Thus $\alpha^2 = 1$. What could α depend on? Since it is just a number, it cannot depend on the momenta p or the spins s, as there are no non-trivlial one-dimensional representations of the Lorentz group. Thus $\alpha = \pm 1$. We call $\alpha = 1$ bosons and $\alpha = -1$ fermions. As we have shown above, by Lorentz invariance, integer spin particles must be bosons and half-integer spin particles must be fermions.

There is one exception, the factor α can depend on the way p_1 and p_2 are interchanged. For example, in the Aharonov-Bohm effect, you can pick up a phase by rotating a particle around a magnetic flux. This only works in $2+1$ dimensions (2 space and one 1 dimension) where we can have things like fluxes or vortices, and the state can remember how many times it has been rotated around it. Think of a string wrapping around a point in 2D. Then you can get α to depend on the path, that is, the number of times it has gone around the point (the winding number) which leads to particles called *anyons* with *parastatics*. But in 3+1 dimensions, you can't wrap a string around a point – you can always unwind such path-dependence. So only Fermi and Bose statistics are possible.

12.5 Causality

Let's go back to a real scalar field,

$$\phi(x) = \int \frac{d^3q}{(2\pi)^3}\Big(a_q^\dagger e^{iqx} + a_q e^{-iqx}\Big) \tag{12.46}$$

where we saw

$$\phi(x)\phi(y) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_p}e^{-ip(x-y)} \tag{12.47}$$

$$[\phi(x),\phi(y)] = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_p}\Big(e^{-ip(x-y)} - e^{ip(x-y)}\Big) \tag{12.48}$$

Let's work in 2 dimensions, where causality is simple. Then for $y-x = (t,z)$ we get

$$[\phi(x),\phi(y)] = \int \frac{dp}{2\pi}\frac{1}{2\omega_p}\Big(e^{i(\omega_p t - pz)} - e^{-i(\omega_p t - pz)}\Big) \tag{12.49}$$

$$= \int_{-\infty}^{\infty}\frac{dp}{2\pi}\frac{i}{\sqrt{p^2+m^2}}\sin(\sqrt{p^2+m^2}t - pz) \tag{12.50}$$

Now suppose the interval is spacelike, for example $y-x=(0,z)$. Then

$$[\phi(x),\phi(y)]=\int\frac{dp}{2\pi}\frac{-i}{\sqrt{p^2+m^2}}\sin(pz)=0 \tag{12.51}$$

Since $\sin(pz)$ is an odd function of p, this integral vanishes.

Next, take $y-x=(t,0)$ = timelike. Then,

$$[\phi(x),\phi(y)]=\int_{-\infty}^{\infty}\frac{dp}{2\pi}\frac{i}{\sqrt{p^2+m^2}}\sin(\sqrt{p^2-m^2}t) \tag{12.52}$$

Now the integrand is an *even* function of p, so in general the integral will not vanish. It's not so easy to evaluate the integral, but the point is simply that it is not zero. Then we can use the fact that $[\phi(x),\phi(y)]$ is Lorentz covariant. So we find that in general commutators vanish at spacelike separations but not at timelike separation.

$$[\phi(x),\phi(y)]\begin{cases}=0, & x-y\,\text{spacelike}\\ \neq 0, & x-y\,\text{timelike}\end{cases} \tag{12.53}$$

This is true in any number of dimensions.

What does this commutator have to do with physics? Remember, we are just doing quantum mechanics here. So when to operators commute they are simultaneously observable. Another way to say that is that if the operators commute they are uncorrelated and can't influence each other. So here we find that if the we measure the field (remember ϕ measures the field) at $x=0$ it doesn't influence the measurement at x at the same time $t=0$. But if we measure the field at $t=0$ it does affect the field at $t=t$ at the same position x. This is a clear statement of causality.

If you try to do the same thing in non-relativistic quantum mechanics, it will fail. For example, the wavefunction is spread out in x but fixed with respect to t. So if we collapse the wavefunction by measuring it at $x=0$ at $t=0$ it will immediately affect what we can see at $x=x$ and $t=0$.

We can try to do the same calculation for spinors. We find that it is the anticommutator which vanishes

$$\{\psi(x),\psi(0)\}=0 \tag{12.54}$$

In fact,

$$[\psi(x),\psi(0)]\neq 0 \tag{12.55}$$

What do we make of this? It implies if we could ever measure a fermionic wavefunction, we could violate causality. Thus fermions are not observable. Only bosons are observable. After all, all measurements in the end are numbers which are bosons. You would have to measure a Grassman number to see a Fermion. So causality implies that only fermion bilinears have

$$[\bar{\psi}(x)\psi(x),\psi(y)\psi(y)]\begin{cases}=0, & x-y\,\text{spacelike}\\ \neq 0, & x-y\,\text{timelike}\end{cases} \tag{12.56}$$

That is simply because $\bar{\psi}(x)\psi(x)$ is a boson.

12.6 QED Feynman rules

Now we have everything we need for QED. The Lagrangian is

$$\mathcal{L}=-\frac{1}{4}F_{\mu\nu}^2+i\bar{\psi}\not{D}\psi-m\bar{\psi}\psi \tag{12.57}$$

$$=-\frac{1}{4}F_{\mu\nu}^2+\bar{\psi}(i\gamma^\mu\partial_\mu+e\gamma^\mu A_\mu-m)\psi \tag{12.58}$$

We know the propagators. A photon is squiggly line

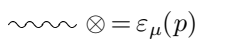
$$= \frac{-i}{k^2+i\varepsilon}\left[\eta_{\mu\nu} - (1-\xi)\frac{k_\mu k_\nu}{k^2}\right] = \frac{-i\eta_{\mu\nu}}{k^2+i\varepsilon}(\text{Feynman gauge})$$

We will usually just choose Feynman gauge, in less we are explicitly checking gauge invariance.

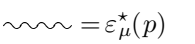
A fermion is a solid line with an arrow indicating particle flow (momentum flow times charge)

$$= \frac{i(\not{p}+m)}{p^2-m^2+i\varepsilon}$$

External photon lines get polarization vectors



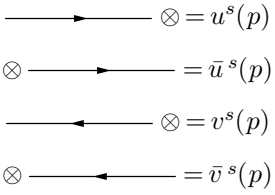
$$= \varepsilon_\mu(p) \quad \text{(incoming)} \tag{12.59}$$



$$= \varepsilon^\star_\mu(p) \quad \text{(outgoing)} \tag{12.60}$$

Here the blob means the rest of the diagram.

External fermion lines get spinors, with us for electrons and vs for positrons.



$$= u^s(p)$$

$$= \bar{u}^s(p)$$

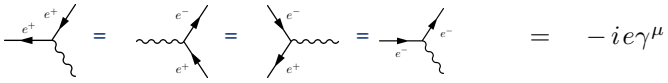
$$= v^s(p)$$

$$= \bar{v}^s(p)$$

The interactions can be read directly off the vertex.

$$V = e\bar{\psi}\gamma^\mu\psi A_\mu \tag{12.61}$$

Since there is no factor of momentum, the sign is the same no matter what is going on at the vertex



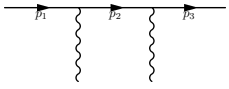
$$= -ie\gamma^\mu$$

The μ on the γ^μ will get contracted with the μ of the photon. This is either in the $\eta_{\mu\nu}$ of the photon propagator, if the photon is internal, or the ϵ_μ polarization tensor if the photon is external.

The $\gamma^\mu = \gamma^\mu_{\alpha\beta}$ as a matrix will also get sandwiched by the two fermion lines, such as

$$\bar{u}\gamma^\mu u = \bar{u}_\alpha \gamma^\mu_{\alpha\beta} u_\beta \tag{12.62}$$

if it's e^-e^- scattering or $\bar{v}\gamma^\mu v$ if it's e^+e^- annihilation. The barred spinor always goes on the left, since the interaction is $\bar{\psi}A_\mu\gamma^\mu\psi$. If there is an internal fermion line between the ends, its propagator goes between the end spinors:



$$= (-ie)^2 \bar{u}(p_3)\gamma^\mu \frac{i(\gamma^\alpha p_2^\alpha + m)}{p_2^2 - m^2 + i\varepsilon}\gamma^\nu u(p_1)\varepsilon_\mu\varepsilon_\nu \tag{12.63}$$

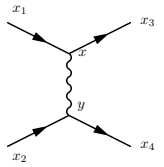
In this example, the 3 γ matrices get multiplied and then sandwiched between the spinors.

If the fermion is a closed loop, then the the spinor indices of the multiplied product get contracted with each other. If we closed the loop in the diagram above, it would be

$$\varepsilon_\mu\varepsilon_\nu\left[\gamma^\mu \frac{i(\gamma^\alpha p_2^\alpha + m)}{p_2^2 - m^2 + i\varepsilon}\gamma^\nu\right]_{\alpha\alpha} = \text{Tr}\left[\gamma^\mu \frac{i(\gamma^\alpha p_2^\alpha + m)}{p_2^2 - m^2 + i\varepsilon}\gamma^\nu\right]\varepsilon_\mu\varepsilon_\nu \tag{12.64}$$

Often computing Feynman diagrams in QED will involve taking the trace of products of γ matrices.

Let us finally be careful about minus signs. A t-channel diagram will have, in terms of position space propogators

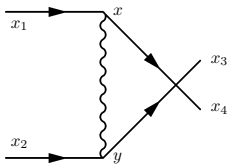


$$= D(x_1, x)D(x, x_3)D(x_2, y)D(y, x_4)\Pi_A(x, y)$$

This corresponds to a particular contraction from the time-ordered product

$$= \langle 0|T\{\Pi_A(x, y)\psi(x_1)\bar{\psi}(x)\psi(x)\bar{\psi}(x_3)\psi(x_1)\bar{\psi}(x)\psi(x)\bar{\psi}(x_4)||0\rangle$$

The u-channel diagram gives the same thing with x_3 and x_4 interchanged



$$= D(x_1, x)D(x, x_4)D(x_2, y)D(y, x_3)$$

which gives

$$= \langle 0|T\{\Pi_A(x, y)\psi(x_1)\bar{\psi}(x)\psi(x)\bar{\psi}(x_4)\psi(x_1)\bar{\psi}(x)\psi(x)\bar{\psi}(x_3)\}|0\rangle$$

This is supposed to correspond to a different contraction of the same time-ordered product. So, using the anti-commutation, we find

$$\cdots\bar{\psi}(x_4)\psi(x_1)\bar{\psi}(x)\psi(x)\bar{\psi}(x_3) = \cdots\bar{\psi}(x_3)\bar{\psi}(x_4)\psi(x_1)\bar{\psi}(x)\psi(x)$$

$$= -\bar{\psi}(x_3)\psi(x_1)\bar{\psi}(x)\psi(x)\bar{\psi}(x_4)$$

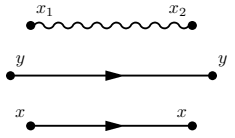
So we get a minus sign in exchanging x_3 and x_4. This is just Fermi statistics. Thus we get an relative minus sign for exchanging two fermions. The overall sign is unimportant, as matrix elements will always be squared.

For loops, we have something like (ignoring the photons)

$$= D(x, x_4)D(x_2, y)$$

$$= \langle 0|T\{\bar{\psi}(x)\psi(y)\bar{\psi}(x)\psi(y)\}|0\rangle \tag{12.65}$$

To get the sign we need to compare it to the product without the loop, which contributes to the vacuum matrix element, but comes from the same time-ordered product



$$= D(x, x)D(y, y)$$

$$= \langle 0|T\{\bar{\psi}(x)\psi(x)\bar{\psi}(y)\psi(y)\}|0\rangle \tag{12.66}$$

$$= -\langle 0|T\{\bar{\psi}(x)\psi(y)\bar{\psi}(y)\psi(x)\}|0\rangle \tag{12.67}$$

So we get an over all minus sign for the loop. This will always happen for any loop with any number of fermions.

Thus

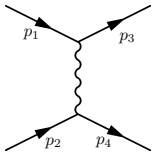
- -1 for interchange of external fermions

- -1 for each fermion loop

12.6.1 Example

For example, let's write down a t-channel diagram for Moller scattering $e^-e^- \to e^-e^-$

$$i\mathcal{M}_t = \quad = (-ie)\bar{u}(p_3)\gamma^\mu u(p_1)\frac{-i[g_{\mu\nu}+(1-\xi)\frac{k_\mu k_\nu}{k^2}]}{k^2}(-ie)\bar{u}(p_4)\gamma^\nu u(p_2)$$



where $k^\mu = p_3^\mu - p_1^\mu$

In Feynman gauge this simplifies too

$$\mathcal{M}_t = (-ie)^2\frac{[\bar{u}(p_3)\gamma^\mu u(p_1)][\bar{u}(p_4)\gamma^\mu u(p_2)]}{t}$$

Which is just a number for particular electron spin states.

Chapter 13
Tree-level QED

13.1 Matrix identities

Before begging the QED calculations, let's derive some useful identities about γ matrices. We will often need to take traces of products of γ matrices. These can be done simply using the cyclic property of the trace

$$\mathrm{Tr}[AB\cdots C] = \mathrm{Tr}[B\cdots CA] \tag{13.1}$$

We will often also use $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$ which satisfies

$$\gamma_5^2 = 1, \quad \gamma_5\gamma_\mu = -\gamma_\mu\gamma_5 \tag{13.2}$$

To keep the spinor indices straight, we will often write $\mathbb{1}$ for the identity on spinor indices. So then

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}\mathbb{1} \tag{13.3}$$

Also,

$$\mathrm{Tr}[\eta^{\mu\nu}\mathbb{1}] = \eta^{\mu\nu}\mathrm{Tr}[\mathbb{1}] = 4\eta^{\mu\nu} \tag{13.4}$$

The $\eta^{\mu\nu}$ are just numbers for each μ and ν and pull out of the trace.

Then

$$\mathrm{Tr}[\gamma^\mu] = \mathrm{Tr}[\gamma_5\gamma_5\gamma^\mu] = \mathrm{Tr}[\gamma_5\gamma^\mu\gamma_5] = -\mathrm{Tr}[\gamma_5\gamma_5\gamma^\mu] = -\mathrm{Tr}[\gamma^\mu] \tag{13.5}$$

$$\Rightarrow \mathrm{Tr}[\gamma^\mu] = 0 \tag{13.6}$$

And

$$\mathrm{Tr}[\gamma^\mu\gamma^\nu] = \mathrm{Tr}[-\gamma^\nu\gamma^\mu + 2\eta^{\mu\nu}] = -\mathrm{Tr}[\gamma^\mu\gamma^\nu] + 8\eta^{\mu\nu} \tag{13.7}$$

$$\Rightarrow \mathrm{Tr}[\gamma^\mu\gamma^\nu] = 4\eta^{\mu\nu} \tag{13.8}$$

In a similar way, you can show

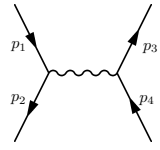
$$\mathrm{Tr}[\gamma^\alpha\gamma^\beta\gamma^\mu] = 0 \tag{13.9}$$

$$\mathrm{Tr}[\gamma^\alpha\gamma^\mu\gamma^\beta\gamma^\nu] = 4(\eta^{\alpha\mu}\eta^{\beta\nu} - \eta^{\alpha\beta}\eta^{\mu\nu} + \eta^{\alpha\nu}\eta^{\mu\beta}) \tag{13.10}$$

you will use this last one a lot! You can remember the signs because adjacent indices give $+$ and the other one gives $-$.

13.2 $e^+e^- \to \mu^+\mu^-$

The simplest tree level scattering process in QED is $e^+e^- \to \mu^+\mu^-$.

$i\mathcal{M} =$ 

$$= (-ie)\bar{v}(p_2)\gamma^\mu u(p_1)\frac{-i[g_{\mu\nu} + (1-\xi)\frac{k_\mu k_\nu}{k^2}]}{(p_1+p_2)^2}(-ie)\bar{u}(p_3)\gamma^\nu v(p_4)$$

Each of these spinors has a spin, such as $u_s(p_1)$ but we are leaving the spin implicit. In Lorentz gauge, this is

$$\mathcal{M} = \frac{e^2}{s}\bar{v}(p_2)\gamma^\mu u(p_1)\bar{u}(p_3)\gamma_\mu v(p_4) \tag{13.11}$$

To calculate $|\mathcal{M}|^2$ we need the conjugate amplitude. To get this, we first recall that

$$\gamma_\mu^\dagger\gamma_0 = \gamma_0\gamma_\mu \quad \text{and} \quad \gamma_0^\dagger = \gamma_0 \tag{13.12}$$

So,

$$(\bar{\psi}_1\gamma^\mu\psi_2)^\dagger = (\psi_1^\dagger\gamma_0\gamma^\mu\psi_2)^\dagger = \psi_2^\dagger\gamma^{\mu\dagger}\gamma_0^\dagger\psi_1 = \psi_2^\dagger\gamma_0\gamma^\mu\psi_1 = \bar{\psi}_2\gamma^\mu\psi_1 \tag{13.13}$$

This nice transformation property is another reason using $\bar{\psi}$ instead of $\psi^\dagger$ is useful. Then,

$$\mathcal{M}^\dagger = \frac{e^2}{s}\bar{v}(p_4)\gamma^\mu u(p_3)\bar{u}(p_1)\gamma_\mu v(p_2) \tag{13.14}$$

Thus,

$$|\mathcal{M}|^2 = \frac{e^4}{s^2}[\bar{v}(p_2)\gamma^\mu u(p_1)][\bar{u}(p_3)\gamma_\mu v(p_4)][\bar{v}(p_4)\gamma^\nu u(p_3)][\bar{u}(p_1)\gamma_\nu v(p_2)] \tag{13.15}$$

I have grouped things in brackets to emphasize that each term is just a number for each μ and each set of spins. Then $|\mathcal{M}|^2$ is a product of these numbers. For example, we could also have written

$$|\mathcal{M}|^2 = \frac{e^4}{s^2}[\bar{v}(p_2)\gamma^\mu u(p_1)][\bar{u}(p_1)\gamma_\nu v(p_2)][\bar{u}(p_3)\gamma_\mu v(p_4)][\bar{v}(p_4)\gamma^\nu u(p_3)] \tag{13.16}$$

Which makes it clear that this is the product of two tensors, one depending on the initial state, and the other depending on the final state.

13.2.1 unpolarized scattering

The easiest things to calculate from this is the cross section for scattering assuming you are insensitive to spin. First, let's sum over the μ^+ spins, using

$$\sum_s v_s(p_4)\bar{v}_s(p_4) = \not{p}_4 - m_\mu \tag{13.17}$$

This contracts two of the brackets:

$$\sum_s [\bar{u}(p_3)\gamma_\mu v_s(p_4)][\bar{v}_s(p_4)\gamma^\nu u(p_3)] = [\bar{u}(p_3)\gamma_\mu(\not{p}_4 - m_\mu)\gamma^\nu u(p_3)] \tag{13.18}$$

Now we sum over the μ^- spins using

$$\sum_s u_s(p_3)\bar{u}_s(p_3) = \not{p}_3 + m_\mu \tag{13.19}$$

This contracts the same bracket again, into a trace:

$$\sum_s \bar{u}(p_3)\gamma_\mu(\not{p}_4 - m)\gamma^\nu u(p_3) = \mathrm{Tr}[(\not{p}_3 + m_\mu)\gamma_\mu(\not{p}_4 - m_\mu)\gamma^\nu] \tag{13.20}$$

We can also assume that we don't have sensitivity to the spins of our initial states. Thus if we do the measurement many times, we will get the *average* over each possibility. This leads to contractions and traces of the electron terms. Then we have

$$\frac{1}{4}\sum_{\text{spins}} |\mathcal{M}|^2 = \frac{e^4}{4s^2}\big(\mathrm{Tr}[(\not{p}_1 + m_e)\gamma_\nu(\not{p}_2 - m_e)\gamma^\mu]\mathrm{Tr}[(\not{p}_3 + m_\mu)\gamma_\mu(\not{p}_4 - m_\mu)\gamma^\nu]\big) \tag{13.21}$$

Now we can simplify using our trace identities

$$\mathrm{Tr}[(\not{p}_3 + m_\mu)\gamma_\mu(\not{p}_4 - m_\mu)\gamma^\nu] = p_3^\alpha p_4^\beta \mathrm{Tr}[\gamma_\alpha\gamma_\mu\gamma_\beta\gamma_\nu] - m_\mu^2 \mathrm{Tr}[\gamma_\mu\gamma_\nu] \tag{13.22}$$

$$= 4(p_3^\mu p_4^\nu + p_4^\mu p_3^\nu - (p_3 p_4)\eta^{\mu\nu}) - 4m_\mu^2\eta^{\mu\nu} \tag{13.23}$$

So,

$$\frac{1}{4}\sum_{\text{spins}} |\mathcal{M}|^2 = \frac{4e^4}{s^2}\big(p_1^\mu p_2^\nu + p_2^\mu p_1^\nu - ((p_1p_2) + m_e^2)\eta^{\mu\nu}\big)\big(p_3^\mu p_4^\nu + p_4^\mu p_3^\nu - ((p_3p_4) + m_\mu^2)\eta^{\mu\nu}\big) \tag{13.24}$$

$$= \frac{8e^4}{s^2}\big[p_{13}p_{24} + p_{14}p_{23} + m_\mu^2 p_{12} + m_e^2 p_{34} + m_e^2 m_\mu^2\big] \tag{13.25}$$

with $p_{\text{ij}} = p_i \cdot p_j$. Now we can simplify this with Mandlestam variables

$$s = (p_1 + p_2)^2 = (p_3 + p_4)^2 = 2m_e^2 + 2p_{12} = 2m_\mu^2 + 2p_{34} \tag{13.26}$$

$$t = (p_1 - p_3)^2 = (p_2 - p_4)^2 = m_e^2 + m_\mu^2 - 2p_{13} = m_e^2 + m_\mu^2 - 2p_{24} \tag{13.27}$$

$$u = (p_1 - p_4)^2 = (p_2 - p_3)^2 = m_e^2 + m_\mu^2 - 2p_{14} = m_e^2 + m_\mu^2 - 2p_{23} \tag{13.28}$$

Thus,

$$p_{12} = \frac{s}{2} - m_e^2, \quad p_{34} = \frac{s}{2} - m_\mu^2 \tag{13.29}$$

$$p_{13} = p_{24} = \frac{1}{2}(m_e^2 + m_\mu^2 - t)$$

$$p_{14} = p_{23} = \frac{1}{2}(m_e^2 + m_\mu^2 - u) \tag{13.30}$$

So,

$$\boxed{\frac{1}{4}\sum_{\text{spins}} |\mathcal{M}|^2 = \frac{2e^4}{s^2}\big[t^2 + u^2 + 4s(m_e^2 + m_\mu^2) - 2(m_e^4 + 4m_e^2m_\mu^2 + m_\mu^4)\big]} \tag{13.31}$$

13.2.2 differential cross section

Since the particles are not all of the same mass, we have to use the more general formula for the differential cross section

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{CM}} = \frac{1}{64\pi^2 E_{\text{CM}}^2}\frac{|p_f|}{|p_i|}|\mathcal{M}|^2 \tag{13.32}$$

In the center of mass frame, there are only two degrees of freedom: energy and angle. We can take

$$p_1 = (E, \vec{k}), \quad p_2 = (E, -\vec{k}), \quad |\vec{k}| = \sqrt{E^2 - m_e^2} \tag{13.33}$$

$$p_3 = (E, \vec{p}) \quad p_4 = (E, -\vec{p}), \quad |\vec{p}| = \sqrt{E^2 - m_\mu^2} \tag{13.34}$$

Then,

$$s = 4E^2 = E_{\text{CM}}^2 \tag{13.35}$$

$$t = (p_1 - p_3)^2 = m_e^2 + m_\mu^2 - 2E^2 + 2\,\vec{k}\,\vec{p} \tag{13.36}$$

$$u = -(\vec{k} + \vec{p})^2 = m_e^2 + m_\mu^2 - 2E^2 - 2\,\vec{k}\,\vec{p} \tag{13.37}$$

And so

$$\frac{d\sigma}{d\Omega} = \frac{e^4}{32\pi^2 E_{\text{CM}}^2 s^2}\frac{|\vec{p}|}{|\vec{k}|}\big[t^2 + u^2 + 4s(m_e^2 + m_\mu^2) - 2(m_e^4 + 4m_e^2m_\mu^2 + m_\mu^4)\big] \tag{13.38}$$

$$= \frac{\alpha^2}{16E^6}\frac{|\vec{p}|}{|\vec{k}|}\left(E^4 + (\vec{k}\cdot\vec{p})^2 + E^2(m_e^2 + m_\mu^2) - \frac{m_e^2m_\mu^2}{2}\right) \tag{13.39}$$

The only angular dependence comes from the $\vec{k}\cdot\vec{p}$ term

$$\vec{k}\cdot\vec{p} = |\vec{k}\,||\vec{p}\,|\cos\theta \tag{13.40}$$

So,

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{16E^6}\frac{|\vec{p}\,|}{|\vec{k}\,|}\left(E^4 + |\vec{k}\,|^2|\vec{p}\,|^2\cos^2\theta + E^2(m_e^2+m_\mu^2) - \frac{m_e^2 m_\mu^2}{2}\right) \tag{13.41}$$

$$|\vec{k}\,| = \sqrt{E^2 - m_e^2}, \quad |\vec{p}\,| = \sqrt{E^2 - m_\mu^2} \tag{13.42}$$

Which is the general result for the $e^+e^-\to\mu^+\mu^-$ rate in the center of mass frame.

Taking $m_e = 0$ for simplicity, gives $|\vec{k}\,| = E$ and this reduces to

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E_{\rm CM}^2}\sqrt{1-\frac{m_\mu^2}{E^2}}\left(1+\frac{m_\mu^2}{E^2}+\left(1-\frac{m_\mu^2}{E^2}\right)\cos^2\theta\right) \tag{13.43}$$

If in addition we take $m_\mu = 0$, which is the ultra high energy limit, we find

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E_{\rm CM}^2}\big(1+\cos^2\theta\big) \tag{13.44}$$

which is the same thing we had from the naive sum over spin sates. Recall that scattering with spins transverse the plane gave $\mathcal{M} = 1$ and scattering with spins in the plane gave $\mathcal{M} = \cos^2\theta$, so this agrees with our previous analysis. You can check explicitly by choosing explicit spinors that our intuition with spin scattering agrees with QED even for the polarized cross-section.

13.3 Rutherford scattering $e^-p^+ \to e^-p^+$

Now let's go back to the problem we considered long ago, scattering of an electron by a Coulomb potential. Recall the classical Rutherford scattering formula

$$\frac{d\sigma}{d\Omega} = \frac{m_e^2 e^4}{4k^4\sin^4\frac{\theta}{2}} \tag{13.45}$$

where $\vec{k}$ is the electron momentum. This is elastic scattering, so $|\vec{k}_i| = |\vec{k}_f| \equiv |\vec{k}\,|$. In Quantum Mechanics, we learned that this is reproduced in the Born approximation, by Fourier transforming the Coulomb potential $V(r) = \frac{e^2}{r}$:

$$\left(\frac{d\sigma}{d\Omega}\right)_{\rm born} = m_e^2\tilde{V}(q)^2 = m_e^2\left[\int d^3x\, e^{-iqx}\frac{e^2}{r}\right]^2 = \frac{m_e^2 e^4}{|\vec{k}\,|^4\sin^4\frac{\theta}{2}} \tag{13.46}$$

where $\vec{q} = \vec{k_i} - \vec{k_f}$ is the momentum transfer (which we have learned to think of as the virtual photon momentum). In field theory, we observed that this is given by a simple t-channel diagram.

$$\left(\frac{d\sigma}{d\Omega}\right)_{\rm QFT} = \frac{1}{64\pi^2 E_{\rm CM}^2}\frac{(2m_e)^2(2m_p)^2}{t^2} \tag{13.47}$$

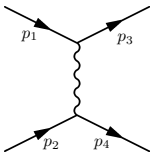
where the $2m_e^2$ and $2m_p^2$ factors are just the non-relativistic normalization of the electron and proton states. Since $t = q^2$ where $q^\mu = p_3^\mu - p_1^\mu$ is the photon momentum. Then $|\vec{q}\,| = \vec{k}\,\sin\frac{\theta}{2}$ and $E_{\rm CM} = m_p$ so we reproduce the Rutherford formula again.

Now let's do it again, properly, in QED.

13.3.1 QED amplitude

As far as QED is concerned, a proton and a muon are the same thing, up to the charge, which gets squared anyway. So let's start with $e^-\mu^- \to e^-\mu^-$. The amplitude is given by a t-channel diagram

$$i\mathcal{M}_t = \quad = (-ie)\bar{u}(p_3)\gamma^\mu u(p_1)\frac{-ig_{\mu\nu}}{q^2}(-ie)\bar{u}(p_4)\gamma^\nu u(p_2)$$



which simplifies to

$$\mathcal{M}_t = \frac{e^2}{t}\bar{u}(p_3)\gamma^\mu u(p_1)\bar{u}(p_4)\gamma_\mu v(p_2) \tag{13.48}$$

The spin averaged answer is is

$$\frac{1}{4}\sum_{\text{spins}}|\mathcal{M}|^2 = \frac{e^4}{4t^2}\big(\text{Tr}[(\not{p}_1 + m_e)\gamma_\nu(\not{p}_3 + m_e)\gamma^\mu]\text{Tr}[(\not{p}_4 + m_\mu)\gamma_\mu(\not{p}_2 + m_\mu)\gamma^\nu]\big) \tag{13.49}$$

This is remarkably similar to what we had for $e^+e^- \to \mu^+\mu^-$

$$\frac{1}{4}\sum_{\text{spins}}|\mathcal{M}|^2 = \frac{e^4}{4s^2}\big(\text{Tr}[(\not{p}_1 + m_e)\gamma_\nu(\not{p}_2 - m_e)\gamma^\mu]\text{Tr}[(\not{p}_3 + m_\mu)\gamma_\mu(\not{p}_4 - m_\mu)\gamma^\nu]\big) \tag{13.50}$$

In fact, the two are identical if we take the $e^+e^- \to \mu^+\mu^-$ formula and replace

$$p_1 \to p_1 \tag{13.51}$$

$$p_2 \to -p_3 \tag{13.52}$$

$$p_3 \to p_4 \tag{13.53}$$

$$p_4 \to -p_2 \tag{13.54}$$

There are two sign flips which cancel. Also, these changes sent $s \to t$, or more generally

$$s = (p_1 + p_2)^2 \to (p_1 - p_3)^2 = t \tag{13.55}$$

$$t = (p_1 - p_3)^2 \to (p_1 - p_4)^2 = u \tag{13.56}$$

$$u = (p_1 - p_4)^2 \to (p_1 + p_2)^2 = s \tag{13.57}$$

These momenta replacement are not physical, since $p_2 \to -p_3$ produces a momentum with negative energy, which cannot be an external particle. It's just an algebraic trick, called a *crossing symmetry* which lets us just use the same algebra to write down the answer. You can prove crossing symmetries in general, even for polarized cross sections with particular spins. But changing the spin of an electron to that of a positron is confusing, because the bases are different. So it is easier just to write down the amplitude that you want and inspect it to find the right transformation. The point of crossing symmetries is not to memorize the formula for how to transform your answer, but to know that they are there, which can save you a lot of work.

So now we can just skip to the final answer. For $e^+e^- \to \mu^+\mu^-$ we had

$$\frac{1}{4}\sum_{\text{spins}}|\mathcal{M}|^2 = \frac{8e^4}{s^2}\big[p_{13}p_{24} + p_{14}p_{23} + m_\mu^2 p_{12} + m_e^2 p_{34} + m_e^2 m_\mu^2\big] \tag{13.58}$$

$$= \frac{2e^4}{s^2}\big[t^2 + u^2 + 4s(m_e^2 + m_\mu^2) - 2(m_e^4 + 4m_e^2 m_\mu^2 + m_\mu^4)\big] \tag{13.59}$$

Therefore, for $e^-e^- \to p^+p^+$we get

$$\frac{1}{4}\sum_{\text{spins}} |\mathcal{M}|^2 = \frac{8e^4}{t^2}\big[p_{14}p_{23} + p_{12}p_{34} - m_p^2 p_{13} - m_e^2 p_{24} + 2m_e^2 m_p^2\big] \tag{13.60}$$

$$= \frac{2e^4}{t^2}\big[u^2 + s^2 + 4t(m_e^2 + m_p^2) - 2(m_e^4 - 4m_e^2 m_p^2 + m_p^4)\big] \tag{13.61}$$

13.3.2 corrections to Rutherford

Now let's take the limit $m_p \gg m_e$ to get the Rutherford answer, but including relativistic corrections. This means we can treat the proton mass as infinite, but have to treat the electron mass as finite to go from the non-relativistic to the relativistic limit. Then the center of mass frame is roughly the protons rest frame, so

$$p_1 = (E, \vec{k}_i), \quad p_3 = (E, \vec{k}_f) \tag{13.62}$$

$$p_2 = (m_p, 0), \quad p_4 \approx (m_p, 0) \tag{13.63}$$

where p_4 can't be exactly at rest, but we discard the corrections. The scattering angle is defined by

$$\vec{k}_i \cdot \vec{k}_f = k^2 \cos\theta = v^2 E^2 \cos\theta \tag{13.64}$$

where

$$v = \frac{|\vec{k}\,|}{E} = \sqrt{1 - \frac{m_e^2}{E^2}} \tag{13.65}$$

is the electron's relativistic velocity. Then,

$$p_{13} = E^2(1 - v^2\cos\theta) \tag{13.66}$$

$$p_{12} = p_{23} = p_{34} = p_{14} = Em_p \tag{13.67}$$

$$p_{24} = m_p^2 \tag{13.68}$$

$$t = (p_1 - p_3)^2 = -(\vec{k}_i - \vec{k}_f)^2 = -2k^2(1 - \cos\theta) = -2v^2E^2(1 - \cos\theta) \tag{13.69}$$

Then,

$$\frac{1}{4}\sum_{\text{spins}} |\mathcal{M}|^2 = \frac{8e^4}{t^2}\big[p_{14}p_{23} + p_{12}p_{34} - m_p^2 p_{13} - m_e^2 p_{24} + 2m_e^2 m_p^2\big] \tag{13.70}$$

$$= \frac{8e^4}{4v^4E^4(1 - \cos\theta)^2}\big[E^2m_p^2 + E^2m_p^2v^2\cos\theta + m_e^2m_p^2\big] \tag{13.71}$$

$$= \frac{2e^4m_p^2}{v^4E^2(1 - \cos\theta)^2}\big[2 - v^2(1 - \cos\theta)\big] \tag{13.72}$$

$$= \frac{e^4m_p^2}{v^4E^2\sin^4\frac{\theta}{2}}\left[1 - v^2\sin^2\frac{\theta}{2}\right] \tag{13.73}$$

Since the COM frame is essentially the lab frame, the differential cross section is

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2E_{\text{CM}}^2}\left[\frac{1}{4}\sum_{\text{spins}} |\mathcal{M}|^2\right] \tag{13.74}$$

$$\boxed{\frac{d\sigma}{d\Omega} = \frac{e^4}{64\pi^2v^2k^2\sin^4\frac{\theta}{2}}\left(1 - v^2\sin^2\frac{\theta}{2}\right)} \tag{13.75}$$

This is known as the Mott formula for relativistic Rutherford scattering. Note that the proton mass has dropped out. For slow velocities $v \ll 1$ and $k \ll E \sim m_e$ so $v \sim \frac{k}{m_e}$. Thus,

$$\frac{d\sigma}{d\Omega} = \frac{e^4 m_e^2}{64\pi^2 k^4 \sin^4\frac{\theta}{2}} \tag{13.76}$$

which is the Rutherford formula. In particular, note that the normalization factors m_e^2 worked out correctly.

13.4 Compton scattering

The next process worth studying is the QED prediction for Compton scattering $\gamma e^- \to \gamma e^-$. By simple relativistic kinematics, Compton was able to predict the shift in wavelength of the scattered light as a function of angle

$$\Delta\lambda = \frac{1}{m}(1 - \cos\theta) \tag{13.77}$$

But he could not predict the intensity of radiation at each angle.

In the classical limit, for scattering soft radiation against electrons, JJ Thompson had derived the formula

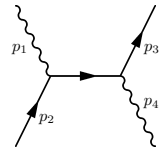
$$\frac{d\sigma}{\mathrm{d}\cos\theta} = \pi r_e^2\big(1 + \cos^2\theta\big) = \frac{\pi\alpha^2}{m^2}(1 + \cos^2\theta) \tag{13.78}$$

r_e is the classical electron radius $r_e = \frac{\alpha}{m}$. The 1 comes from radiation polarized in the plane of scattering and the $\cos\theta$ from polarization out of the plane. From QED we should be able to reproduce this formula, plus the radiative corrections.

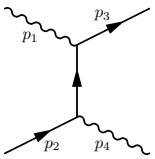
13.4.1 Klein-Nishina formula

There are two diagrams:

$$i\mathcal{M}_s = \quad = (-ie)^2 \varepsilon^\mu_{\text{in}} \varepsilon^\nu_{\text{out}} \bar{u}(p_3)\gamma^\nu \frac{i(\not{p}_1 + \not{p}_2 + m)}{(p_1 + p_2)^2 - m^2}\gamma^\mu u(p_2)$$



$$i\mathcal{M}_t = \quad = (-ie)^2 \varepsilon^\mu_{\text{in}} \varepsilon^\nu_{\text{out}} \bar{u}(p_3)\gamma^\mu \frac{i(\not{p}_2 - \not{p}_4 + m)}{(p_2 - p_4)^2 - m^2}\gamma^\nu u(p_2)$$



So the sum is

$$\mathcal{M} = e^2 \varepsilon^{\text{in}}_\mu \varepsilon^{\text{out}}_\nu \bar{u}(p_3)\left[\frac{\gamma^\nu(\not{p}_1 + \not{p}_2 + m)\gamma^\mu}{s - m^2} + \frac{\gamma^\mu(\not{p}_2 - \not{p}_4 + m)\gamma^\nu}{t - m^2}\right] u(p_2) \tag{13.79}$$

We would like to calculate the unpolarized cross-section.

13.4.2 photon polarization sums

To square this and sum over polarizations, we need to use a trick for the photon polarization sum. There is no way to write the sum over transverse modes in a Lorentz invariant way. But if we sum over any old random basis of 4-vectors, we would have

$$\sum_{i=1}^{4} \varepsilon_\mu^i \varepsilon_\nu^{i\star} = \delta_{\mu\nu} \tag{13.80}$$

The two unphysical polarizations can be taken to be forward and backwards with respect to the photon's momentum k_μ:

$$\varepsilon_f^\mu = k^\nu = (E, \vec{k}) \tag{13.81}$$

$$\varepsilon_b^\mu = \tilde{k}^\mu = (E, -\vec{k}) \tag{13.82}$$

Then,

$$\sum_{i=1}^{2} \varepsilon_\mu^i \varepsilon_\nu^{i\star} = -\eta_{\mu\nu} + \frac{1}{2E^2}\Big(k^\mu \tilde{k}^\nu + \tilde{k}^\nu k^\mu\Big) \tag{13.83}$$

It's easy to check this relation by choosing explicit transverse modes.

So suppose we have an amplitude

$$\mathcal{M} = \epsilon_\mu^i M_\mu \tag{13.84}$$

$$\sum_{\text{pols}} \mathcal{M}^2 = \epsilon_\mu M_\mu M_\nu \epsilon_\nu = M_\mu M_\mu + \frac{1}{E^2}\Big(k_\mu M_\mu M_\nu \tilde{k}_\nu + \tilde{k}_\mu M_\mu M_\nu k_\nu\Big) \tag{13.85}$$

But by the Ward identity, $k^\mu M_\mu = 0$. Therefore, we can simply replace

$$\sum_{i=1}^{2} \varepsilon_\mu^i \varepsilon_\nu^{i\star} = -\eta_{\mu\nu} \tag{13.86}$$

in any physical matrix element.

13.4.3 t-channel term

We want to calculate $\mathcal{M}^2$, which includes the 4 terms, from the t-channel and u-channel diagrams squared and their cross term. Let's just evaluate one piece, in the high energy limit, to see what's involved. Take

$$\mathcal{M}_t = \frac{e^2}{t} \varepsilon_\mu^{\text{in}} \varepsilon_\nu^{\text{out}} \bar{u}(p_3)\gamma^\mu(\not{p}_2 - \not{p}_4)\gamma^\nu u(p_2) \tag{13.87}$$

$$\sum_{\text{spins/pols}} \mathcal{M}_t^2 = \frac{e^4}{t^2} \text{Tr}[\not{p}_3 \gamma^\mu (\not{p}_2 - \not{p}_4)\gamma^\nu \not{p}_2 \gamma_\nu (\not{p}_2 - \not{p}_4)\gamma_\mu] \tag{13.88}$$

Now use

$$\gamma^\nu \not{p} \gamma_\nu = -2\not{p} \tag{13.89}$$

$$q^\mu = p_2^\mu - p_4^\mu = p_3^\mu - p_1^\mu \tag{13.90}$$

To get

$$\sum \mathcal{M}_t^2 = 4\frac{e^4}{t^2} \text{Tr}[\not{p}_3 \not{q} \not{p}_2 \not{q}] \tag{13.91}$$

$$= 16\frac{e^4}{t^2}\big(2(p_3 \cdot q)(p_2 \cdot q) - p_{23} q^2\big) = 16\frac{e^4}{t^2}(2p_{13}p_{24} + 2p_{23}p_{13}) = 8\frac{e^4}{t^2}\big(t^2 + ut\big) \tag{13.92}$$

$$= -8e^4 \frac{s}{t} = 8e^4 \frac{p_{12}}{p_{24}} \tag{13.93}$$

Note that one of the factors of t canceled, so the divergence at $t=0$ is not $\frac{1}{t^2}$ but simply $\frac{1}{t}$.

13.4.4 Klein-Nishina formula.

Including all the terms

$$\mathcal{M}=e^2\varepsilon_\mu^{\rm in}\varepsilon_\nu^{\rm out}\bar{u}\,(p_3)\left[\frac{\gamma^\nu\big(\not{p}_1+\not{p}_2+m\big)\gamma^\mu}{s-m^2}+\frac{\gamma^\mu\big(\not{p}_2-\not{p}_4+m\big)\gamma^\nu}{t-m^2}\right]u(p_2) \tag{13.94}$$

Then summing/averaging over spins and polarizations

$$\frac{1}{4}\sum_{\rm pols}\mathcal{M}^2=e^4{\rm Tr}\Bigg\{(\not{p}_3+m)\left[\frac{\gamma^\nu\big(\not{p}_1+\not{p}_2+m\big)\gamma^\mu}{s-m^2}+\frac{\gamma^\mu\big(\not{p}_2-\not{p}_4+m\big)\gamma^\nu}{t-m^2}\right] \tag{13.95}$$

$$\times(\not{p}_2+m)\left[\frac{\gamma^\mu\big(\not{p}_1+\not{p}_2+m\big)\gamma^\nu}{s-m^2}+\frac{\gamma^\nu\big(\not{p}_2-\not{p}_4+m\big)\gamma^\mu}{t-m^2}\right]\Bigg\} \tag{13.96}$$

This is a huge mess. But it can be evaluated. The result is rather simple

$$\boxed{\frac{1}{4}\sum_{\rm pols}\mathcal{M}^2=2e^4\left[\frac{p_{24}}{p_{12}}+\frac{p_{12}}{p_{24}}+2m^2\left(\frac{1}{p_{12}}-\frac{1}{p_{24}}\right)+m^4\left(\frac{1}{p_{12}}-\frac{1}{p_{24}}\right)^2\right]} \tag{13.97}$$

13.4.5 low-energy limit

Let's start with the low energy limit, where it makes sense to work in the lab frame. Then

$$p_1=(\omega,0,0,\omega)\quad p_2=(m,0,0,0) \tag{13.98}$$

$$p_3=(E',\vec{p}\,')\quad p_4=(\omega',\omega'\sin\theta,0,\omega'\cos\theta) \tag{13.99}$$

Note that these relations imply

$$\frac{1}{\omega'}-\frac{1}{\omega}=\frac{1}{m}(1-\cos\theta) \tag{13.100}$$

and

$$\omega'=\frac{\omega}{1+\frac{\omega}{m}(1-\cos\theta)} \tag{13.101}$$

which is the formula for the shifted frequency as a function of angle. There is no QED in this relation – it is just momentum conservation – but it is still a very important formula!

Then

$$p_{12}=\omega m \tag{13.102}$$

$$p_{24}=\omega' m \tag{13.103}$$

Thus we get a simple formula for $\mathcal{M}^2$

$$\frac{1}{4}\sum_{\rm pols}\mathcal{M}^2=2e^4\left[\frac{\omega'}{\omega}+\frac{\omega}{\omega'}-2(1-\cos\theta)+(1-\cos\theta)^2\right] \tag{13.104}$$

$$=2e^4\left[\frac{\omega'}{\omega}+\frac{\omega}{\omega'}-\sin^2\theta\right] \tag{13.105}$$

But the phases space is a mess.

In the lab frame, we have to go back to our general formula

$$d\sigma = \frac{1}{(2E_1)(2E_2)|\vec{v}_1 - \vec{v}_2|}|\mathcal{M}|^2 d\Pi_{\text{LIPS}} \tag{13.106}$$

$$= \frac{1}{4\omega m}|\mathcal{M}|^2 d\Pi_{\text{LIPS}} \tag{13.107}$$

Now we need

$$\int d\Pi_{\text{LIPS}} = \int \frac{d^3p_3}{(2\pi)^3}\frac{1}{2E'}\int \frac{d^3p_4}{(2\pi)^3}\frac{1}{2\omega'}\Big[(2\pi)^4\delta^{(4)}(p_1^\mu + p_2^\mu - p_3^\mu - p_4^\mu)\Big] \tag{13.108}$$

The delta function fixes the 3-momenta when we integrate over d^3p_4, leaving the energy constraint

$$\int d\Pi_{\text{LIPS}} = \frac{1}{4(2\pi)^2}\int \omega'^2 d\Omega d\omega' \frac{1}{\omega' E'}\delta(\sum E) \tag{13.109}$$

$$= \frac{1}{8\pi}\int \text{dcos}\theta d\omega' \frac{\omega'}{E'}\delta(\sum E) \tag{13.110}$$

Now we want to integrate over ω' to enforce the energy constraint $E' + \omega' = m + \omega$. But we have to be a little careful because E' and ω are already constrained by the electron's on-shell condition

$$E'^2 = m^2 + p'^2 = m^2 + (\omega' \text{sin}\theta)^2 + (\omega' \text{cos}\theta - \omega)^2 \tag{13.111}$$

$$= m^2 + \omega'^2 + \omega^2 - 2\omega\omega' \text{cos}\theta \tag{13.112}$$

$$E' \frac{dE'}{d\omega'} = \omega' - \omega \text{cos}\theta \tag{13.113}$$

Thus

$$\int d\Pi_{\text{LIPS}} = \frac{1}{8\pi}\int \text{dcos}\theta d\omega' \frac{\omega'}{E'}\delta(\omega' + E'(\omega') - m - \omega) \tag{13.114}$$

$$= \frac{1}{8\pi}\int \text{dcos}\theta \frac{\omega'}{E'}(1 + \frac{dE'}{d\omega'})^{-1} \tag{13.115}$$

$$= \frac{1}{8\pi}\int \text{dcos}\theta \frac{\omega'}{E'}\left(1 + \frac{\omega' - \omega \text{cos}\theta}{E'}\right)^{-1} \tag{13.116}$$

$$= \frac{1}{8\pi}\int \text{dcos}\theta \frac{(\omega')^2}{\omega m} \tag{13.117}$$

And so

$$\frac{d\sigma}{\text{dcos}\theta} = \frac{1}{4\omega m}\frac{1}{8\pi}\frac{(\omega')^2}{\omega m}2e^4\left[\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \text{sin}^2\theta\right] \tag{13.118}$$

$$\boxed{\frac{d\sigma}{\text{dcos}\theta} = \frac{\pi\alpha^2}{m^2}\left(\frac{\omega'}{\omega}\right)^2\left[\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \text{sin}^2\theta\right]} \tag{13.119}$$

This is the Klein-Nishina formula. This is first calculated by Klein and Nishina in 1929 and was one of the first tests of QCD.

In other variables,

$$\frac{d\sigma}{\text{dcos}\theta} = \frac{\pi\alpha^2}{m^2}\left(1 + \text{cos}^2\theta - \frac{2\omega}{m}(1 + \text{cos}^2\theta)(1 - \text{cos}\theta) + \mathcal{O}(\frac{1}{m^2})\right) \tag{13.120}$$

Note that in the limit $m \to \infty$

$$\frac{d\sigma}{\text{dcos}\theta} = \frac{\pi\alpha^2}{m^2}\big[1 + \text{cos}^2\theta\big] \tag{13.121}$$

This is the Thompson scattering cross sections for classical electromagnetic radiation by a free electron. We have calculated the full relativistic corrections.

13.4.6 high-energy behavior

Now let's consider the opposite limit, $\omega \gg m$. Then the COM frame makes the most sense. In this case

$$p_1 = (\omega, 0, 0, \omega), \quad p_2 = (E, 0, 0, -\omega) \tag{13.122}$$

$$p_3 = (E, -\omega\sin\theta, 0, -\omega\cos\theta) \quad p_4 = (\omega, \omega\sin\theta, 0, \omega\cos\theta) \tag{13.123}$$

$$p_{12} = \omega(E + \omega) \tag{13.124}$$

$$p_{24} = \omega(E + \omega\cos\theta) \tag{13.125}$$

For $\omega \gg m_e$,

$$E \approx m \tag{13.126}$$

$$s \approx 4\omega^2, \quad t \approx -2p_{24} = 2\omega^2(1 + \cos\theta) \tag{13.127}$$

Then,

$$\frac{1}{4}\sum_{\text{pols}} \mathcal{M}^2 = 2e^4\left[\frac{p_{24}}{p_{12}} + \frac{p_{12}}{p_{24}}\right] = 2e^4\left[\frac{t}{s} + \frac{s}{t}\right] \tag{13.128}$$

$$= 2e^4\left[1 + \cos\theta + \frac{1}{1 + \cos\theta}\right] \tag{13.129}$$

This has a singularity at $\theta = \pi$. That is where the photon and electron bounce off each other and go backwards. The mass of the electron cuts off the singularity, but still the scattering amplitude gets very large as $\theta \sim \pi$. The cross section is

$$\frac{d\sigma}{d\cos\theta} = \frac{e^4}{32\pi\omega^2}\left[1 + \cos\theta + \frac{1}{1 + \cos\theta}\right] \tag{13.130}$$

which blows up at $\theta = \pi$.

The obvious place to expect the pole at $\theta = \pi$ is from the t-channel diagram. This is easy to see because

$$t = \omega^2(1 + \cos\theta) \tag{13.131}$$

Since the t-channel matrix element is

$$\mathcal{M}_t = \frac{e^2}{t}\varepsilon_\mu^{\text{in}}\varepsilon_\nu^{\text{out}}\bar{u}(p_3)\gamma^\mu(\not{p}_2 - \not{p}_4)\gamma^\nu u(p_2) \tag{13.132}$$

we should naively expect a worse divergence, as $\frac{1}{t^2}$. In fact this would happen in a scalar field theory, such as one with interaction $g\phi^3$

$$\mathcal{M}_t = \frac{g^2}{t}, \quad \mathcal{M}^2 = \frac{g^4}{t^2} \tag{13.133}$$

which has a strong t^2 pole.

In QED, we calculated the t-channel diagram in the massless limit

$$\mathcal{M}_t = \frac{e^2}{t}\varepsilon_\mu^{\text{in}}\varepsilon_\nu^{\text{out}}\bar{u}(p_3)\gamma^\mu(\not{p}_2 - \not{p}_4)\gamma^\nu u(p_2) \tag{13.134}$$

and found

$$\frac{1}{4}\sum \mathcal{M}_t^2 = 2e^4\frac{s}{t} = 2e^4\frac{1}{1 + \cos\theta} \tag{13.135}$$

This gives the entire $\frac{1}{t}$ pole, so we don't have to worry about interference for the purpose of understanding the singularity. Where did the other factor of t come from to cancel the pole?

For $\theta = \pi - \phi$ with $\phi \sim 0$, the momenta become

$$p_1 = (\omega, 0, 0, \omega), \quad p_2 = (\omega, 0, 0, -\omega) \tag{13.136}$$

$$p_3 = (\omega, -\omega\phi, 0, \omega) \quad p_4 = (\omega, \omega\phi, 0, -\omega) \tag{13.137}$$

$$t = \omega^2\phi^2 \tag{13.138}$$

So a $\frac{1}{t}$ pole goes like $\frac{1}{\phi^2}$, but $\frac{1}{t^2}$ goes like $\frac{1}{\phi^4}$. But notice the momentu factor in te matrix element also fanishes as $p_2 \to p_4$:

$$\not{p}_2 - \not{p}_4 = -\omega\phi\not{k}, \quad k = (0,1,0,0) \tag{13.139}$$

So,

$$\mathcal{M}_t = \frac{e^2}{\omega^2\phi^2}\big[-\bar{u}(p_3)\not{\epsilon}_{\text{in}}(\not{p}_2 - \not{p}_4)\not{\epsilon}_{\text{out}}u(p_2)\big] = \frac{e^2}{\omega\phi}\big[\bar{u}(p_3)\not{\epsilon}_{\text{in}}\not{k}\not{\epsilon}_{\text{out}}u(p_2)\big] \tag{13.140}$$

Thus one factor of ϕ is canceling. This factor came from the spinors, and is an effect of spin conservation.

To evaluate this product, observe that it has 3 gamma matrices, so it preserves helicity. We can see this by inserting factors of $1 \pm \gamma_5$, and using

$$(1 \pm \gamma_5)\gamma_\mu = \gamma_\mu(1 \pm \gamma_5) \tag{13.141}$$

$$(1+\gamma_5)u_L = 0 \quad \bar{u}_L(1-\gamma_5) = 0 \tag{13.142}$$

So,

$$\bar{\psi}\psi = \bar{\psi}_L u_R + \bar{\psi}_R\psi_L \tag{13.143}$$

$$\bar{\psi}\gamma^\mu\psi = \bar{\psi}_L\gamma^\mu\psi_L + \bar{\psi}_R\gamma^\mu\psi_R \tag{13.144}$$

$$\bar{\psi}\gamma^\mu\gamma^\nu\psi = \bar{\psi}_L\gamma^\mu\gamma^\nu\psi_L + \bar{\psi}_R\gamma^\mu\gamma^\nu\psi_L \tag{13.145}$$

$$\bar{\psi}\gamma^\mu\gamma^\alpha\gamma^\beta\psi = \bar{\psi}_L\gamma^\mu\gamma^\alpha\gamma^\beta\psi_L + \bar{\psi}_R\gamma^\mu\gamma^\alpha\gamma^\beta\psi_R \tag{13.146}$$

Thus

$$\mathcal{M}_t = \frac{e^2}{\omega\phi}\big[\bar{u}_L(p_3)\not{\epsilon}_{\text{in}}\not{k}\not{\epsilon}_{\text{out}}u_L(p_2) + \bar{u}_R(p_3)\not{\epsilon}_{\text{in}}\not{k}\not{\epsilon}_{\text{out}}u_R(p_2)\big] \tag{13.147}$$

This was obvious anyway since in QED for massless electrons, there's not coupling between right and left handed states. Thus $u(p_2)$ and $\bar{u}(p_3)$ are either both right handed or both left handed. So say right handed. Since p_2 is in the z-direction, this means it's spin is up $u(p_2) = |\uparrow\rangle$. Then $\not{\epsilon}_{\text{out}}u(p_2) = 0$ unless $\epsilon_\mu^{\text{out}} = \Big|\begin{matrix}\downarrow\end{matrix}\Big\rangle$. That is, $\frac{1}{2} - 1 = -\frac{1}{2}$. Similarly, $\bar{u}(p_3) = |\downarrow\rangle$. Thus $\not{\epsilon}_{\text{in}} = \Big|\begin{matrix}\uparrow\end{matrix}\Big\rangle$. This means that the incoming photon and the outgoing photon are both right handed.

This has an important physical implication. Consider shooting a laser beam at an high energy beam of electrons. Since the laser is polarized, it will dominantly scatter backwards only certain electron helicities. This can be a useful way to polarize your electrons.

13.4.7 singularity cancellation from gauge invariance

We have seen that in ϕ^3 theory, the matrix element does have a $\frac{1}{t^2}$ pole

$$\mathcal{M}_t^2 = \frac{g^4}{t^2} \tag{13.148}$$

but in QED the singularity is only $\frac{1}{t}$.

The cancellation of the leading divergence is a consequence of gauge invariance. We will see this again and again. It is easier to explore this in scalar QED. First of all, we expect that Compton scattering in scalar QED should be simple. The intermediate scalar cannot carry angular information, so the matrix elements should be totally isotropic. There certainly should be no singularity. Let's see how this comes out of the calculation.

The t-channel amplitude is

$$\mathcal{M}_t = \frac{e^2}{t}\varepsilon_\mu^{\text{in}}(2p_3^\mu - p_1^\mu)\varepsilon_\nu^{\text{out}}(2p_2^\nu - p_4^\nu) \tag{13.149}$$

which is not zero. While $\mathcal{M}_t^2$ looks like it should have a singularity, in fact

$$\mathcal{M}_t^2 = \frac{e^4}{t^2}(2p_3 - p_1)^2(2p_2 - p_4)^2 = 16\frac{e^4}{t^2}p_{13}^2 p_{24}^2 = 4e^4 \tag{13.150}$$

So this is not vanishing, even though we *know* the intermediate particle is a scalar so it should vanish. To be precise, it should vanish as the intermediate particle goes on shell, which it does as the propogator approaches the singularity.

Even though the $\mathcal{M}_t^2$ diagram was totally regular, there can still be a singularity from the cross terms between $\mathcal{M}_t$ and the s-channel and the 4-point vertex:

$$\mathcal{M}_s = \frac{e^2}{s}\varepsilon_\mu^{\text{in}}(p_1^\mu + 2p_2^\mu)\varepsilon_\nu^{\text{out}}(p_4^\nu + 2p_3^\nu) \tag{13.151}$$

$$\mathcal{M}_4 = -2e^2\varepsilon_\mu^{\text{in}}\varepsilon_\mu^{\text{out}} \tag{13.152}$$

But we find

$$\mathcal{M}_s\mathcal{M}_t = \frac{e^4}{st}[(p_1^\mu + 2p_2^\mu)(2p_3^\mu - p_1^\mu)][(p_4^\nu + 2p_3^\nu)(2p_2^\nu - p_4^\nu)] \tag{13.153}$$

$$= \frac{e^4}{st}(2p_{13} + 4p_{23} - 2p_{12})(2p_{24} + 4p_{23} - 2p_{43}) \tag{13.154}$$

$$= \frac{e^4}{st}(t + 2u + s)^2 = e^4\left[\frac{s}{t} + \frac{t}{s} + 2\right] \tag{13.155}$$

And

$$\mathcal{M}_4\mathcal{M}_t = -2\frac{e^4}{t}(2p_3^\mu - p_1^\mu)(2p_2^\mu - p_4^\mu) = -2\frac{e^4}{t}(4p_{23} - 2p_{12} - 2p_{34} + p_{14}) \tag{13.156}$$

$$= -e^4(\frac{s}{t} + 5) \tag{13.157}$$

So these cancel too. In fact, working out the rest of the amplitudes

$$\mathcal{M}_4\mathcal{M}_s = -e^4(\frac{t}{s} + 5) \tag{13.158}$$

$$\mathcal{M}_s^2 = 4e^4 \tag{13.159}$$

$$\mathcal{M}_4^2 = 16e^4 \tag{13.160}$$

We find

$$\mathcal{M}^2 = \mathcal{M}_t^2 + \mathcal{M}_s^2 + \mathcal{M}_4^2 + 2\mathcal{M}_s\mathcal{M}_t + 2\,\mathcal{M}_s\mathcal{M}_t + 2\mathcal{M}_4\mathcal{M}_s = (4 + 4 + 16 + 4 - 10 - 10)e^4 \tag{13.161}$$

$$\boxed{\mathcal{M}^2 = 8e^4} \tag{13.162}$$

So there is no pole, and no angular information, as expected.

Actually, this process in scalar QED is a great place to learn an important lesson about polarization sums. When we replace $\varepsilon_\mu\varepsilon_\nu$ by $-\eta_{\mu\nu}$, we are assuming the Ward identity. The Ward identity does not hold for each individual diagram, but only for the sum of all diagrams for the same physical process. Thus doing the replacement in the t-channel does not tell us anything physical about the contribution of the t-channel diagram. In fact, there is no way at all to talk about the contribution of a single diagram to the amplitude.

To get closer to the physics, we can try to sum over the physical transverse modes, instead of doing the sum where we replace $\varepsilon_\mu\varepsilon_\nu$ by $g_{\mu\nu}$. For transverse modes, the amplitudes simplify to

$$\mathcal{M}_t = 4\frac{e^2}{t}(\varepsilon^{\text{in}} \cdot p_3)(\varepsilon_\nu^{\text{out}} \cdot p_2) \tag{13.163}$$

$$\mathcal{M}_s = 4\frac{e^2}{s}(\varepsilon^{\text{in}} \cdot p_2)(\varepsilon_\nu^{\text{out}} \cdot p_3) \tag{13.164}$$

$$\mathcal{M}_4 = -2e^2\varepsilon_\mu^{\text{in}}\varepsilon_\mu^{\text{out}} \tag{13.165}$$

We are not using the Ward identity, just the fact that our polarizations are transverse.

In the high energy frame we had

$$p_1 = (\omega, 0, 0, \omega), \quad p_2 = (\omega, 0, 0, -\omega) \tag{13.166}$$

$$p_3 = (\omega, -\omega\sin\theta, 0, -\omega\cos\theta) \quad p_4 = (\omega, \omega\sin\theta, 0, \omega\cos\theta) \tag{13.167}$$

So the polarizations are

$$\epsilon^1_{\rm in} = (0, 1, 0, 0), \quad \epsilon^2_{\rm in} = (0, 0, 1, 0) \tag{13.168}$$

$$\epsilon^1_{\rm out} = (0, \cos\theta, 0, -\sin\theta), \quad \epsilon^2_{\rm out} = (0, 0, 1, 0) \tag{13.169}$$

Also,

$$s = 4\omega^2, \quad t = -2\omega^2(1+\cos\theta), \quad u = -2\omega^2(1-\cos\theta) \tag{13.170}$$

Now let's see what vanishes.

The s-channel is easy, since the particles are back to back, $\mathcal{M}_s = 0$ for every polarization.

$$\mathcal{M}_s^{11} = \mathcal{M}_s^{21} = \mathcal{M}_s^{12} = \mathcal{M}_s^{22} = 0 \tag{13.171}$$

For the t-channel the only non-vanishing amplitude has $\epsilon^1_{\rm in}$ and $\epsilon^1_{\rm out}$ giving

$$\mathcal{M}_t^{11} = 4\frac{e^2}{t}(\omega\sin\theta)(-\omega\sin\theta) = e^2(4 + 4\frac{t}{s}) \tag{13.172}$$

$$\mathcal{M}_t^{21} = \mathcal{M}_t^{12} = \mathcal{M}_t^{22} = 0 \tag{13.173}$$

Finally, for $\mathcal{M}_4$ we get two non-vanishing amplitudes

$$\mathcal{M}_4^{11} = 2e^2\cos\theta = e^2(-2 - 4\frac{t}{s}), \quad \mathcal{M}_4^{22} = 2e^2 \tag{13.174}$$

$$\mathcal{M}_4^{21} = \mathcal{M}_4^{12} = 0 \tag{13.175}$$

So there is no pole at all in the t-channel anywhere. In fact the scattering amplitudes for each transverse polarization are separately finite

$$1 \to 1: \quad (\mathcal{M}_t^{11} + \mathcal{M}_4^{11})^2 = 4e^4 \tag{13.176}$$

$$2 \to 2: \quad (\mathcal{M}_4^{22})^2 = 4e^4 \tag{13.177}$$

So the total cross section is just

$$\mathcal{M}^2 = 8e^4 \tag{13.178}$$

which is what we found by the spin sums above.

But note the difference. When we used the spin sum relation with the Ward identity, we got $\mathcal{M}_t^2 = 4e^4$ but summing over transverse modes by had got $\mathcal{M}_t^2 = 16e^4(1+\frac{t}{s})^2$.

13.5 Processes in QED

If we only consider $2 \to 2$ scattering and can choose from electrons, positrons, muons, anti-muons and photons, we get the following tree-level QED processes.

- $e^+e^- \to e^+e^-$: Bhabha scattering (1936).

  First calculated by Homi Jehengir Bhabha (Indian, 1909-1966). The positron was not discovered until 1932, so it would be a while before the differential cross section that Bhabha predicted could be measured in the lab. However, the total cross section for $e^+e^- \to e^+e^-$ was important for cosmic ray physics from the 30s onward.

- $e^+e^- \to \mu^+\mu^-$: Electron-positron annihilation into muons.

- $e^-e^- \to e^-e^-$: Møller Scattering (1931).

  First calculated in the ultralativistic regime by Christian Møller (Danish, 1904-1980). In the non-relativistic regime it is called Coulomb scattering or Rutherford scattering. Moller calculated the cross section based on some guesses and consistency requirements, not using QED. The cross section was calculated in QED only a few years later by Bethe and Fermi. Moller scattering was not measured until 1950 (by a Canadian, Lorne Albert Page). This is partly because people didn't consider it interesting until renormalization was understood and radiative corrections could be measured.

- $e^-\mu^- \to e^-\mu^-$: Electron-muon scattering, the same as Rutherford scattering. Crossing of electron-positron annihilation into muons.

- $\gamma e^- \to \gamma e^-$: Compton Scattering.

  Observed in 1923 by Arthur Holly Compton. The differential scattering formula is called the Klein-Nishina formula (Oskar Klein and Yoshio Nioshina, 1929). This was one of the first results obtained from QED. Before this, all that was known was the classical Thomson scattering formula, which was already in disagreement with experiment by the early 1920s.

- $e^-e^+ \to \gamma\gamma$: Pair annihilation into photons. Crossing of Compton scattering

- $\gamma\gamma \to e^+e^-$: Photon annihilation into e^+e^-.

- $\gamma\gamma \to \gamma\gamma$: Halpern Scattering (1933) Otto Halpern. Also known as light-by-light scattering. No tree-level contribution. Begins at 1-loop.

- $e^-e^- \to \mu^-\mu^-$: forbidden in QED. In the standard model, electron number *and* muon number are separately conserved.

Chapter 14
Path Integrals

14.1 Introduction

We have learned quantum field theory using the canonical quantization approach, which is based on creation and annihilation operators. There is a completely different way to do QFT called the Path Integral formulation. It says

$$\langle 0|T\{\phi(x_1)\cdots\phi(x_n)\}|0\rangle = \int \mathcal{D}\phi\phi(x_1)\cdots\phi(x_n)e^{iS[\phi]} \tag{14.1}$$

The left-hand side is exactly the kind of time-ordered product we use to calculate S-matrix elements. The $\mathcal{D}\phi$ on the right-hand side means integrate over all possible *classical* field configurations $\phi(x,\ t)$ with a phase given by the *classical* action evaluated in that field configuration.

The intuition for the path integral comes from a simple thought experiment you can do in quantum mechanics. Recall the double slit experiment: the amplitude for field to propagate from a source through a screen with two slits to a detector is the sum of the amplitudes to propagate through each slit separately. We add up the two amplitudes separate and then square to get the probability. Now try three slits, you get the sum of 3 paths. Now, keep slitting until the screen is gone. The final answer is that the amplitude is the sum of all possible different paths. That's all the path integral is calculating.

There's something very similar classically, that you might remember from your E&M class, called Huygens's principle. It says that to calculate the propagation of electromagnetic waves, you can treat each point in the wavefront as the center of a fresh disturbance and a new source for the weaves. This is useful, for example, in thinking about diffraction, where you can propagate the plane wave along to the slits, and then start the waves propagating anew from each slit. It's totally intuitive and works for any waves. Just think of a pond with some sticks in it where you can really see the waves moving along. If the wave goes through a hole, a new wave starts from the hole and keeps going. So the path integral is doing just this kind of wave-propagation analysis, but for quantum-mechanical particle-waves, rather than classical waves.

There are a number of amazing things about path integrals. For example, they imply that by dealing with only classical field configurations you get the quantum amplitude. This is really crazy if you think about it – these classical fields all commute, so you are also getting the non-commutativity for free somehow. Time ordering also just seems to drop out. And where are the particles? What happened to second quantization? In some ways, path integrals take the wave nature of matter to be primary, while the canonical method starts with particles.

Path integrals are in many ways simpler and more beautiful the canonical quantization, but they obscure some of the physics. Nowadays, people often just start with the path integral, using it to *define* the theory (even though, the integral over field configurations $\mathcal{D}\phi$ is not exactly well-defined mathematically). The canonical procedure, where we start with free fields and add interactions, is inherently perturbative. Path integrals on the other hand, don't mention free fields at all, so they can be used to study non-perturbative things, like lattice QCD, instantons, black holes, etc. Thus path integrals are more general, but they are equivalent to canonical quantization for perturbative calculations, and provide drastically different physical intuition.

To begin, we'll derive equation 1, then we'll re-derive the Feynman rules and continue discussing who to interpret the path integral.

14.2 Gaussian integrals

Before we begin, let's review how to compute the Gaussian integral

$$\mathcal{I} = \int_{-\infty}^{\infty} dp e^{-\frac{1}{2}ap^2+Jp} \tag{14.2}$$

First of all, complete the square.

$$\mathcal{I} = \int_{-\infty}^{\infty} dp e^{-\frac{1}{2}a(p-\frac{J}{a})^2+\frac{J^2}{2a}} \tag{14.3}$$

Then shift $p \to p + \frac{J}{a}$. The measure doesn't change. So

$$\mathcal{I} = e^{\frac{J^2}{2a}} \int_{-\infty}^{\infty} dp e^{-\frac{1}{2}ap^2} = \frac{1}{\sqrt{a}} e^{\frac{J^2}{2a}} \int dp e^{-\frac{1}{2}p^2} \tag{14.4}$$

Now use a trick:

$$[\int dp e^{-\frac{1}{2}p^2}]^2 = \int dx \int dy e^{-\frac{1}{2}x^2} e^{-\frac{1}{2}y^2} = 2\pi \int_0^{\infty} r dr e^{-\frac{1}{2}r^2} = \pi \int_0^{\infty} dr^2 e^{-\frac{1}{2}r^2} = 2\pi \tag{14.5}$$

So,

$$\int_{-\infty}^{\infty} dp e^{-\frac{1}{2}ap^2+Jp} = \sqrt{\frac{2\pi}{a}} e^{\frac{J^2}{2a}} \tag{14.6}$$

We can also generalize it for many p_i, so that $ap^2 = p_i a_{ij} p_j = \vec{p} A \vec{p}$, with A a matrix. If we diagonalize A then the integral is just a product of integrals over the p_i, so we just get the product of all the factors above. That is,

$$\boxed{\int_{-\infty}^{\infty} d\vec{p}\, e^{-\frac{1}{2}\vec{p}A\vec{p}+\vec{J}\vec{p}} = \sqrt{\frac{(2\pi)^n}{\det A}} e^{\frac{1}{2}\vec{J}A^{-1}\vec{J}}} \tag{14.7}$$

This formula is really useful for path integrals.

14.3 The Path Integral

The easiest way to derive the path integral is to start with non-relativistic quantum mechanics.

14.3.1 path integral in QM

In non-relativistic quantum mechanics, the Hamiltonian is

$$H(x,t) = \frac{p^2}{2m} + V(x,t) \tag{14.8}$$

Suppose our initial state $|i\rangle$ is a particle at x_i at time t_i and we want to project it onto the final state $|f\rangle$ at x_f and time t_f. We find

$$\langle f|i\rangle = \langle x_f t_f | x_i t_i\rangle \sim \left\langle x_f \middle| e^{-iH(t_f - t_i)} \middle| x_i \right\rangle \tag{14.9}$$

This is only heuristic because $H(t)$ and $H(t')$ don't commute, so we can't just solve the Schrodinger equation in this simple way with an exponential. However, we *can* solve it this way for infinitesimal time intervals. So let's break this down into small time intervals δt with times $t_j = t_i + j\delta_t$ and $t_n = t_f$.

$$\langle f|i\rangle = \int dx_n \cdots dx_1 \left\langle x_f \middle| e^{-iH(x_f,t_f)\delta t} | x_n \right\rangle \langle x_n | \cdots | x_2 \rangle \left\langle x_2 | e^{-iH(x_2,t_2)\delta t} | x_1 \right\rangle \left\langle x_1 | e^{-iH(x_1,t_1)\delta t} \middle| x_i \right\rangle$$

If we use a basis of free fields

$$\langle p|x\rangle = e^{ipx} \tag{14.10}$$

to give

$$\left\langle x_{j+1}\right|e^{-iH\delta t}\left|x_j\right\rangle = \int \frac{dp}{2\pi}\langle x_{j+1}|p\rangle\langle p|e^{-i\left[\frac{p^2}{2m}+V(x_j,t_j)\right]\delta t}|x_j\rangle \tag{14.11}$$

$$= e^{-iV(x_j,t_j)\delta t}\int \frac{dp}{2\pi}e^{-i\frac{p^2}{2m}\delta t}e^{-ip[x_{j+1}-x_j]} \tag{14.12}$$

Now we can use the Gaussian integral formula we computed above.

$$\int_{-\infty}^{\infty} dp e^{-\frac{1}{2}ap^2+Jp} = \sqrt{\frac{2\pi}{a}}e^{\frac{J^2}{2a}} \tag{14.13}$$

For our integral, $a = i\frac{\delta t}{m}$, $J = -i(x_{j+1}-x_j)$ so we get:

$$\left\langle x_{j+1}\right|e^{-iH\delta t}\left|x_j\right\rangle = Ne^{-iV(x,t)\delta t}e^{i\frac{m}{2}\delta t\frac{(x_{j+1}-x_j)^2}{(\delta t)^2}} \tag{14.14}$$

$$= Ne^{i[\frac{1}{2}m\dot{x}^2-V(x,t)]\delta t} \tag{14.15}$$

$$= Ne^{i\mathcal{L}(x,\dot{x})\delta t} \tag{14.16}$$

where N is the normalization, which we'll ignore. All that happened here is that the Gaussian integral performed the Legendre transform for us, to go from $H(x,p)\to L(x,\dot{x})$.

So,

$$\langle f|i\rangle = N\int dx_n\cdots dx_1\, e^{i\mathcal{L}(x_n,\dot{x}_n)\delta t}\cdots e^{i\mathcal{L}(x_1,\dot{x}_1)\delta t} \tag{14.17}$$

and taking the limit $\delta t\to 0$ we get

$$\langle f|i\rangle = N\int \mathcal{D}x(t)e^{iS(x,\dot{x})} \tag{14.18}$$

where we sum over all paths $x(t)$ and the action is $S=\int dt\mathcal{L}$.

14.4 Path integral in QFT

The field theory derivation in field theory, is very similar, but the set of intermediate states is more complicated. Let's start with just the vacuum matrix element

$$\langle 0,t_f|0,t_i\rangle \sim \left\langle 0\right|e^{-iH(t_f-t_i)}\left|0\right\rangle \tag{14.19}$$

In QM we broke this down into integrals over $|x\rangle\langle x|$ for intermediate times. The states $|x\rangle$ are eigenstates of the $\hat{x}$ operator. In field theory, the equivalent of $\hat{x}$ is $\hat{\phi}(x)$, where, in case you've forgotten,

$$\hat{\phi}(\vec{x}) = \int \frac{d^3k}{(2\pi)^3}(a_k^\dagger e^{i\vec{k}\vec{x}}+a_k e^{-i\vec{k}\vec{x}}) \tag{14.20}$$

This is an infinite number of operators, one at each point x. We put the hat on ϕ to remind you that it's an operator. Then the equivalent of $|x\rangle\langle x|$ is a complete set of eigenstates of $\hat{\phi}$

$$\hat{\phi}(\vec{x})|\Phi\rangle = \Phi(\vec{x})|\Phi\rangle \tag{14.21}$$

The eigenvalues are functions of space $\Phi(\vec{x})$. Of course, $\Phi(\vec{x})$ will have to satisfy some kind of normalization condition, so that it has a fixed number of particles, but we will worry about that later.

Now our basis is over states $|\Phi\rangle$. So our matrix element can be chopped up into

$$\langle 0,t_f|0,t_i\rangle = \int \mathcal{D}\Phi_1(x)\cdots\mathcal{D}\Phi_n(x)\left\langle 0\right|e^{-iH(t_n)\delta t}|\Phi_n\rangle\langle\cdots\rangle\left\langle\Phi_2\right|e^{-iH(t_2)\delta t}|\Phi_1\rangle\left\langle\Phi_1\right|e^{-iH(t_1)\delta t}|0\rangle$$

Now each of these pieces becomes

$$\left\langle \Phi_{j+1}|e^{iH(t_j)\delta t}|\Phi_j\right\rangle = e^{i\int d^3x\mathcal{L}[\Phi,\dot{\Phi}]\delta t} \tag{14.22}$$

And taking $\delta t \to 0$ gives

$$\langle 0,t_f|0,t_i\rangle = \int \mathcal{D}\Phi(x,t)e^{iS[\Phi]} \tag{14.23}$$

where $S[\Phi] = \int d^4x\mathcal{L}$.

So the path integral tells us to integrate over all *classical* field configurations Φ. Note that Φ is not just the one-particle states, it can have 2-particle states, etc. We can remember this by drawing pictures for the paths – including disconnected bubbles. Actually, we really sum over all kinds of discontinuous, disconnected random fluctuations, but the ones that dominate the path integral will have a nice physical interpretation as we can see.

14.5 Classical limit

Let's put back $\hbar$ for a moment. Since $\hbar$ has dimensions of action, it is simply

$$\langle 0,t_f|0,t_i\rangle = \int \mathcal{D}\Phi(x,t)e^{\frac{i}{\hbar}S[\Phi]} \tag{14.24}$$

Let's forget about the i for a moment and suppose we need to calculate

$$\int \mathcal{D}\Phi(x,t)e^{-\frac{1}{\hbar}S[\Phi]} \tag{14.25}$$

In this case, the integral would clearly be dominated by Φ_0 which is where $S[\Phi]$ has a minimum – everything else would give a bigger $S[\Phi]$ and be infinitely more suppressed as $\hbar \to 0$.

Now, when we put the i back in, the same thing happens, not because the non-minimal terms are zero, but because away from the minimum you have to sum over phases swirling around infinitely fast. When you sum infinitely swirling phases, you also get something which goes to zero. There's a theorem that says this happens, related to the *method of stationary descent* or sometimes *method of steepest descent*. Another way to see it is to use the more intuitive case with $e^{-S[\Phi]/\hbar}$. Since we expect the answer to be well defined, it should be an analytic function of Φ_0. So we can take $\hbar \to 0$ in the imaginary direction, showing that the integral is still dominated by $S[\Phi_0]$.

In any case, the classical limit is dominated by a field configuration which minimizes the action. This is the classical path that a particle would take. So the path integral has a nice intuitive classical correspondence.

14.6 Time-ordered products

Suppose we insert a field at fixed x and t into the path integral

$$\mathcal{I} = \int \mathcal{D}\Phi(x,t)e^{iS[\Phi]}\Phi(x_j,t_j) \tag{14.26}$$

What does this represent?

Going back through our derivation, we have

$$\mathcal{I} = \int \mathcal{D}\Phi_1(x)\cdots\mathcal{D}\Phi_n(x)\left\langle 0\middle|e^{iH(t_n)\delta t}|\Phi_n\right\rangle\cdots\left\langle \Phi_2|e^{iH(t_2)\delta t}|\Phi_1\right\rangle\left\langle \Phi_1|e^{iH(t_1)\delta t}|0\right\rangle\Phi_j(x_j)$$

since the subscript on Φ is just it's point in time, we know which Φ_i's these correspond to. Let's take the part with just Φ_j

$$\int \mathcal{D}\Phi_j(x)\left\{e^{iH(t_n)\delta t}|\Phi_j\rangle\Phi_j(x_j)\langle\Phi_j|\right\} = \hat{\phi}(x_j)\int \mathcal{D}\Phi_j(x)|\Phi_j\rangle\langle\Phi_j| \tag{14.27}$$

So we get to replace $\Phi(x_j)$ by the operator $\hat{\phi}(x_j)$ stuck in at the time t_j. Then we can collapse up all the integrals to give

$$\int \mathcal{D}\Phi(x,t) e^{iS[\Phi]} \Phi(\vec{x}_j, t_j) = \left\langle 0 \middle| \hat{\phi}(\vec{x}_j, t_j) \middle| 0 \right\rangle \tag{14.28}$$

If you find the collapsing-up-the-integrals confusing, just think about the derivation backwards. An insertion of $\hat{\phi}(\vec{x}_j, t_j)$ will end up by $|\Phi_j\rangle\langle\Phi_j|$, and then pull out a factor of $\Phi(x_j, t_j)$.

Now say we insert two fields

$$\int \mathcal{D}\Phi(\vec{x}\,, t) e^{iS[\Phi]} \Phi(\vec{x}_1, t_1) \Phi(\vec{x}_2, t_2) \tag{14.29}$$

The fields will get inserted in the appropriate matrix element. But if you check, you will see that the earlier field will always come out on the right of the later field. So we get

$$\int \mathcal{D}\Phi(x) e^{iS[\Phi]} \Phi(x_1)\Phi(x_2) = \left\langle 0 \middle| T\left\{ \hat{\phi}(x_1)\hat{\phi}(x_2) \right\} \middle| 0 \right\rangle \tag{14.30}$$

We get time ordering for free in the path integral. In general

$$\int \mathcal{D}\Phi(x) e^{iS[\Phi]} \Phi(x_1)\cdots\Phi(x_n) = \left\langle 0 \middle| T\left\{ \hat{\phi}(x_1)\cdots\hat{\phi}(x_n) \right\} \middle| 0 \right\rangle \tag{14.31}$$

Why does this work? There are a few cross checks you can do. As an easy one, suppose the answer were

$$\int \mathcal{D}\Phi(x) e^{iS[\Phi]} \Phi(x_1)\Phi(x_2) = \left\langle 0 \middle| \hat{\phi}(x_1)\hat{\phi}(x_2) \middle| 0 \right\rangle \tag{14.32}$$

Well, the left hand side doesn't care whether I write $\Phi(x_1)\Phi(x_2)$ or $\Phi(x_2)\Phi(x_1)$, since these are classical fields. So what would determine what order I write the fields on the right? We see it must be something that makes the fields effectively commute, like the time-ordering operator. We'll do another check once we get a little more experience playing with the path integral.

From now on, we'll just use $\phi(x)$ instead of $\Phi(x)$, for the classical fields.

14.6.1 Current shorthand

There's a great way to calculate path integrals using currents. Let's add a source term to our action. So define

$$Z[J] = \int \mathcal{D}\phi \exp\left\{ iS[\phi] + i\int d^4x\, J(x)\phi(x) \right\} \tag{14.33}$$

Then,

$$Z[0] = \int \mathcal{D}\phi e^{i\int d^4x \mathcal{L}[\phi]} = \langle 0|0\rangle \tag{14.34}$$

Next, observe that

$$\frac{d}{dJ(x_1)} \int d^4x\, J(x)\phi(x) = \phi(x_1) \tag{14.35}$$

So,

$$-i\frac{dZ}{dJ(x_1)} = \int \mathcal{D}\phi \exp\left\{ iS[\phi] + i\int d^4x\, J(x)\phi(x) \right\} \phi(x_1) \tag{14.36}$$

And thus

$$-i\frac{dZ}{dJ(x_1)}|_{J=0} = \int \mathcal{D}\phi \exp\{iS[\phi]\}\phi(x_1) = \langle 0|\phi(x_1)|0\rangle \tag{14.37}$$

Similarly,

$$(-i)^n \frac{d^n Z}{dJ(x_1)\cdots dJ(x_n)}|_{J=0} = \langle 0|T\{\phi(x_1)\cdots\phi(x_n)\}|0\rangle \tag{14.38}$$

So this is a nice way of calculating time-ordered products.

14.7 Solving the free Path Integral

First, let's look at the free theory.

$$\mathcal{L} = -\frac{1}{2}\phi(\Box + m^2)\phi \tag{14.39}$$

Then

$$Z_0[J] = \int \mathcal{D}\phi \exp\left\{ i \int d^4x \left(-\frac{1}{2}\phi(\Box + m^2)\phi + J(x)\phi(x) \right) \right\} \tag{14.40}$$

Now we can solve this exactly, since we already have our relation

$$\int_{-\infty}^{\infty} d\,\vec{p}\, e^{-\frac{1}{2}\vec{p}A\vec{p} + \vec{J}\vec{p}} = \sqrt{\frac{(2\pi)^n}{\det A}} e^{\frac{1}{2}\vec{J}A^{-1}\vec{J}} \tag{14.41}$$

The path integral is just an infinite number of p_i components.

This is exactly what we have, with $A = (\Box + m^2)$. We have already studied the inverse of $(\Box + m^2)$,

$$(\Box + m^2)\phi(x) = J(x) \quad \Rightarrow \quad \phi\,(x) = \int d^4x\, \Pi(x-y) J(y) \tag{14.42}$$

where Π is the Green's function satisfying

$$(\Box_x + m^2)\Pi(x-y) = \delta(x-y) \tag{14.43}$$

Explicitly,

$$\Pi(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{1}{p^2 - m^2} e^{ip(x-y)} \tag{14.44}$$

up to boundary conditions. Thus

$$Z_0[J] = N \exp\left\{ i \int d^4x \int d^4y \frac{1}{2} J(x)\Pi(x-y)J(y) \right\} \tag{14.45}$$

And so,

$$\langle 0|T\{\phi(x)\phi(y)\}|0\rangle = (-i)^2 \frac{d^2 Z}{dJ(x)dJ(y)}|_{J=0} = i\Pi(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2} e^{ip(x-y)} \tag{14.46}$$

What happened to the time-ordering?

The standard answer is that we need to put a convergence factor in the path integral

$$Z_0[J] = \int \mathcal{D}\phi \exp\left\{ i \int d^4x - \frac{1}{2}\phi(\Box + m^2)\phi + J(x)\phi(x) \right\} \exp\left\{ -\varepsilon \int d^4x \phi^2) \right\} \tag{14.47}$$

$$= \int \mathcal{D}\phi \exp\left\{ i \int d^4x \frac{1}{2}\phi(-\Box - m^2 + i\epsilon)\phi + J(x)\phi(x) \right\} \tag{14.48}$$

$$= \exp\left\{ -\frac{1}{2} \int d^4x \int d^4y J(x) D_F(x-y) J(y) \right\} \tag{14.49}$$

Where $D_F(x-y)$ is the Feynman propagator

$$\langle 0|T\{\phi(x)\phi(y)\}|0\rangle = (-i)^2 \frac{d^2 Z_0}{dJ(x)dJ(y)}|_{J=0} = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\varepsilon} e^{ip(x-y)} = D_F(x-y) \tag{14.50}$$

I'm not sure how legitimate this is. It's really hard to define the path integral mathematically, but at least this convergence factor sounds believable. In any case, we knew we had to get the time-ordered product out, so this is the correct answer.

So, for for the free field case we have solved the free path integral *exactly*. We get the same thing as just using the classical equations of motion and plugging back in. That is what we have been calling integrating out a field. But note that we are not integrating out the field when we do the Gaussian integral, we are actually computing the path integral. This only works for *free* fields, where we can solve it exactly. So there is no difference between a classical and quantum system of free fields. If there are interactions, we can approximate the integral by expanding around the minimum and then integrating out the field. This works if the quantum corrections are small, or the interactions are weak.

14.7.1 4-point function

We can also compute higher order products

$$\Big\langle 0|T\{\hat{\phi}(x_1)\hat{\phi}(x_2)\hat{\phi}(x_3)\hat{\phi}(x_4)\Big\}|0\Big\rangle = (-i)^4 \frac{d^4 Z_0}{dJ(x_1)\cdots dJ(x_4)}\Big|_{J=0} \tag{14.51}$$

$$= \frac{d^4}{dJ(x_1)\cdots dJ(x_4)}\, e^{-\frac{1}{2}\int d^4x \int d^4y J(x) D_F(x-y) J(y)}\Big|_{J=0} \tag{14.52}$$

$$= \frac{d^4}{dJ(x_1)dJ(x_2)dJ(x_3)}\Big(- \int d^4z D_F(x_4 - z) J(z)\Big) e^{-\frac{1}{2}\int d^4x \int d^4y J(x) D_F(x-y) J(y)}\Big|_{J=0} \tag{14.53}$$

Before we continue, let's simplify the notation

Let us abbreviate this as

$$\langle 0|T\{\phi_1\phi_2\phi_3\phi_4\}|0\rangle = \frac{d^4}{dJ_1 dJ_2 dJ_3 dJ_4}\, e^{-\frac{1}{2}J_x D_{\mathrm{x\,y}} J_y}\Big|_{J=0} \tag{14.54}$$

$$= \frac{d^3}{dJ_1 dJ_2 dJ_3}(-J_z D_{z4})\, e^{-\frac{1}{2}J_x D_{\mathrm{x\,y}} J_y}\Big|_{J=0} \tag{14.55}$$

$$= \frac{d^2}{dJ_1 dJ_2}(-D_{34} + J_z D_{z3} J_w D_{w4})\, e^{-\frac{1}{2}J_x D_{\mathrm{x\,y}} J_y}\Big|_{J=0} \tag{14.56}$$

$$= \frac{d}{dJ_1}(D_{34} J_z D_{z2} + D_{23} J_w D_{w4} + J_z D_{z3} D_{24} - J_z D_{z3} J_w D_{w4} J_r D_{r2})\, e^{-\frac{1}{2}J_x D_{\mathrm{x\,y}} J_y}\Big|_{J=0} \tag{14.57}$$

$$= D_{34}D_{12} + D_{23}D_{14} + D_{13}D_{24} \tag{14.58}$$

So we get

$$\Big\langle 0|T\{\hat{\phi}(x_1)\hat{\phi}(x_2)\hat{\phi}(x_3)\hat{\phi}(x_4)\Big\}|0\Big\rangle = \tag{14.59}$$

$$D_F(x_3-x_4)D_F(x_1-x_2) + D_F(x_2-x_3)D_F(x_1-x_3) + D_F(x_1-x_3)D_F(x_2-x_4) \tag{14.60}$$

These are the 3 possible contractions. This is exactly what we found from time-dependent perturbation theory.

14.8 Interactions

Now suppose we have interactions

$$\mathcal{L} = -\frac{1}{2}\phi(\Box + m^2)\phi + \lambda\phi^4 \tag{14.61}$$

Then, we can write

$$Z[J] = \int \mathcal{D}\phi e^{i\int d^4x \left[\frac{1}{2}\phi(-\Box - m^2 + i\varepsilon)\phi + J(x)\phi(x) + \lambda\phi^4\right]} \tag{14.62}$$

$$= \int \mathcal{D}\phi e^{i\int d^4x \frac{1}{2}\phi(-\Box - m^2 + i\varepsilon)\phi + J(x)\phi(x)} e^{i\int d^4x \frac{\lambda}{4!}\phi^4} \tag{14.63}$$

$$= \int \mathcal{D}\phi e^{i\int d^4x \frac{1}{2}\phi(-\Box - m^2 + i\varepsilon)\phi + J(x)\phi(x)} \left[1 + i\lambda \int d^4z \phi^4(z) + \frac{(i\lambda)^2}{2!} \int d^4z \int d^4w \phi^4(z)\phi^4(w) + \cdots \right]$$

Then

$$\left\langle 0|T\{\hat{\phi}(x_1)\hat{\phi}(x_2)\}|0 \right\rangle = \left\langle 0|T\{\hat{\phi}_I(x_1)\hat{\phi}_I(x_2)\}|0 \right\rangle + i\lambda \int d^4z \left\langle 0|T\{\hat{\phi}_I(x_1)\hat{\phi}_I(x_2)\hat{\phi}_I(z)^4\}|0\rangle + \cdots \right.$$

$$= \left\langle 0|T\{\hat{\phi}_I(x_1)\hat{\phi}_I(x_2)e^{i\int d^4z \lambda \hat{\phi}_I(z)^4}\}|0 \right\rangle \tag{14.64}$$

I have added the I subscript to remind us that these are fields in the free theory, evolving with the free Hamiltonian (what we called the interaction picture before). So we have reproduced exactly the expression for the perturbation expansion the generated the Feynman rules. That is, have reproduced the Feynman rules from the path integral.

Another occasionally useful notation is

$$Z[J] = e^{i\int d^4x V[\frac{d}{dJ(x)}]} \int \mathcal{D}\phi e^{-\frac{1}{2}\int d^4x \int d^4y J(x) D_F(x-y) J(y)} \tag{14.65}$$

$$= e^{i\int d^4x V[\frac{d}{dJ(x)}]} Z_0[J]\Big|_{J=0} \tag{14.66}$$

This emphasizes that everything is determined by Z_0.

Finally, remember that what we were interested in for S-matrix elements was

$$\langle \Omega|T\{\phi(x_1)\cdots\phi(x_n)\}|\Omega\rangle = \frac{\left\langle 0|T\{\phi(x_1)\cdots\phi(x_n)e^{i\int d^4z \lambda \hat{\phi}_I(z)^4}|0\right\rangle}{\left\langle 0|e^{i\int d^4z \lambda \hat{\phi}_I(z)^4}|0\right\rangle} \tag{14.67}$$

$$= \frac{1}{Z[J]} \frac{d^n Z[J]}{dJ^n}\Big|_{J=0} \tag{14.68}$$

Since the same Z appears in the numerator and the denominator, the normalization of Z drops out, which is why we have been able to ignore it so far.

14.9 The ward identity

One of the key things that makes path integrals useful is that we can do field redefinitions.

For example, say we have the path integral for some gauge theory

$$\mathcal{L}[A_\mu, \phi_i] = -\frac{1}{4}F_{\mu\nu}^2 + |D_\mu \phi_i|^2 - m^2|\phi_i|^2 + \cdots \tag{14.69}$$

Then the path integral is

$$Z[0] = \int \mathcal{D}\phi_i \mathcal{D}\phi_i^\star \mathcal{D}A_\mu e^{i\int d^4x \mathcal{L}[A,\phi_i]} \tag{14.70}$$

It is just the integral over all the fields.

Now suppose we change variables

$$A_\mu = A'_\mu(x) + V_\mu(x) \tag{14.71}$$

$$\phi_i = \phi'_i(x) + \Delta_i(x) \tag{14.72}$$

Since we are integrating over *all* fields, the path integral is automatically invariant, for any $V_\mu(x)$ and any $\Delta_i(x)$.

$$\int \mathcal{D}\phi'_i \mathcal{D}\phi_i'^\star \mathcal{D}A'_\mu e^{i\int d^4x \mathcal{L}[A',\phi'_i]} = \int \mathcal{D}\phi_i \mathcal{D}\phi_i^\star \mathcal{D}A_\mu e^{i\int d^4x \mathcal{L}[A,\phi_i]} \tag{14.73}$$

This is not very impressive.

Now consider a change of variables which is a gauge transformation

$$A_\mu(x) = A'_\mu(x) + \partial_\mu \alpha(x) \tag{14.74}$$

$$\phi_i(x) = e^{i\alpha(x)}\phi'_i \tag{14.75}$$

For some function $\alpha(x)$. Then we can make stronger statements. The Lagrangian by itself is invariant, because this is a gauge transformation.

$$e^{i\int d^4x \mathcal{L}[A',\phi'_i]} = e^{i\int d^4x \mathcal{L}[A,\phi_i]} \tag{14.76}$$

The measure $\mathcal{D}A$ is invariant because this is just a linear shift. The measure $\mathcal{D}\phi_i$ changes by a factor $e^{i\alpha(x)}$ but $\mathcal{D}\phi_i^\star$ changes by $e^{-i\alpha(x)}$ and these two factors cancel. Since for every field we need to integrate over it and its conjugate, the cancellation will always occur. So

$$\int \mathcal{D}\phi'_i \mathcal{D}\phi'^\star_i \mathcal{D}A'_\mu = \int \mathcal{D}\phi_i \mathcal{D}\phi_i^\star \mathcal{D}A_\mu \tag{14.77}$$

Still not very impressive.

But now consider the correlation function of something involving an A_μ field

$$\langle 0|A_\mu(x_1)|0\rangle = \int \mathcal{D}\phi_i \mathcal{D}\phi_i^\star \mathcal{D}A_\mu e^{i\int d^4x \mathcal{L}[A,\phi_i]} A_\mu(x_1) \tag{14.78}$$

Now make the same replacement. Then any change of variables will not affect the path integral,

$$\int \mathcal{D}\phi_i \mathcal{D}\phi_i^\star \mathcal{D}A_\mu e^{i\int d^4x \mathcal{L}[A,\phi_i]} A_\mu(x_1) = \int \mathcal{D}\phi'_i \mathcal{D}\phi'^\star_i \mathcal{D}A'_\mu e^{i\int d^4x \mathcal{L}[A',\phi'_i]} A'_\mu(x) \tag{14.79}$$

But also, since the measure and action are separately invariant, this simplifies to

$$= \int \mathcal{D}\phi_i \mathcal{D}\phi_i^\star \mathcal{D}A_\mu e^{i\int d^4x \mathcal{L}[A,\phi_i]} [A_\mu(x_1) + \partial_\mu \alpha(x_1)] \tag{14.80}$$

which means

$$\int \mathcal{D}\phi_i \mathcal{D}\phi_i^\star \mathcal{D}A_\mu e^{i\int d^4x \mathcal{L}[A,\phi_i]} \partial_\mu \alpha(x_1) = 0 \tag{14.81}$$

This is an example of the Ward identity.

It follows more generally as well. Suppose we add in some more fields

$$\langle 0|A_\mu(x_1)\phi_i(x_2)\cdots\phi_j(x_n)|0\rangle = \int \mathcal{D}\phi_i \mathcal{D}\phi_i^\star \mathcal{D}A_\mu e^{i\int d^4x \mathcal{L}[A,\phi_i]} A_\mu(x_1)\phi_i(x_2)\cdots\phi_j(x_n) \tag{14.82}$$

$$= \int \mathcal{D}\phi_i \mathcal{D}\phi_i^\star \mathcal{D}A_\mu e^{i\int d^4x \mathcal{L}[A,\phi_i]} [A_\mu(x_1) + \partial_\mu \alpha(x_1)] e^{i\alpha(x_2)}\phi_i(x_2)\cdots e^{i\alpha(x_n)}\phi_j(x_n) \tag{14.83}$$

These phase factors are just numbers, which don't matter. For example, just taking the absolute value, we find

$$|\langle 0|A_\mu(x_1)\phi_i(x_2)\cdots\phi_j(x_n)|0\rangle| = |\langle 0|A_\mu(x_1)\phi_i(x_2)\cdots\phi_j(x_n)|0\rangle + \langle 0|\partial_\mu\alpha(x_1)\phi_i(x_2)\cdots\phi_j(x_n)|0\rangle|$$

This implies that $\langle 0|\partial_\mu\alpha(x_1)\phi_i(x_2)\cdots\phi_j(x_n)|0\rangle = 0$ which is the ward identity with one A_μ.

Explicitly, if we write

$$\langle 0|A(x_1)\phi_i(x_2)\cdots\phi_j(x_n)|0\rangle = \int \frac{d^4k}{(2\pi)^4}\varepsilon_\mu(k) M^\mu(k, x_2, \cdots x_n) e^{ikx_1} \tag{14.84}$$

we have shown

$$\int \frac{d^4k}{(2\pi)^4}\partial_\mu\alpha(x_1) M^\mu(k, x_2, \cdots x_n) e^{ikx_1} = -\alpha(x_1)\int \frac{d^4k}{(2\pi)^4} k_\mu M^\mu(k, x_2, \cdots x_n) e^{ikx_1} = 0 \tag{14.85}$$

$$\Rightarrow \boxed{k_\mu M^\mu(k, x_2, \cdots x_n) = 0} \tag{14.86}$$

which is the Ward identity in momentum space.

Finally, suppose there are lots of A_μ's. We find

$$\langle 0|A_\mu(x_1)\cdots A_\mu(x_n)\cdots|0\rangle = \langle 0|(A_\mu(x_1)+\partial_\mu\alpha(x_1))\cdots(A_\mu(x_n)+\partial_\mu\alpha(x_n)\cdots|0\rangle \tag{14.87}$$

For α infinitesimal, this implies

$$\langle 0|\partial_\mu(x_1)\cdots A_\mu(x_n)|0\rangle + \langle 0|A_\mu(x_1)\partial_\mu(x_2)\cdots|0\rangle + \langle 0|A_\mu\cdots\partial_\mu\alpha(x_n)\cdots|0\rangle = 0 \tag{14.88}$$

which is the statement that we have to sum over all possible ways that a photon can be replaced with a k_μ. Because the photons are identical particles, we would have to sum over all these graphs to get the amplitude for the polarization $\epsilon_\mu = k_\mu$, so this is exactly what we would calculate.

Thus, the path integral makes the derivation of the Ward identity easy. It is also completely non-perturbative.

14.10 Gauge invariance: Fadyeev-Popov

Another thing that's easy to do with path integrals is to prove gauge invariance, meaning ξ dependence of the amplitudes.

Recall that vector field Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu} = \frac{1}{2}A_\mu(-k^2\eta_{\mu\nu}+k_\mu k_\nu)A_\nu + J_\mu A_\mu \tag{14.89}$$

And the equations of motion are

$$(-k^2\eta_{\mu\nu}+k_\mu k_\nu)A_\nu = J_\mu \tag{14.90}$$

which is not invertible because this matrix has 0 determinant (it has an eigenvector k_μ with eigenvalue 0). The physical reason it's not invertible is because we can't uniquely solve for A_μ in terms of J_μ because of gauge invariance:

$$A_\mu \to A_\mu + \partial_\mu\alpha(x) \tag{14.91}$$

That is, many currents correspond to the same vector field. Our previous solution was to gauge fix by adding the term $\frac{1}{\xi}(\partial_\mu A_\mu)^2$ to the Lagrangian. Now we will justify that prescription, and prove gauge invariance in general, through the Fadyeev-Popov procedure.

With a general set of fields ϕ_i and interactions we are interested in computing

$$Z_{\mathcal{O}} = \int \mathcal{D}A_\mu \mathcal{D}\phi_i e^{i\int d^4x \mathcal{L}[A,\phi_i]}\mathcal{O} \tag{14.92}$$

where $\mathcal{O}$ refers to whatever we're taking the correlation function of (for example, $\mathcal{O} = \phi(x)\phi(y)$)

$$\langle\Omega|T\{\mathcal{O}\}|\Omega\rangle = \frac{Z_{\mathcal{O}}}{Z_1} \tag{14.93}$$

Recall that under a gauge transformation $\partial_\mu A_\mu = \Box\alpha$. Since we can always go to a gauge where $\partial_\mu A_\mu = 0$, this means we can find a function $\alpha = \frac{1}{\Box}\partial_\mu A_\mu$. Now consider the following function

$$f(\xi) = \int \mathcal{D}\pi e^{i\int d^4x \frac{1}{\xi}(\Box\pi)^2} \tag{14.94}$$

This is just some function of ξ, probably infinite. Now let's shift

$$\pi(x) \to \pi(x) - \alpha(x) = \pi(x) - \frac{1}{\Box}\partial_\mu A_\mu \tag{14.95}$$

This is just a shift, so the integration measure doesn't change. Then

$$f(\xi) = \int \mathcal{D}\pi e^{i\int d^4x \frac{1}{\xi}(\Box\pi - \partial_\mu A_\mu)^2} \tag{14.96}$$

This is still just the same function of ξ, which despite appearances, is independent of A_μ. So,

$$Z_{\mathcal{O}} = \frac{1}{f(\xi)} \int \mathcal{D}\pi \mathcal{D}A_\mu \mathcal{D}\phi_i e^{i\int d^4x \mathcal{L}[A,\phi_i] + \frac{1}{\xi}(\Box\pi - \partial_\mu A_\mu)^2} \mathcal{O} \tag{14.97}$$

Now let's do the gauge transformation shift, with $\pi(x)$ as our gauge parameter.

$$A_\mu = A'_\mu + \partial_\mu \pi \tag{14.98}$$

$$\phi_i = e^{i\pi}\phi'_i \tag{14.99}$$

Again the measure $\mathcal{D}\pi\mathcal{D}A_\mu\mathcal{D}\phi_i$ and the action $\mathcal{L}[A,\phi_i]$ don't change. $\mathcal{O}$ may change. For example, if it's

$$\mathcal{O} = A_\mu(x_1)\cdots A_\mu(x_n)\phi_i(x_{n+1})\cdots\phi_j(x_m)$$

The A_μ transform to $A_\mu + k_\mu$ and the k_μ parts vanish by the Ward identity. The fields ϕ_i will give phases $e^{i\pi(x_i)}$, which don't don't depend on ξ. Then we get

$$\int \mathcal{D}\pi e^{i\pi(x_{n+1})}\cdots e^{i\pi(x_m)} = \int \mathcal{D}\pi' \tag{14.100}$$

where $\pi'(x) = e^{i\pi(x_{n+1})}\cdots e^{i\pi(x_m)}\pi(x)$. So this factor is actually independent of all the $x's$. Therefore, overall, we get some constant normalization times the gauge-fixed path integral

$$Z_{\mathcal{O}} = N_\xi \int \mathcal{D}A_\mu \mathcal{D}\phi_i e^{i\int d^4x \mathcal{L}[A,\phi_i] + \frac{1}{\xi}(\partial_\mu A_\mu)^2} \mathcal{O} \tag{14.101}$$

where $N_\xi = \frac{1}{f(\xi)}$. But since we are always computing

$$\langle\Omega|T\{\mathcal{O}\}|\Omega\rangle = \frac{Z_{\mathcal{O}}}{Z_1} \tag{14.102}$$

The N_ξ drops out. The point is that $Z_{\mathcal{O}}$ only depends on ξ through it's normalization. Thus the time-ordered products are ξ independent.

Thus all S-matrix elements of gauge invariant operators are guaranteed to be independent of ξ, which is the statement of gauge invariance. This proof is completely non-perturbative.

14.11 Fermionic path integral

For a path integral over fermions, it's basically the same, but we have to allow for the fact that the fermions anticommute. In the quantum theory, we saw that the quantum fields must anticommute to have a Lorentz-invariant S-matrix. But even classically, a Lorentz invariant spinor inner product must be antisymmetric. For example, the only Lorentz invariant constructed from a single Weyl spinor is

$$\psi_L^T\sigma_2\psi_L = \begin{pmatrix} \psi_1 & \psi_2 \end{pmatrix}\begin{pmatrix} & i \\ -i & \end{pmatrix}\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = i(\psi_1\psi_2 - \psi_2\psi_1) = |\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle \tag{14.103}$$

From this we concluded that $\{\psi_1, \psi_2\} = 0$ which means $\psi(x)$ is valued in the anti-commuting Grassman algebra.

The Grassman numbers is simply a set of objects which add commutatively and can by multiplied by real numbers. Consider a general function $f(\theta)$ of Grassman numbers θ. Since $\theta^2 = \{\theta, \theta\} = 0$, the Taylor expansion of $f(\theta)$ around $\theta = 0$ stops at first order

$$f(\theta) = a + b\theta \tag{14.104}$$

The basic thing we want to do with these functions is integrate them over all Grassman numbers to get a real number out. Since we're making up the rules for Grassman numbers as we go along, let's just define

$$\int d\theta(a+b\theta) = b \tag{14.105}$$

That is,

$$\int d\theta = 0, \qquad \int d\theta\theta = 1 \tag{14.106}$$

We also need a convention for multiple integrals, so we say

$$\int d\theta_1 d\theta_2 \theta_2 \theta_1 = \int d\theta_1 \theta_1 = 1 \tag{14.107}$$

Thus,

$$\int d\theta_2 d\theta_1 \theta_2 \theta_1 = -1 \tag{14.108}$$

And so on. These rules are internally consistent, and have some proper mathematical justification, which is totally uninteresting.

Now, let's try our Gaussian integral

$$\mathcal{I} = \int d\theta e^{\frac{1}{2}a\theta^2} = \int d\theta = 0 \tag{14.109}$$

Since $\theta^2 = 0$. Let's try with two $\theta's$. Let $\psi = (\theta_1\ \theta_2)^T$ and consider $\psi^T A\psi$. Then

$$\psi^T A\psi = (A_{12} - A_{21})\theta_1\theta_2 = \det(A)\theta_1\theta_2 \tag{14.110}$$

Where we have taken A antisymmetric to write it as a determinant. The symmetric part of A simply doesn't contribute. Then this is

$$\int d\theta_1 d\theta_2 e^{\psi^T A\psi} = \int d\theta_1 d\theta_2 (1 + \det(A)\theta_1\theta_2) = \det(A) \tag{14.111}$$

The generalization to n variables is simply

$$\int d\vec{\theta} e^{\vec{\theta}A\vec{\theta}} = \det(A) \tag{14.112}$$

This is very different from the bosonic integral

$$\int d\,\vec{p}\, e^{-\frac{1}{2}\vec{p}A\,\vec{p}} \sim \sqrt{\frac{1}{\det A}} = (\det A)^{-1/2} \tag{14.113}$$

But that's what it is. This det factor comes up from time to time, but mostly it is just the infinite normalization of the path integral, which we can ignore in taking ratios.

To calculate the generating functional $Z[J]$ we need a fermionic source η. Let's just take 2 fields ψ and $\bar{\psi}$, and complete the square

$$\bar{\psi}A\psi + \bar{\eta}\psi + \bar{\psi}\eta = (\bar{\psi} + \bar{\eta}A^{-1})A(\psi + A^{-1}\eta) - \bar{\eta}A^{-1}\eta \tag{14.114}$$

Then

$$\int d\psi d\bar{\psi} e^{\bar{\psi}A\psi + \bar{\eta}\psi + \bar{\psi}\eta} = e^{-\bar{\eta}A^{-1}\eta} \int d\psi d\bar{\psi} e^{(\bar{\psi} + \bar{\eta}A^{-1})A(\psi + A^{-1}\eta)} \tag{14.115}$$

$$= e^{-\bar{\eta}A^{-1}\eta} \int d\psi d\bar{\psi} e^{\bar{\psi}A\psi} = \det(A)\, e^{-\bar{\eta}A^{-1}\eta} \tag{14.116}$$

$$= \mathcal{N} e^{-\bar{\eta}A^{-1}\eta} \tag{14.117}$$

for some normalization $\mathcal{N}$.

In particular the two point function is

$$\langle 0|T\{\bar{\psi}(0)\psi(x)\}|0\rangle = \frac{d^2}{d\eta(0)d\bar{\eta}(x)}\int \mathcal{D}\psi\mathcal{D}\bar{\psi}e^{\bar{\psi}A\psi+\bar{\eta}\psi+\bar{\psi}\eta}\Big|_{\eta=0} = NA^{-1} \tag{14.118}$$

For a free Dirac field,

$$Z[\eta] = \int \mathcal{D}\psi\mathcal{D}\bar{\psi}e^{i\int d^4x[\bar{\psi}(i\not{\partial}-m+i\varepsilon)\psi+\bar{\eta}\psi+\bar{\psi}\eta} \tag{14.119}$$

where again, the $i\varepsilon$ comes in to make the path integral converge. Then $A = i\not{\partial} - m + i\varepsilon$ and we have

$$\langle 0|T\{\bar{\psi}(0)\psi(x)\}|0\rangle = \frac{1}{Z[\eta]}\frac{d^2Z[\eta]}{d\eta(0)d\bar{\eta}(x)}\Big|_{\eta=0} = i(i\not{\partial}-m+i\varepsilon)^{-1} \tag{14.120}$$

$$= \int \frac{d^4p}{(2\pi)^4}\frac{i}{\not{p}-m+i\varepsilon}e^{ipx} \tag{14.121}$$

This simplifies using $(\not{p}-m)(\not{p}+m) = p^2-m^2$, which implies

$$\frac{1}{\not{p}-m+i\varepsilon} = \frac{\not{p}+m}{p^2-m^2+i\varepsilon} \tag{14.122}$$

So

$$\langle \Omega|T\{\bar{\psi}(0)\psi(x)\}|\Omega\rangle = \frac{\langle 0|T\{\bar{\psi}(0)\psi(x)\}|0\rangle}{\langle 0|0\rangle} = \int \frac{d^4p}{(2\pi)^4}\frac{i(\not{p}+m)}{p^2-m^2+i\varepsilon}e^{ipx} \tag{14.123}$$

which is the Dirac propagator.

So Fermionic path integrals seem like they would be really hard and confusing, but in the end they are quite simple and you can mostly just forget about the fact that there this a lot of weird mathematics going into them.

14.12 Schwinger-Dyson equations

An important feature of the path integral is that it tells us how the classical equations of motion are modified in the quantum theory. These are known as Schwinger-Dyson equations.

Consider the correlation function of $\bar{\psi}(x_1)\psi(x_2)$:

$$Z_{12} = \int \mathcal{D}A_\mu\mathcal{D}\psi\mathcal{D}\bar{\psi}\exp\Big(i\int d^4x\Big[-\frac{1}{4}F_{\mu\nu}^2 + \bar{\psi}(i\not{\partial}+eA_\mu+m)\psi\Big]\Big)\bar{\psi}(x_1)\psi(x_2) \tag{14.124}$$

Under a field redefinition,

$$\psi \to e^{i\alpha}\psi \tag{14.125}$$

the measure is invariant. The Lagrangian is not invariant, since we have not transformed A_μ. Instead,

$$\bar{\psi}(x)\not{\partial}\psi(x) \to \bar{\psi}(x)\not{\partial}\psi(x) + i\partial_\mu\alpha(x)\bar{\psi}(x)\gamma^\mu\psi(x) \tag{14.126}$$

$$\bar{\psi}(x_1)\psi(x_2) \to e^{-i\alpha(x_1)}e^{i\alpha(x_2)}\bar{\psi}(x_1)\psi(x_2) \tag{14.127}$$

But since we are integrating over all ψ, the path integral is invariant. Thus, expanding to first order in α

$$0 = \int \mathcal{D}A_\mu\mathcal{D}\psi\mathcal{D}\bar{\psi}e^{iS} \tag{14.128}$$

$$\times\Big[\int d^4x\, i\partial_\mu\alpha(x)\bar{\psi}(x)\gamma^\mu\psi(x)\Big]\bar{\psi}(x_1)\psi(x_2) - i\alpha(x_1)\bar{\psi}(x_1)\psi(x_2) + i(x_2)\bar{\psi}(x_1)\psi(x_2) \tag{14.129}$$

We can simplify this by adding δ functions: $\alpha(x_1) = \int d^4x\delta(x-x_1)\alpha(x)$. Also, we define

$$j_\mu(x) = \bar{\psi}(x)\gamma^\mu\psi(x) \tag{14.130}$$

and integrate the $\partial_\mu \alpha j^\mu$ term by parts. Then,

$$0 = \int \mathcal{D}A_\mu \mathcal{D}\psi \mathcal{D}\bar{\psi} e^{iS} \int d^4x\, \alpha(x)[\partial_\mu j^\mu(x) - \delta(x-x_1) + \delta(x-x_2)]\bar{\psi}(x_1)\psi(x_2) \tag{14.131}$$

Dividing by $Z[0]$, this implies

$$\boxed{\partial_\mu \langle 0|T\{j^\mu(x)\bar{\psi}(x_1)\psi(x_2)\}\rangle = \delta(x-x_1)\langle 0|T\{\bar{\psi}(x_1)\psi(x_2)\}\rangle - \delta(x-x_2)\langle 0|T\{\bar{\psi}(x_1)\psi(x_2)\}\rangle}$$

This is known as a Schwinger-Dyson equation.

14.12.1 Ward identity from Schwinger-Dyson equations

Recall that classically, $\partial_\mu j^\mu = 0$. So this says that that equation still holds within correlation functions, up to contact interactions where $x \to x_j$, where x_j is another point appearing in the correlation function. Note that this equation is exact. We derived it by expanding to first order in α, but that just means we take infinitesimal transformations, which is not the same as calculating the correlation function to first order in e.

While this equation looks messy, it is actually just a generalization of the Ward identity. To see that, Fourier transform

$$\langle 0|T\{j^\mu(x)\bar{\psi}(x_1)\psi(x_2)\}\rangle = \int \frac{d^4p}{(2\pi)^4}\frac{d^4q_1}{(2\pi)^4}\frac{d^4q_2}{(2\pi)^4} e^{ipx}e^{iq_1x_1}e^{iq_2x_2} M_\mu(p,q_1,q_2) \tag{14.132}$$

$$\langle 0|T\{\bar{\psi}(x_1)\psi(x_2)\}\rangle = \int \frac{d^4q_2}{(2\pi)^4}\frac{d^4q_2}{(2\pi)^4} e^{iq_1x_1}e^{iq_2x_2} M_0(q_1,q_2) \tag{14.133}$$

So,

$$\int d^4x\, d^4x_1 d^4x_2 e^{ipx}e^{iq_1x_2}e^{iq_2x_2}\delta(x-x_1)\langle 0|T\{\bar{\psi}(x_1)\psi(x_2)\}\rangle = M_0(q_1+p,q_2) \tag{14.134}$$

Then we have

$$p_\mu M^\mu(p,q_1,q_2) = M_0(q_1+p,q_2) - M_0(q_1,q_2+p) \tag{14.135}$$

This is known as a generalized Ward-Takahashi identity.

This equation actaully implies the Ward identity. To see it, first, observe that $M_0(q,q')$ is just the two point function for an electron propogator with incoming momentum q and outgoing momentum q'. By momentum conservation, it will vanish unless $q = q'$. Thus $M_0(q,q') = M_0(q)$. The Ward identity, with $q = q_1$ incoming is then

$$\boxed{p_\mu M^\mu(p,q) = M_0(q+p) - M_0(q)} \tag{14.136}$$

Now if the put the electron on shell, as we must for S-matrix elements by LSZ, then $\not{q} = m$ and $M_0(q+p) = M_0(q)$, since neither depend on the momentum factors. So we find $p_\mu M^\mu = 0$.

The function M^μ is directly related to matrix elements of photons. Indeed, the interaction $j^\mu(x)\bar{\psi}(x_1)\psi(x_2)$ term is the matrix element of $\bar{\psi}(x_1)\psi(x_2)$ in the presence of an external current. The current interacts through the photon, so j^μ can be thought of as a place holder for the photon polarization. So,

$$\langle 0|T\{A_\mu(p)\bar{\psi}(q_1)\gamma^\mu\psi(q_2)\}\rangle = \varepsilon_\mu M^\mu(p,q_1,q_2) = \varepsilon_\mu M^\mu(p,q) \tag{14.137}$$

So $p_\mu M^\mu$ is the regular Ward identity for S-matrix elements we have been using all along. This is another way to prove it.

The more general Ward-Takahashi identity is especially useful for proving non-perturbative statements about matrix elements when the electron is not on-shell: $M_0(q) \neq M_0(q+p)$. One example is the non-renormalization of electric charge, which we will discuss shortly.

14.13 BRST invariance

There is a beautiful symmetry called BRST invariance, which is a residual exact symmetry of the Lagrangian even after gauge fixing. It is particularly useful for studying more complicated gauge theories, but it is a little easier to understand in the QED case.

Notice that the gauge fixing term $(\partial_\mu A_\mu)^2$ is invariant under gauge transformations $A_\mu \to A_\mu + \partial_\mu \alpha$ for $\Box\alpha = 0$. So this is a residual symmetry of the entire Lagrangian. As we have already seen, it is awkward to deal with constrained theories, and constraints on symmetries are no different. So consider adding to the Lagrangian two new free fields c and d

$$\mathcal{L} = \mathcal{L}[A, \phi_i] + \frac{1}{\xi}(\partial_\mu A_\mu)^2 - d\Box c \tag{14.138}$$

So the path integral becomes

$$Z_{\mathcal{O}} = \int \mathcal{D}c\mathcal{D}b\mathcal{D}A_\mu \mathcal{D}\phi_i e^{i\int d^4x \mathcal{L}[A,\phi_i] + \frac{1}{\xi}(\partial_\mu A_\mu)^2 - d\Box c} \tag{14.139}$$

The fields d and c are free, so their path integral can be solved exactly. The equations of motion are $\Box c = \Box d = 0$.

This path integral has an unconstrained global symmetry, parametrized by θ, for which

$$\Delta\phi_i = \theta c \phi_i \tag{14.140}$$

$$\Delta A_\mu = \theta \partial_\mu c(x) \tag{14.141}$$

$$\Delta d = \theta \frac{2}{\xi^2}(\partial_\mu A_\mu) \tag{14.142}$$

$$\Delta c = 0 \tag{14.143}$$

We can check

$$\Delta(\frac{1}{\xi}\partial_\mu A_\mu)^2 = \frac{2}{\xi^2} e(\partial_\mu A_\mu)\Box c = (\Delta d)\Box c \tag{14.144}$$

All we did was replace the constrained transformation with $\Box\alpha = 0$ with an unconstrained one. But it is the same, since $\Box c = 0$ is exact because c is a free field.

So we have isolated a residual symmetry in the gauge fixed Lagrangian. It is called BRST invariance. Actually, for technical reasons, BRST requires that θ be an anti-commuting Grassman number. You can check that if θ and b and c are fermionic, then the transformation is nilpotent $\Delta^2 = 0$. But none of this is relevant for QED.

To see why BRST is powerful, consider free solutions

$$A_\mu(x) = \sum_i a_i \int \frac{d^4k}{(2\pi)^4} \epsilon_\mu^i(k) e^{ikx} \tag{14.145}$$

$$c(x) = a_c \int \frac{d^4k}{(2\pi)^4} e^{ikx} \tag{14.146}$$

$$d(x) = a_d \int \frac{d^4k}{(2\pi)^4} e^{ikx} \tag{14.147}$$

So the free solutions are characterized by 6 numbers. For the A_μ field, there are two physical transverse modes $\epsilon_\mu^{1,2}$ which are normalizable and transverse $k_\mu \epsilon_\mu^{1,2} = 0$. The two unphysical modes are the backwards polarized mode ϵ_μ^b, which is not transverse $k_\mu \epsilon_\mu^b = 1 \neq 0$, and the forward polarized mode $\epsilon_\mu^f = k_\mu$, which is not normalizable. So we can think of free particles as being characterized by one big 6-component multiplet

$$(a_1, a_2, a_f, a_b, a_c, a_d) \tag{14.148}$$

Under BRST transformations, $\Delta A_\mu = \theta k_\mu c = \theta a_c \epsilon_f$ so $\Delta a_f = \theta a_c$. Since $\Delta d = \theta \frac{2}{\xi^2} k_\mu A_\mu = \frac{2}{\xi^2}\theta$ so $\Delta a_d = \theta \frac{2}{\xi^2} a_b$. Thus

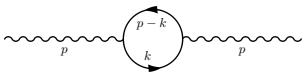
$$(a_1, a_2, a_f, a_b, a_c, a_d) \to (a_1, a_2, a_f + \theta a_c, a_b, a_c, a_d + \theta \frac{2}{\xi^2} a_b) \tag{14.149}$$

The point is that the only fields involved in this transformation of the unphysical polarizations and the "ghost" fields a and b. Since the global symmetry corresponds to a conserved current, by Noether's theorem, only the unphysical particles are charged under the symmetry. So if we start with a state with no unphysical particles, by charge conservation, we will end up in a state with no unphysical particles. This means that the S-matrix restricted to physical states is unitary. We knew this already, since QED has a Ward identity, but for more complicated gauge theories, the extra symmetry is an important tool.

Chapter 15
The Casimir Effect

15.1 Introduction

Now we come to the real heart of quantum field theory: loops. Loops generically are infinite. For example, the "vacuum polarization diagram in scalar QED is



$$= (-ie)^2 \int \frac{d^4k}{(2\pi)^4} \frac{2k^\mu - p^\mu}{(p-k)^2 - m^2} \frac{-2k^\mu + p^\mu}{k^2 - m^2} \tag{15.1}$$

In the region of the integral at large $k \gg p, m$ this is

$$\sim 4e^2 \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^2} \sim \int k\, dk = \infty \tag{15.2}$$

This is the same quadratically divergent integral Oppenheimer found when he tried to calculate the Lamb shift.

So what are we supposed to do about it? The basic answer is very simple: this loop is not by itself measurable. As long as we are always computing physical measurable quantities, the answer will come out finite. However, in practice the way it works is a bit more complicated. Instead of computing a physical thing all along, we regulate the infinite to get something finite in the intermediate steps. Then when all the pieces are put together, the answer for the observable turns out to be independent of the regulator. This is the program of renormalization. Why it's called *renormalization* will become clear soon enough.

15.2 Casimir effect

Let's start with the most obvious divergence, that in the Hamiltonian. Remember that

$$\mathcal{H} = \int \frac{d^3k}{(2\pi)^3} \omega_k \left(a_k^\dagger a_k + \frac{1}{2} \right) \tag{15.3}$$

where $\omega_k = \hbar k$. So the contribution to the vacuum energy of the photon zero modes is

$$E = \langle 0 | \mathcal{H} | 0 \rangle = \int \frac{d^3k}{(2\pi)^3} \frac{\omega_k}{2} \approx \frac{1}{4\pi^2} \int k^3 dk = \infty \tag{15.4}$$

Is this a totally meaningless quantity?

Well, first of all, we expect that zero point energy is meaningless. For example, with potential energy, only energy differences matter and the absolute energy is irrelevant. What about here, can we use this zero-point energy to calculate an energy difference? We can!

Consider asking what is the zero-point energy in a box of size a. If the energy changes with a then we can calculate $F = \frac{dE}{da}$ which will be a force on the walls of the box. In this case, we have a natural infrared (low energy) cutoff on k: $k > \frac{1}{a}$. Of course, this doesn't cut off the ultraviolet (high energy) divergence at large k, but if we're careful there will be a finite residual dependence on a. To be careful, we realize immediately that if we change a then the energy inside *and* outside the box will change, which messes things up. So let's put in a third wall on our box far away, at $L \gg a$. Then the energy completely outside the box is independent of r, so we can immediately drop it.

We'll work with a one dimensional box for simplicity, and use a scalar field instead of the photon. In one dimension of size r the frequencies are just $\omega_n = \frac{\pi}{r} n$ so this integral becomes a sum

$$E(r) = \sum_n \frac{\omega_n}{2}, \quad \omega_n = \frac{\pi}{r} n \tag{15.5}$$

still infinite, of course.

The total energy is the sum of the energy on the right side $r = (L - a)$ plus the energy on the left side, $d = a$:

$$E_{\text{tot}}(a) = E(a) + E(L-a) = \left(\frac{1}{a} + \frac{1}{L-a} \right) \frac{\pi}{2} \sum_{n=1}^{\infty} n \tag{15.6}$$

We don't expect the total energy to be finite, but we *do* hope to find a finite value for the force:

$$F(a) = -\frac{dE_{\text{tot}}}{da} = \left(\frac{1}{a^2} - \frac{1}{(L-a)^2} \right) \frac{\pi}{2} \sum_{n=1}^{\infty} n \tag{15.7}$$

However, this is $\infty - \infty$, which is undefined. For $L \to \infty$ this becomes

$$F(a) = \frac{\pi}{2} \frac{1}{a^2} (1 + 2 + 3 + \cdots) = \infty \tag{15.8}$$

So the plates are infinitely repulsive. That's obviously wrong.

What are we missing? Physics! These boundaries at $0, r, L$ are forcing the electromagnetic waves to be quantized. That is from an interaction of the photons with the boundary plates. These plates are made of atoms. Now think about these superhigh energy radiation modes, with supersmall wavelengths. They are are going to just plow through the walls of the box. Since we are only interested in the modes which are affected by the walls, these ultrahigh frequency modes should be irrelevant.

15.3 Hard Cutoff

Say we put in a high frequency cutoff, $\omega < \pi\Lambda$. We can think of Λ as $\frac{1}{\text{atomic size}}$, or some other natural way to limit high-frequency light from contributing. Then

$$n_{\max}(r) = \Lambda r \tag{15.9}$$

So,

$$E(r) = \frac{1}{r} \frac{\pi}{2} \sum_{n=1}^{n_{\max}} n = \frac{\pi}{2r} \frac{n_{\max}(n_{\max}+1)}{2} = \frac{\pi}{4r} (\Lambda r)(\Lambda r + 1) = \frac{\pi}{4} (\Lambda^2 r + \Lambda) \tag{15.10}$$

Then

$$E_{\text{tot}} = E(L-a) + E(a) \tag{15.11}$$

$$= \frac{\pi}{4} \big(\Lambda^2 L + 2\Lambda \big) \tag{15.12}$$

So we get some infinite constant, but one which is independent of a. Thus $F(a) = -\frac{dE_{\text{tot}}}{da} = 0$. Now the force is no longer infinite, but vanishes. Is that the right answer?

Yes, to leading order. But we were a little two quick with this calculation. The hard cutoff means a mode is either included or not. Thus even though we change r continuously, $n_{\max} = \Lambda r$ can only change by discrete amounts. We can write this mathematically with a floor function

$$n_{\max}(r) = \lfloor \Lambda r \rfloor \tag{15.13}$$

So the sum is

$$E(r) = \frac{\pi}{4r} \lfloor \Lambda r \rfloor (\lfloor \Lambda r \rfloor + 1) \tag{15.14}$$

And then

$$E_{\text{tot}}(a) = E(L-a) + E(a) = \frac{\pi}{4} \left\{ \frac{\lfloor \Lambda L - \Lambda a \rfloor (\lfloor \Lambda L - \Lambda a \rfloor + 1)}{L-a} + \frac{\lfloor \Lambda a \rfloor (\lfloor \Lambda a \rfloor + 1)}{a} \right\} \tag{15.15}$$

Above we found that if we remove the floors, $E_{\rm tot}$ is constant but now there's clearly some a dependence.

We can always write $\lfloor \Lambda r \rfloor = \Lambda r - x$ with $0 < x < 1$. Then

$$E(r) = \frac{\pi}{4}\big[\Lambda^2 r + \Lambda - 2\Lambda x\big] - \frac{\pi}{4}\frac{1}{r}x(1-x) \tag{15.16}$$

For simplicity, we can assume ΛL is an integer, so $\Lambda L = \lfloor \Lambda L \rfloor$. Then $E_{\rm tot}(a)$ simplifies to

$$E_{\rm tot}(a) = \frac{\pi}{4}L\Lambda^2 - \frac{\pi}{4}\left[\frac{1}{a} + \frac{1}{L-a}\right]x(1-x) \tag{15.17}$$

Note that for $x = 0$, we find $E_{\rm tot} = \frac{\pi}{4}L\Lambda^4$, which is infinite as $\Lambda \to \infty$, but not the same infinity as without the floor functions, $E_{\rm tot} = \frac{\pi}{4}(L\Lambda^4 + 2\Lambda)$. This can be traced to the fact that $y - y = 0$ but $\lfloor y \rfloor + \lfloor -y \rfloor = -1$. It is of no consequence physically, because the infinity is not physical.

Now we can take $L \to \infty$ giving

$$E_{\rm tot}(a) \approx \frac{\pi}{4}L\Lambda^2 - \frac{\pi}{4a}x(1-x) \tag{15.18}$$

Since $x = \Lambda a - \lfloor \Lambda a \rfloor$, then we should find a periodic function of period $\frac{1}{L}$ if we subtract the constant and multiply by a:

$$f(a) = \frac{2}{\pi}a\Big[E_{\rm tot}(a) - \frac{\pi}{4}\Lambda^2 L\Big] = -\frac{1}{2}x(1-x) \tag{15.19}$$

We can check by plotting $f(a)$ (with the exact $E_{\rm tot}(a)$), we can see the periodicity explicitly

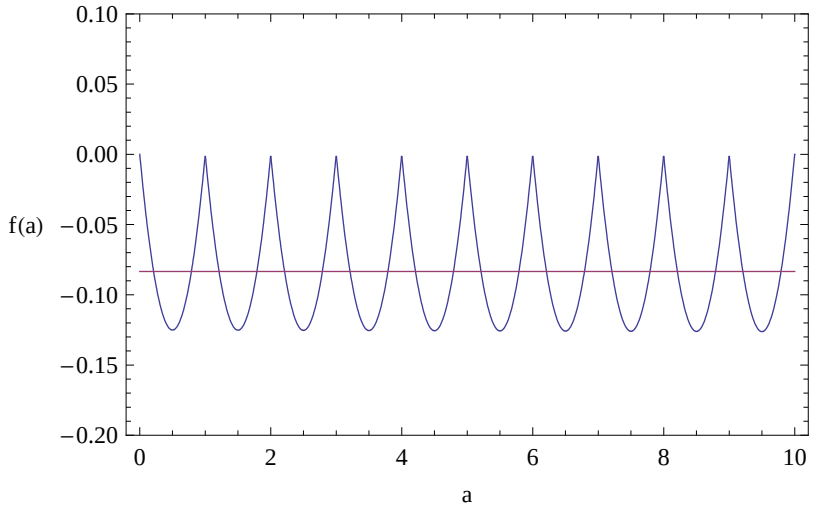


Figure 15.1. Casimir energy function $f(a) = \frac{2}{\pi}a\Big[E_{\rm tot}(a) - \frac{\pi}{4}\Lambda^2 L\Big]$ for a hard cutoff $\Lambda \equiv 1$, with $L = 1000$. The line is $-\frac{1}{12}$. Thanks to David Simmons-Duffin for this plot.

Thus the total energy is not smooth! We can smooth it out by averaging the energy over x between 0 and 1, using $\int x(1-x) = \frac{1}{6}$ so,

$$E_{\rm tot}(a) \approx \frac{\pi}{4}L\Lambda^2 - \frac{\pi}{24a} \tag{15.20}$$

Then we get a non-trivial result for the force:

$$F(a) = -\frac{dE_{\rm tot}}{da} = -\frac{\pi}{24a^2} \tag{15.21}$$

Putting back in the $\hbar$ and c, we found that the Casimir force in 1 dimension is

$$\boxed{F(a) = -\frac{\pi\hbar c}{24a^2}} \tag{15.22}$$

This is an attractive force. We can see that the force is purely quantum mechanical because it is proportional to $\hbar$.

In 3 dimensions, with the 2 photon polarizations, the answer is

$$F(a) = -\frac{\pi^2 \hbar c}{240 a^4} A \tag{15.23}$$

where A is the area of the walls of the box. The Casimir force has been observed (only conclusively in 1997!).

15.4 Regulator Dependence

I hope you find this calculation of the Casimir effect incredibly disconcerting. We found the force to be independent of Λ. But the calculation seemed to be very sensitive to what type of cutoff we actually used. What if we took a different model besides the hard cutoff for regulating the UV modes? It turns out you will get the same answer no matter what. That's a pretty amazing fact, so let's try a couple more regulators.

15.4.1 heat kernel regularization

Another reasonable physical assumption besides a hard cutoff would be that there is some penetration depth of the modes into the walls, with high frequency modes getting further. This means that the contribution of high frequency modes to the relevant energy sum is exponentially suppressed. Thus we can try

$$E(r) = \frac{1}{2} \sum_n \omega_n e^{-\omega_n/(\pi\Lambda)} \tag{15.24}$$

This is called heat kernel regularization.

Expanding with $\omega_n = \frac{\pi}{r} n$

$$E(r) = \frac{1}{r}\frac{\pi}{2} \sum_{n=1}^{\infty} n e^{-n/(\Lambda r)} = \frac{1}{r}\frac{\pi}{2} \sum_{n=1}^{n_{\max}} n e^{-\varepsilon n}, \quad \varepsilon = \frac{1}{\Lambda r} \ll 1 \tag{15.25}$$

Now we can calculate

$$\sum_{n=1}^{\infty} n e^{-\varepsilon n} = -\partial_\varepsilon \sum_{n=1}^{\infty} e^{-\varepsilon n} = -\partial_\varepsilon \frac{1}{1-e^{-\varepsilon}} = \frac{e^{-\varepsilon}}{(1-e^{-\varepsilon})^2} = \frac{1}{\varepsilon^2} - \frac{1}{12} + \frac{\varepsilon^2}{240} + \cdots \tag{15.26}$$

Already, we see the factor $-\frac{1}{12}$ appearing.

So

$$E(r) = \frac{1}{r}\frac{\pi}{2}\left[\Lambda^2 r^2 - \frac{1}{12} + \frac{1}{r^2\Lambda^2} \cdots\right] = \frac{\pi}{2} r\Lambda^2 - \frac{\pi}{24 r} + \cdots \tag{15.27}$$

Then,

$$F(a) = -\frac{d}{da}[E(L-a) + E(a)] = -\frac{d}{da}\left[\frac{\pi}{2} L\Lambda^2 - \frac{\pi}{24}\left(\frac{1}{L-a} + \frac{1}{a}\right) + \cdots\right] \tag{15.28}$$

$$= \frac{\pi}{24}\left(\frac{1}{(L-a)^2} - \frac{1}{a^2}\right) + \cdots \tag{15.29}$$

Now take $L \to \infty$ and we get

$$F(a) = -\frac{\pi \hbar c}{24 a^2} \tag{15.30}$$

which is the same thing.

15.4.2 other regulators

What else can we try? We can use a Gaussian regulator

$$E(r) = \frac{1}{2}\sum_n \omega_n e^{-(\frac{\omega_n}{\pi\Lambda})^2} \tag{15.31}$$

Or a zeta-function regulator

$$E(r) = \frac{1}{2}\sum_n \omega_n \left(\frac{\omega_n}{\mu}\right)^{-s} \tag{15.32}$$

where we take $s \to 0$ instead of $\omega_{\max} \to \infty$ and have added an arbitrary scale μ to keep the dimensions correct. μ does not have to be large – it should drop out for any μ.

Let's work out the zeta-function case. Substituting in for ω_n we get

$$E(r) = \frac{1}{2}\left(\frac{\pi}{r}\right)^{1-s}\mu^s \sum_n n^{1-s} \tag{15.33}$$

This sum is the definition of the Riemann zeta-function:

$$\sum n^{1-s} = \zeta(s-1) = -\frac{1}{12} - 0.165\,s + \cdots \tag{15.34}$$

So we get

$$E(r) = \frac{1}{r}\frac{\pi}{2}[\zeta(s-1)] = \frac{1}{r}\frac{\pi}{2}\left[-\frac{1}{12} + O(s)\cdots\right] \tag{15.35}$$

the same constant from this approach as before. Note also that, we don't get the divergent piece at all.

The energy comes out as

$$E(r) = \frac{1}{r}\frac{\pi}{2}\left[-\frac{1}{12} + \cdots\right] \tag{15.36}$$

This is the same as what the heat-kernel regularization gave, but with a different divergent term. In this case, the divergent term is just 0. That's ok, because the divergent term is not physical, but it is reassuring that the physical term, the $-\frac{1}{12}$ comes out the same.

All four of the regulators agree:

$$E(r) = \frac{1}{2}\sum_n \omega_n \theta(\pi\Lambda - \omega_n) \quad (\text{hard cutoff}) \tag{15.37}$$

$$E(r) = \frac{1}{2}\sum_n \omega_n e^{-\frac{\omega_n}{\pi\Lambda}} \quad (\text{heat kernel}) \tag{15.38}$$

$$E(r) = \frac{1}{2}\sum_n \omega_n e^{-(\frac{\omega_n}{\pi\Lambda})^2} \quad (\text{Gaussian}) \tag{15.39}$$

$$E(r) = \frac{1}{2}\sum_n \omega_n \left(\frac{\omega_n}{\mu}\right)^{-s} \quad (\text{zeta}-\text{function}) \tag{15.40}$$

That these regulators all agree is reasuring, but seems somewhat mysterious.

15.4.3 regulator independent derivation

Actaully, Casimir himself showed in his original paper a way to calculate this sum in a regulator independent way.

$$E(a) = \frac{\pi}{2}\sum_n \frac{n}{a} f(\frac{n}{a\Lambda}) \tag{15.41}$$

Then

$$E(L-a) = \frac{\pi}{2}(L-a)\Lambda^2 \sum \frac{n}{(L-a)\Lambda^2} f\left[\frac{n}{(L-a)\Lambda}\right] \tag{15.42}$$

we can take the continuum limit with $x = \frac{n}{(L-a)\Lambda}$. Then

$$E(L-a) = \frac{\pi}{2} L\Lambda^2 \int x\,dx\,f(x) - \frac{\pi}{2} a\Lambda^2 \int x\,dx\,f(x) \tag{15.43}$$

The first integral is just the extrinsic vecuum energy density

$$\rho = \frac{\pi}{2}\Lambda^2 \int x\,dx\,f(x) \tag{15.44}$$

The second integral simplifies with $x \to \frac{n}{a\Lambda}$. Then

$$E_{\text{tot}} = E(a) + E(L-a) = \rho L + \frac{\pi}{2a}\left[\sum_n n\,f(\frac{n}{a\Lambda}) - \int n\,dn\,f(\frac{n}{a\Lambda}) \right] \tag{15.45}$$

The difference between an infinite sum and an infinite integral is given by a Euler-McClaurin series

$$\int_0^N F(n)dn - \sum_{n=1}^{N} F(n) = \frac{F(0)+F(N)}{2} + \frac{F'(N)-F'(0)}{12} + \cdots + B_j \frac{F^{(j+1)}(N) - F^{(j+1)}(0)}{j!}$$

where B_j are the Bernoulli numbers. $B_2 = \frac{1}{6}$.

In our case, $F(n) = n\,f(\frac{n}{a\Lambda})$. So, assuming that

$$\lim_{x\to\infty} x\,f(x) = 0 \tag{15.46}$$

then

$$E_{\text{tot}} = \rho L - \frac{\pi f(0)}{24a} + \frac{B_3}{3!}\frac{\pi}{2a^2\Lambda} f(0) + \cdots \tag{15.47}$$

For example, if $f(x) = e^{-x}$, then

$$E_{\text{tot}} = \frac{\pi}{2}\Lambda^2 L - \frac{\pi}{24a} + \mathcal{O}\left(\frac{1}{a^2\Lambda}\right) \tag{15.48}$$

The important point is that this only depends on $f(0)$, that is, the propertires of the energy spectrum near $\omega \sim 0$, i.e. in the infrared. So,

The Casimir force is an infrared effect

Of course, we knew this should be true, becuase only modes of size $\frac{1}{a}$ can reach both walls of the box to transmit the force. The regulator must only satisfy

$$f(0) = 1 \quad \text{and} \quad \lim_{x\to\infty} x\,f(x) = 0 \tag{15.49}$$

meaning that $f(x)$ does not affect the energies as $E \to 0$ and that $f(x)$ dies faster than $\frac{1}{x}$ at high energy. You can see that all 4 of the regulators satisfy these requirements. The hard cutoff is a mess because it is not differentiable.

15.4.4 string theory aside

A funny way to write what we have done is to replace

$$\frac{\pi}{2r}\sum n \quad \to \quad -\frac{\pi}{24r} \tag{15.50}$$

So that we take

$$1 + 2 + 3 + \cdots = -\frac{1}{12} \tag{15.51}$$

This relation has an important use in string theory. In string theory, the mass of particles is determined by the string tension α':

$$m^2 = \frac{1}{\alpha'} n + E_0 \tag{15.52}$$

where c is the constant Casimir energy and n is the excitation number. So there is a whole tower of particles with different masses. In d dimensions,

$$E_0 = \frac{1}{\alpha'}\left(\frac{2-d}{2}\right)\left(-\frac{1}{12}\right) \tag{15.53}$$

where the $-\frac{1}{12}$ comes from the same series we have just summed. Now, you can show in string theory that the $n = 1$ excitations have two polarizations. So they must be massless. Then solving for $m = 0$ you find $d = 26$. That's why string theory takes place in 26 dimensions. If you do the same calculation for the superstring, you find $d = 10$.

15.5 Counterterms

In the above analysis, we not only took $\omega_{\max} \to \infty$ but also $L \to \infty$. Why did we need this third boundary at $r = L$ at all? Let's suppose we didn't have it, but just calculated the energy inside the box of size r. Then we would get

$$E(r) = \frac{\pi}{2} r\Lambda^2 - \frac{\pi}{24r} + \cdots \tag{15.54}$$

This first term is linear in the volume. It can be interpreted as saying there is some finite energy density $\rho = \frac{E}{r} = \frac{\pi}{2}\Lambda^2$. Thus suppose instead of just the free field we were using to calculate the ground state energy, we took

$$\mathcal{L} = \cdots + \rho_c \tag{15.55}$$

where ρ_c is constant. This new term gives an infinite contribution from $\int dx \rho_c$ in the action. Now if we choose $\rho_c = -\frac{\pi}{2}\Lambda^2$ it exactly cancels the $\frac{\pi}{2} r\Lambda^2$ term. ρ_c is just some other unmeasurable zero-point energy. We see it is convenient to choose it to be something to cancel the finite term.

In the zeta-function regulator, where no divergent terms come out of the calculation, we could take $\rho_c = 0$. So what we choose for ρ_c will depend on what regulator we want. ρ_c is called a counterterm. Counterterms give purely infinite contributions to *intermediate steps* in calculations, but when we compute *physical quantities* they drop out.

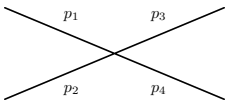
15.6 Scalar field theory example

Before we do any physical calculation, let's get an overview of the way things are going to work out. It's easiest to do this in a scalar field theory. So, start with ϕ^4 theory

$$\mathcal{L} = -\frac{1}{2}\phi\Box\phi + \frac{\lambda}{4!}\phi^4 \tag{15.56}$$

The reason we use a $\lambda\phi^4$ interaction, instead of $g\phi^3$ is because λ is dimensionless, while g has dimensions of mass. Thus dealing with renormalization is slightly simpler because there are fewer mass scales to deal with, but we could have done ϕ^3 instead. We have also set the mass of ϕ to zero, for simplicity.

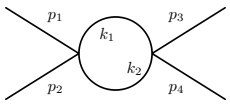
Then at tree-level, $\phi\phi \to \phi\phi$ scattering is given by the simple cross diagram

$i\mathcal{M}_1 =$ 

$= i\lambda$ (15.57)

The leading correction comes from loops, such as this s-channel one

$$i\mathcal{M}_2 = \quad (15.58)$$



There is also a t and u channel diagram, but let's forget about them (for example, if we had 3 fields with a $\frac{\lambda}{2!}(\phi_1^2\phi_2^2+\phi_2^2\phi_3^2)$ interaction, there would only be an s-channel contribution to $\phi_1\phi_1\to\phi_3\phi_3$).

Let $p=p_1+p_2=p_3+p_4$, then $k_1+k_2=p$ so we can set $k_1=k$ and $k_2=p-k$ and integrate over k. The diagram becomes

$$i\mathcal{M}_2=(i\lambda)^2\int\frac{d^4k}{(2\pi)^4}\frac{i}{k^2}\frac{i}{(p-k)^2} \quad (15.59)$$

Now change variables to $k=\frac{1}{2}(q+p)$ so $k-p=\frac{1}{2}(q-p)$. Then this simplifies to

$$i\mathcal{M}_2=\lambda^2\int\frac{d^4q}{(2\pi)^4}\frac{1}{(q^2-s)^2} \quad (15.60)$$

Here we see the divergence: at large q this becomes $\int\frac{d^4q}{q^4}$ which blows up logarithmically. To characterize the divergence, let's take a derivative

$$\frac{d}{ds}i\mathcal{M}_2=\lambda^2\int\frac{d^4q}{(2\pi)^4}\frac{2}{(q^2-s)^3} \quad (15.61)$$

It's not too hard to work out this integral, since it's convergent. By dimensional analysis, the result has mass dimension -2 and since s is the only scale in the integral, the answer must be $c\frac{1}{s}$ for some number c. The exact integral is not too hard to do, but we will just quote the answer

$$\frac{d}{ds}i\mathcal{M}_2=\lambda^2\Big[\frac{-i}{16\pi^2 s}\Big] \quad (15.62)$$

Every QFT textbook will have formulas and derivations for integrals of this type, which are totally standard.

This means that

$$\mathcal{M}_2=-\frac{\lambda^2}{16\pi^2}\log(s)+c \quad (15.63)$$

where c is an integration constant. Since the integral is divergent, c is going to be infinity. Also, since s has dimension of mass squared, it's nice to write the constant as $c=-\log\Lambda^2$ where Λ has dimension of mass. Then we have

$$\mathcal{M}_2=-\frac{\lambda^2}{16\pi^2}\log(\frac{s}{\Lambda^2}) \quad (15.64)$$

So the total matrix element is

$$\mathcal{M}(s)=\lambda-\frac{\lambda^2}{16\pi^2}\Big[\log(\frac{s}{\Lambda^2})\Big] \quad (15.65)$$

Now we must renormalize.

15.6.1 renormalizion of λ

First of all, notice that while $\mathcal{M}(s)$ is infinite, the difference between $\mathcal{M}(s_1)$ and $\mathcal{M}(s_0)$ at two different scales is finite:

$$\mathcal{M}(s_1)-\mathcal{M}(s_2)=\frac{\lambda^2}{\pi^2}\Big[\log(\frac{s_2}{s_1})\Big] \quad (15.66)$$

Should we also expect that $\mathcal{M}(s)$ itself be finite? After all, $\mathcal{M}^2$ is supposed to be a physical cross section.

To answer this, let's think more about λ. λ should be measuring the strength of the ϕ^4 interaction. So to measure λ we would simple measure the cross section for $\phi\phi \to \phi\phi$ scattering, or equivalently, $\mathcal{M}$. But this matrix element is not just proportional to λ but also has the λ^2 correction above. Thus, it is impossible to simply extract λ from this scattering process. Instead, let us just define a "renormalized" λ as the value of the matrix element at a particular $s = s_0$.

So

$$\lambda_R \equiv \mathcal{M}(s_0) = \lambda - \frac{\lambda^2}{16\pi^2}\log(\frac{s_0}{\Lambda^2}) + \cdots \tag{15.67}$$

This equation relates the parameter λ of the Lagrangian to the value of the observed scatting amplitude λ_R at a particular center-of-mass energy s_0. We can also conclude that λ_R is finite, while λ must be infinite, to cancel the infinity from $\log\Lambda^2$.

Next, solve for λ in terms of λ_R:

$$\lambda_R = \lambda + a\lambda^2 + \cdots \quad \Leftrightarrow \quad \lambda = \lambda_R - a\lambda^2 - \cdots = \lambda_R - a\lambda_R^2 + \cdots \tag{15.68}$$

Note that to invert this equation, we work order by order in λ, which is only a formal perturbation expansion, as $\lambda = \infty$, but it equivalent to a perturbation expansion in $\hbar$, since $\hbar$ counts the loop factors. So,

$$\lambda = \lambda_R + \frac{\lambda_R^2}{16\pi^2}\Big[\log(\frac{s_0}{\Lambda^2})\Big] + \cdots \tag{15.69}$$

Now, suppose we measure the cross section at a different center-of-mass energy s. Then

$$\mathcal{M}(s) = \lambda - \frac{\lambda^2}{16\pi^2}\log(\frac{s}{\Lambda^2}) \tag{15.70}$$

$$= \left[\lambda_R + \frac{\lambda_R^2}{16\pi^2}\log(\frac{s_0}{\Lambda^2})\right] - \frac{\lambda_R^2}{16\pi^2}\log(\frac{s}{\Lambda^2}) + \cdots \tag{15.71}$$

$$= \lambda_R - \frac{\lambda_R^2}{16\pi^2}\log(\frac{s}{s_0}) + \cdots \tag{15.72}$$

This equation now gives us a general expression for $\mathcal{M}$ which is finite order-by-order in perturbation theory.

15.6.2 counterterm interpretation

Another way of getting the same result is to add a counterterm to the Larganian. That means adding another interaction just like the first, but infinite. So we take as our Lagrangian

$$\mathcal{L} = -\frac{1}{2}\phi\Box\phi + \frac{\lambda_R}{4!}\phi^4 + \frac{\lambda_{ct}}{4!}\phi^4 \tag{15.73}$$

where the counterterm λ_{ct} is infinite, but formally of order λ_R^2. Then, working to order λ_R^2, the amplitude is

$$\mathcal{M}(s) = \lambda_R + \lambda_{ct} - \frac{\lambda_R^2}{16\pi^2}\log(\frac{s}{\Lambda^2}) \tag{15.74}$$

Now we can choose $\lambda_{\rm ct}$ to be whatever we want. If we take it to be

$$\lambda_{\rm ct} = \frac{\lambda_R^2}{16\pi^2}\log(\frac{s_0}{\Lambda^2}) \tag{15.75}$$

Then,

$$\mathcal{M}(s) = \lambda_R - \frac{\lambda_R^2}{16\pi^2}\log(\frac{s}{s_0}) \tag{15.76}$$

which is finite. In particular, this choice of $\lambda_{\rm ct}$ makes $\mathcal{M}(s_0) = \lambda_R$, which was our definition of λ_R above.

Doing things this way, with counterterms but as a perturbative expansion in the physical coupling λ_R is known as *renormalized perturbation theory*. The previous way, where we compute physical quantities like $\mathcal{M}(s_1) - \mathcal{M}(s_2)$ directly is sometimes called *physical pertrubation theory*. The two are equivalent. But for complicated calculations, renormalized perturbation theory is much easier.

Chapter 16
Regularization Schemes and Tricks

16.1 Feynman Parameters

One of the best tricks is the Feynman Parameter trick. Here's a bunch of integrals that we will need (see Peskin and Schroeder for the derivations)

$$\frac{1}{AB} = \int_0^1 dx \frac{1}{[A+(B-A)x]^2} = \int_0^1 dx dy \delta(x+y-1) \frac{1}{[xA+yB]^2} \tag{16.1}$$

$$\frac{1}{AB^n} = \int_0^1 dx dy \delta(x+y-1) \frac{n y^{n-1}}{[xA+yB]^{n+1}} \tag{16.2}$$

$$\frac{1}{ABC} = \int_0^1 dx dy dz \delta(x+y+z-1) \frac{2}{[xA+yB+zC]^3} \tag{16.3}$$

The reason these are useful is that they let us complete the square in the denominator.

16.2 Wick Rotations

A nice way to do Minkowski integrals, like the ones above, is to observe that once the $i\varepsilon$ factors are put in, they look like

$$\int \frac{d^4k}{(2\pi)^4} \frac{1}{(k^2-\Delta+i\varepsilon)^3} \tag{16.4}$$

Which has poles at $k_0 = \sqrt{\vec{k}^2+\Delta} - i\epsilon$ and $k_0 = -\sqrt{\vec{k}^2+\Delta} + i\epsilon$. After all, the location of these poles comes from time ordering, which was the whole reason we use $i\varepsilon$ to begin with. So, the poles are in the top left and bottom right quadrants of the k_0 complex plane for. Thus, we can rotate the contour of integration counterclockwise from the real line to the imaginary line. That is, the substitution $k_0 \to i\,k_0$ is allowed. Then $k^2 \to -k_0^2 - \vec{k}^2 = -k_E^2$, where $k_E^2 = k_0^2 + \vec{k}^2$ is the Euclidean momentum.

The volume of a 3D Euclidean sphere is $2\pi^2$, so

$$\int \frac{d^4k_E}{(2\pi)^4} f(k_E^2) = \frac{1}{8\pi^2} \int_0^\infty k_E^3 dk_E f(k_E^2) \tag{16.5}$$

This makes the integral much easier

$$\int \frac{d^4k}{(2\pi)^4} \frac{1}{(k^2-\Delta)^3} = i \int \frac{d^4k_E}{(2\pi)^4} \frac{1}{(-k_E^2-\Delta)^3} \tag{16.6}$$

$$= (-1)^3 \frac{i}{8\pi^2} \int_0^\infty dk_E \frac{k_E^3}{(k_E^2+\Delta)^3} \tag{16.7}$$

$$= -\frac{i}{8\pi^2} \frac{1}{4\Delta} \tag{16.8}$$

So,

$$\int \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2-\Delta)^3} = \frac{-i}{32\pi^2\Delta} \tag{16.9}$$

Other formulas following from Wick rotations are

$$\int \frac{d^4k}{(2\pi)^4}\frac{k^2}{(k^2-\Delta)^4} = \frac{-i}{48\pi^2}\frac{1}{\Delta} \tag{16.10}$$

$$\int \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2-\Delta)^r} = i\frac{(-1)^r}{(4\pi)^2}\frac{1}{(r-1)(r-2)}\frac{1}{\Delta^{r-2}} \tag{16.11}$$

$$\int \frac{d^4k}{(2\pi)^4}\frac{k^2}{(k^2-\Delta)^r} = i\frac{(-1)^{r-1}}{(4\pi)^2}\frac{2}{(r-1)(r-2)(r-3)}\frac{1}{\Delta^{(r-3)}} \tag{16.12}$$

and so on. These formulas are all derived in section 6.3 of Peskin and Schroeder.

The Wick rotation can only be justified if the integral is convergent. Well, if it's not convergent, it is useless to Wick rotate anyway, since you will still get infinity. But the technical problem is the divergence can indicate the presence of new poles that invalidate the contour rotation. Keep in mind that the Wick rotation is just a trick to do these integrals. There is nothing physical about it.

16.3 Quadratic integrals – derivative method

We have already seen that the quadratic integral is logarithmically divergent

$$\mathcal{I}(\Delta) = \int \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2-\Delta+i\varepsilon)^2} = \infty \tag{16.13}$$

But if we take the derivative, the integral can be done

$$\frac{d}{d\Delta}\mathcal{I}(\Delta) = \int \frac{d^4k}{(2\pi)^4}\frac{2}{(k^2-\Delta+i\varepsilon)^3} = -\frac{i}{16\pi^2\Delta} \tag{16.14}$$

So

$$\mathcal{I}(\Delta) = -\frac{i}{16\pi^2}\log\frac{\Delta}{\Lambda^2} \tag{16.15}$$

where Λ is some parameter. This trick is not a real regulator. Since Λ is just a constant of integration, we wouldn't know how to use the same regulator with a different integral, which will be necessary to cancel divergences in more complicated problems. But with any proper regulator, the integral is finite and this trick is totally valid, thus the result will always be of this form.

16.4 Pauli-Villars regularization

One useful regulator was invented by Pauli and Villars. A normal propagator looks like

$$\Pi_m = \frac{i}{p^2-m^2+i\varepsilon} \tag{16.16}$$

They add a particle with very large mass, but whose propagator lacks the factor of i

$$\Pi_M = \frac{1}{p^2-M^2+i\varepsilon} \tag{16.17}$$

Keep in mind that the sign of the residue of the propagator is dictated by unitarity – a real particle with this sign would have negative norm, and violate conservation of probability. So this particle is purely fictitious, we call it a *ghost*.

An amplitude will sum over the real particle, with mass m and the ghost particle, with fixed large mass $M \gg m$. Without the i, we get a relative minus sign in the loop. So, we get something like

$$\mathcal{I} = \int \frac{d^4k}{(2\pi)^4}\left[\frac{1}{(k^2-m^2+i\varepsilon)^2} - \frac{1}{(k^2-M^2+i\varepsilon)^2}\right] \tag{16.18}$$

You can now see that for $k \gg M \gg m$ these terms are both $\frac{1}{k^4}$ and so they cancel.

To do this integral, since it is convergent, we can Wick rotate. Then we get

$$\mathcal{I} = \frac{i}{8\pi^2}(-1)^2 \int_0^\infty dk_E \left[\frac{k_E^3}{(k_E^2-m^2)^2} - \frac{k_E^3}{(k_E^2-M^2)^2}\right] \tag{16.19}$$

$$= \frac{i}{16\pi^2}\left[\frac{M^2}{k^2-M^2} - \frac{m^2}{k^2-m^2} + \log\frac{k^2-m^2}{k^2-M^2}\right]_0^\infty \tag{16.20}$$

$$= -\frac{i}{16\pi^2}\log\frac{m^2}{M^2} \tag{16.21}$$

This is the same answer we got from the derivative trick above.

A useful trick in Pauli villars is to note that

$$\frac{1}{k^2-m^2} - \frac{1}{k^2-M^2} = \int_{m^2}^{M^2} \frac{-1}{(k^2-z)^2} dz \tag{16.22}$$

which allows us to simply square the propagator and add another Feynmanlike paramter z. This can be a useful trick for computing PV regulated integrals in pracice. Also note that

$$\int dm^2 \frac{d}{dm^2}\frac{1}{k^2-m^2} = \int dm^2 \frac{-1}{(k^2-m^2)^2} \tag{16.23}$$

So in fact PV is fancy, but systematic, way of implementing the derivatie trick.

Pauli-Villars is occasionally useful. But for complicated multiloop diagrams, you need to introduce many fictitious particles (one for each real particle will not do it), and it quickly becomes impractical. You also can't give the photon a ghost partner with a large mass because this mass breaks gauge invariance (actually a PV photon does work in QED but it will not it more general gauge theories).

16.5 Dimensional Regularization

The most important regulator you will encounter is dimensional regularization, invented by 't Hooft. The basic idea is that an integral like

$$\int \frac{d^dk}{(2\pi)^d}\frac{1}{(k^2-\Delta+i\varepsilon)^2} \tag{16.24}$$

is divergent only if $d > 4$, when there are more powers of k upstairs than downstairs. But if $d < 4$, then it will converge. If it's convergent we can Wick rotate, and the answer comes from just analytically continuing all our formulas above to d dimensions.

The volume of a sphere in various dimensions is

$$\int d\Omega_2 = 2\pi\,(\text{circle}), \qquad \int d\Omega_3 = 4\pi\,(\text{sphere}), \qquad \int d\Omega_4 = 2\pi^3\,(3-\text{sphere}) \tag{16.25}$$

The general formula is

$$\int d\Omega_d = 2\frac{\pi^{d/2}}{\Gamma(d/2)} \tag{16.26}$$

where $\Gamma(x)$ is the Gamma function. $\Gamma(x)$ comes up a lot. It is the analytic continuation of the factorial: $\Gamma(x+1) = x!$, and $\Gamma(1) = \Gamma(2) = 1$. $\Gamma(z)$ has simple poles at 0 and all the negative integers.

The Wick rotation gives

$$\int \frac{d^dk}{(2\pi)^d}\frac{1}{(k^2-\Delta+i\varepsilon)^2} = 2i\frac{\pi^{d/2}}{\Gamma(d/2)}(2\pi)^{-d}\int dk_E \frac{k_E^{d-1}}{(k_E^2+\Delta)^2} \tag{16.27}$$

$$= \frac{i}{(4\pi)^{d/2}}\Gamma(\frac{4-d}{2})\left(\frac{1}{\Delta}\right)^{2-\frac{d}{2}} \tag{16.28}$$

So here we see the Γ function emerging again.

We will often need to expand $\Gamma(x)$ around the pole at $x=0$:'

$$\Gamma(\epsilon) = \frac{1}{\epsilon} - \gamma_E + \mathcal{O}(\epsilon) + \cdots \tag{16.29}$$

Here, γ_E is the Euler-Mascheroni constant $\gamma_E \approx 0.577$. The pole comes from going back to $d=4$, so we will often write $d = 4-\epsilon$ [note: sometimes people use $d = 4-2\epsilon$, so if you're ever off by a factor of 2 in comparing to the literature, check this.]

Finally, we like to have our matrix elements in d dimensions have the same mass dimension as our matrix elements in $d=4$. Thus we analytically continue our integral as

$$\int \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2-\Delta+i\varepsilon)^2} = \lim_{d\to 4}\mu^{4-d}\int \frac{d^dk}{(2\pi)^d}\frac{1}{(k^2-\Delta+i\varepsilon)^2} \tag{16.30}$$

μ is some regulator dependant constant with dimension of mass . In contrast to the other regulators (like M in Pauli-Villars, or Λ in general), μ is *not* a large mass scale. That is, dim reg agrees with our original integral not when $\mu\to\infty$ but when $d\to 4$. It is very important to keep this in mind as we go along.

Then we have,

$$\mathcal{I} = \mu^{4-d}\int \frac{d^dk}{(2\pi)^d}\frac{1}{(k^2-\Delta+i\varepsilon)^2} = \mu^{4-d}\frac{i}{(4\pi)^{d/2}}\Gamma(\frac{4-d}{2})\left(\frac{1}{\Delta}\right)^{2-\frac{d}{2}} \tag{16.31}$$

Now letting $d=4-\epsilon$ we expand this around $\epsilon=0$ and get

$$\mathcal{I} = \frac{1}{16\pi^2}\left[\frac{2}{\epsilon} + (-\gamma_E + \log 4\pi + \log\mu^2 - \log\Delta) + \mathcal{O}(\varepsilon)\right] \tag{16.32}$$

$$= \frac{1}{16\pi^2}\left[\frac{2}{\epsilon} + \log\frac{4\pi e^{-\gamma_E}\mu^2}{m^2} + \mathcal{O}(\varepsilon)\right] \tag{16.33}$$

The γ_E comes from the integral $\int \frac{d^dk}{k^4}$, the 4π comes from the phase space $\frac{1}{(2\pi)^d}$ and the μ comes from the μ^d. This combination $4\pi e^{-\gamma_E}\mu^2$comes up a lot, so we give it a name

$$\tilde{\mu}^2 = 4\pi e^{-\gamma_E}\mu^2 \tag{16.34}$$

Thus

$$\mathcal{I}(m) = \frac{i}{16\pi^2}\left[\frac{2}{\epsilon} + \log\frac{\tilde{\mu}^2}{m^2} + \mathcal{O}(\varepsilon)\right] \tag{16.35}$$

Sometimes people omit the twiddle and just write μ for $\tilde{\mu}$. Note that there is still a divergence in this expression as $\epsilon\to 0$. There are two unphysical ingredients in dim reg, factors of ϵ and factors of $\tilde{\mu}$. Both must drop out in physical quantities, for example $\mathcal{I}(m_1) - \mathcal{I}(m_2)$ is independent of $\tilde{\mu}$ and ϵ.

An important feature of dimensional regularization is that it helps characterize the degree to which integrals diverge. For example, an integral like

$$\int d^4k \frac{1}{k^4} \tag{16.36}$$

is logarithmically divergent. In d dimensions, it has a simple pole at $d=4$, and no other poles for $d<4$. You can see this from the $\Gamma(\frac{4-d}{2})$ term in Eq. (16.28).

A quadratically divergent integral, like

$$\int d^4k \frac{1}{k^2} \tag{16.37}$$

has a strong singularity at $k=\infty$. With a cutoff, it goes like $\int \frac{d^4k}{k^2} \sim \Lambda^2$ which is quadratically divergent. However, in d dimensions, it is not divergent at all, it just has a funny dependence on d. If $d=2$, the integral would be logarithmically divergent, so there is a corresponding pole in the dim reg amplitude at $d=2$. This shows up in a $\Gamma(\frac{2-d}{2})$ term, which has poles at $d=2$ and $d=4$, but no poles for $d<2$. So dim reg translates the degree of divergence into the singularity structure of the amplitudes in d dimensions. This can be really helpful for seeing the singularity structure of more complicated integrals.

Some generally useful dim-reg formula are given in Appendix A of P&S.

16.6 k^μ integrals

If we have an integral with factors of momenta, such as $k^\mu k^\nu$ in the numerator:

$$\int \frac{d^4k}{(2\pi)^4} \frac{k^\mu k^\nu}{(k^2-\Delta)^4} \tag{16.38}$$

we have another trick for simplifying them. Since the denominator only depends on the scalar p^2, it is impossible for the integral to produce a piece proportional to p^μ or p^ν anymore. Thus the $k^\mu k^\nu$ term must give something proportional to $g^{\mu\nu}$. That is, we must get the same thing as having replaced $k^\mu k^\nu$ by $ck^2 g^{\mu\nu}$ for some number c. Contracting with $g^{\mu\nu}$, we see that $c=\frac{1}{4}$ or more generally $c=\frac{1}{d}$.

$$\int d^dk \frac{k^\mu k^\nu}{(k^2-\Delta)^4} = \int d^dk \frac{\frac{1}{d}k^\mu k^\nu}{(k^2-\Delta)^4} \tag{16.39}$$

If there is just one factor of k^μ in the numerator

$$\int \frac{d^4k}{(2\pi)^4} \frac{k^\mu}{(k^2-\Delta)^4} \tag{16.40}$$

Then the integrand is antisymmetric under $k \to -k$. Since we are integrating over all k, the integral therefore vanishes. So we will only need to keep terms with even powers of k in the numerator.

16.7 Summary

The main advantage of dimensional regularization is that it preserves gauge invariance, since we can have gauge fields in any dimension. Note that Pauli-Villars breaks gauge invariance because a massive gauge boson, ghost or not, cannot be gauge invariant. On the other hand, Pauli-Villars works nicely for fermions, while fermions in d dimensions can be messy because we have to analytically continue the γ matrix algebra. When you have fermions *and* gauge bosons in the loop, people generally use dimensional regularization. In fact, dim reg is practically the only regulator used for serious computations these days.

You will also encounter other regulators from time to time, such as

- Hard cutoff – $k_E < \Lambda$. This breaks translation invariance, and usually every symmetry in the theory. But it is perhaps the most intuitive regularization procedure.
- Point splitting (Dirac 1934). Divergences at $k \to \infty$ correspond to two fields approaching each other $x_1 \to x_2$. Point splitting puts a lower bound on this $x_1^\mu - x_2^\mu > \epsilon^\mu$. This also breaks translation invariance and is impractical for gauge theories, but is useful in theories with composite operators (such as a current $j^\mu(x) = \bar{\psi}(x)\gamma^\mu\psi(x)$).
- Lattice regularization. Breaks translation invariance and Lorentz invariance. But it is possible to construct a lattice such that translation and Lorentz invariance are restored at large distances
- Analytic regularization. Change propagator from $(k^2 - m^2)^{-1} \to (k^2 - m^2)^{-\alpha}$. This also breaks Gauge invariance, but is often easy to use.

Chapter 17
Vacuum Polarization

17.1 Introduction

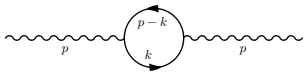
In QED, the Coulomb potential $V(r) = \frac{e^2}{r}$ is given by the exchange of a single photon

$$\sim\!\sim\!\sim = \frac{e^2}{p^2} \tag{17.1}$$

We saw this by Fourier transforming the propagator

$$V(r) = \int \frac{d^3p}{(2\pi)^3} \frac{e^2}{p^2} e^{i\,\vec{p}\vec{x}} = \frac{e^2}{r} \tag{17.2}$$

A one-loop correction to this comes from a e^+e^- loop inside the photon line



(17.3)

This will give us a correction to $V(r)$ proportional to e^4. We will show that the charge e is infinite, but it can be replaced by a renormalized charge, order by order in perturbation theory. And this has a physical effect: a correction to Coulomb's law, predicted by quantum field theory. We want to work out that prediction and compare to experiment.

You can think of the graph as the creation of virtual e^+e^- pairs, that is, a virtual dipole. In the same way that a dielectric material such as water, would get polarized if we put it in a electric field, this graph tells us how the vacuum itself gets polarized when it interacts with electromagnetic radiation. So this loop graph is called the vacuum polarization amplitude.

The loop is a mess because it involves photons and spinors. So let's build it up in pieces, starting with ϕ^3 theory, then scalar QED, then finally real QED.

Even this loop is a mess, but we can do it with some tricks and a lot of hard work!

17.2 Scalar ϕ^3 theory

As a warm-up for the vacuum polarization calculation in QED, we will start with scalar ϕ^3 theory with Lagrangian

$$\mathcal{L} = -\frac{1}{2}\phi(\Box + m^2)\phi + \frac{g}{3!}\phi^3 \tag{17.4}$$

So now we want to compute

$$i\mathcal{M} = \quad \text{(p − k, k, p, p)} \quad = \frac{1}{2}(ig)^2 \int \frac{d^4k}{(2\pi)^4} \frac{i}{(p-k)^2 - m^2 + i\varepsilon} \frac{i}{k^2 - m^2 + i\varepsilon} \tag{17.5}$$

which will tell us the virtual correction to the Yukawa potential. We will allow the initial and final line to be off-shell: $p^2 \neq m^2$, since we are calculating the correction to the initial ϕ propagator $\frac{i}{p^2 - m^2}$ which also must have $p^2 \neq m^2$ to make any sense. For example, in a t-channel scattering process, we would have $t = -p^2$ so p^2 characterizes the scattering angle.

First, we can use the Feynman Parameter trick

$$\frac{1}{AB}=\int_0^1 dx\,\frac{1}{[A+(B-A)x]^2} \tag{17.6}$$

with $A=(p-k)^2-m^2+i\varepsilon$ and $B=k^2-m^2+i\varepsilon$,

$$A+[B-A]x=(p-k)^2-m^2+i\varepsilon+\left[k^2-(p-k)^2\right]x \tag{17.7}$$

$$=k^2-2kp(1-x)+p^2-p^2x-m^2+i\varepsilon \tag{17.8}$$

$$=(k-p(1-x))^2+p^2x(1-x)-m^2+i\varepsilon \tag{17.9}$$

So,

$$\int \frac{d^4k}{(2\pi)^4}\frac{1}{(p-k)^2-m^2+i\varepsilon}\frac{1}{k^2-m^2+i\varepsilon}$$

$$=\int \frac{d^4k}{(2\pi)^4}\int_0^1 dx\frac{1}{[(k-p(1-x))^2-p^2x(1-x)-m^2+i\varepsilon]^2} \tag{17.10}$$

Now shift $k\to k+p(1-x)$ in the integral. The measure is unchanged, and we get

$$i\mathcal{M}=\frac{g^2}{2}\int \frac{d^4k}{(2\pi)^4}\int_0^1 dx\frac{1}{[k^2-(m^2-p^2x(1-x))+i\varepsilon]^2} \tag{17.11}$$

At this point, we can use our formulas above (say from Pauli-Villars with regulator mass Λ):

$$\int \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2-\Delta+i\varepsilon)^2}=-\frac{i}{16\pi^2}\log\frac{\Delta}{\Lambda^2} \tag{17.12}$$

So, with $\Delta=p^2x(1-x)-m^2$, we get

$$i\mathcal{M}=-\frac{ig^2}{32\pi^2}\int_0^1 dx\log\frac{m^2+tx(1-x)}{\Lambda^2},\quad t=-p^2 \tag{17.13}$$

We have replaced $-p^2$ by t to emphasize the fact that the momentum of a virtual scalar, for example, in t channel exchange will be spacelike, so $t=-p^2>0$.

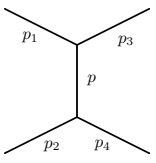
This integral can be done – the integrand is perfectly well behaved between $x=0$ and $x=1$. You get some ugly arctan functions. But let's just take $m=0$. Then we get

$$\mathcal{M}=\frac{g^2}{32\pi^2}\left[2-\log\frac{t}{\Lambda^2}\right] \tag{17.14}$$

The 2 cannot be physical, because we can remove it by redefining $\Lambda^2\to\Lambda^2e^{-2}$. But we can't remove the t dependence because M is just a number. So, how is this logarithmically divergent quantity related to anything physical?

17.2.1 interpretation

The diagram we computed is a correction to the tree level ϕ propagator. To see this, observe that the propagator is essentially the same as the t-channel scattering diagram

$$i\mathcal{M}_t^0=$$ 

$$=(ig)^2\frac{i}{p^2}=i\frac{g^2}{t} \tag{17.15}$$

If we insert our scalar bubble in the middle, we get

$$i\mathcal{M}_t^1 = \quad = (ig)^2 \frac{i}{p^2} i\mathcal{M} \frac{i}{p^2} = ig^2 \frac{1}{p^2}\left[-\frac{g^2}{32\pi^2}\log\frac{t}{M^2}\right]\frac{1}{p^2} \tag{17.16}$$

So that

$$i\mathcal{M}_t = i\frac{g^2}{t}\left(1 - \frac{1}{32\pi^2}\frac{g^2}{t}\log\frac{t}{M^2}\right) \tag{17.17}$$

If this expression looks a little funny, keep in mind that g is not a number in ϕ^3 theory but has dimensions of $\sqrt{\text{mass}}$. This actually makes ϕ^3 a little more confusing than QED, but not insurmountably so. Let's just simplify things by writing the whole thing in terms a new t-dependent variable $\tilde{g}^2 \equiv \frac{g^2}{t}$ which is dimensionless. Then

$$\mathcal{M}_t = \tilde{g}^2 - \frac{1}{32\pi^2}\tilde{g}^4 \log\frac{t}{M^2} \tag{17.18}$$

We'll come back to this t dependent scaling of the coupling when we discuss renormalizable vs non-renormalizable theories, but for now, let's just make this substitution and press on so we can get to vacuum polarization in QED.

Then we can *define* a renormalized coupling $\tilde{g}_R$ at some fixed scale t_0 by

$$\tilde{g}_R^2 \equiv \mathcal{M}_{t_0} \tag{17.19}$$

This is called a *renormalization condition*. It is a definition, and by definition, it holds to all orders in perturbation theory. It follows that

$$\tilde{g}_R^2 = \mathcal{M}_{t_0} = \tilde{g}^2 - \frac{1}{32\pi^2}\frac{\tilde{g}^4}{t_0}\log\frac{t_0}{M^2} \tag{17.20}$$

and we can invert this too

$$\tilde{g}^2 = \tilde{g}_R^2 + \frac{1}{32\pi^2}\frac{\tilde{g}_R^4}{t_0}\log\frac{t_0}{M^2} \tag{17.21}$$

Which leads to a prediction, for the matrix element at a different scale t

$$\mathcal{M}_t = \tilde{g}^2 + \frac{1}{32\pi^2}\tilde{g}^4\log\frac{t}{M^2} \tag{17.22}$$

$$= \tilde{g}_R^2 + \frac{1}{32\pi^2}\tilde{g}_R^4\log\frac{t}{t_0} \tag{17.23}$$

Thus we can measure $\mathcal{M}$ at one t and then make a nontrivial prediction at another value of t.

17.3 Vacuum polarization in scalar QED

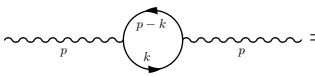
Now let's move on to the next simplest example, scalar QED

We found that in ϕ^3 theory,

$$i\mathcal{M}_{\phi^3} = \quad = (ig)^2 \int \frac{d^4k}{(2\pi)^4}\frac{i}{(p-k)^2 - m^2 + i\epsilon}\frac{i}{k^2 - m^2 + i\varepsilon} \tag{17.24}$$

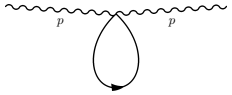
$$= -\frac{ig^2}{16\pi^2}\int_0^1 dx \log\frac{m^2 + tx(1-x)}{M^2} \tag{17.25}$$

The equivalent calculation in scalar QED is



$$= (-ie)^2 \int \frac{d^4k}{(2\pi)^4} \frac{i}{(p-k)^2 - m^2 + i\varepsilon} \frac{i}{k^2 - m^2 + i\varepsilon} (2k^\mu - p^\mu)(2k^\nu - p^\nu)$$

This looks basically the same, except for the numerator factors. In scalar QED there is also another diagram



$$= 2ie^2 g_{\mu\nu} \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m^2} \tag{17.26}$$

So we can add them together to get

$$i\mathcal{M}_{\mu\nu} = -e^2 \int \frac{d^4k}{(2\pi)^4} \frac{-4k^\mu k^\nu + 2p^\mu k^\nu + 2p^\nu k^\mu - p^\mu p^\nu + 2g^{\mu\nu}((p-k)^2 - m^2)}{[(p-k)^2 - m^2 + i\varepsilon][k^2 - m^2 + i\varepsilon]} \tag{17.27}$$

We can simplify things by anticipating the answer. By Lorentz invariance, the only 4-vector not integrated over is p^μ, so

$$M_{\mu\nu} = \Delta_1(p^2, m^2) p^2 g_{\mu\nu} + \Delta_2(p^2, m^2) p_\mu p_\nu \tag{17.28}$$

The reason we want to do this calculation is that it provides a correction to the Coulomb potential, or equivalently the photon propagator. At leading order, this propagator is

$$\Pi^0_{\mu\nu}(p) = -i \frac{g_{\mu\nu} + (1-\xi)\frac{p_\mu p_\nu}{p^2}}{p^2} \tag{17.29}$$

So for $\xi = 1$ the correction will be of the general form

$$\Pi_{\mu\nu}(p) = \Pi^0_{\mu\beta}(p) M_{\alpha\beta} \Pi^0_{\beta\nu}(p) = \frac{i}{p^2}\left(\Delta_1 \frac{g_{\mu\nu}}{p^2} + \Delta_2 \frac{p_\mu p_\nu}{p^2} \right) \tag{17.30}$$

So we can write for the full propagator

$$\Pi_{\mu\nu}(p) = \Pi^0_{\mu\nu}(p) + \Pi^1_{\mu\nu}(p) = -i \frac{(1+\Delta_1) g_{\mu\nu} + \Delta_2 \frac{p_\mu p_\nu}{p^2}}{p^2 + i\varepsilon} \tag{17.31}$$

So once we calculate Δ_1 and Δ_2 we can just use $\Pi_{\mu\nu}$ instead of $\Pi^0_{\mu\nu}$ in our QED calculations, to include the loop effect. But the Δ_2 term is just a change of gauge, and we have already proved with the path integral that QED is gauge invariant no matter what, so we really only need to compute Δ_1. This means extracting the term proportional to $g^{\mu\nu}$ in $M^{\mu\nu}$.

Looking at our amplitude

$$i\mathcal{M}_{\mu\nu} = -e^2 \int \frac{d^4k}{(2\pi)^4} \frac{-4k^\mu k^\nu + 2p^\mu k^\nu + 2p^\nu k^\mu - p^\mu p^\nu + 2g^{\mu\nu}((p-k)^2 - m^2)}{[(p-k)^2 - m^2 + i\varepsilon][k^2 - m^2 + i\varepsilon]} \tag{17.32}$$

we see that it is the sum of a bunch of terms. The $p^\mu p^\nu$ term we can pull out of the integral, and will give us only a contribution to Δ_2, so we can drop it. For the $p^\mu k^\nu$ term, we can pull p^μ out of the integral, so whatever the k integral gives, it must provide a p^ν by Lorentz invariance. So these terms can be dropped too. The $k^\mu k^\nu$ term is important – it may give a $p^\mu p^\nu$ piece, but may also give a $g^{\mu\nu}$ piece, which is what we're looking for. So we only need to consider

$$i\mathcal{M}_{\mu\nu} = -e^2 \int \frac{d^4k}{(2\pi)^4} \frac{-4k^\mu k^\nu + 2g^{\mu\nu}((p-k)^2 - m^2)}{[(p-k)^2 - m^2 + i\varepsilon][k^2 - m^2 + i\varepsilon]} \tag{17.33}$$

The denominator can be manipulated using Feynman parameters just as with the ϕ^3 theory

$$i\mathcal{M}_{\mu\nu} = -e^2 \int \frac{d^4k}{(2\pi)^4} \int_0^1 dx \frac{-4k^\mu k^\nu + 2g^{\mu\nu}((p-k)^2 - m^2)}{[(k - p(1-x))^2 + p^2 x(1-x) - m^2 + i\varepsilon]^2} \tag{17.34}$$

However, now when we shift $k^\mu \to k^\mu + p^\mu(1-x)$ we get a correction to the numerator. We get

$$i\mathcal{M}_{\mu\nu} = -e^2 \int \frac{d^4k}{(2\pi)^4} \int_0^1 dx \frac{-4(k^\mu + p^\mu(1-x))(k^\nu + p^\nu(1-x)) + 2g^{\mu\nu}((xp-k)^2 - m^2)}{[k^2 + p^2x(1-x) - m^2 + i\varepsilon]^2} \tag{17.35}$$

As we have said, we don't care about $p^\mu p^\nu$ pieces, or pieces linear in p^ν. Also, pieces like $p \cdot k$ are odd under $k \to -k$, for which the rest of the integral, including the measure is invariant. So these terms must give zero by symmetry. Thus we have only

$$i\mathcal{M}_{\mu\nu} = 2e^2 \int \frac{d^4k}{(2\pi)^4} \int_0^1 dx \frac{2k^\mu k^\nu - g^{\mu\nu}(k^2 - m^2 + x^2p^2)}{[k^2 + p^2x(1-x) - m^2 + i\varepsilon]^2} \tag{17.36}$$

It looks like this integral is much more badly divergent than the ϕ^3 theory – it is now quadratically instead of logarithmically divergent. That is, if we cut off at $k = \Lambda$ we will get something proportional to Λ^2. This is not a problem – the same thing happened for the Casimir energy, where we saw we could regulate with a heat-kernel regulator or a zeta-function regulator. Instead, we will use dimension regularization, which shows nicely why the integral isn't as badly divergent as it seems.

In d dimensions (using $k^\mu k^\nu \to \frac{1}{d} k^2 g^{\mu\nu}$), the integral becomes

$$i\mathcal{M}_{\mu\nu} = -2e^2\, \mu^{4-d} g^{\mu\nu} \int_0^1 dx \int \frac{d^dk}{(2\pi)^d} \frac{(1-\frac{2}{d})k^2 + x^2p^2 - m^2}{[k^2 + p^2x(1-x) - m^2]^2} \tag{17.37}$$

Using the formulas (A.46 of P&S)

$$\int \frac{d^4k}{(2\pi)^4} \frac{k^2}{[k^2-\Delta]^2} = -\frac{d}{2} \frac{i}{(4\pi)^{d/2}} \Gamma(1-\frac{d}{2}) \left(\frac{1}{\Delta}\right)^{1-\frac{d}{2}} \tag{17.38}$$

$$\int \frac{d^4k}{(2\pi)^4} \frac{k^\mu k^\nu}{[k^2-\Delta]^2} = -\frac{g^{\mu\nu}}{2} \frac{i}{(4\pi)^{d/2}} \Gamma(1-\frac{d}{2}) \left(\frac{1}{\Delta}\right)^{1-\frac{d}{2}} \tag{17.39}$$

$$\int \frac{d^dk}{(2\pi)^d} \frac{1}{[k^2-\Delta]^2} = \frac{i}{(4\pi)^{d/2}} \Gamma(2-\frac{d}{2}) \left(\frac{1}{\Delta}\right)^{2-\frac{d}{2}} \tag{17.40}$$

with $\Delta = m^2 - p^2x(1-x)$ we find

$$i\mathcal{M}_{\mu\nu} = -2\frac{ie^2}{(4\pi)^{d/2}}\, g^{\mu\nu}\mu^{4-d} \int_0^1 dx \left[(1-\frac{d}{2})\Gamma(1-\frac{d}{2}) \left(\frac{1}{\Delta}\right)^{1-\frac{d}{2}} + (x^2p^2 - m^2)\Gamma(2-\frac{d}{2}) \left(\frac{1}{\Delta}\right)^{2-\frac{d}{2}} \right]$$

Note that $\Gamma(2-\frac{d}{2}) = (1-\frac{d}{2})\Gamma(1-\frac{d}{2})$, so this simplifies to

$$i\mathcal{M}_{\mu\nu} = -2\frac{ie^2}{(4\pi)^{d/2}}\, g^{\mu\nu}\mu^{4-d} \int_0^1 dx \left[p^2x(2x-1)\Gamma(2-\frac{d}{2}) \left(\frac{1}{\Delta}\right)^{2-\frac{d}{2}} \right]$$

Now expanding as $d = 4 - \varepsilon$ we get

$$i\mathcal{M}_{\mu\nu} = -\frac{ie^2}{8\pi^2} p^2 g_{\mu\nu} \int_0^1 dx\, x(2x-1) \left[\frac{1}{\varepsilon} + \log\left(\frac{4\pi e^{-\gamma_E}\mu^2}{m^2 - p^2x(1-x)}\right) + \mathcal{O}(\varepsilon) \right]$$

The $\frac{1}{\varepsilon}$ gives the infinite constant that we will subtract off. For small m and $t = -p^2$, we find

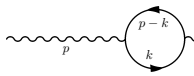
$$\mathcal{M}_{\mu\nu} = -\frac{\alpha}{12\pi} p^2 g_{\mu\nu} \left(\log\frac{\tilde{\mu}^2}{t} + C\right)$$

where $C \sim \frac{1}{\varepsilon} + \text{const}$ and $\tilde{\mu}^2 = 4\pi e^{\gamma_E}\mu^2$ as before.

Before we interpret this, let's just do the QED calculation, as it's almost exactly the same.

17.4 Spinor QED

Now let's do the real deal, QED. In this case, the loop is



$$= -(-ie)^2 \int \frac{d^4k}{(2\pi)^4} \frac{i}{(p-k)^2-m^2} \frac{i}{k^2-m^2} \mathrm{Tr}[\gamma^\mu(\not{p}-\not{k}+m)\gamma^\nu(\not{k}+m)]$$

where the -1 in front comes from the fermion loop. Note that there is only 1 diagram in this case.

Using our trace theorems, we find

$$\mathrm{Tr}[\gamma^\mu(\not{p}-\not{k}+m)\gamma^\nu(\not{k}+m)] = 4[(p^\mu-k^\mu)k^\nu + k^\mu(p^\nu-k^\nu) - g^{\mu\nu}(pk-k^2-m^2)] \tag{17.41}$$

$$= 4[p^\mu k^\nu + k^\mu p^\nu - 2k^\mu k^\nu + g^{\mu\nu}(k^2-pk+m^2)] \tag{17.42}$$

Thus we can drop the $p^\mu p^\nu$, p^μ and p^ν terms as before, and we find

$$i\mathcal{M}_{\mu\nu} = -e^2 \int \frac{d^4k}{(2\pi)^4} \frac{-4k^\mu k^\nu + 2g^{\mu\nu}(k^2-pk-m^2)}{[(p-k)^2-m^2+i\varepsilon][k^2-m^2+i\varepsilon]} \tag{17.43}$$

Recall that in scalar QED we had the same thing but the $g^{\mu\nu}$ term was $2g^{\mu\nu}(k-p)^2$. After going to Feynman parameters and changing $k^\mu \to k^\mu + p^\mu(1-x)$ we now find

$$i\mathcal{M}_{\mu\nu} = 4e^2 \int \frac{d^4k}{(2\pi)^4} \int_0^1 dx \frac{2k^\mu k^\nu - g^{\mu\nu}(k^2 - x(1-x)p^2 - m^2)}{[k^2+p^2x(1-x)-m^2]^2} \tag{17.44}$$

This integral is almost exactly the same as the one for scalar QED, with only a slight change in the numerator, from x^2 to $-x(1-x)$.

The result, in dim reg, is then

$$i\mathcal{M}_{\mu\nu} = -\frac{ie^2}{4\pi^2} p^2 g_{\mu\nu} \int_0^1 dx\, x(1-x)\left[\frac{1}{\varepsilon} + \log\left(\frac{4\pi e^{-\gamma_E}\mu^2}{m^2 - p^2x(1-x)}\right) + \mathcal{O}(\varepsilon)\right]$$

which is the same as in scalar QED except for $x(1-x)$ instead of $x(2x-1)$ in the integrand. The change can be traced to $\Delta - x(1-x)p^2 - m^2 = -2p^2(1-x)x$ instead of $\Delta + x^2p^2 - m^2 = p^2x(2x-1)$. So we find for large $t = -p^2 \gg m^2$

$$\mathcal{M}_{\mu\nu} = -\frac{\alpha}{3\pi} p^2 g_{\mu\nu} \left(\log\frac{\tilde{\mu}^2}{t} + C\right)$$

Thus an electron loop gives the same thing as a scalar loop multiplied by a factor of 4.

17.5 Physics of vacuum polarization

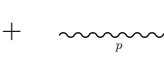
We have found that the fermion loop gives a factor of $M_{\mu\nu}$

$$M_{\mu\nu} = e^2 p^2 g_{\mu\nu} \Pi_2(p^2) \tag{17.45}$$

with

$$\Pi_2(p^2) = -\frac{1}{4\pi^2} \int_0^1 dx\, x(1-x)\left[\frac{1}{\varepsilon} + \log\left(\frac{\tilde{\mu}^2}{m^2 - p^2x(1-x)}\right)\right] \tag{17.46}$$

Thus the correction to the photon propagator is

$\Pi_{\mu\nu} =$  $+$

$$= -i\frac{g_{\mu\nu}}{p^2} + \frac{-i}{p^2} i\mathcal{M}_{\mu\nu} \frac{-i}{p^2} + \cdots$$

$$= -i\frac{(1+e^2\pi(p^2))g_{\mu\nu}}{p^2} + p^\mu p^\nu \text{ terms} \tag{17.47}$$

So then the scattering amplitude is just

$$\tilde{V}(p) = e^2 \frac{1 + e^2\pi(p^2)}{p^2} \tag{17.48}$$

We could define the renormalized electric change by saying that the force between two particles goes like $\frac{e_R^2}{r}$ at some fixed r. Equivalently, at fixed $p = p_0$. So,

$$p_0^2 \tilde{V}(p_0) = e_R^2 = e^2 + e^4\pi(p_0^2) \tag{17.49}$$

$$\Rightarrow e^2 = e_R^2 - e_R^4\pi(p_0^2) \tag{17.50}$$

Of course $\pi(p_0^2)$ is still infinite. However, now at another scale

$$p^2\tilde{V}(p) = e^2 + e^4\pi(p^2) \tag{17.51}$$

$$= e_R^2 + e_R^4(\pi(p^2) - \pi(p_0^2)) \tag{17.52}$$

Which is finite. It is conventional to take $p_0 = 0$, corresponding to $r = \infty$. This is a reasonable scale at which to expect the quantum corrections to Coulomb's law are finite. Then

$$\pi(p^2) - \pi(0) = \frac{1}{4\pi^2}\int_0^1 dx\, x(1-x)\log\left[1 - \frac{p^2}{m^2}x(1-x)\right] \tag{17.53}$$

Thus

$$\boxed{\tilde{V}(p) = \frac{e_R^2}{p^2} + \frac{e_R^4}{p^2}\frac{1}{4\pi^2}\int_0^1 dx\, x(1-x)\log\left[1 - \frac{p^2}{m^2}x(1-x)\right]} \tag{17.54}$$

which is a totally finite correction to the Coulomb potential.

17.5.1 small momentum – Lamb shift

First, let's look at the small-momentum, large distance limit. For $p \ll m$

$$\int_0^1 dx\, x(1-x)\log\left[1 - \frac{p^2}{m^2}x(1-x)\right] \approx \int_0^1 dx\, x(1-x)\left[-\frac{p^2}{m^2}x(1-x)\right] = -\frac{p^2}{30m^2} \tag{17.55}$$

Thus,

$$\tilde{V}(p) = \frac{e_R^2}{p^2} - \frac{e_R^4}{30\pi^2 m^2} \tag{17.56}$$

The Fourier transform of this constant piece is a $\delta(r)$ term, so

$$V(r) = -\frac{e^2}{4\pi r} - \frac{e^4}{60\pi^4 m^2}\delta(r) \tag{17.57}$$

This is known as the Uehling term.

How do we see the effect of this extra term? One way is to plug this into the Schrodinger equation and see the effect on the states of the hydrogen atom. Only the s-waves have support at $r = 0$ and so this shifts them. The energy is negative, so the energy is lowered. This is a contribution to the Lamb shift, the hyperfine structure of the Hydrogen atom (this gives a -27 MHz which is a measurable contribution to the -1028 MHz Lamb shift).

More carefully, the effective potential at 1-loop is (see P&S section 7.5)

$$V(r) = -\frac{e^2}{4\pi r}\left(1 + \frac{2e^2}{12\pi}\int_0^1 dx\, e^{-2mrx}\left(\frac{2x^2+1}{2x^4}\right)\sqrt{x^2-1}\right) \tag{17.58}$$

This is known as the Uehling potential (1935). For $r \gg \frac{1}{m}$,

$$V(r) = -\frac{\alpha}{r}\left[1 + \frac{\alpha}{4\sqrt{\pi}}\frac{1}{(mr)^{3/2}}e^{-2mr}\right], \quad r \gg \frac{1}{m} \tag{17.59}$$

You can see that the finite correction has extent $1/m_e = r_e$, the Compton radius of the electron. This is much smaller than the Bohr radius, $r_e = \alpha a_0$, so it basically only effects the $l = 0$ modes, and the δ function approximation is valid.

For $r \ll \frac{1}{m}$

$$V(r) = -\frac{\alpha}{r}\left[1 + \frac{2\alpha}{3\pi}\left(\log\frac{1}{mr} - \frac{5}{6} - \gamma_E\right)\right], \quad r \ll \frac{1}{m} \tag{17.60}$$

Thus at small distances there is a logarithmic dependence on the electron mass.

The Uehling potential doesn't give the only contribution to the Lamb shift. The Uehling potential applies only in the limit that we can talk about a potential at all – the non-relativistic limit. Otherwise, the Born approximation is not useful. Even when the electron is non-relativistic there are still important relativistic corrections, such as the spin-orbit coupling. To study the radiative corrections in the relativistic limit, we need to look at the large momentum region of the amplitude.

17.5.2 large momentum – logarithms of p

The opposite limit, $p \gg m$ corresponds to short distances. Then

$$\int_0^1 dx\, x(1-x)\log\left[1 - \frac{p^2}{m^2}x(1-x)\right] \approx \log(-\frac{p^2}{m^2})\int_0^1 dx\, x(1-x) = \frac{1}{6}\log(-\frac{p^2}{m^2}) \tag{17.61}$$

Then,

$$\tilde{V}(p) = \frac{e_R^2}{p^2}\left(1 + \frac{e_R^2}{12\pi^2}\log\frac{-p^2}{m^2}\right) \tag{17.62}$$

The $-p^2$ is ok, since Coulomb scattering corresponds to t-channel exchange with $t < 0$.

If we compare the potential at two scales p_1 and p_2 which are both $\gg m$, then

$$p_1^2\tilde{V}(p_1) - p_2^2 V(p_2) = \frac{e_R^4}{12\pi^2}\log\frac{p_1^2}{p_2^2} \tag{17.63}$$

which is independent of m. Note however, that we could not have simply take $m = 0$ at the beginning, because the loop would have been divergent. This kind of divergence is known as a mass singularity or infrared divergence. We will discuss them more in a few days.

So this says that there is an effective charge to the electron at short distances which goes like

$$e_{\rm eff}^2(p^2) = e_R^2\left(1 + \frac{e_R^2}{12\pi^2}\log\frac{-p^2}{m^2}\right) \tag{17.64}$$

Note that the sign of the correction is very important – the correction is positive and the effective charge gets larger at short distances. At large distances, the "bare" charge is screened by the virtual electron-positron dipole pairs. At smaller distances, there is less room for the screening and the effective charge increases.

Note that the charge only gets large very slowly. In fact taking

$$\alpha_R = \frac{e_R^2}{4\pi} = \frac{1}{137} \quad \Leftrightarrow \quad e_R = 0.303 \tag{17.65}$$

Thus

$$\alpha_{\rm eff} = \frac{1}{137}\left(1 + 0.0077\log\frac{-p^2}{m^2}\right) \tag{17.66}$$

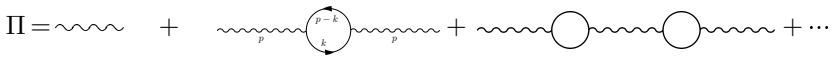
One would have to measure down to exponentially small energies in order to see the effect of this correction.

Nevertheless, there is an *extremely* important qualitative point about this calculation: perturbation theory breaks down. True, it is only at extraordinarily high energies $p \sim 10^{270}$ eV, but it means that QED is not a complete theory. This is known as a Landau Pole.

QED has a Landau Pole: perturbation theory breaks down at short distances

17.5.3 running coupling

We can actually include higher order corrections to the effective electric charge without much more work. We now know the result of the whole series



This gives

$$\tilde{V}(p) = \frac{e_R^2}{p^2}\left[1 + \frac{e_R^2}{12\pi^2}\log\frac{-p^2}{m^2} + \left(\frac{e_R^2}{12\pi^2}\log\frac{-p^2}{m^2}\right)^2 + \cdots\right] \tag{17.67}$$

$$= \frac{1}{p^2}\left[\frac{e_R^2}{1 - \frac{e_R^2}{12\pi^2}\log\frac{-p^2}{m^2}}\right] \tag{17.68}$$

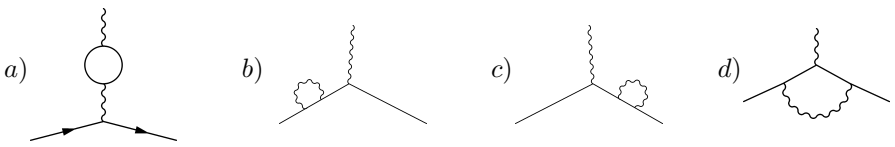
So now the momentum-dependent electric charge

$$e_{\text{eff}}^2(p^2) = \frac{e_R^2}{1 - \frac{e_R^2}{12\pi^2}\log\frac{-p^2}{m^2}}, \quad e_R \equiv e(m_e) \tag{17.69}$$

Note that we moved the reference value of the charge from $p = 0$ to $p = m_e$. This is natural, because the simple log dependence in this formula comes from the limit $p \gg m_e$, so near $p = m_e$ we should match on to the exact formula with the p/m_e corrections. In any case, the difference between the effective charge at $p = m_e$ and $p = 0$ is very small, they are both $e = \sqrt{4\pi\alpha} \sim 0.3$ up to α^2 corrections. We will eventually re-derive this formula for the effective charge using the renormalization group, which is a more general method. But the Landau pole singularity is unavoidable in QED.

17.6 Other graphs

What about the other contributions to the Coulomb potential? It seems that diagrams like



should all contribute. It actually turns out that the contribution of graphs b, c, d to the electric charge cancel. This is intuitively obvious. The electric charge measures the interactions of the photons, independently of what the photon is interacting with.

We can see this explicitly if we rescale the QED Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + \bar{\psi}\,(i\not{\partial} + e\,A_\mu + m)\psi \tag{17.70}$$

by $A_\mu \to \frac{1}{\sqrt{e}}A_\mu$. Then it becomes

$$\mathcal{L} = -\frac{1}{4e^2}F_{\mu\nu}^2 + \bar{\psi}\,(i\not{\partial} + A_\mu + m)\psi \tag{17.71}$$

From which the photon propagator would be

$$\Pi_{\mu\nu} = -\,i\,g_{\mu\nu}\frac{e^2}{p^2} \tag{17.72}$$

So the electric charge is closely connected to the propagator of the photon.

The easiest way to see physically that only graph a is relevant is to consider the Coulomb force between two very heavy charged sources. Then the electron propagator is very nearly on shell and very insensitive to changes in p going through the photon line. If we no longer care about the dynmaics of the electron, we can write the Lagrangian as

$$\mathcal{L} = -\frac{1}{4e^2}F_{\mu\nu}^2 + J_\mu A_\mu \tag{17.73}$$

where J_μ is the non-dynamical source current. Then there are no graphs b c and d to compute!

That's not a proof, it's just an argument. Keep in mind also that the Coulomb force and Coulomb potential are only defined in the non-relativistic approximation. More generally, we will need some more invariant notion of what we mean by the electric charge, which we will come to when we deal with renormalization in general. Then we will calculate graphs b, c and d as well, and show that their contributions to the renormalization of the electric charge all cancel. That will be the proof.

Chapter 18
g-2

18.1 Introduction

We saw that when you take the non-relativistic limit of the Dirac equation in the presence of a magnetic field, you get

$$H = \frac{p^2}{2m} + V(r) + \mu_B \vec{B} \cdot (\vec{L} + g_e \vec{S}), \quad g_e = 2 \tag{18.1}$$

where $\mu_B = \frac{e}{2m}$ is the Bohr magneton and $\vec{S} = \frac{1}{2}\vec{\sigma}$. We found that $g_e = 2$ exactly in the Dirac equation, and this was a prediction of Dirac's theory. Now we would like to calculate radiative corrections to g_e.

First of all, we should ask what can the correction depend on? It can depend on the electron mass, but then by dimensional analysis there must be another scale in the problem. This could be the momentum of the virtual photon, but then this correction would be more complicated than just $\vec{B} \cdot \vec{S}$. Thus the only possibility is that we get a finite numerical correction to g_e.

Let's work backwards. The non-relativistic Hamiltonian came from acting on the Dirac equation with the conjugate operator

$$0 = (i\not{\partial} + e\not{A} + m)(i\not{\partial} + e\not{A} - m)\psi = \Big((i\partial_\mu + e A_\mu)^2 + \frac{e}{2} F_{\mu\nu}\sigma^{\mu\nu} - m^2 \Big)\psi \tag{18.2}$$

where $\sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu]$. The $F_{\mu\nu}$ term can be written as

$$\frac{e}{2} F_{\mu\nu}\sigma^{\mu\nu}\psi = e \begin{pmatrix} (\vec{B} + i\,\vec{E})\vec{S} & \\ & (\vec{B} - i\vec{E})\vec{S} \end{pmatrix} \psi \tag{18.3}$$

$$= (e\,\vec{B}\vec{S} \pm i e \vec{E}\vec{S})\psi \tag{18.4}$$

So we can get a non-relativistic interaction energy from the $F_{\mu\nu}\sigma^{\mu\nu}$ term which looks like

$$\mathrm{Re}[\bar{\psi}\,\frac{e}{2}\,F_{\mu\nu}\sigma^{\mu\nu}\psi] = e\,\vec{B}\vec{S} \tag{18.5}$$

In the non-relativistic normalization, this should be $\mu_B g_e \vec{B}\vec{S} = \frac{e}{2m} g_e \vec{B}\vec{S}$. Thus, we can read g_e off as $2m$ times the coefficient of this $\frac{e}{2}F_{\mu\nu}\sigma^{\mu\nu}$ coupling.

Now suppose we add additional term to the Lagrangian of the form

$$\mathcal{L} = \cdots + c_1 \frac{1}{m}\frac{e}{2}\bar{\psi}\sigma^{\mu\nu}\psi F_{\mu\nu} + c_2 \frac{1}{m^3}\frac{e}{2}\bar{\psi}\sigma^{\mu\nu}\psi \Box F_{\mu\nu} + \cdots \tag{18.6}$$

$$= \cdots + \frac{1}{m}\frac{e}{2}\bar{\psi}\sigma^{\mu\nu}\psi F_2(\frac{\Box}{m^2})F_{\mu\nu} \tag{18.7}$$

These are terms with the $\bar{\psi}\sigma^{\mu\nu}\psi$ coupling and additional derivatives acting on $F_{\mu\nu}$. The factors of m_e are added by dimensional analysis. In the last step, we have combined them into a general function of $\Box$, called a form factor.

This form factor also gives a coupling to the magnetic field in the Hamiltonian

$$\mathcal{H} = \cdots + 2e\,F_2(\frac{p^2}{m^2})\vec{B}\vec{S} \tag{18.8}$$

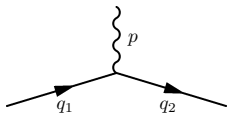
In particular, the mange tic moment, measured a large distance ($p \sim 0$) would get a correction proportional to $F_2(0)$. Comparing to the non-relativistic normalization, we find a correction to g_e which is $2m$ times $\frac{1}{m}F_2(0)$ or

$$g_e = 2 + 2F_2(0) \tag{18.9}$$

We could add this form-factor term directly to the Lagrangian, but let us suppose it is not there and we just have the QED Lagrangian. Then this term can be generated, in a predictive way, from radiative corrections, as we will now see.

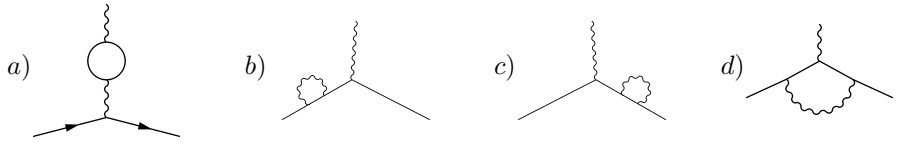
18.2 Vertex correction

The correction to the magnetic moment must come from graphs involving the photon and the electron. The leading graph is simply the vertex



$$= -ie\varepsilon_\mu \bar{u}(q_2)\gamma^\mu u(q_1) \tag{18.10}$$

At next order in α, there are 4 graphs



We don't want to assume that the photon is on-shell, because the magnetic field interaction comes from an external current.

We know we need one μ index to contract with the photon polarization

$$\mathcal{M}_{\text{tot}} = \varepsilon_\mu \bar{u}(q_2)[f_1\gamma^\mu + f_2 p^\mu + f_3 q_1^\mu + f_4 q_2^\mu]u(q_1) \tag{18.11}$$

Using momentum conservation, $p^\mu = q_2^\mu - q_1^\mu$, we can set $f_2 = 0$ and remove the p dependence. These $f's$ might involve spinors, but only if these are contracted with vectors. If there are factors of $\not{q}_1$ or $\not{q}_2$ in the f's, they can be removed by hitting the u or $\bar{u}$ and replacing them with masses. E.g

$$\not{q}_1 u(q_1) = m_e\, u(q_1) \tag{18.12}$$

so we can safely assume the f_i are real valued functions, not spinors. Moreover, they can only depend on $p^2 = 2m_e^2 - 2q_1\cdot q_2$ and on m_e, and we can fix the relative dependence by dimensional analysis.

Next, using the ward identity,

$$\not{p} f_1 + f_3 p_\mu q_1^\mu + f_4 p_\mu q_2^\mu = 0 \tag{18.13}$$

So we can eliminate f_4. We actually need more than the Ward identity, because the photon is not on-shell, but gauge invariance guarantees this "effective" Ward identity anyway.

Then we can use the Gordon identity

$$\bar{u}(q_2)(q_1^\mu + q_2^\mu)u(q_1) = (2m_e)\bar{u}(q_2)\gamma^\mu u(q_1) - \bar{u}(q_2) i\sigma^{\mu\nu} p_\nu u(q_1) \tag{18.14}$$

to write the matrix element as

$$\mathcal{M}_{\text{tot}} = \epsilon_\mu \mathcal{M}^\mu \tag{18.15}$$

$$i\mathcal{M}^\mu = (-ie)\bar{u}(q_2)\left[F_1(\frac{p^2}{m_e^2})\gamma^\mu + \frac{i\sigma^{\mu\nu}}{2m_e} p_\nu F_2(\frac{p^2}{m_e^2})\right]u(q_1) \tag{18.16}$$

Which is our final form. F_1 and F_2 are called form factors. This parametrization holds to all orders in perturbation theory. We will often write

$$i\mathcal{M}^\mu = (-ie)\bar{u}(q_2)\Gamma^\mu u(q_1) \tag{18.17}$$

$$\Gamma^\mu(p) = F_1(\frac{p^2}{m_e^2})\gamma^\mu + \frac{i\sigma^{\mu\nu}}{2m_e}p_\nu F_2(\frac{p^2}{m_e^2}) \tag{18.18}$$

The leading graph gives

$$F_1 = 1, \quad F_2 = 0 \tag{18.19}$$

Which of these two form factors could give an electron magnetic moment? First, F_1 can only modify the original $eA_\mu\bar{\psi}\gamma^\mu\psi$ coupling. This renormalizes the electric charge, as we saw from the vacuum polarization diagram. In fact, the entire effect of these terms are to renormalize the electric charge, so no new effect, like a magnetic moment, can come from this.

F_2, on the other hand, has precisely the structure of a magnetic moment (which is why we put it in this form with the Gordon identity to begin with). The interaction with the electric field comes from replacing the polarization vector with the gauge field. Then the term becomes

$$ip_\nu A_\mu(p)\bar{u}(q_2)\sigma^{\mu\nu}u(q_1) = \frac{1}{2}F_{\mu\nu}\bar{u}(q_2)\sigma^{\mu\nu}u(q_1) \tag{18.20}$$

Where we have used antisymmetry of $\sigma_{\mu\nu}$ to antisymmetrize $p_\mu A_\nu$. This is exactly the form we are looking for for the magnetic moment. Then we can read off the correction as

$$g_e = 2 + 2F_2(0) \tag{18.21}$$

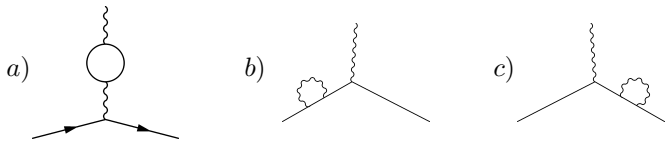
In summary, if we can find the term that looks like $\bar{u}(q_2)(q_1^\mu + q_2^\mu)u(q_1)$, we can pluck off the form factor as the coefficient in

$$i\mathcal{M}^\mu = \frac{e}{2m_e}F_2(p^2)[\bar{u}(q_2)(q_1^\mu + q_2^\mu)u(q_1)] + \cdots \tag{18.22}$$

which is the only part relevant to the magnetic moment.

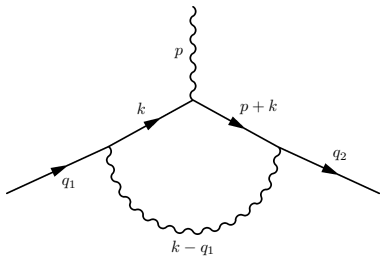
18.3 Evaluating the graphs

The first 3 loop graphs



can only give terms proportional to γ^μ. This is easy to see because these graphs just correct the propagators for the corresponding particles.

Thus the only graph we have to consider is



With p^μ going in, $p^\mu = q_2^\mu - q_1^\mu$.

This is

$$i\mathcal{M}^\mu = (-ie)^3 \int \frac{d^4k}{(2\pi)^4} \frac{-i}{(k-q_1)^2} \bar{u}(q_2)\gamma^\nu \frac{i(\not{p}+\not{k}+m)}{(p+k)^2-m^2}\gamma^\mu \frac{i(\not{k}+m)}{k^2-m^2}\gamma_\nu u(q_1) \tag{18.23}$$

$$= -ie^3 \bar{u}(q_2) \int \frac{d^4k}{(2\pi)^4} \frac{\gamma^\nu(\not{p}+\not{k}+m)\gamma^\mu(\not{k}+m)\gamma_\nu}{(k-q_1)^2[(p+k)^2-m^2][k^2-m^2]} u(q_1) \tag{18.24}$$

Let's start with the denominator. It has 3 terms, so we need to use the ABC Feynman parameters with

$$\frac{1}{ABC} = \int_0^1 dx\,dy\,dz\,\delta(x+y+z-1)\frac{2}{[xA+yB+zC]^3} \tag{18.25}$$

Setting

$$A = (k-q_1)^2 + i\varepsilon \tag{18.26}$$

$$B = (p+k)^2 - m^2 + i\varepsilon \tag{18.27}$$

$$C = k^2 - m^2 + i\varepsilon \tag{18.28}$$

We get

$$D = xA + yB + \mathrm{zC} \tag{18.29}$$

$$= k^2 + 2k(yp - zq_1) + yp^2 + zq_1^2 - (x+y)m^2 + i\varepsilon \tag{18.30}$$

Completing the square

$$D = (k + yp^\mu - zq_1^\mu)^2 - \Delta + i\varepsilon \tag{18.31}$$

$$\Delta = -xyp^2 + (1-z)^2 m^2 \tag{18.32}$$

Thus we want to change variables to

$$k^\mu = l^\mu - yp^\mu + zq_1^\mu \tag{18.33}$$

Now the numerator is:

$$N^\mu = \bar{u}(q_2)\gamma^\nu(\not{p}+\not{k}+m)\gamma^\mu(\not{k}+m)\gamma_\nu u(q_1) \tag{18.34}$$

$$= -2\bar{u}(q_2)\big[\not{k}\gamma^\mu\not{p} + \not{k}\gamma^\mu\not{k} + m^2\gamma^\mu - 2m(2k^\mu + p^\mu)\big]u(q_1) \tag{18.35}$$

Pulling the transformation on k gives

$$-\frac{1}{2}N^\mu = \bar{u}(q_2)\big[\big(\not{l} - y\not{p} + z\not{q}_1\big)\gamma^\mu\not{p} + \big(\not{l} - y\not{p} + z\not{q}_1\big)\gamma^\mu\big(\not{l} - y\not{p} + z\not{q}_1\big)\big]u(q_1) \tag{18.36}$$

$$= \bar{u}(q_2)\big[m^2\gamma^\mu - 2m(2l^\mu - 2yp^\mu + 2zq_1^\mu + p^\mu)\big]u(q_1) \tag{18.37}$$

After a bunch of algebra, and using $l^\mu l^\nu = \frac{1}{4}\gamma^{\mu\nu}l^2$, this simplifies to (I'm going back to k instead of l because it's easier to read)

$$-\frac{1}{2}N^\mu = \left[-\frac{1}{2}k^2 - (1-x)(1-y)p^2 + (1-2z-z^2)m^2\right]\bar{u}(q_2)\gamma^\mu u(q_1) \tag{18.38}$$

$$+ mz(z-1)\bar{u}(q_2)(q_1^\mu + q_2^\mu)u(q_1) + m(z-2)(x-y)\bar{u}(q_2)p^\mu u(q_1) \tag{18.39}$$

And at this point we have

$$i\mathcal{M}^\mu = -2ie^3 \int_0^1 dx\,dy\,dz\,\delta(x+y+z-1) \int \frac{d^4k}{(2\pi)^4}\frac{N^\mu(q,p)}{(k^2-\Delta)^3} \tag{18.40}$$

Now the p^μ term should vanish by the Ward identity. This is easy to see because Δ and the integral measure are symmetric in $x \leftrightarrow y$ but the p^μ term is antisymmetric. And we don't care about the γ^μ term for the magnetic moment, F_2.

The relevant term is

$$i\mathcal{M}^\mu = \bar{u}\,(q_2)(q_1^\mu + q_2^\mu)u(q_1)\left[4e^3\int_0^1 dx\,dy\,dz\,\delta(x+y+z-1)\int \frac{d^4k}{(2\pi)^4}\frac{m\,z(z-1)}{(k^2-\Delta)^3}\right] \tag{18.41}$$

Recalling that $F_2(p^2)$ was defined as the coefficient of this operator, normalized to

$$\mathcal{M}^\mu = \frac{e}{2m_e}F_2(p^2)\bar{u}\,(q_2)(q_1^\mu + q_2^\mu)u(q_1) + \cdots \tag{18.42}$$

Thus, we can just pluck off the form factor

$$F_2(p^2) = i\frac{2m}{e}\,4e^3\int_0^1 dx\,dy\,dz\,\delta(x+y+z-1)\int \frac{d^4k}{(2\pi)^4}\frac{m\,z(z-1)}{(k^2-\Delta)^3} \tag{18.43}$$

Now we use

$$\int \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2-\Delta)^3} = \frac{-i}{32\pi^2\Delta} \tag{18.44}$$

To get

$$F_2(p^2) = \frac{\alpha}{\pi}m^2\int_0^1 dx\,dy\,dz\,\delta(x+y+z-1)\frac{z(z-1)}{(1-z)^2m^2 - xyp^2} \tag{18.45}$$

At $p^2 = 0$ this integral is finite. Explicitly

$$F_2(0) = \frac{\alpha}{\pi}\int_0^1 dz\int_0^1 dy\int_0^1 dx\,\delta(x+y+z-1)\frac{z}{(1-z)} \tag{18.46}$$

$$= \frac{\alpha}{\pi}\int_0^1 dz\int_0^{1-z} dy\,\frac{z}{(1-z)} \tag{18.47}$$

$$= \frac{\alpha}{2\pi} \tag{18.48}$$

Thus

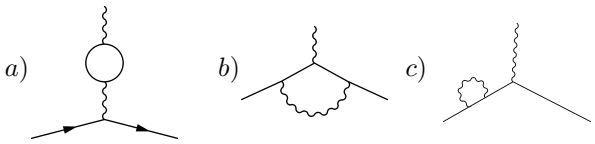
$$g_e = 2 + \frac{\alpha}{\pi} = 2.00232 \tag{18.49}$$

This result was first obtained by Julian Schwinger in 1948. This calculation was critically important historically for convincing people that the loops in QED had physical consequences. At the time of Schwinger's calculation, the best measurement was $g_e = 2.00236 \pm 0.00006$ which is in remarkable agreement with Schwinger's answer.

Chapter 19
Systematic Renormalization

19.1 Measuring Correlation Functions

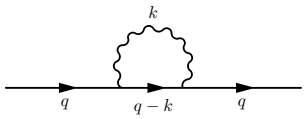
In QED there are basically only three independent 1-loop Feynman diagrams



(19.1)

Graph a) represents the polarization of the vacuum. We saw that this lead to the running (scale-dependent) electric charge. Graph b) contributed to the magnetic moment of the electron. We found $g_e = 2 + \frac{\alpha}{\pi}$ which agrees almost perfectly with experiment. There was also a part of graph b) proportional to γ^μ that we didn't address. And then there's graph c).

Graph c) is a very similar to graph a), but instead of the loop correction the photon propagator, it corrects the electron propagator. So we might naturally attempt to interpret it the same way, however, there is a complication. In general, this graph will give some function of the electron momentum and it's mass



$$= -i\Sigma_2(\not{q}, m_e)$$

This is called the electron self-energy graph.

In the same way that the photon self-energy (vacuum polarization) corrected the Coulomb potential, this graph will correct the potential generated by the exchange of an electron. For example, for Compton scattering. However, in order to use the interpretation of a potential, we need to be able to talk about the classical limit. For the Coulomb potential, this makes sense, as at large distances, $V(q)$ has a noticeable momentum dependence at small momentum transverse q: at small q it leads to Coulomb's law $V(r) = \frac{1}{r}$. However, the classical limit of electron exchange is Thompson scattering. But in Thompson scattering, you are not at all sensitive to the momentum transfer. This is simply because for $q \ll m_e$, the electron propagator is just $\frac{1}{m_e}$ and there is no q dependence. Thus, trying to think of a simple physical observable, like $V(p_1) - V(p_2)$ for photon exchange, is not so easy.

We might imagine it makes more sense to think about short distance effects of the self-energy diagram, keeping in the relativistic regime. For example, Compton scattering at high energy. There we could imagine the relevant energies are $q \gtrsim m_e$. So we could try to evaluate $\Sigma_2(\not{q})$ at $\not{q} = m_e$ and compare to a different value of $\Sigma(\not{q})$ to get a physical effect. This works in principle, but it practice, it is really a pain. The problem is that m_e is not well defined in perturbation theory. For example, including the self-energy graph, the effective electron propagator is

$$\Pi(\not{q}) = \text{—} + \quad \text{(k, q, q − k, q)} \quad = \frac{1}{\not{q} - m_e} + \frac{1}{\not{q} - m_e}\Sigma_2(\not{q})\frac{1}{\not{q} - m_e} \tag{19.2}$$

Since, in general, $\Sigma_2(m_e) \neq 0$, this correction will have a double pole at $\not{q} = m_e$ instead of a simple pole. However, a single electron state is associated with a simple pole in the propagator (we used this in the LSZ theorem). In fact, one way to *define* the physical electron mass is as the location of the pole. That interpretation is immediately suspect when there is a double pole present. (It's actually worse than that – the self-energy graph generates a cut in the complex plane starting at $\not{q} = m_e$). So this Σ_2 correction does more than just correct Compton scattering, it also forces us to be more specific about our definition of the electron mass. To disentangle these two effects it is helpful to have a more general framework for discussing renormalization.

A more general understanding of renormalization will also help us understand how predictive our theory actually is. That is, it will let us count the number of free parameters and number of experiments we have to do to fix them. It will allow us to calculate matrix elements as a perturbation series in the finite small renormalized electric charge e_R, instead of the infinite bare electric charge e we have been using. And it will let us not have to worry about what is observable all the time – we will renormalize parameters in the Lagrangian, and the observables will automatically come out finite. So there are lots of reasons to be systematic about renormalization!

The first step is to generalize our discussion of observables from things we can actually measure to things we could *in principle* measure. Observables in quantum field theory are time-ordered correlation functions of fields at different space time points, also known as Green's functions:

$$G(x_1, \cdots, x_n) = \langle 0|T\{\phi_1(x_1)\cdots\phi_n(x_n)\}|0\rangle \tag{19.3}$$

where ϕ can be any fields (e.g. A_μ or ψ of ϕ or something else). We often are interested in S-matrix elements, which are computed from Green's functions in momentum space, with the external lines removed by LSZ

$$\langle p_f|S|p_i\rangle = (p_f^2 - m^2)\cdots(p_i^2 - m^2)G(p_i, \cdots, p_f) \tag{19.4}$$

It may be difficult, but in principle all of these things can be measured, and so we should have a prescription to make sure they come out finite. Studying general features of these Green's functions is therefore the key to a systematic understanding of renormalization in QED.

19.2 Tadpoles

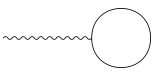
Let's start with the simplest Green's functions

$$\langle 0|\phi(p)|0\rangle \tag{19.5}$$

It is the expectation value of a field in the vacuum, also known as the *vacuum expectation value* (vev).

This correlation function is 0 at tree-level, since there are no 1-point interactions in the Lagrangian. If there were a 1-point interaction in the Lagrangian, so $\mathcal{L} = c\phi + \phi(\Box + m^2)\phi + \cdots$, then we should really complete the square $\phi \to \phi - \frac{c}{2m^2}$ to remove the tadpole. These shifted fields would be our real propagating degrees of freedom. This procedure will be discussed in the future, in the discussion of spontaneous symmetry breaking.

At 1-loop, a function like $\langle 0|A_\mu|0\rangle$ it would get a correction from

 (19.6)

this is called a *tadpole*. A nonzero expectation value for a vector field $\langle 0|A_\mu|0\rangle$ would break Lorentz invariance. This certainly happens, for example, if we cool down a ferromagnet below it's Curie temperature, at some point the lowest energy configuration has the spins aligned. This picks out a particular direction, breaking Lorentz and rotational invariance. This effect can be understood by calculating tadpoles for the photon. It signals spontaneous symmetry breaking.

Even for a charged scalar, if $\langle 0|\phi|0\rangle \neq 0$, some symmetry is broken. In this case invariance under $\phi \to e^{i\alpha}\phi$. This was the symmetry which QED gauged, so this breaks gauge invariance. Thus the photon does not have to remain massless. In fact, one way to understand type II superconductors is through the condensation of cooper pairs which are bosonic bound states of electrons $\phi(x) = e(x)e(x)$. The binding forces are transmitted via phonons. When ϕ gets a vev, the photon picks up a mass proportional to this vev

$$m_\gamma \sim \langle e\, e\rangle \neq 0 \tag{19.7}$$

Then the Coulomb force goes from long range $\frac{1}{r}$ to Yukawa $\frac{1}{r}e^{-m_\gamma r}$. That explains why magnetic fields can't penetrate superconductors – the Meissner effect. All of this is very beautiful stuff, known as the BCS theory of superconductivity, and I hope you will get to study it in detail anyway. The point here is just that tadpoles are very important in general, but they are not so relevant right now for our systematic study of renormalization in QED.

The mass of the W and Z bosons are non-zero. Can these masses be seen as coming from some kind of tadpole diagrams? The answer seems to be yes, but we don't yet know what exactly it is that we are computing the tadpoles for. The Cooper pairs are called the Higgs, but we don't know what the Cooper pairs are made of. What is the equivalent of the electrons that are being bound? And what is the equivalent of the phonons that bind them? That is the biggest mystery in high energy physics!

We will study tadpoles and spontaneous symmetry breaking in great detail later in the course. But it is a distraction from our current task of understanding renormalization. There are no non-vanishing tadpoles in QED, as you will see immediately if you try to compute one.

19.3 Electron self-energy

Next, let us look at the two point functions. Starting with fermions, we have

$$G_2(p_1,p_2)(2\pi)^4\delta^4(p_1+p_2) = \int d^4x \int d^4y e^{i(p_1x+p_2y)}\big\langle 0|T\{\bar\psi(x)\psi(y)||0\big\rangle \tag{19.8}$$

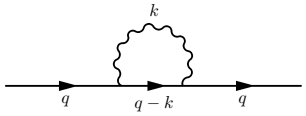
Because we know that all these delta function factors and phase space integrals drop out, we will just abbreviate this as

$$G_2(q) = \big\langle 0|T\big\{\bar\psi\psi\big\}|0\big\rangle \tag{19.9}$$

At tree level, this is just the fermion propagator

$$G_2(q) = \Pi_0(\not q) = \frac{i}{\not q - m} \tag{19.10}$$

At 1-loop it gets a correction due to the *electron self-energy* graph



$$= \Pi_0(\not q)\big[-i\Sigma_2(\not q)\big]\Pi_0(\not q)$$

where the $-i$ is just a convention. This is the equivalent of the photon self-energy graph, which we called *vacuum polarization*.

19.3.1 self-energy loop graph

The expression for the loop is

$$-i\Sigma_2(\not q) = (ie)^2\int \frac{d^4k}{(2\pi)^4}\gamma^\mu \frac{i(\not q - \not k + m_e)}{(q-k)^2 - m_e^2 + i\varepsilon}\gamma^\mu \frac{-i}{k^2+i\varepsilon} \tag{19.11}$$

Here we are seeing a new feature – the integrand vanishes when $k=0$. This is known as an infrared divergence. We didn't get such a feature in the vacuum polarization graph because only the electron was going around the loop. Whenever you have massless particles in a loop, you get infrared divergences.

Infrared divergences can be regulated just like UV divergences, and the regulator drops out of physical calculations. In fact, if you remember the expression for the Coulomb potential was

$$\tilde{V}(p)=\frac{e_R^2}{p^2}\left(1+\frac{e_R^2}{12\pi^2}\log\frac{-p^2}{m^2}\right) \tag{19.12}$$

but that for two scale $p_1, p_2 \gg m$ the difference between the potentials doesn't depend on m

$$p_1^2\tilde{V}(p_1)-p_2^2V(p_2)=\frac{e_R^2}{12\pi^2}\log\frac{p_1^2}{p_2^2} \tag{19.13}$$

The same thing will happen for the self energy graph.

So let us regulate the self energy by giving the photon a mass m_γ.

$$\frac{1}{k^2}\to\frac{1}{k^2-m_\gamma^2} \tag{19.14}$$

For now, don't worry about the photon mass dropping out of calculations – think of it as really being physical. We want to deal with one problem at a time, and now it is how to remove the UV divergences.

To regulate the UV divergence, we have to choose a regularization scheme. It turns out that dimensional regularization also removes infrared divergences. This is great for doing calculations, but not so useful if you want to see what's going on. So we will use Pauli-Villars.

It is enough to add only a ghost photon. You have to be very careful doing this, since it breaks gauge invariance, but in this case it turns out it's ok. So we will replace

$$\frac{1}{k^2}\to\frac{1}{k^2-m_\gamma^2}\to\frac{1}{k^2-m_\gamma^2}-\frac{1}{k^2-\Lambda^2}=\int_{m_\gamma^2}^{\Lambda^2}\frac{-1}{(k^2-z)^2} \tag{19.15}$$

Thus our self-energy graph becomes

$$-i\Sigma_2(\not q)=e^2\int_{m_\gamma^2}^{\Lambda^2}dz\int\frac{d^4k}{(2\pi)^4}\frac{-2(\not q-\not k)+m_e}{(q-k)^2-m_e^2+i\varepsilon}\frac{-1}{(k^2-z+i\varepsilon)^2} \tag{19.16}$$

Now we can introduce Feynman parameters, Wick rotate, and evaluate the integral, giving (see PS section 7.1)

$$\boxed{\Sigma_2(\not q)=\frac{e^2}{16\pi}\int_0^1 dx(4m_e-2x\not q)\log\frac{x\Lambda^2}{(1-x)m_e^2+x\,m_\gamma^2-x(1-x)q^2}} \tag{19.17}$$

There are 4 scales here: Λ, m_e, q and m_γ. We'll deal with them one at a time.

19.3.2 wave-function renormalization at $m_e=0$

First we want to get rid of Λ. So take $m_e=0$ but still assume the photon has mass m_γ. Then

$$\Sigma_2(\not q)=-\not q\frac{e^2}{8\pi}\int_0^1 dx\,x\log\frac{\Lambda^2}{m_\gamma^2-(1-x)q^2} \tag{19.18}$$

We would like to take $q\to 0$ here, to put the electron on shell for it's propagator. This gives

$$\Sigma_2(\not q)=-\not q\frac{\alpha}{4\pi}\log\frac{\Lambda^2}{m_\gamma^2}+\mathcal{O}(q^2) \tag{19.19}$$

So now the $O(e^2)$ electron propagator is

$$\Pi(\not q) = \text{—} + \quad \underset{q \qquad q-k \qquad q}{\overset{k}{\text{—}\!\!\circlearrowright\!\!\text{—}}} \quad = \frac{i}{\not q} + \frac{i}{\not q}\big[-i\Sigma_2(\not q)\big]\frac{i}{\not q} \tag{19.20}$$

$$= \frac{i}{\not q}\left(1 - \frac{\alpha}{4\pi}\log\frac{\Lambda^2}{m_\gamma^2}\right) \tag{19.21}$$

As expected we have found an infinite correction to the two point function

$$G_2(q) = \langle 0|T\{\bar\psi\psi\}|0\rangle = \frac{i}{\not q}\left(1 - \frac{\alpha}{4\pi}\log\frac{\Lambda^2}{m_\gamma^2}\right) \tag{19.22}$$

What are we to do about this infinity? For the Coulomb potential, we were able to absorb the infinity into the electric charge. But here, the leading order expression doesn't seem to have any parameters into which we can absorb the infinity.

In fact, there is a parameter: the normalization of the fermion wavefunction. Nobody ever told us what the wavefunctions were supposed to be! Say we had defined

$$\psi(x) = \frac{1}{\sqrt{Z_2}}\int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}\Big(a_p^\dagger e^{ipx} + a_p e^{-ipx}\Big) \tag{19.23}$$

for some number Z_2. This is the eponymous *renormalization*: $\psi^R = \frac{1}{\sqrt{Z_2}}\psi^0$. We say R for renormalized, physical quantities, and 0 for infinite bare quantities.

Sticking a bunch of these Z factors in the Lagrangian, it becomes

$$\mathcal{L} = -\frac{1}{4}Z_3F_{\mu\nu}^2 + iZ_2\bar\psi_R\not\partial\psi_R + Z_2m_0\bar\psi_R\psi_R + e_0Z_2\sqrt{Z_3}\,\bar\psi_R\not A_R\psi_R \tag{19.24}$$

What kind of observables would this change? Well, we are in the process of cataloging all the observables, so we are bound to find out. The tadpoles vanish no matter what, and we are working on the two point function.

For the two-point function, the tree level piece gets a $\frac{1}{Z_2}$. For the loop:

$$\frac{i}{\not q - m}\left[e^2\frac{1}{\not q - k - m}\frac{1}{k^2}\right]\frac{i}{\not q - m} \to \frac{i/Z_2}{\not q - m}\left[(eZ_2\sqrt{Z_3})^2\frac{1/Z_2}{\not q - k - m}\frac{1/Z_3}{k^2}\right]\frac{i/Z_2}{\not q - m} \tag{19.25}$$

$$= \frac{1}{Z_2}\frac{i}{\not q - m}\left[e^2\frac{1}{\not q - k - m}\frac{1}{k^2}\right]\frac{i}{\not q - m} \tag{19.26}$$

So,

$$G_2(q) = \frac{1}{Z_2}\frac{i}{\not q}\left(1 - \frac{\alpha}{4\pi}\log\frac{\Lambda^2}{m_\gamma^2}\right) \tag{19.27}$$

We know already that $Z_2 = 1$ to leading order. If we take $Z_2 = 1 + \delta_2$ with

$$\delta_2 = -\frac{\alpha}{4\pi}\log\frac{\Lambda^2}{m_\gamma^2} \tag{19.28}$$

then

$$G_2(q) = \frac{i}{\not q}\big(1 + \mathcal{O}(\alpha^2)\big) \tag{19.29}$$

which renormalizes the propagator back to what we started with.

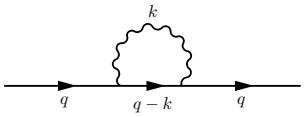
There is no physical prediction in this calculation, because, as we said, it is hard to come up with observables. The important point is that we have gotten rid of the UV divergence in the 2-point function.

The fact that only one factor of Z_2 comes through the expression is because the only thing that actually depends on Z are the external states in the Green's function $G_2 = \langle \bar{\psi}\psi \rangle$.

19.3.3 wave-function renormalization with a mass

Now let's put the electron mass back in. Then the two point function becomes

$$G_2(\not q) = \text{—} + \quad = \frac{i}{\not q - m_e} + \frac{i}{\not q - m_e}[-i\Sigma_2(\not q)]\frac{i}{\not q - m_e} \tag{19.30}$$



For $m_e \gg m_\gamma$, we can set $m_\gamma = 0$. Then,

$$\Sigma_2(\not q) = \frac{\alpha}{4\pi}\int_0^1 dx(4m_e - 2x\not q)\log\frac{x\Lambda^2}{(1-x)m_e^2 - x(1-x)q^2} \tag{19.31}$$

We expect that this to be zero when $\not q = m_e$, to cancel one of the poles in the expansion of G_2. However,

$$\Sigma_2(m_e) = \frac{3\alpha}{8\pi}m_e\left(1 - 2\log\frac{m_e^2}{\Lambda^2}\right) \tag{19.32}$$

which is not zero. So that means that $G_2(\not q)$ has a double pole at $\not q = m_e$. What is the physical interpretation of that?

Perhaps we could remove it with the wave function renormalization. Putting in the Z factor, and writing

$$Z_2 = 1 + \delta_2 \tag{19.33}$$

with $\delta_2 = \mathcal{O}(e^2)$, the only effect is from the leading propagator:

$$\frac{1}{Z_2}\frac{i}{\not q - m_e} = \frac{i}{\not q - m_e} + \frac{i}{\not q - m_e}\big[i\delta_2(\not q - m_e)\big]\frac{i}{\not q - m_e} \tag{19.34}$$

Then we get

$$\Sigma(\not q) = \Sigma_2(\not q) - \delta_2\big(\not q - m_e\big) \tag{19.35}$$

And we can see that there is no way to choose a number δ_2 so that $\Sigma(m_e) = 0$.

The resolution is that the electron mass in the Lagrangian is not necessarily the physical electron mass (the location of the pole). So let us treat the Lagrangian mass as a free parameter and use it's renormalization to remove the double pole. We write

$$m_e = \frac{1}{Z_2}(m_R + \delta_m) \tag{19.36}$$

where m_e is taken to be infinite, but m_R is the physical finite, renormalized electron mass.

Then

$$\frac{1}{Z_2}\frac{i}{\not q - m_e} = \frac{1}{1+\delta_2}\frac{i}{\not q - m_R + \delta_2 m_R - \delta_m} \tag{19.37}$$

$$= \frac{i}{\not q - m_e} + \frac{i}{\not q - m_e}\big[-\delta_2(\not q - m_R) - m_R\delta_2 + \delta_m\big]\frac{1}{\not q - m_e} \tag{19.38}$$

$$= \frac{i}{\not q - m_e} + \frac{i}{\not q - m_e}\big[-i(-\delta_2\not q + \delta_m)\big]\frac{i}{\not q - m_e} \tag{19.39}$$

Note that the $m_R\delta_2$ pieces have canceled, which is why we defined m_e with a factor of Z_2.

So,

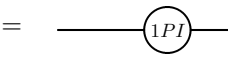
$$\Sigma(\not q) = \Sigma_2(\not q) - \delta_2\not q + \delta_m \tag{19.40}$$

Now there is a value of δ_2 and δ_m for which $\Sigma(m_R) = 0$. We can use this to enforce that there is only a single pole in the propagator. Let us press on a bit before setting the other condition to determine δ_m and δ_2.

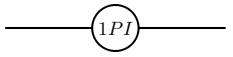
19.3.4 one-particle irreducible graphs

We only included one particular self-energy correction in the series we summed to correct the electron propagator. There are lots more graphs, and the full non-perturbative propagator, should properly be defined as the sum of all these graphs. To make sure we are not counting anything twice, we must only include graphs which cannot be cut in two by cutting a single propagator. Those would form a different series, which we could also sum. So anything which cannot be cut that connects one incoming electron to one outgoing electron are called one-particle irreducible graphs.

Then we define

$\Sigma(\not{q}) =$  $= \Sigma_2(\not{q}) + \cdots$ (19.41)

Each 1PI insertion will give a contribution to the Green's function

$G_2(\not{q}) =$ — +  $+ \cdots$ (19.42)

$$= \frac{i}{\not{q} - m} + \frac{i}{\not{q} - m}[-i\Sigma(\not{q})]\frac{i}{\not{q} - m} + \cdots \tag{19.43}$$

$$= \frac{i}{\not{q} - m}\left(1 + \frac{\Sigma(\not{q})}{\not{q} - m} + \left(\frac{\Sigma(\not{q})}{\not{q} - m}\right)^2 + \cdots\right) = \frac{i}{\not{q} - m}\frac{1}{1 - \frac{\Sigma(\not{q})}{\not{q} - m}} \tag{19.44}$$

$$= \frac{i}{\not{q} - m - \Sigma(\not{q})} \tag{19.45}$$

For the full non-perturbative propagator.

We found that if we shift the bare mass m_e and add a renormalization factor to ψ, then

$$\Sigma(\not{q}) = \Sigma_2(\not{q}) - \delta_2\not{q} + \delta_m + \mathcal{O}(e^4) \tag{19.46}$$

19.3.5 pole-mass renormalization conditions

At this point, there appears a natural condition which we can impose to fix the δ_m and δ_2: we can demand that the pole in the propagator remain at $\not{q} = m_R$. Recall that to use (or really, to interpret) the LSZ theorem, we needed to project out the 1-particle states by multiplying by terms like $\not{p} - m$. Since the physical one particle states have poles at their physical masses, this should be projecting out $\not{p} - m_R$. This gives zero unless there is a pole in the corresponding matrix element, so it is critical that the pole remain there. So the renormalized Green's function had better have a pole at $\not{p} = m_R$.

So we can define the renormalized mass by

$$\boxed{\Sigma(m_R) = 0} \tag{19.47}$$

This implies to order e_R^2 that

$$\Sigma_2(m_R) = m_R\delta_2 - \delta_m \tag{19.48}$$

Thus, at order e_R^2 (with PV regulator)

$$m_R\delta_2 - \delta_m = \frac{3\alpha}{8\pi}m_e\left(1 - 2\log\frac{m_R^2}{\Lambda^2}\right) \quad \text{(Pauli} - \text{Villars)} \tag{19.49}$$

Another reasonable requirement we might impose is that the residue of the pole of the propagator be 1 (or i). In some sense this is just a convention for what we multiply by in LSZ, but we have to be consistent with that convention! So,

$$1 = \lim_{\not{q}\to m_R} \frac{\not{q} - m_R}{\not{q} - m_R - \Sigma(\not{q})} = \lim_{\not{q}\to m_R} \frac{1}{1 - \frac{d}{d\not{q}}\Sigma(\not{q})} \tag{19.50}$$

where we have used L'Hospital's rule. This implies

$$\boxed{\frac{d}{d\not{q}}\Sigma(\not{q})\Big|_{\not{q}=m_R} = 0} \tag{19.51}$$

which means, to order e_R^2

$$\delta_2 = \frac{d}{d\not{q}}\Sigma_2(\not{q})\Big|_{\not{q}=m_R} \tag{19.52}$$

Recalling

$$\Sigma_2(\not{q}) = \frac{\alpha}{4\pi}\int_0^1 dx(4m_e - 2x\not{q})\log\frac{x\Lambda^2}{(1-x)m_e^2 + x m_\gamma^2 - x(1-x)q^2} \quad \text{(PV)} \tag{19.53}$$

This leads to

$$\delta_2 = \frac{d}{d\not{q}}\Sigma(\not{q})\Big|_{\not{q}=m_R} = \frac{\alpha}{2\pi}\Bigg[\int dx\Bigg[- x\log\frac{x\Lambda^2}{(1-x)m_e^2 + x m_\gamma^2} + 2\frac{(2-x)x(1-x)m_R^2}{(1-x)^2 m_R^2 + x m_\gamma^2}\Bigg] \quad \text{(PV)} \tag{19.54}$$

$$= -\frac{\alpha}{4\pi}\log\Lambda^2 + \cdots \tag{19.55}$$

Note that the divergent $\log\Lambda^2$ part is the same as we found for a massless electron, which makes sense as the UV divergences shouldn't be sensitive to any finite mass. And

$$\delta m = \frac{\alpha}{4\pi} m_e \int_0^1 dx(4-2x)\log\frac{x\Lambda^2}{(1-x)^2 m_e^2 + x m_\gamma^2} \quad \text{(PV)} \tag{19.56}$$

If we repeat the calculation in dimensional regularization, we would find

$$\Sigma_2(\not{q}) = \frac{e_R^2}{(4\pi)^{d/2}}\mu^{4-d}\Gamma(2-\frac{d}{2})\int_0^1 dx \frac{(4-\varepsilon)m_R - (2-\varepsilon)x\not{q}}{((1-x)(m_R^2 - x q^2) + x m_\gamma^2)^{2-d/2}}, \quad \text{(DR)} \tag{19.57}$$

So,

$$\delta_2 = \frac{d}{d\not{q}}\Sigma(\not{q})\Big|_{\not{q}=m_R} = \frac{e_R^2}{16\pi^2}\Bigg[- \frac{4}{\varepsilon} - 1 - \log\frac{\tilde{\mu}^2}{m_R^2} + 4\int dx \frac{2-x}{1-x} x \frac{(1-x)^2 m_R^2}{(1-x)^2 m_R^2 + x m_\gamma^2}\Bigg], \quad \text{(DR)} \tag{19.58}$$

$$\delta_m = m_R\delta_2 + \Sigma_2(m_R) = m_R\delta_2 + \frac{e_R^2}{16\pi^2} m_R\Big(\frac{4}{\varepsilon} + \frac{3}{2} + 3\log\frac{\tilde{\mu}^2}{m_R^2}\Big), \quad \text{(DR)} \tag{19.59}$$

19.3.6 Summary of electron self-energy renormalization

In summary, the two conditions

$$\Sigma(m_R) = 0 \tag{19.60}$$

$$\frac{d}{d\not{q}}\Sigma(\not{q})\Big|_{\not{q}=m_R} = 0 \tag{19.61}$$

allow us to uniquely fix the counterterms δ_2 and δ_m. These are known as *renormalization conditions*. The conditions we have chosen force the physical electron mass m_R and the residue of the propagator to be constant to all orders in perturbation theory. The values of the counterterms depend on the UV regularization scheme and can depend on an infrared regularization scheme (*e.g.* photon mass) as well. The counterterms δ are formally infinite.

By the way, notice that in this scheme we cannot ask about radiative corrections to the electron mass. That question has no meaning in this pole-mass scheme, which defines mass as the location of the pole in the propagator. There are other ways to define mass, such as the $\overline{\text{MS}}$ mass, for which you can ask about radiative corrections, which we discuss below, or the Pythia mass, which is related to a particular Monte Carlo event generator. In those cases, radiative corrections to masses do have meaning. If you find this confusing, think about how you could measure the radiative correction to a mass. Not so easy!

19.4 Amputation

The two point function constructed out of renormalized fields is

$$G_2^R(\not p) = \langle 0|T\{\psi^R\psi^R\}|0\rangle = \frac{i}{\not p - m_R} + \text{regular at } \not p = m_R \tag{19.62}$$

In terms of the bare fields $\psi^0 = \sqrt{Z_2}\psi^R$ it is

$$G_2^{\text{bare}}(\not p) = \langle 0|T\{\psi^0\psi^0\}|0\rangle = Z_2\frac{i}{\not p - m_R} = \frac{i}{\not p - m_R - \Sigma(\not p)} \tag{19.63}$$

where Σ includes counterterm corrections from the expansion of Z_2.

Recall that the LSZ theorem projected out the physical one-particle asymptotic states in the S matrix by multiplying by factors of $p^2 - m_e^2$:

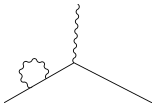
$$\langle f|S|i\rangle = (\not p_f - m_e)\cdots(\not p_i - m_e)\langle 0|T\{\psi^0\cdots\psi^0\}|0\rangle \quad (\text{original LSZ}) \tag{19.64}$$

Which mass should this be? At tree level it doesn't matter, but including the loop effects, we had better make sure we are projecting out the physical electron mass, by adding factors of $p^2 - m_R^2$ with m_R being the location of the pole. So

$$\langle f|S|i\rangle = (\not p_f - m_R)\cdots(\not p_i - m_R)\langle 0|T\{\psi^0\cdots\psi^0\}|0\rangle \quad (\text{renormalized LSZ}) \tag{19.65}$$

Note that this is still expressed in terms of time ordered products of the bare fields.

Now, if we are to use this exact propagator, we must also *truncate* external lines. That is, we should compute the 1PI bubbles on external lines separately, then add them to the diagrams without 1PI bubbles. This is also known as *amputation*. Also, not a big deal. It actually makes things easier. For example, consider



(19.66)

This graph just sets the counterterms δ_2 and δ_m but doesn't change the physical electron mass e_R or the residue of the electron propagator. Thus as long as we use the physical electron mass and canonically normalized interaction in Coulomb's law, we can ignore this graph. Another way to see it, is that the external electron here is on shell. So $\not p = m_R$ and then $\Sigma(\not p) = 0$ by the renormalization condition. The result is that the external lines in the graph would give factors of

$$\langle 0|T\{\psi^0\psi^0\}|0\rangle = Z_2\frac{i}{\not p - m_R} \tag{19.67}$$

So it is these poles with factors of Z_2 that get canceled by LSZ. Then if we look at amputated diagrams, LSZ simplifies

$$\boxed{\langle f|S|i\rangle = Z_2^n\langle 0|T\{\psi^0\cdots\psi^0\}|0\rangle^{\text{amputated}} = \langle 0|T\{\psi^R\cdots\psi^R\}|0\rangle^{\text{amputated}}} \tag{19.68}$$

So to compute S-matrix elements, we can just use the renormalized propagators, and drop all corrections to external lines.

S-matrix elements can be thought of as Green's functions in which the external particles are on-shell. Since they are on shell, there are poles, and these poles must be removed to make sense of the answer. The point of amputation is

- To remove the poles so the external lines can be taken on-shell
- Since the external lines are on-shell, they are always evaluated *at* the renormalization point – there is never any additional measurable effect of the renormalization of external lines.

There are more general Green's functions we can calculate, with off-shell external particles. Usually, these can be extracted from on-shell states. For example, the Green's function for an off-shell photon with momentum p^μ ($p^2 \neq 0$) can be calculated from the Green's function for an on-shell e^+e^- whose center of mass energy is $\sqrt{p^2}$.

19.5 Renormalized perturbation theory

The idea behind renormalization is that for every infinity, there should be a free parameter to absorb it. We started by absorbing infinities into a constant vacuum energy density (the Casimir force), and then into the electric charge (renormalized Coulomb potential). Now we have begun a systematic study of correlation functions and found we had two new parameters, the electron mass and the normalization of the electron field which we could also use. Are we always going to get a new parameter for every calculation?

Looking at the QED path integral

$$Z = \int \mathcal{D}A_\mu \mathcal{D}\psi \mathcal{D}\bar{\psi} \exp\left(i \int d^4x \left[-\frac{1}{4} F_{\mu\nu}^2 + \bar{\psi}\,(i\not{\partial} + e\,A_\mu + m)\psi + \rho \right] \right) \tag{19.69}$$

it seems that there are only three parameters m, e, and the vacuum energy density ρ. However, the path integral will give the same answer for correlation functions if we rescale the fields. That is, the normalization of the path integral drops out of any physical calculation.

Let us add a 0 to the fields to denote their *bare* (unrenormalized) values:

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}^2 + \bar{\psi}^0 (i\not{\partial} + e^0\, A_\mu^0 + m^0)\psi^0 + \rho \tag{19.70}$$

We saw that Green's functions of these bare fields are ambiguous up to normalization

$$\left\langle 0 | T\{\psi^0 \psi^0\} | 0 \right\rangle = \frac{i\,Z_2}{\not{q} - m_R} \tag{19.71}$$

So if we rescale the fields by

$$\psi^0 = \sqrt{Z_2}\,\psi^R \tag{19.72}$$

$$A_\mu^0 = \sqrt{Z_3} A_\mu^R \tag{19.73}$$

then

$$\left\langle 0 | T\{\psi^R \psi^R\} | 0 \right\rangle = \frac{i}{\not{q} - m_R} \tag{19.74}$$

which is nice.

The Lagrangian becomes

$$\mathcal{L} = -\frac{1}{4} Z_3 F_{\mu\nu}^2 + i\,Z_2 \bar{\psi}_R \not{\partial} \psi_R + Z_2 m_0 \bar{\psi}_R \psi_R + e_0 Z_2 \sqrt{Z_3}\; \bar{\psi}_R \not{A}_R \psi_R + \rho_R \tag{19.75}$$

So we see that QED has 5 terms in it's Lagrangian and 5 parameters: m_0, e_0, Z_2, Z_3 and ρ. All of these are formally infinite, and we can absorb additional infinities into any of them as required. The vacuum energy ρ is largely external to QED, so we will just drop it from now on. The important point is that all we have are these 5 parameters, and they must be sufficient to absorb every infinity. There are many more than 5 correlation functions we can compute!

Also, it is conventional to define

$$Z_1 \equiv \frac{e_0}{e_R} Z_2 \sqrt{Z_3} \tag{19.76}$$

and to use Z_1 instead of e_0. Then (dropping the subscript R on fields)

$$\mathcal{L} = -\frac{1}{4} Z_3 F_{\mu\nu}^2 + i Z_2 \bar{\psi} \not{\partial} \psi + Z_2 m_0 \bar{\psi}\psi + e_R Z_1 \bar{\psi} \not{A} \psi \tag{19.77}$$

At leading order, we know what all the constants are: $Z_3 = Z_2 = 1$ and the mass and charge are determined by their physical, renormalized, values $m_0 = m_R$ and $e_0 = e_R$. The 0 subscripts are called *bare* parameters, and the R subscript is for renormalized. At next-to-leading order, these constants are all infinite. But if we do our perturbation theory expansion in powers of e_R, this is a calculable series, and physical quantities will come out finite.

So we can expand perturbatively

$$Z_3 = 1 + \delta_3 \tag{19.78}$$

$$Z_2 = 1 + \delta_2 \tag{19.79}$$

$$Z_1 = 1 + \delta_1 = 1 + \delta_e + \delta_2 + \frac{1}{2}\delta_3 \tag{19.80}$$

$$Z_2 m_0 = m_R + \delta_m \tag{19.81}$$

where all the $\delta' s$ start at $O(e^2)$. Then the Lagrangian is

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}^2 + i\,\bar{\psi}\not{\partial}\psi - m_R \bar{\psi}\psi + e_R \bar{\psi}\not{A}\psi \tag{19.82}$$

$$-\frac{1}{4}\delta_3 F_{\mu\nu}^2 + i\delta_2 \bar{\psi}\not{\partial}\psi - \delta_m \bar{\psi}\psi + e_R \delta_1 \bar{\psi}\not{A}\psi \tag{19.83}$$

A nice feature of this approach is that since the $\delta' s$ are all infinite, but higher order in powers of e_R they can be simply used as vertices in the Lagrangian. These counterterms subtract off the infinities directly, leaving finite quantities. Moreover, we can do the perturbation expansion directly in terms of e_R instead of e_0, which is nice because e_0 is infinite.

Note that the counterterms δ must be numbers (or functions of e_R and m_R) – they cannot depend on derivatives or factors of momentum. We assumed this when we commuted the factors of Z through everything. If they did depend on momentum, that would correspond to the renormalization of a different tree level operator. For example, if $\delta_m = p^2$ then that would have renormalized $\bar{\psi}\Box\psi$, which is not what we started with. You can have counterterms like this, but then you need to start from a different theory than QED (*i.e.* one whose Lagrangian has a term like $\bar{\psi}\Box\psi$). We will come back to this point when we discuss renormalizable vs non-renormalizable theories.

The original Lagrangian Eq (19.70) is considered bare and the one with the $Z' s$ Eq (19.75) considered renormalized. That is, in the bare Lagrangian the fields are infinite. In the renormalized Lagrangian, the fields are finite and the infinities are all in the counterterms. Sometimes we drop the subscript R.

19.5.1 2-point function in renormalized perturbation theory

Let's compare the calculation of the two point function in bare perturbation theory and in renormalized perturbation theory.

In renormalized perturbation theory, the Feynman diagrams give

$$\langle 0|T\{\psi^R \psi^R\}|0\rangle = \text{—} + \quad = \frac{i}{\not{q} - m_R} + \frac{i}{\not{q} - m_R}[-i\Sigma_2(\not{q})]\frac{i}{\not{q} - m_R} \tag{19.84}$$

and there is also a counterterm graph

$$\quad = \frac{i}{\not{q} - m_R} i(\not{q}\delta_2 - \delta_m)\frac{i}{\not{q} - m_R} \tag{19.85}$$

So,

$$\langle 0|T\{\psi^R\psi^R\}|0\rangle = \frac{i}{\not q - m_R} + \frac{i}{\not q - m_R}\frac{\Sigma_2(\not q) - \not q\delta_2 + \delta_m}{\not q - m_R} \tag{19.86}$$

The way we did it originally (bare perturbation theory),

$$\langle 0|T\{\psi^R\psi^R\}|0\rangle = \frac{1}{Z_2}\langle 0|T\{\psi^0\psi^0\}|0\rangle \tag{19.87}$$

and

$$\langle 0|T\{\psi^0\psi^0\}|0\rangle = \text{—} + \quad = \frac{i}{\not q - m_0} + \frac{i}{\not q - m_0}[-i\Sigma_2(\not q)]\frac{i}{\not q - m_0} \tag{19.88}$$

Then expanding $m_0 = m_R + \delta_m + m_R\delta_2$ and $Z_2 = 1 + \delta_2$

$$\langle 0|T\{\psi^R\psi^R\}|0\rangle = \frac{1}{1+\delta_2}\frac{i}{\not q - m_R + \delta_m + m_R\delta_2} + \frac{i}{\not q - m_R}[-i\Sigma_2(\not q)]\frac{i}{\not q - m_R} \tag{19.89}$$

$$= \frac{i}{\not q - m_R} + \frac{i}{\not q - m_R}\frac{\Sigma_2(\not q) - \not q\delta_2 + \delta_m}{\not q - m_R} \tag{19.90}$$

A note on signs. Let's double check the signs here: in bare perturbation theory the factor $\frac{1}{Z_2} = \frac{1}{1+\delta_2}$ leads to $-\delta_2\frac{i\not q}{i\not q} = +\delta_2 i\not q\frac{i}{\not q}$. In renormalized perturbation theory, the kinetic term gives $Z_2 i\not\partial = (1+\delta_2)\not q$, which leads to an $\delta_2 i\not q$ vertex, and to the same correction $+\delta_2 i\not q\frac{i}{\not q}$. This works for the 2-point function, which is a special case – it is the thing we amputate

In general, with an amputated amplitude, we use

$$\langle 0|T\{\psi^R\cdots\psi^R\}|0\rangle^{\text{amputated}} = Z_2^n\langle 0|T\{\psi^0\cdots\psi^0\}|0\rangle^{\text{amputated}} \tag{19.91}$$

Then the Z factors on the right are all that's left over. For example, for an amputated three point function $\langle 0|T\{A_\mu^R\bar\psi^R\psi^R\}|0\rangle = \sqrt{Z_3}Z_2\langle 0|T\{A_\mu^0\bar\psi^0\psi^0\}|0\rangle$. Recalling that $Z_1 = \sqrt{Z_3}Z_2$ we see we will get the same counterterms automatically using the two methods. Thus they are really two ways of saying the same thing.

The difference is

- Renormalized perturbation theory: all counterterms come from the Lagrangian
- Bare perturbation theory: field Z-factor counterterms come from external states in correlation functions, mass/charge/interaction counterterms comes from replacing bare parameters by their physical values.

The renormalized perturbation theory approach is simpler in many ways, but we will go back and forth a bit to check consistency.

We have already found reasonable conditions to fix δ_2 and δ_m, so let's continue our study of correlation function to fix δ_1 and δ_3.

19.6 Photon self-energy 2-point function: vacuum polarization

The other non-vanishing 2-point function in QED is the photon propagator

$$\Pi^{\mu\nu}(p) = \langle 0|T\{A_R^\mu A_R^\nu\}|0\rangle \tag{19.92}$$

We already calculated the vacuum polarization graph

$$\sim\sim\sim\sim\sim\sim\sim\sim = -i\,(g^{\mu\nu} - \frac{p^\mu p^\nu}{p^2})\Pi_2(p^2)$$

with

$$\Pi_2(p^2) = -8\frac{e^2}{(4\pi)^{d/2}}\Gamma(2-\frac{d}{2})\mu^{4-d}\int_0^1 dx\, x(1-x)\Big(\frac{1}{m_e^2 - p^2x(1-x)}\Big)^{2-\frac{d}{2}} \tag{19.93}$$

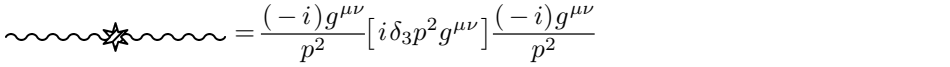
$$= -\frac{2\alpha}{3\pi}\frac{1}{\varepsilon} - \frac{2\alpha}{\pi}\int_0^1 dx\, x(1-x)\log\Big(\frac{\tilde{\mu}^2}{m_e^2 - p^2x(1-x)}\Big) \tag{19.94}$$

Actually, we only computed the $p^2 g^{\mu\nu}$ part of the expression. Peskin and Schroeder compute the rest. In any case, we could have guessed the form by the Ward identity. Note that

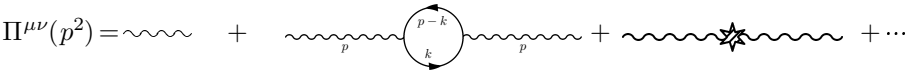
$$(p^\mu p^\nu - p^2 g^{\mu\nu})^2 = p^2(p^\mu p^\nu - p^2 g^{\mu\nu}) \tag{19.95}$$

So that this is a projector. For simplicity, we will just drop the $p^\mu p^\nu$ terms, since we can always put them back in the end by asking that the Ward identity be restored.

Before, we had claimed we claimed we could absorb infinities of the vacuum polarization graph into the renormalized electric coupling e_R^2. That was for Coulomb scattering, where the photon was an internal line. Here, this is just the 2 point function of the photon. For the 2-point function, the leading order piece is not proportional to e_R^2, thus we cannot absorb the infinity into e_R. However, since the Green's function does involve the external states, we do have the Z_3 counterterm to play with. It gives:



Now resumming the series to all orders, including the other 1PI graphs gives



$$= -i\frac{g^{\mu\nu}}{p^2(1-\Pi(p^2))} \tag{19.96}$$

with

$$\Pi(p^2) = \Pi_2(p^2) + \delta_3 + \cdots \tag{19.97}$$

We have dropped gauge dependent terms. (You can actually treat the gauge as a parameter in the Lagrangian, introduce a counterterm for it, and calculate radiative corrections. This is silly in practice, as they always drop out of physical calculations. But it is useful for proving the renormalizability of gauge theories.)

Note that the photon automatically has a pole at $p=0$. This is because the function vacuum polarization loop is proportional to p^2. Thus, the only condition we require is that the residue of the pole at $p=0$ be $-ig^{\mu\nu}$. Which means

$$\boxed{\Pi(0) = 0} \tag{19.98}$$

At order e_R^2, this implies

$$\delta_3 = -\Pi_2(0) = \frac{4}{3}\frac{e^2}{(4\pi)^{d/2}}\Gamma(2-\frac{d}{2})\Big(\frac{\mu}{m_e}\Big)^{4-d} \tag{19.99}$$

$$= \frac{2\alpha}{3\pi}\frac{1}{\varepsilon} + \frac{\alpha}{3\pi}\log\frac{\tilde{\mu}^2}{m_e^2} \tag{19.100}$$

So then the complete Π-function is

$$\Pi(p^2) = -\frac{2\alpha}{\pi}\int_0^1 dx\, x(1-x)\log\Big(\frac{m_e^2}{m_e^2 - p^2x(1-x)}\Big) \tag{19.101}$$

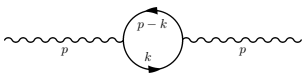
which is totally finite (and also $\tilde{\mu}$ independent!).

Although there is no infrared divergences here, there would be if we took $m_e \to 0$. So this is very similar to the infrared divergence in the electron self-energy diagram.

19.6.1 gauge invariance and counterterms

For the electron propagator, we needed two conditions, one to fix the location of the pole, and the other to fix its residue. For the photon, the location is fixed by gauge invariance, so we only need one condition. That is especially convenient because there is only one free parameter, the normalization Z_3 for the A_μ field.

It would have been possible, *a priori*, that the loop might have given



$$= -i\,(p^\mu p^\nu - p^2 g^{\mu\nu}) g^{\mu\nu}\Pi_2(p^2) + M^2\Pi_M(p^2)$$

With the additional term proportional to some dimensionful quantity M (presumably proportional to the electron mass). This would have led to

$$\Pi^{\mu\nu} = -i\frac{g^{\mu\nu}}{p^2(1-\Pi(p^2)) + M^2\Pi_M(p^2)} \tag{19.102}$$

Then we would have needed a counterterm so that we could renormalize the photon mass back to it's physical location. However, there is no such counterterm available. To get one, we would have needed

$$\mathcal{L} = -\frac{1}{4}Z_3 F_{\mu\nu}^2 + Z_3 m_\gamma^2 A_\mu^2 \tag{19.103}$$

Which allows for the counterterm to appear in the redefinition of m_γ. Instead, no M^2 term is generated, and not coincidentally this addition to the Lagrangian is forbidden by gauge invariance. We are seeing a very important and general phenomenon: the connection between symmetries and Feynman diagrams. We will be coming back to this theme again and again.

It is also worth pointing out that the generation of a photon mass is something that you have to be very careful about. If we didn't use dimensional regularization, we might have ended up with a photon mass. Dimensional regularization is nice because you can define gauge invariance in any dimension. In contrast with a Pauli-Villars regulator, or a hard cutoff, the explicit scale destroys gauge invariance and in fact a spurious M^2 term does appear.

19.7 3-point functions

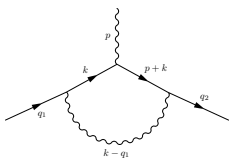
So much for 2-point functions. We have made everything finite. Now let's look at 3-point functions. The simplest one is simply

$$G_3 = \langle 0|T\{\bar{\psi}(q_1)A_\mu(p)\psi(q_2)\}|0\rangle \tag{19.104}$$

This is non-zero at tree level due to the QED vertex.

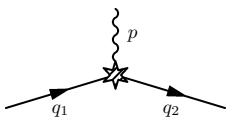
$$G_3 = ie\gamma^\mu \tag{19.105}$$

At next order, we need to compute a loop



$$= \Gamma_2^\mu = ie_R\left(F_1(p^2)\gamma^\mu + \frac{i\sigma^{\mu\nu}}{2m_e}p_\nu F_2(p^2)\right) \tag{19.106}$$

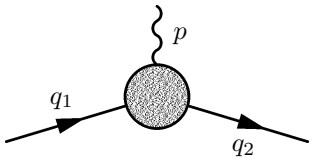
There is also the counterterm contribution to the vertex correction,



$$= ie_R\,\delta_1\gamma^\mu \tag{19.107}$$

which we can use to absorb the infinities in $F_1(p^2)$.

The decomposition into form factors just relied on Lorentz invariance, so it holds even non-perturbatively. Thus we can write, in general



$$= \Gamma^\mu(p) = ie_R\left(F_1(p^2)\gamma^\mu + \frac{i\sigma^{\mu\nu}}{2m_e}p_\nu F_2(p^2) \right) \tag{19.108}$$

where the blob refers to all amputated vertex corrections. Only amputated corrections contribute because the corrections to the external lines get absorbed in the external line propagators, and then truncated by LSZ.

At order e_R^2, $F_1(p^2) = 1$ and $F_2(p^2) = 0$. In fact, through our study of g-2, we observed that $F_2(0)$ contributes to the magnetic moment of the electron. $F_2(0)$ is finite, and at the one loop level, we found

$$F_2(p^2) = \alpha \int_0^1 dz \int_0^{1-z} dy \frac{z(z-1)m_R^2}{(1-z)^2 m_R^2 - xyp^2} \tag{19.109}$$

F_1 on the other hand controls the gross interactions of the electron with the photon. It is the part that leads to Coulomb's law. Thus the electric charge measured in Coulomb's law is proportional to $e_R F_1(0)$. So we have to set $F_1(0) = 1$. That is, we demand

$$\Gamma^\mu(0) = ie_R\gamma^\mu \tag{19.110}$$

This is the final renormalization condition. It implies that the renormalized electric charge is what is measured by Coulomb's law, and *by definition*, does not get radiative corrections.

F_1 is UV and IR divergent. In dim-reg, with a photon mass, it comes out to

$$F_1(p^2) = \frac{e_R^2}{(4\pi)^{d/2}}\mu^{4-d}\int_0^1 dxdydz\delta(x+y+z-1)\left\{ \frac{\Gamma(2-\frac{d}{2})(2-\varepsilon)^2}{\Delta^{2-d/2}} \right. \tag{19.111}$$

$$\left. + \frac{\Gamma(3-\frac{d}{2})}{\Delta^{3-d/2}}\big(p^2[2(1-x)(1-y) - \varepsilon xy] + m_R^2[2(1-4z+z^2) - \varepsilon(1-z)^2]\big) \right\} \tag{19.112}$$

where

$$\Delta = (1-z)^2 m_R^2 + z m_\gamma^2 - xyp^2 \tag{19.113}$$

Expanding in $d = 4-\varepsilon$, we get

$$F_1(p^2) = \frac{e_R^2}{16\pi^2} \times \tag{19.114}$$

$$\left(\frac{2}{\varepsilon} + \int_0^1 dxdydz\delta(x+y+z-1)\left[\frac{2m_R^2(1-4z+z^2) + 2p^2(1-x)(1-y)}{\Delta} + 2\log\frac{\tilde{\mu}^2}{\Delta} \right] \right) \tag{19.115}$$

At $p = 0$, we can perform the x and y integrals giving (after some algebra)

$$F_1(0) = \frac{e_R^2}{16\pi^2}\left[\frac{4}{\varepsilon} + 1 + \log\frac{\tilde{\mu}^2}{m_R^2} - 4\int dx \frac{2-x}{1-x} x \frac{(1-x)^2 m_R^2}{(1-x)^2 m_R^2 + x m_\gamma^2} \right], \tag{19.116}$$

At order e_R^2 the renormalization condition $F_1(0) = 1$ implies

$$\delta_1 = -F_1(0) \tag{19.117}$$

In fact, comparing with Equation (?), we find

$$\delta_1 = \delta_2 \tag{19.118}$$

To order e_R^2. This is a very important relation, that we will discuss a bit more shortly.

Keep in mind that even though $\delta_1 = \delta_2$, we still need both counterterms to absorb all the infinities!

With a Pauli-Villars regulator, and a photon mass

$$F_1(p^2) = \frac{\alpha}{2\pi}\int_0^1 dx\,dy\,dz\,\delta(x+y+z-1)\left\{\log\frac{z\Lambda^2}{\Delta}+\frac{1}{\Delta}\left[(1-x)(1-y)p^2+(1-4z+z^2)m^2\right]\right\}$$

with the same Δ as above. F_2 is the same in dim-reg as in PV as it is UV finite (and IR finite).

19.8 QED renormalization conditions

To summarize, a set of renormalization conditions in QED which is sufficient to fix all the counterterms is

$$\Sigma(\not{q})\Big|_{\not{q}=m_R} = 0 \tag{19.119}$$

$$\frac{d}{d\not{q}}\Sigma(\not{q})\Big|_{\not{q}=m_R} = 0 \tag{19.120}$$

$$\Gamma^\mu(p)\Big|_{p=0} = \gamma^\mu \tag{19.121}$$

$$\Pi(p)\Big|_{p=0} = 0 \tag{19.122}$$

where $\Sigma(\not{q})$ are the 1PI electron self energy corrections, Γ^μ are the amputated corrections to the QED vertex, and $\Pi(p)$ are the 1PI vacuum polarization corrections.

The first two conditions fix the electron propagator to

$$\langle 0|T\{\bar\psi\psi\}|0\rangle = \frac{i}{\not{q}-m_R} + \text{regular at } \not{q}=m_R \tag{19.123}$$

the second condition fixes the renormalized electric charge e_R to be what is measured by Coulomb's law at large distance. The final condition forces the photon propagator to be

$$\Pi_\gamma^{\mu\nu}(p) = \frac{-g^{\mu\nu}}{p^2} + \text{regular at } p=0 + \text{gauge}-\text{dependent piesces} \tag{19.124}$$

19.9 Minimal Subtraction

We have discussed one set of rather physical renormalization conditions which were enough to fix the counterterms to all orders. There is another way to fix these counterterms, which is not as physical, but turns out to be much more useful in practice for complicated calculations. The prescription is to remove just the infinities. This is appropriately called *minimal subtraction.*

The convention for the Lagrangian with minimal subtraction is slightly different. We write

$$\mathcal{L} = -\frac{1}{4}Z_A F_{\mu\nu}^2 + i\,Z_\psi\bar\psi\not{\partial}\psi + m_R Z_\psi Z_m\bar\psi\psi + e_R Z_e\sqrt{Z_A}Z_\psi\bar\psi\not{A}\psi \tag{19.125}$$

$$Z_A = 1+\delta_A \tag{19.126}$$

$$Z_\psi = 1+\delta_\psi \tag{19.127}$$

$$Z_e = 1+\delta_e \tag{19.128}$$

$$Z_m = 1+\delta_m \tag{19.129}$$

The electron self energy graph plus counterterms gives, with Pauli-Villars

$$\Sigma(\not{q}) = \frac{e^2}{16\pi}\int dx(4m_R - 2\not{q}x)\log\frac{x\Lambda^2}{(1-x)(m_R^2 - x\,q^2)} + \not{q}\delta_\psi - m_R(\delta_m+\delta_\psi) \tag{19.130}$$

The infinity is the $\log\Lambda$ term. So we choose

$$\delta_\psi = \frac{e^2}{16\pi}\int dx(2x)\log\frac{\Lambda^2}{\mu^2} = \frac{e^2}{16\pi}\Big[\log\frac{\Lambda^2}{\mu^2}\Big], \quad (\text{PV}, \overline{\text{MS}}) \tag{19.131}$$

$$\delta_m = \frac{e^2}{16\pi}4\log\frac{\Lambda^2}{\mu^2} - \delta_\psi = \frac{e^2}{16\pi}\Big[3\log\frac{\Lambda^2}{\mu^2}\Big], \quad (\text{PV}, \overline{\text{MS}}) \tag{19.132}$$

We have introduced a scale μ to keep this dimensionally correct. μ can be anything (e.g. $\mu = m_R$) – it is not infinite like Λ.

We then get

$$\Sigma(\not{q}) = \frac{e^2}{16\pi}\int dx(4m_R - 2\not{q}x)\log\frac{x\mu^2}{(1-x)(m_R^2 - xq^2)} \tag{19.133}$$

One nice thing about this scheme is that there are never infrared divergences in the counterterms, only in the finite remainder for $\Sigma(\not{q})$.

Actually, minimal subtraction is hardly ever used with Pauli-Villars. In fact, Pauli-Villars is hardly ever used. It's really not so useful, because it breaks gauge invariance. In dimensional regularization,

$$\Sigma(\not{q}) = \frac{e_R^2}{16\pi^2}\Big[\frac{-8m_R + 2\not{q}}{\varepsilon} + 2(\not{q} - m_R) + \int dx(4m_R - 2\not{q}x)\log\frac{4\pi e^{-\gamma_E}\mu^2}{(1-x)(m_R^2 - xq^2)}\Big] + \not{q}\delta_\psi - m_R(\delta_m + \delta_\psi) \tag{19.134}$$

So now we take

$$\delta_\psi = \frac{\alpha}{4\pi}\Big[-\frac{2}{\varepsilon}\Big] \tag{19.135}$$

$$\delta_m = \frac{\alpha}{4\pi}\Big[-\frac{8}{\varepsilon}\Big] - \delta_\psi = \frac{\alpha}{4\pi}\Big[-\frac{6}{\varepsilon}\Big] \tag{19.136}$$

leaving

$$\Sigma(\not{q}) = \frac{e_R^2}{16\pi^2}\Big[2(\not{q} - m_R) + \int dx(4m_R - 2\not{q}x)\log\frac{4\pi e^{-\gamma_E}\mu^2}{(1-x)(m_R^2 - xq^2)}\Big]$$

So you see this scheme is extraordinarily simple.

The scheme is used almost exclusively in complicated calculations. It is often convenient to also subtract off the $4\pi e^{-\gamma_E}$ in the counterterm, by choosing

$$\delta_\psi = \frac{\alpha}{4\pi}\Big[-\frac{2}{\varepsilon} + \log(4\pi e^{-\gamma_E})\Big] \tag{19.137}$$

$$\delta_m = \frac{\alpha}{4\pi}\Big[-\frac{6}{\varepsilon} - 2\log(4\pi e^{-\gamma_E})\Big] \tag{19.138}$$

Then

$$\Sigma(\not{q}) = \frac{e_R^2}{16\pi^2}\Big[2(\not{q} - m_R) + \int dx(4m_R - 2\not{q}x)\log\frac{\mu^2}{(1-x)(m_R^2 - xq^2)}\Big]$$

This is known as $\overline{\text{MS}}$ (em-ess-bar). It is equivalent to just using MS counterterms but replacing $\mu^2 \to \tilde{\mu}^2 = 4\pi e^{-\gamma_E}\mu^2$, which is totally fine since μ was arbitrary anyway.

For completeness, we can also compute from the vacuum polarization and the vertex correction graphs,

$$\delta_A = \frac{e_R^2}{16\pi^2}\Big[-\frac{8}{3\epsilon}\Big] \tag{19.139}$$

$$\delta_e = \frac{e_R^2}{16\pi}\Big[\frac{4}{3\varepsilon}\Big] \tag{19.140}$$

In terms of the old variables $Z_1 = Z_e\sqrt{Z_A}Z_\psi$ and so.

$$\delta_1 = \delta_e + \frac{\delta_A}{2} + \delta_\psi = \delta_\psi = \delta_2 \tag{19.141}$$

Which we showed using the other subtraction scheme.

An important point, which is often confused, is that there are two scales involved: the infinite UV scale Λ, which is taken to infinity, and a finite low-energy scale μ. Both PV and DR have these two scales: in PV μ is obscure, and in DR Λ is obscure. But they are both there. μ is completely arbitrary, and the physical calculations should not depend on it. However, $\Lambda = \infty$ so the calculation *cannot* depend on Λ. As we will see, at each order in perturbation theory, there can be some residual dependence on μ, but as we go to higher orders in perturbation theory, the dependence on μ should disappear. In fact, by demanding that the μ dependence disappear, one is lead to the renormalization group.

19.9.1 propagator in $\overline{\text{MS}}$

With minimal subtraction, the propagator is no longer guaranteed to have a pole at $\not{p} = m_R$, and the residue is not guaranteed to be 1. We are still doing an expansion in physical parameters, and the electron propagator is still

$$\Pi_e = \frac{1}{\not{p} - m_R + \Sigma(\not{p})} \tag{19.142}$$

but we no longer have $\Sigma(m_R) = 0$. Explicitly,

$$\Sigma(m_R) = \frac{e_R^2}{16\pi^2} m_R \left(5 + 3\log\frac{\mu^2}{m_R^2} \right) \tag{19.143}$$

In fact what we mean by the renormalized mass m_R is not the location of the pole, but the term in the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + i\,\bar\psi\not\partial\psi + m_R\bar\psi\psi + e_R\bar\psi\not A\psi + \text{counterterms} \tag{19.144}$$

This is called the $\overline{\text{MS}}$ mass. Let us call the pole mass m_P. Then m_P satisfies

$$0 = m_P - m_R - \Sigma(m_P) \tag{19.145}$$

$$m_P = m_R - \frac{e_R^2}{16\pi^2}\left[2(m_P - m_R) + \int dx (4m_R - 2m_P x)\log\frac{\mu^2}{(1-x)(m_R^2 - x m_P^2)} \right] \tag{19.146}$$

This equation has some solution which depends on μ. If we choose $\mu = m_R$ it leads to a calculable shift in the pole mass to first order in perturbation theory.

There is a very good reason that $\overline{\text{MS}}$ masses are better than pole masses – many fermions, such as quarks, are not free. So they do not correspond to asymptotic states, and thus talking about a free propagator for these fields does not correspond to anything physical. This is a particularly troublesome confusion for the top quark mass, where the difference between the pole mass and $\overline{\text{MS}}$ mass definitions differ by a few GeV.

19.10 $\delta_1 = \delta_2$: implications and proof

The relation $\delta_1 = \delta_2$ is a very important result of QED, so it is worth a discussion. $\delta_1 = \delta_2$ implies $Z_1 = Z_2$. Recalling that $e_R Z_1 = e_0 Z_2\sqrt{Z_3}$, it follows that

$$e_R = e_0\sqrt{Z_3} \tag{19.147}$$

In other words $\delta_e = -\frac{1}{2}\delta_3$. That is, the renormalization of the electric charge is determined completely by the renormalization of the photon field strength. That is why we were able to calculate the renormalization of the electric charge from only the vacuum polarization graphs.

There is an important physical implication of $Z_1 = Z_2$. For example, suppose we have a theory with two different kinds of particles say a quark of charge $\frac{1}{3}$ and an electron of charge 1. Then the Lagrangian is

$$\mathcal{L} = -\frac{1}{4}Z_3 F_{\mu\nu}^2 + i\,Z_{2e}\bar{\psi}_e \not{\partial}\psi + e_R Z_{1e}\bar{\psi}_e \not{A}\psi_e + i\,Z_{2q}\bar{\psi}_q \not{\partial}\psi_q + \frac{1}{3}e_R Z_{1q}\bar{\psi}\not{A}\psi_q \tag{19.148}$$

$$= -\frac{1}{4}Z_3 F_{\mu\nu}^2 + Z_{2e}\bar{\psi}_e(i\not{\partial} + e_R\not{A})\psi_e + Z_{2q}\bar{\psi}_q(i\not{\partial} + \frac{1}{3}e_R\not{A})\psi_q \tag{19.149}$$

where $Z_{1e} = Z_{2e}$ and $Z_{1e} = Z_{2q}$ have been used (but only shown at 1-loop). This rescaling implies the relationship between the coefficient of $i\not{\partial}$ and of $e_R\not{A}$ is not corrected at 1-loop. In other words, the relative charges of the electron and the quark are not corrected.

This is pretty remarkable. It explains why the observed charge of the proton and the charge of the electron can be exactly opposite, even in the presence of vastly different interactions of the two particles. *A priori*, we might have suspected that because of strong interactions and virtual mesons surrounding the proton, the types of radiative corrections for the proton would be vastly more complicated than for the electron. But, as it turns out, this does not happen – it is the renormalization of the photon field strength rescales the electric charge, but the corrections to the relative charges of the proton the electron cancel.

To see that it works to all orders, first rescale $A_\mu \to \frac{1}{e_R}A_\mu$. Then the Lagrangian becomes

$$\mathcal{L} = -\frac{1}{4e_R^2}Z_3 F_{\mu\nu}^2 + Z_{2e}\bar{\psi}_e(i\not{\partial} + \frac{Z_{1e}}{Z_{2e}}\not{A})\psi_e + Z_{2q}\bar{\psi}_q(i\not{\partial} + \frac{Z_{1q}}{Z_{2q}}\frac{1}{3}\not{A})\psi_q \tag{19.150}$$

At tree level, with $Z_i = 1$ and the Lagrangian is invariant under

$$\psi_q \to e^{\frac{1}{3}i\alpha}\psi_q \tag{19.151}$$

$$\psi_e \to e^{i\alpha}\psi_e \tag{19.152}$$

$$A_\mu \to A_\mu + \partial_\mu\alpha \tag{19.153}$$

Observe that the charges (1 and $\frac{1}{3}$) appear in the transformation law but e_R does not. Also, the transformation has nothing to do with perturbation theory. Since this symmetry holds in the tree-level theory, and the regulator preserves gauge invariance, then the loop corrections should also preserve the gauge symmetry, and therfore the counterterms respect the symmetry too. In other words, the form of the covariant derivatives $\not{D}\psi$ is fixed by gauge invariance. Thus $Z_1 = Z_2$ to all orders.

We will now show that this relation implies that the current $\bar{\psi}\gamma^\mu\psi$ does not get renormalized, which is an important, general, and useful result. And, for completeness, we will derive $Z_1 = Z_2$ from the generalized Ward-Takahashi identity.

19.10.1 charge is not renormalized

The classical electromagnetic current density is

$$j_\mu(x) = \bar{\psi}\,(x)\gamma_\mu\psi(x) \tag{19.154}$$

This couples to the photon as $e\,A_\mu j^\mu$. The 0 component of e times the current j_0, integrated over all space times, gives the total charge.

$$Q^{\rm EM} = e\int d^3x\, j_0 = -\,e\int d^3x\psi^\dagger\psi = \text{total charge} \tag{19.155}$$

It does not change with time, since there the current vanishes at infinity and

$$\partial_t Q = e\int d^3x\,\partial_0 j_0 = e\int d^3x\partial_i j_i = e\,j_i(\infty) = 0 \tag{19.156}$$

The 0-component of the current itself simply gives the number of particles minus the number of anti-particles. This is also called lepton number.

$$Q = \int d^3x\, j_0 = \#\text{particles} - \#\text{antipartles} \equiv \text{lepton number} \tag{19.157}$$

The charge is conserved.

As a check, we can calculate that for the free fields

$$\psi(x)Q = \int d^3y \psi(x)\psi^\dagger(y)\psi(y) = Q\psi(x) + \int d^3y \delta^3(x-y)\psi(y) = Q\psi(x) + \psi(x) \tag{19.158}$$

$$Q\psi^\dagger(x) = \int d^3y \psi^\dagger(y)\psi(y)\psi^\dagger(x) = \psi^\dagger(x)Q + \int d^3y \delta^3(x-y)\psi^\dagger(y) = \psi^\dagger(x)Q + \psi^\dagger(x) \tag{19.159}$$

So,

$$[Q\,,\psi] = -\psi \tag{19.160}$$

$$[Q\,,\psi^\dagger] = \psi^\dagger \tag{19.161}$$

That is, Q generates particle number. Since the current is conserved non-perturbatively, this should be true to all orders in perturbation theory. Let us check.

Now consider $j_\mu(x) = \bar\psi\,(x)\gamma^\mu\psi(x)$ as an operator in the theory in its own right. It is an example of a *composite operator* – ψ and $\bar\psi$ are taken at the same point. We can calculate Green's functions of j_μ just like anything else. The leading non-vanishing amputated Green's function for this operator is

$$\big\langle 0|T\{j_\mu(p)\bar\psi\,(q_1)\psi(q_2)\}|0\big\rangle = \big\langle 0|T\{(\bar\psi\gamma_\mu\psi)(p)\bar\psi\,(q_1)\psi(q_2)\}|0\big\rangle = \gamma_\mu \tag{19.162}$$

There is only one relevant one-loop graph, whose divergent part we already know.

$$= -\delta_1\gamma^\mu + \text{finite} \tag{19.163}$$

But now we do not have the δ_1 counterterm to cancel this δ_1 infinity.

To cancel this infinity, we have to allow that the current may be renormalized. So we write

$$j^R_\mu = Z_j j^0_\mu \tag{19.164}$$

with

$$Z_j = 1 + \delta_j \tag{19.165}$$

It is critical that we allow for the composite operator to be renormalized separately from its constituent fields.

Let's compute δ_j first in bare perturbation theory, then in renormalized perturbation theory. In bare perturbation theory, the Z factors compensate for the renormalization of the external states in the amputated amplitude:

$$\big\langle 0|T\{j^R_\mu(p)\bar\psi^{\,R}(q_1)\psi^R(q_2)\}|0\big\rangle = \frac{1}{Z_j}Z_2\big\langle 0|T\{j_\mu(p)\bar\psi\,(q_1)\psi(q_2)\}|0\big\rangle \tag{19.166}$$

Now we add the counterterm contributions from the $Z_j^{-1}Z_2$ factor in front

$$\text{ct} = (-\delta_j + \delta_2)\gamma^\mu \tag{19.167}$$

So the sum is $-\delta_1 - \delta_j + \delta_2$. But since $\delta_1 = \delta_2$, all the divergences have canceled in $\delta_2 = \delta_1$, we can simply set $\delta_j = 0$. In other words

The current is not renormalized

In renormalized perturbation theory, we have to work at the level of the Lagrangian. So we add our new operator to the Lagrangian as

$$j_\mu \bar{\psi}\gamma^\mu\psi = Z_2 Z_j j_\mu^R \bar{\psi}^R \gamma^\mu \psi^R \tag{19.168}$$

This leads to a $i(\delta_2 + \delta_j)\gamma^\mu$ vertex as before.

We could have anticipated that the current must be conserved by looking back at the conservation of charge equation. Putting in the renormalized fields.

$$[Z_j Q^R, \psi^R] = -\psi^R \tag{19.169}$$

Thus unless $Z_j = 1$ to all orders, this equation cannot be satisfied.

Note that the current not being renormalized is *not* the same as saying that the divergences in the graphs cancel. That is,

(19.170)

In fact, the sum of these three graphs is divergent. The correct statement is that the fermion field strength renormalization is enough to cancel all the divergences, we do not need to renormalize the current separately. It also does not say we do not need both counterterms δ_1 and δ_2. We do need them both, but they are both equal.

19.10.2 All-orders proof of $\delta_1 = \delta_2$

That $\delta_1 = \delta_2$ to all orders follows from the generalized Ward-Takahashi identity, which was introduced in our discussion of path integrals:

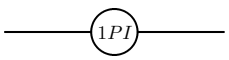
$$p_\mu M^\mu(p,q) = M_0(q+p) - M_0(q) \tag{19.171}$$

Where,

$$M_0(q) = \left\langle 0|T\{\bar{\psi}(q)\psi(-q)\}\right\rangle \tag{19.172}$$

$$M^\mu(p,q) = \left\langle 0|T\{A^\mu(p)\bar{\psi}(q_1)\psi(q_2)\}\right\rangle \tag{19.173}$$

We called the general non-perturbative form of the electron self-energy diagrams $\Sigma(\not{q})$



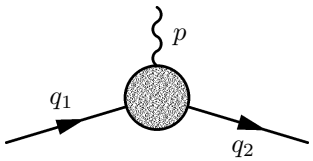
$$i\Sigma(\not{q}) \tag{19.174}$$

So,

$$M_0(q) = \bar{u}(q)\frac{1}{\not{q} - m - \Sigma(\not{q})}u(q) = \bar{u}(q)\Pi(\not{q})u(q) \tag{19.175}$$

where $\Pi(\not{q}) = \frac{1}{\not{q} - m - \Sigma(\not{q})}$ is the full electron propagator.

Similarly, M_μ should sum all contributions to the photon vertex. All contributions are given by all amputated contributions plus external line corrections. The amputated contributions can be written as



$$= \bar{u}(q_1)\Gamma_\mu(p,q_1)u(q_2) \tag{19.176}$$

So,

$$M^\mu(p,q) = \bar{u}(q_1)\Pi(\not{q})\Gamma^\mu(p,q)\Pi(\not{p}+\not{q})u(q_2) \tag{19.177}$$

Thus, writing $q_2 = q_1 + p$, the Ward-Takahashi identity implies

$$p_\mu \Pi(\not{q}_1)\Gamma^\mu(p,q_1)\Pi(\not{q}_1 + \not{p}) = \Pi_2(\not{q}_1 + \not{p}) - \Pi_2(\not{q}_1) \tag{19.178}$$

Dividing by $\Pi(\not{q}_1)\Pi(\not{q}_1+\not{p})$, this implies

$$p_\mu\Gamma^\mu(p,q) = \frac{1}{\Pi(\not{q}+\not{p})} - \frac{1}{\Pi(\not{q})} \tag{19.179}$$

$$= \not{p} + \Sigma(\not{q}+\not{p}) - \Sigma(\not{q}) \tag{19.180}$$

Now recall that the amputated graphs are expressed conventionally in terms of form factors

$$\Gamma^\mu(p,q) = F_1(p^2)\gamma^\mu + \frac{i\sigma^{\mu\nu}}{2m_e}p_\nu F_2(p^2) \tag{19.181}$$

Since the F_2 term has a factor ofp_ν then

$$\lim_{p\to 0} p_\mu\Gamma^\mu(p,q) = F_1(0)\not{p} \tag{19.182}$$

Then taking $p\to 0$ holding q_1 fixed, the Ward-Takahashi identity implies

$$F_1(0) = \lim_{p\to 0}\frac{\not{p}+\Sigma(\not{q}_1+\not{p})-\Sigma(\not{q}_1)}{\not{p}} = 1 + \frac{d}{d\not{q}}\Sigma(\not{q}) \tag{19.183}$$

Putting the electron on shell, and recalling that $F_1(0) = 1+\delta_1$ and $\delta_2 = \frac{d}{d\not{q}}\Sigma(\not{q})\Big|_{\not{q}=m_R}$ this implies

$$\delta_1 = \delta_2 \tag{19.184}$$

to all orders.

We can express this more generally, by looking back at our renormalized Lagrangian

$$\mathcal{L} = -\frac{1}{4}Z_3F_{\mu\nu}^2 + i\,Z_2\bar{\psi}\not{\partial}\psi + Z_2m_0\bar{\psi}\psi + e_RZ_1\bar{\psi}\not{A}\psi \tag{19.185}$$

The electron propagator following from this is

$$\Pi(\not{q}) = \frac{1}{Z_2}\frac{i}{\not{q}-m} \tag{19.186}$$

And the photon vertex is simply $Z_1\bar{\psi}\gamma^\mu\psi$ so $\lim_{p\to 0}\Gamma^\mu(0) = Z_1\gamma^\mu$. Then

$$\lim_{p\to 0} p_\mu\Gamma^\mu(p,q_1) = Z_1\not{p} \tag{19.187}$$

$$= \lim_{p\to 0}\left[\frac{1}{\Pi(\not{q}_1+\not{p})} - \frac{1}{\Pi(\not{q}_1)}\right] = Z_2\not{p} \tag{19.188}$$

Thus $Z_1 = Z_2$ exactly.

Chapter 20
Infrared Divergences

20.1 Introduction

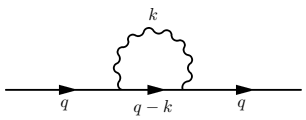
We have shown that the 1, 2 and 3 point functions in QED are UV finite at one loop. We were able to introduce 4 counterterms $(\delta_m, \delta_1, \delta_2, \delta_3)$ which canceled all the infinities. Now let us move on to four point functions, such as

$$\langle 0|T\{\bar{\psi}(p_1)\psi(p_2)\bar{\psi}(p_3)\psi(p_4)\}|0\rangle \tag{20.1}$$

For example, this can represent Moller scattering, or Bhabha scattering. We will take it to be $e^+e^- \to \mu^+\mu^-$ for simplicity (this process only has an s-channel diagram). Looking at these 4-point functions at one-loop will help us understand how to combine previous loop calculations and counterterms into new observables, and will also illustrate a new feature: cancellation of infrared divergences.

Recall that δ_1 and δ_2 depend on the fictitious photon mass, m_γ. This mass was introduced to make the loops finite and is an example of an infrared regulator. In the $\overline{\text{MS}}$ subtraction scheme, none of the counterterms depend on the infrared regulator, but there is still m_γ dependence from the loops. In fact, the existence of infrared singularities is not related to the particular choice of UV regulator or subtraction scheme. Of course, that makes sense as the UV and IR are non-intersecting regions of phase space.

We have seen two graphs which are infrared divergent: the electron self energy graph



$$= \int \frac{d^4k}{(2\pi)^4}\gamma^\mu \frac{i(\not{q}-\not{k}+m)}{(q-k)^2-m^2+i\varepsilon}\gamma^\mu \frac{1}{k^2} \tag{20.2}$$

and the QED vertex correction graph, which is similar. Both blow up as $k \to 0$, These infrared divergences comes from the photon being massless, so we were inspired to add a fictitious photon mass m_γ. Then, the integral becomes

$$\int \frac{d^4k}{(2\pi)^4}\gamma^\mu \frac{i(\not{q}-\not{k}+m)}{(q-k)^2-m^2+i\varepsilon}\gamma^\mu \frac{1}{k^2-m_\gamma^2}$$

which no longer blows up as $k \to 0$.

Infrared divergences are funny things. They occur when there are massless particles in a theory. The electron is not massless, but at high energies it is effectively massless. We would often like to just set $m_e = 0$ but this will often induce an infrared divergence. In fact the running of the electric charge,

$$e^2_{\text{eff}}(p^2) = e^2_R\left(1 + \frac{e^2_R}{12\pi^2}\log\frac{-p^2}{m_e^2}\right) \tag{20.3}$$

is singular as $m_e \to 0$.

However, we can suspect that there is no real problem with $m_e \to 0$ or $m_\gamma \to 0$ by going back to considering physical quantities. For example, if we look at the difference between the Coulomb potential at two scales, we have

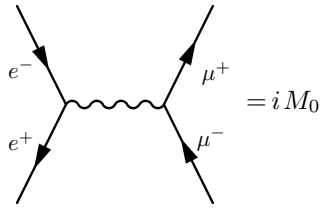
$$p_1^2\tilde{V}(p_1^2) - p_2^2\tilde{V}(p_2^2) = \frac{e^2_R}{12\pi^2}\log\frac{p_1^2}{p_2^2} \tag{20.4}$$

which is in fact independent of m_e. So this observable difference is finite as $m_e \to 0$. The same would be true of the "effective Compton potential", coming from t channel electron exchange in $e^-\gamma \to e^-\gamma$, and it's loop correction, the self-energy contribution Eq. (1). There the difference between Compton scattering matrix elements at two scales would be finite as $m_\gamma \to 0$ for exactly the same reason as the difference between Coulomb scattering matrix elements at two scales is finite as $m_e \to 0$.

The general resolution of these infrared diverges is based on a similar physical principle to the resolution of UV divergences: make sure you are calculating physical quantities. But there is an added complication. Sometimes, instead of just relating the same measurement at different scales, infrared divergences will often only cancel if we add different contributions to the *same* measurement. We will explore this first with an explicit example, and then talk about general features of infrared divergences.

20.2 Total cross section for $e^+e^- \to \mu^+\mu^-$

The simplest 4-point function to consider is for $e^+e^- \to \mu^+\mu^-$



We had found the tree level differential cross section, with $m_e = m_\nu = 0$ is

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4Q^2}\left(1+\cos^2\theta\right) \tag{20.5}$$

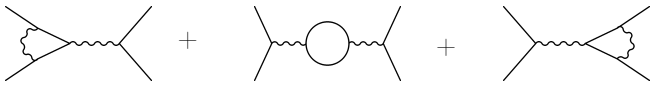
where we are calling the COM energy Q (instead of $\sqrt{s}$)

Then the total cross section is

$$\sigma_0 = \int_0^{2\pi} d\phi \int_{-1}^{1} d\cos\theta \frac{d\sigma}{d\Omega} = \frac{4\pi\alpha^2}{3Q^2} \tag{20.6}$$

We would like to calculate the $\mathcal{O}(\alpha^3)$ correction to this. This is sometimes known as the Born cross section, because it's calculated in the first Born approximation: $\sigma_B = \frac{4\pi\alpha^2}{3Q^2}$.

First, there are the loop graphs. These contribute to the same Green's function at order α^3. Considering only amputated graphs there are 3



(20.7)

The next-to-leading order $\mathcal{O}(\alpha^3)$ result is the interference between these graphs (of order α^2) and the original graph (of order α).

We will break down this calculation into pieces. For example, if the photon only couples to muons, the first two graphs would be irrelevant. Or we can imagine that instead of $e^+e^- \to \mu^+\mu^-$, we simply have a massive vector boson (for example, the Z) decaying to $\mu^+\mu^-$, or even just an off-shell photon, $\gamma^\star$. We can represent these possibilities by, instead of calculating the Green's function

$$G_4 = \left\langle 0|T\{\bar{\psi}_{e^+}(p_1)\psi_{e^-}(p_2)\bar{\psi}_{\mu^+}(p_3)\psi_{\mu^-}(p_4)\}|0\right\rangle \tag{20.8}$$

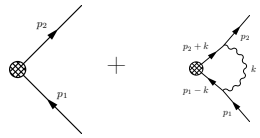
we calculate instead

$$G_J = \left\langle 0|T\{j_\mu(p)\bar{\psi}_{\mu^+}(p_1)\psi_{\mu^-}(p_2)\}|0\right\rangle \tag{20.9}$$

where

$$j_\mu(x) = \bar{\psi}\,(x)\gamma^\mu\psi(x) \tag{20.10}$$

This current differs from the thing $\bar{\psi}\,(x_1)\psi(x_2)$ in the 4-point function because the current is evaluated a single space-time point. We can represent this with an operator insertion

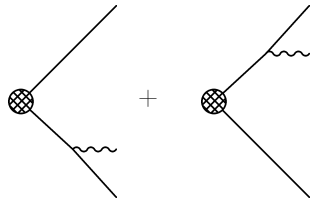


If you find this operator language confusing, don't worry about it. In the end, we will have calculated the full $e^+e^- \to \mu^+\mu^-$. The current language just lets us break the problem into pieces, so we can see what's important at each stage, but if you prefer to think about doing it all at once, go right ahead!

In addition to the loop corrections to the 4-point function, to cancel the infrared divergences, we will need to calculate also real emission graphs. These are the tree level contributions to a 5-point function

$$G_5 = \big\langle 0|T\{\bar{\psi}\psi\bar{\psi}\psi A_\mu\}|0\big\rangle \tag{20.11}$$

Or with the current, simply the current matrix elements G_J with an extra photon. The graphs are



We will do the loops first, then the real emission graphs, and then show that we can take $m_\gamma \to 0$ after then are combined into the full cross section $d\sigma(e^+e^- \to \mu^+\mu^-(+\gamma))$.

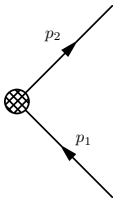
We take $m_e = m_\mu$ for simplicity.

20.3 Loops

20.3.1 Tree level graph

The tree-level graph that we need to interfere with is

$$i\,M_0 = \qquad\qquad = i\,\bar{u}\,(p_2)\gamma^\mu v(p_1) \tag{20.12}$$



Then the normalization is set by

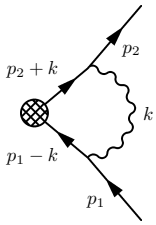
$$\sum_{\text{pols}} |M_0|^2 = -\,\text{Tr}[\not{p}_2\gamma^\mu\not{p}_1\gamma_\mu] = 8p_1p_2 = 4Q^2 \tag{20.13}$$

So writing $\sigma_0 = 4Q^2$ lets us rescale our results to match the full $e^+e^- \to \mu^+\mu^-$.

20.3.2 Vertex Correction

The vertex correction is

$$iM_\Gamma = \quad = i\bar{u}(p_2)\Gamma^\mu v(p_1) \tag{20.14}$$



with

$$\Gamma^\mu = (-i)(i)^2(-ie)^2 \int \frac{d^4k}{(2\pi)^4} \frac{g^{\alpha\beta} - (1-\xi)\frac{k_\alpha k_\beta}{k^2}}{k^2} \gamma^\alpha \frac{\not{p}_2 + \not{k}}{(p_2+k)^2} \gamma^\mu \frac{\not{p}_1 - \not{k}}{(p_1-k)^2} \gamma^\beta \tag{20.15}$$

which is the same amplitude we've already encountered. We wrote it in terms of form factors

$$\Gamma^\mu = F_1(p^2)\gamma^\mu + \frac{i\sigma^{\mu\nu}}{2m_e} p_\nu F_2(p^2) \tag{20.16}$$

where $p = p_1 + p_2$ is the momentum coming into the current vertex. Recall, only F_1 is divergent, and the divergence is given by $F_1(0) = -\delta_1 + \text{finite}$.

The interference with M_0 will give

$$M_\Gamma^\dagger M_0 + M_0^\dagger M_\Gamma = 2\text{Tr}[\not{p}_2 \Gamma^\mu \not{p}_1 \gamma_\mu] + c.c. \tag{20.17}$$

$$= 2F_1(p^2)\text{Tr}\left[\not{p}_2 \gamma^\mu \not{p}_1 \gamma_\mu\right] + \frac{1}{m_e} p_\nu F_2(p^2) \text{Tr}[\not{p}_2 \sigma^{\mu\nu} \not{p}_1 \gamma_\mu \tag{20.18}$$

The first term just gives $2F_1(Q^2)\sigma_0$. The second term has an odd number of γ matrices in the trace, so it vanishes (NB: this would not be true if the fermions were massive, or if we looked at the NNLO, α^2 correction). Thus the total correction at order α is given by $2F_1(Q^2)\sigma_0$.

With a Pauli-Villars regulator for the UV and a photon mass m_γ for the IR

$$F_1(Q^2) = \frac{\alpha}{2\pi} \int_0^1 dx\, dy\, dz\, \delta(x+y+z-1) \left\{ \log \frac{z\Lambda^2}{z m_\gamma^2 - xyQ^2} + \frac{(1-x)(1-y)Q^2}{z m_\gamma^2 - xyQ^2} \right\} \tag{20.19}$$

This is actually considerably simpler than the general expression, because we are taking massless fermions.

The first term is IR finite, and it gives

$$F_1(Q^2) = \frac{\alpha}{2\pi} \left\{ \frac{3}{4} + \frac{1}{2} \log \frac{\Lambda^2}{Q^2} + \cdots \right\} \tag{20.20}$$

The second term is IR divergent but UV finite. The integrals are hard to do. Expanding near $m_\gamma = 0$ gives

$$F_1(Q^2) = \frac{\alpha}{2\pi} \left\{ \frac{1}{2} \log \frac{\Lambda^2}{m_\gamma^2} - \frac{1}{2} \log^2 \frac{m_\gamma^2}{Q^2} - \frac{3}{2} \log \frac{m_\gamma^2}{Q^2} - \frac{7}{4} + \frac{\pi^2}{6} + \mathcal{O}(m_\gamma^2) \right\} \tag{20.21}$$

Note at least that the $\log\Lambda^2$ term agrees with what we had above.

So the result of this loop's contribution to the total rate is

$$\sigma_\Gamma = \frac{\alpha}{2\pi} \sigma_0 \left\{ -\log^2 \frac{m_\gamma^2}{Q^2} - 3\log \frac{m_\gamma^2}{Q^2} - \frac{7}{2} + \frac{\pi^2}{3} + \log \frac{\Lambda^2}{m_\gamma^2} \right\} \tag{20.22}$$

Note the double logs. Double logs are diagnostic of infrared divergences. They have a name: Sudakov double logarithms. There is a lot of fun physics associated with them.

20.3.3 counterterms

Next, consider the counterterm contribution.

In bare perturbation theory, we are trying to calculate the Green's function

$$\langle 0|j_\mu^R \bar{\psi}^R \psi^R|0\rangle = \frac{Z_2}{Z_j}\langle 0|j_\mu^0 \bar{\psi}^0 \psi^0|0\rangle \tag{20.23}$$

so we get the counterterms $\delta_2 - \delta_j$. We already know that $\delta_j = 0$ by current conservation, so all that's left is δ_2.

Recall that in PV with a massive photon and massless electron

$$\delta_2 = -\frac{\alpha}{4\pi}\log\frac{\Lambda^2}{m_\gamma^2} \tag{20.24}$$

Thus including the interference of this counterterm with the tree-level graph, we get a factor of two from the cross-terming, and therefore

$$\sigma_{\rm ct} = \sigma_0\left(-\frac{\alpha}{2\pi}\log\frac{\Lambda^2}{m_\gamma^2}\right) \tag{20.25}$$

Thus the sum of the radiative corrections is

$$\boxed{\sigma_V = \sigma_{\rm ct} + \sigma_\Gamma = \frac{\alpha}{2\pi}\sigma_0\left\{-\log^2\frac{m_\gamma^2}{Q^2} - 3\log\frac{m_\gamma^2}{Q^2} - \frac{7}{2} + \frac{\pi^2}{3}\right\}} \tag{20.26}$$

which is UV finite.

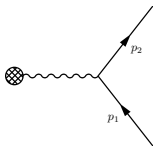
The same calculation in renormalized perturbation theory gets a factor of $\delta_2 - \delta_j$ from the vertex

$$\mathcal{L} = \cdots + \frac{Z_2}{Z_j} j_\mu \bar{\psi}\gamma^\mu\psi \tag{20.27}$$

so it's the same thing.

20.3.4 expanding out the process

Let's take in more of the full $e^+e^- \to \mu^+\mu^-$ process to see what's going on a little better. Suppose we work more more step in. That is, we try to compute



Where now the current source is creating a photon. Let me call this current Z_μ. So we are calculating

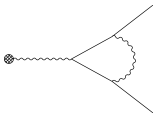
$$\langle 0|T\{Z_\mu\bar{\psi}\psi\}|0\rangle \tag{20.28}$$

Note this operator Z_μs is a different operator from the lepton number current $j_\mu = \bar{\psi}\gamma^\mu\psi$ we were using before. You can think of this current as a classical radiation source. You can also think of is simply as an external photon, however, we want the photon to be off-shell, so that's not quite right. Instead, try thinking of it like a heavy gauge boson, say the Z. Ok, just think of it as whatever you like. I'm going to think of it as an operator.

The tree level contribution is now

$$\sum_{\text{spins, } Z \text{ pols}} |\langle 0|T\{Z_\mu\bar{\psi}\psi\}|0\rangle|^2 = \frac{e^2}{Q^4}\mathrm{Tr}[\not{p}_2\gamma^\mu\not{p}_1\gamma^\nu] = 4\frac{e^2}{Q^2} \tag{20.29}$$

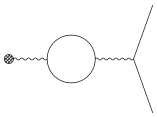
It's the same, but we get an extra factor of e from the vertex and a couple of $\frac{1}{Q^2}$'s from the photon propagator. The loop on the muon side is the same as before



(20.30)

It's divergent part is still $-\delta_1$.

Since the current caps the end of the photon line, so the photon should be considered internal. So there are other diagrams and counterterms as well. First, there is the internal photon loop



$\sim \Pi_2(p^2)$

This is just the vacuum polarization diagram. We computed it in the t-channel. In the s-channel, it's the same up to a crossing symmetry ($t \to Q^2$), so

$$\Pi_2(Q^2) = -\frac{2\alpha}{\pi}\int_0^1 dx\, x(1-x)\log\left(\frac{\Lambda^2}{m_e^2 + Q^2\, x(1-x)}\right) \tag{20.31}$$

This function's IR divergences is cut off naturally by the electron mass. The divergence is

$$\Pi_2(0) = -\frac{\alpha}{3\pi}\log\left(\frac{\Lambda^2}{m_e^2}\right) \tag{20.32}$$

which is just $-\delta_3$.

The counterterms are a little different. Including the field strength renormalization in the matrix element (bare perturbation theory) gives

$$\left\langle 0|T\left\{Z_\mu^R \bar{\psi}^R \psi^R\right\}|0\right\rangle = Z_2\sqrt{Z_Z}\left\langle 0|T\left\{Z_\mu^0 \bar{\psi}^0 \psi^0\right\}|0\right\rangle \tag{20.33}$$

(the $\sqrt{}$ in $\sqrt{Z_Z}$ is just a convention). So we get a factor of $\delta_2 - \frac{1}{2}\delta_Z$ for the counterterms from the field strengths. We also get a factor of $-\delta_e$ from the renormalization of the bare electric charge in the QED vertex: $e_0 = e_R - \delta_e$.

Thus the total UV divergences are $\delta_2 - \frac{1}{2}\delta_Z$(Green's function) $-\delta_e$(electron) $-\delta_3$(vac pol loop) $-\delta_1$(vertex loop) $= -\frac{1}{2}\delta_Z - \delta_e - \delta_3$. But since $\delta_e = -\frac{1}{2}\delta_3$ we get $\delta_Z = \delta_3$. If we had included this current in the Lagrangian (renormalized perturbation theory), it would be

$$\mathcal{L} = \cdots + \frac{\sqrt{Z_3}}{\sqrt{Z_Z}} Z_\mu A^\mu \tag{20.34}$$

So we get a factor of $\frac{1}{2}\delta_3 - \frac{1}{2}\delta_Z$ from this term. We also have the real QED vertex counterterm, δ_1 and the vertex loop $-\delta_1$. Then the vacuum polarization gives $-\delta_3$ and its counterterm δ_3. So we are left with just $\frac{1}{2}(\delta_3 - \delta_Z)$. Thus we find $\delta_Z = \delta_3$. The finite parts of the loops are of course the same. We are quickly confirming that bare and renormalized perturbation theory are quite trivially the same thing.

Add the counterterm to the vacuum polarization graph gives, for $Q \gg m_e$,

$$\Pi_2(Q^2) + \delta_3 = \frac{2\alpha}{\pi}\int_0^1 dx\, x(1-x)\log\left(1 + \frac{Q^2}{m_e^2}\, x(1-x)\right) \tag{20.35}$$

$$\approx \frac{\alpha}{3\pi}\log\frac{Q^2}{m_e^2} \tag{20.36}$$

This gives an additional contribution to the total cross section. Note that the charge appearing here is the renormalized charge, defined by what is measured in the Coulomb force at $p = 0$.

Also note that the electron mass appears, but we wanted to set $m_e = 0$. Recall that our definition of the electric charge was thing thing measured at $p = m_e \sim 0$. So, the m_e in $\frac{\alpha}{3\pi}\log\frac{Q^2}{m_e}$ is really just the renormalization point. That is, more generally, we can write

$$\Pi_2(Q^2) = \frac{\alpha_R}{3\pi}\log\frac{Q^2}{p_0^2} \tag{20.37}$$

where α_R is defined as the electric charge measured at $p = p_0$. For comparison with other experiments, it's easier just to leave $m_e \neq 0$ for this effect.

First, note that the tree level cross section becomes (still normalization the phase space to 1) $\sigma_0 = e^2 4Q^2$. So

$$\sigma_{\text{tot}} = 4\,\frac{e_R^2}{Q^2} + 4\,\frac{e_R^2}{Q^2}\frac{e_R^2}{12\pi^2}\log\frac{Q^2}{m_e^2} \tag{20.38}$$

This is nothing more than the running of the electric charge in the bare cross section

$$\sigma_{\text{tot}} = 16\pi Q^2 \alpha_Q \tag{20.39}$$

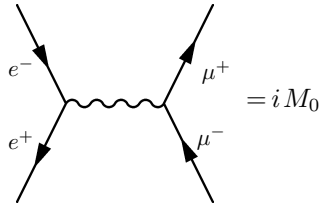
with

$$\alpha_Q = \alpha_{\text{eff}}(Q^2) = \alpha_R\left(1 + \frac{\alpha_R}{3\pi}\log\frac{Q^2}{m_e}\right) \tag{20.40}$$

This is the same scale-dependent effective coupling we were using for the corrections to Coulomb's law.

20.3.5 full $e^+e^- \to \mu^+\mu^-$

Now what happens when we calculate the full matrix elements



Corresponding to matrix elements

$$\langle 0|e^+e^-\mu^+\mu^-|0\rangle \tag{20.41}$$

with no weird operators anymore (phew!). The total cross section is

$$\sigma_0 = \frac{4\pi\alpha_Q^2}{3Q^2} \tag{20.42}$$

where I have included the vacuum-polarization loop in writing α_Q instead of α_R for this cross section. Remember $\alpha_Q = \alpha_R + \mathcal{O}(\alpha_R^2)$.

For the loops, there is the vertex correction on the left and the vacuum polarization graph, plus a new graph – the vertex correction on the left hand side. Thus the total loop divergences are

$$\text{divergences} = -\delta_1 - \delta_3 - \delta_1 \tag{20.43}$$

For counterterms, we get just the two vertex counterterms and the photon propagator counterterm. The external lines are amputated so they do not contribute. Thus we have

$$\text{counterterms} = \delta_1 + \delta_3 + \delta_1 \tag{20.44}$$

So all the divergences cancel.

Thus the total loop contribution is

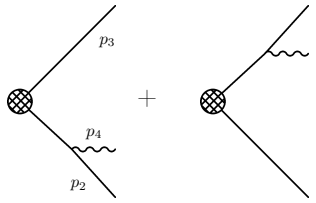
$$\sigma_V = \sigma_0\frac{\alpha}{2\pi}2\left[-\log^2\frac{m_\gamma^2}{Q^2} - 3\log\frac{m_\gamma^2}{Q^2} - \frac{7}{2} + \frac{\pi^2}{3}\right]$$

and α_R is defined in terms of the charge measured at $p \to 0$. So we just get twice what we had for just the loop on the muon side, plus the vacuum polarization effect.

20.4 Real-emission: Final state radiation

The real emission graphs are easy to compute. However, the phase space integrals with a massive photon can be somewhat tedious. This analysis is taken from Rick Field's book "Applications of perturbative QCD".

First, take massless photon emission from the muon side only. We need to calculate these diagrams



Which are

$$\bar{u}(p_2)\gamma^\mu \frac{i}{\not p_4 + \not p_2}\gamma^\nu v(p_3) + \bar{u}(p_2)\gamma^\mu \frac{i}{\not p_4 + \not p_3}\gamma^\nu v(p_3) \tag{20.45}$$

note $p_4 + p_3 = p - p_2$.

Let us define the Mandelstam variables

$$s = (p_3 + p_4)^2 = Q^2(1 - x_2) \tag{20.46}$$

$$t = (p_2 + p_4)^2 = Q^2(1 - x_1) \tag{20.47}$$

$$u = (p_2 + p_3)^2 = Q^2(1 - x_3) \tag{20.48}$$

Note that now $s + t + u = Q^2$ since the current is off-shell.

The calculation is straightforward and we find

$$\frac{d^2\sigma}{ds\,dt} = \sigma_0 \frac{\alpha}{2\pi}\left[\frac{s}{t} + \frac{t}{s} + \frac{2u}{st}\right] \tag{20.49}$$

or equivalently

$$\frac{d^2\sigma}{dx_1 dx_2} = \sigma_0 \frac{\alpha}{2\pi}\frac{x_1^2 + x_2^2}{(1 - x_1)(1 - x_2)} \tag{20.50}$$

With a massive photon, this gets modified. There are both explicit factors of m_γ that come into the matrix elements, but also factors of m_γ in the phase space integrals to make a cross section. The result is

$$\frac{d^2\sigma}{dx_1 dx_2} = \sigma_0 \frac{\alpha}{2\pi}\frac{1}{(1 - x_1)(1 - x_2)}\left[x_1^2 + x_2^2 + \beta\left[2(x_1 + x_2) + \frac{(1 - x_1)^2 + (1 - x_2)^2}{(1 - x_1)(1 - x_2)}\right] + 2\beta\right] \tag{20.51}$$

with $\beta = \frac{m_\gamma^2}{Q^2}$. Now, $s + t + u = Q^2 + Q^2\beta^2$.

To calculate the total cross section, we have to integrate over phase space. Since

$$x_1 + x_2 + x_3 = 2 - \beta \tag{20.52}$$

$$0 < x_i < 1 \tag{20.53}$$

We get restricted phase space. For example, if $x_3 = 0$ then $x_1 < 1 - \beta$ or else x_2 would have to be larger than 1. The limits are

$$0 \leqslant x_i \leqslant 1 - \beta \tag{20.54}$$

$$1 - \beta - x_1 \leqslant x_2 \leqslant \frac{1 - x_1 - \beta}{1 - x_1} \tag{20.55}$$

Thus,

$$\sigma_R = \int_0^{1-\beta} dx_1 \int_{1-\beta-x_1}^{\frac{1-\beta-x_1}{1-x_1}} dx_2 \frac{d^2\sigma}{dx_1 dx_2} = \sigma_0 \frac{\alpha}{2\pi}\left\{ \log^2\frac{m_\gamma^2}{Q^2} + 3\log\frac{m_\gamma^2}{Q^2} - \frac{\pi^2}{3} + 5 \right\} \tag{20.56}$$

Recalling

$$\sigma_V = \sigma_{\rm ct} + \sigma_\Gamma = \frac{\alpha}{2\pi}\sigma_0\left\{ -\log^2\frac{m_g^2}{Q^2} - 3\log\frac{m_g^2}{Q^2} - \frac{7}{2} + \frac{\pi^2}{3} \right\} \tag{20.57}$$

We see that all the $\log m_\gamma$ terms precisely cancel, and we are left with

$$\sigma_R + \sigma_V = \frac{3\alpha}{4\pi}\sigma_0 \tag{20.58}$$

So we see that if we include the virtual contribution *and* the real emission, the IR divergences cancel.

We have, at this point, only included real emission from the muon side, and the vertex correction loop from the muon side. We must now also be able to cancel the infrared divergence from the vertex correction on the electron side.

20.5 Initial state radiation

That's all well and good for the muon side, where it makes sense to sum over the possibility of emitting muons and photons in calculating the total cross section.

If we integrate over all initial state photons as well, we would find

$$\sigma_{\rm tot}[e^+e^-(+\gamma) \to \mu^+\mu^-(+\gamma)] = \sigma_0\left(1 + \frac{3\alpha_{\rm eff}}{2\pi}\right) \tag{20.59}$$

which is perfectly finite. However, shouldn't we be able to collide the e^+e^- without summing also over initial state photons?

Actually, no. While it does not make sense to integrate over *all* possible initial-state photons, it is impossible to make sure you just have an electron and not an electron and a bunch of photons. So at least you must sum over the soft stuff, which will be enough to cut-off the divergence. One reason this might seem confusing is because our language has been imprecise. All along we have pretended we've been colliding momentum eigenstates, *i.e* plane waves. In fact, we are colliding wave packets. The uncertainty in the wave-packet has to do with not knowing how much energy there is, and also not known how much energy is in the electron or in its soup of surrounding soft photons, which are also in the wave-packet.

For example, suppose the wave packet has some energy uncertainty $E_i \ll Q$. E_i can be very small, even $E_i \ll m_e$, but it must be greater than zero. Then integrating the photon energy up to E_i will give

$$\sigma_R = \sigma_0 \frac{\alpha}{2\pi}\left\{ \log^2\frac{m_\gamma^2}{E_i^2} + 3\log\frac{m_\gamma^2}{E_i} + c \right\} \tag{20.60}$$

for some constant c. Then adding this to our correction will cancel the logs of m_γ leaving

$$\sigma_{\rm tot}[e^+e^-(+\gamma) \to \mu^+\mu^-(+\gamma)] = \sigma_0\left[1 + \frac{3\alpha}{4\pi}\left(\log^2\frac{Q^2}{E_i^2} + 3\log\frac{Q^2}{E_i^2} + c\right)\right] \tag{20.61}$$

which is finite – it doesn't depend on the fictitious photon mass.

How do we deal with this in practice? First of all, we could just choose E_i of some scale so that $\alpha\log^2\frac{Q^2}{E_i^2} \ll 1$. But do we have to really know what the uncertainty is on our wavepackets and include that in the calculation? No we don't. In fact, there are many uncertainties which are much larger than wave-packet width, such as the impact parameter of the e^+ and e^- we are colliding. Averaging over impact parameters completely overwhelms any uncertainty-relation bounded wave-packet uncertainty. The practical resolution is that we should just stick to observables that are not sensitive to the electron side. Let me give some examples.

20.6 Applications of infrared divergences

A process like $e^+e^- \to \mu^+\mu^-$ is not well-defined in perturbation theory because of the infrared divergence. We found that we need to sum over additional photon emissions to get a finite quantity. Quantities which are insensitive to the infrared regulator are called *infrared safe*. There is a theorem due to Kinoshita-Lee and Nauenberg (KLN theorem) which says that any physically observable quantity is infrared safe.

Here are a few examples of the importance of infrared safety in practice:

- Decay widths

For situations where the cross section is important, such as Z boson production, it's usually resonance production. In fact, the soft photons help bring the Z to the resonance peak, a process called *radiative return*. In this case, it's just the decay width you measure, so you only need the final state loops. The decay width is calculable, finite, and does not depend on whether it was e^+e^- or something else that produced the Z. Thus the correction to the width is a way to test for the $\frac{3\alpha}{2\pi}$ correction.

For example, suppose the Z decays to quarks and to muons. Quarks have charge $\frac{2}{3}$ or $-\frac{1}{3}$, so the photon loop correction will be smaller for the width to quarks than the width to muons. Thus the ratio of partial widths directly measures α and the charges of the particles in the decay. The electron side drops out completely. So we don't need to bother with the E_i cutoff or the vertex correction loop.

- Protons in QCD

For QCD, where you don't have free quarks, the wave-packets have a natural size – the proton. The proton is a bundle of quarks and gluons, where the gluons are mostly soft, and can be thought of to a first approximation like the photons we have been using in this discussion. These virtual and soft gluons strongly influence the distribution of energies of quarks inside the proton, in a calculable way. This is the physical basis of *parton distribution functions*. So there's a lot of physics in initial state radiation, it just doesn't have much to do with total cross sections in QED.

- Final state energy cuts.

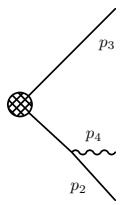
Swapping a $\log m_\gamma$ for a $\log E$ works for final state radiation too, and there it *is* important. We don't have to integrate over all photons on the $\mu^+\mu^-$ side. We can do something less *inclusive* (more *exclusive*), and integrate only over photons up to an energy E_f. If we put a cut on the photon energy, we will get logs of that energy divided by Q in the cross section.

$$\frac{d\sigma_{\text{tot}}}{dE_f} = \sigma_{\text{tot}}\left(\alpha^2 \log^2 \frac{E_f^2}{Q^2} + \cdots\right) \tag{20.62}$$

This E_f is often a realistic quantity related to actual parameters of an experiment. For example, the experiment BABAR at SLAC measures the decays of B mesons to kaons and photons ($B \to K\gamma$). They are only sensitive to photons harder than 1.6 GeV. Softer photons are below the threshold sensitivity of their detector. Thus this log is a very important quantitative feature of the cross section they measure, and must be included in attempts to constrain new physics (eg. supersymmetry) using $B \to K + \gamma$ decays.

- Soft/Collinear singularities

Note that for small $E_f \ll Q$ there will be large logarithms which may invalidate your perturbation expansion. In fact, it seems that if we look at $d\sigma/dE_f$ it will be completely dominated by small E_f. Another way of saying this is that the dominant radiation will be soft. We can see this from the Feynman diagrams directly



The propagator is $\sim \frac{1}{(p_4+p_2)^2} \sim \frac{1}{E_\gamma E_2}$. So the divergence is coming from integrating $\int_0^{E_f} \frac{1}{E_\gamma} dE_\gamma$. More generally, this propagator will vanish when

$$p_4 \cdot p_2 = 0 \tag{20.63}$$

Writing $p_4 = E_\gamma(1,0,0,1)$, this happens when either $E_\gamma \to 0$, which is called a soft divergence, or when p_3 is also in the z direction, that is $\vec{p}_3 \cdot \bar{p}_4 = 0$. That is called a collinear singularity. So the cross section is dominated by soft and collinear singularities.

- Jets

It is not hard to see that when the radiation is soft enough, the enhancement in the cross section will be substantial. This is compensated for by the fact that the soft radiation is hard to see. But what it means is that the outgoing particle, say the muon, will be accompanied by lots of soft or collinear photons. This broadens the outgoing particle from a plane wave into something like a wave-packet. This wave-packet is not a Gaussian wave-packet, and not easy to describe theoretically, but it has a calculable shape.

While these final-state muon-photon wave-packets are are hard to see in QED, they are easy to see in QCD. In QCD, the muon is replaced by a quark and the photon replaced by a gluon. The quark itself, and the additional soft gluons turn into separate observable particles, such as pions and kaons. Thus a quark in QCD turns into a *jet* of *hadrons*.

These jets are a very real and characteristic phenomenon of all high-energy collisions. We have explained their existence by studying the infrared singularity structure of Feynman diagrams in quantum field theory.

- Lamb shift

The dominant contribution to the Lamb shift is given by an infrared divergent effect. The loop gives a $\log\frac{m_\gamma^2}{m_e^2}$ and the real emission contribution replaces this m_γ with the binding energy of the orbital E_B. The result is $\Delta E \sim \log\frac{m_\gamma^2}{E_B^2} \sim 1000\,\text{MHz}$. We will discuss this in detail in the next lecture.

20.7 $e^+e^- \to \mu^+\mu^-$ in dim reg

The calculation of the total cross section can be done also in dimensional regularization with minimal subtraction or $\overline{\text{MS}}$. Repeating the calculation this way helps illustrate regulator independence not just of UV quantities, but also of IR quantities. This part is taken from Muta "Foundations of Quantum Chromodynamics", which you should look at to get more details.

20.7.1 loops

The important loop in dimensional regularization is the F_1 form factor. We computed it already with an electron and photon mass

$$F_1(p^2) = \frac{e_R^2}{(4\pi)^{d/2}}\mu^{4-d}\int_0^1 dx\,dy\,dz\,\delta(x+y+z-1)\left\{\frac{\Gamma(2-\frac{d}{2})(2-\varepsilon)^2}{\Delta^{2-d/2}}\right. \tag{20.64}$$

$$\left. + \frac{\Gamma(3-\frac{d}{2})}{\Delta^{3-d/2}}\big(p^2[2(1-x)(1-y)-\varepsilon xy] + m_e^2[2(1-4z+z^2)-\varepsilon(1-z)^2]\big)\right\} \tag{20.65}$$

where

$$\Delta = (1-z)^2 m_e^2 + z m_\gamma^2 - xyp^2 \tag{20.66}$$

A nice thing about dimensional regularization is that can be used to regulate infrared as well as ultraviolet divergences. This is because

$$\int \frac{d^d k}{(2\pi)^4}\frac{1}{k^4} \tag{20.67}$$

is finite in at small k for $d>4$. That means should expand for $d=4+\varepsilon$. However, as long as we keep d as a variable, we will always be able to do the integrals and expand in ε, not worrying about whether then sign of ε is positive or negative. So, we can just set $m_e=m_\gamma=0$ *before* doing the integrals, then expand near $d=4$.

For $m_e=m_\gamma=0$ in dim reg the form factor at $p^2=Q^2$ becomes

$$F_1(Q^2)=\frac{e^2}{(4\pi)^{d/2}}\left(\frac{\mu^2}{-Q^2}\right)^{4-d}\int_0^1 dx\,dy\,dz\,\delta(x+y+z-1) \tag{20.68}$$

$$\times\left\{\Gamma(2-\frac{d}{2})\frac{(2-\varepsilon)^2}{2}\frac{1}{(xy)^{2-\frac{d}{2}}}+\Gamma(3-\frac{d}{2})\frac{2(1-x)(1-y)-\varepsilon xy}{(xy)^{3-d/2}}\right\} \tag{20.69}$$

$$=\frac{\alpha}{4\pi}\left(\frac{4\pi\mu^2}{-Q^2}\right)^{\varepsilon}\Gamma(3-\frac{d}{2})\frac{\Gamma(\frac{d}{2}-1)\Gamma(\frac{d}{2})}{\Gamma(d-1)}\left(-\frac{8}{\varepsilon^2}+\frac{2}{\varepsilon}-2\right) \tag{20.70}$$

Expanding out the $\frac{1}{\varepsilon}(-Q^2)^\varepsilon$ gives $\log(-Q^2)=\log Q^2+\pi i$. At order α we only need $\mathrm{Re}[M_\Gamma M_0]$ and M_0 is real, so we can drop these imaginary parts. However, note that

$$\log^2(-Q^2)=\log^2 Q^2+2\pi i\log \mathrm{Q}^2-\pi^2 \tag{20.71}$$

So we get very important π^2 terms which affect the final answer.

Then,

$$\mathrm{Re}[F_1(Q^2)]=\frac{\alpha}{4\pi}\left[-\frac{8}{\varepsilon^2}-\frac{6+4\log\frac{\tilde\mu^2}{Q^2}}{\varepsilon}-2\log^2\frac{\tilde\mu^2}{Q^2}-\log\frac{\tilde\mu^2}{Q^2}-8+\frac{7\pi^2}{6}\right] \tag{20.72}$$

Now, recall that with the massive electron and photon, these masses regulated the infrared divergences. There, we found

$$F_1(p^2)=\frac{\alpha}{4\pi}\left[\frac{2}{\varepsilon}+\cdots\right] \tag{20.73}$$

Since the UV divergences have to do with high energy, they are independent of any small mass the electron or photon might have. Thus we can figure out which of the ε's in F_1 are UV divergences and which are IR divergences

$$\mathrm{Re}[F_1(Q^2)]=\frac{\alpha}{4\pi}\left[\frac{2}{\varepsilon_{\mathrm{UV}}}-\frac{8}{\varepsilon_{\mathrm{IR}}^2}-\frac{8+4\log\frac{\tilde\mu^2}{Q^2}}{\varepsilon_{\mathrm{IR}}}-2\log^2\frac{\tilde\mu^2}{Q^2}-\log\frac{\tilde\mu^2}{Q^2}-8+\frac{7\pi^2}{6}\right] \tag{20.74}$$

We already know that this UV divergence will get canceled by the counterterm. But in order to appreciate the IR content of the counterterm, we need to understand how to compute dimensionless integrals in dim reg.

20.7.2 dimensionless integrals in dim-reg

The counterterm we need is δ_1. Recall that δ_1 cancels the divergence of the vertex correction and δ_2 cancels the divergence of the self energy diagram. Since $\delta_1=\delta_2$, it is easier just to look at the self-energy diagram which is simpler.

The self energy diagram with massless fields is

$$\Sigma(\not{p}) = e^2 \not{p} \int \frac{d^dk}{(2\pi)^d} \frac{1}{k^4} \tag{20.75}$$

What are we to make of the k integral, which doesn't have any structure in it besides the infinity? If we wick rotate and do the phase space integral, we find

$$\int \frac{d^dk}{(2\pi)^d} \frac{1}{k^4} = \frac{i}{8\pi^2} \frac{\pi^{d/2}}{\Gamma(\frac{d}{2})} \int_0^\infty k^{3-d} dk = \frac{i}{8\pi^2} \left[\frac{1}{4-d} k^{4-d} \right]_0^\infty \tag{20.76}$$

The upper limit of integration can only be used if $d < 4$ and the lower limit if $d > 4$. But d is fixed, so what do we make of this?

Let us split the integral into two parts

$$\int_0^\infty k^{3-d} dk = \int_0^\Lambda k^{3-d} dk + \int_\Lambda^\infty k^{3-d} dk \tag{20.77}$$

In the first integral, we write $d = 4 - \varepsilon_{\rm IR}$, with $\varepsilon_{\rm IR} < 0$. Then

$$\int_0^\Lambda k^{3-d} dk = \frac{1}{\varepsilon_{\rm IR}} \Lambda^{-\varepsilon_{\rm IR}} = \frac{1}{\varepsilon_{\rm IR}} - \log\Lambda \tag{20.78}$$

For the second, we write $d = 4 - \varepsilon_{\rm UV}$ with $\varepsilon_{\rm UV} > 0$. Then

$$\int_\Lambda^\infty k^{3-d} dk = -\frac{1}{\varepsilon_{\rm UV}} \Lambda^{-\varepsilon_{\rm UV}} = -\frac{1}{\varepsilon_{\rm UV}} + \log\Lambda \tag{20.79}$$

So the sum of these is

$$\int_0^\infty k^{3-d} dk = \frac{1}{\varepsilon_{\rm IR}} - \frac{1}{\varepsilon_{\rm UV}} \tag{20.80}$$

Thus,

$$\Sigma(\not{p}) = \frac{\alpha}{4\pi} \not{p} \left[\frac{2}{\varepsilon_{\rm IR}} - \frac{2}{\varepsilon_{\rm UV}} \right] \tag{20.81}$$

And the counterterm with minimal subtraction is just $\delta_2 = \frac{\alpha}{4\pi} \left[-\frac{2}{\varepsilon_{\rm UV}} \right] = \delta_1$.

If we are using minimal subtraction, then we have to make sure to use the full propagators for the external lines which include IR divergent loop corrections. This gives a factor of

$$\frac{\alpha}{4\pi} \left[\frac{2}{\varepsilon_{\rm IR}} \right] \tag{20.82}$$

Instead, it is conventional to add the IR divergence into your subtraction terms so that

$$\delta_1 = \delta_2 = \frac{\alpha}{4\pi} \left[-\frac{2}{\varepsilon_{\rm IR}} + \frac{2}{\varepsilon_{\rm UV}} \right] \tag{20.83}$$

This is just 0. In general, dimensionless integrals in dim reg are 0. When you add them to a process with their counterterms, they to convert UV divergences to IR divergences. Now we've seen how to make sense of $\infty - \infty$ and $0 - 0$.

20.7.3 total virtual part

Adding the counterterm

$$\delta_1 = \frac{\alpha}{4\pi} \left[-\frac{2}{\varepsilon_{\rm UV}} + \frac{2}{\varepsilon_{\rm IR}} \right] \tag{20.84}$$

to the loop gives

$$\mathrm{Re}[F_1(Q^2)] = \frac{\alpha}{4\pi}\left[-\frac{8}{\varepsilon_{\rm IR}^2} - \frac{6+4\log\frac{\tilde\mu^2}{Q^2}}{\varepsilon_{\rm IR}} - 2\log^2\frac{\tilde\mu^2}{Q^2} - \log\frac{\tilde\mu^2}{Q^2} - 8 + \frac{7\pi^2}{6}\right] \tag{20.85}$$

where all the UV divergences cancel (are replaced by IR divergences).

This separation between UV and IR divergences in dim reg is somewhat ill-defined. That's why we did the Pauli-Villars regulator with a finite photon mass above.

So the contribution to the cross section is

$$\sigma_V = \sigma_0(F_1(Q^2) + F_1(Q^2)^\star) \tag{20.86}$$

$$= \frac{\alpha}{2\pi}\left[-\frac{8}{\varepsilon_{\rm IR}^2} - \frac{6+4\log\frac{\tilde\mu^2}{Q^2}}{\varepsilon_{\rm IR}} - 2\log^2\frac{\tilde\mu^2}{Q^2} - \log\frac{\tilde\mu^2}{Q^2} - 8 + \frac{7\pi^2}{6}\right] \tag{20.87}$$

20.7.4 real emission

Next, we have to compute the real emission diagrams.

The matrix elements in 4D give

$$|M^2| = \frac{\alpha}{2\pi}\frac{x^2+y^2-\varepsilon(2-x-y)^2}{(1-x)^{1+\varepsilon}(1-y)^{1+\varepsilon}(1-(2-x-y))^2} = \frac{x^2+y^2}{(1-x)(1-y)} + \mathcal{O}(\varepsilon) \tag{20.88}$$

Integrating this gives

$$\int_0^1 dx \int_{1-x}^1 dy |M^2| = \frac{\Gamma(1-\frac{\varepsilon}{2})\Gamma(-\frac{\varepsilon}{2})}{\Gamma(3-\frac{3}{2}\varepsilon)}\left(-4 + 6\varepsilon - \frac{5}{2}\varepsilon^2\right) \tag{20.89}$$

For the cross section, we also have to include a factor from d dimensional $2\to 5$ phase space

$$\frac{d\sigma_R}{dx\,dy} = \sigma_B \frac{\alpha}{2\pi}\left(\frac{4\pi\mu^2}{Q^2}\right)^\varepsilon \frac{3(1-\varepsilon)}{(3-2\varepsilon)\Gamma(2-2\varepsilon)}|M^2| \tag{20.90}$$

But now the Born cross section changes too, because of d dimension phase space (we have to keep track of all the ε^2 terms, because they give finite parts on top of the $\frac{1}{\varepsilon^2}$ poles).

$$\sigma_B^d = \sigma_0\left(\frac{4\pi\mu^2}{Q^2}\right)^\varepsilon \frac{3(1-\varepsilon)\Gamma(2-\varepsilon)}{(3-2\varepsilon)\Gamma(2-2\varepsilon)} = \sigma_0 + \mathcal{O}(\varepsilon) \tag{20.91}$$

All these ε are $\varepsilon = \varepsilon_{\rm IR}$ because there are no loops.

$$\sigma_B = \sigma_0\left(\frac{4\pi\mu^2}{Q^2}\right)^{\varepsilon/2} \frac{\Gamma(\frac{d}{2})}{\Gamma(d-2)}\frac{1-\frac{\varepsilon}{2}}{1-\frac{\varepsilon}{3}} = \sigma_0 + \cdots \tag{20.92}$$

Thus

$$\sigma_R = \sigma_0\left(\frac{\alpha}{2\pi}\right)\left(\frac{8}{\varepsilon_{\rm IR}^2} + \frac{6+4\log\frac{\tilde\mu^2}{Q^2}}{\varepsilon_{\rm IR}} + 2\log^2\frac{\tilde\mu^2}{Q^2} + \log\frac{\tilde\mu^2}{Q^2} + \frac{19}{2} - \frac{7\pi^2}{6}\right) \tag{20.93}$$

where there are no UV divergences in the real emission, so we have written it as $\varepsilon_{\rm IR}$ to be explicit.

Thus all of the IR divergences, and all of the logs cancel, and we are left with

$$\sigma_{\rm tot} = \sigma_V + \sigma_R = \sigma_0\left(\frac{3\alpha}{4\pi}\right) \tag{20.94}$$

which is exactly what we got with PV and m_γ. This is beautiful confirmation of the consistency of all these ideas.

20.8 Compton scattering

The infrared divergences also effect other process, like Compton scattering (which as we have noted is just the crossed version of $\gamma \to e^+e^-\gamma$). There

$$\frac{d\sigma(\gamma e^- \to \gamma e^-)}{\mathrm{dcos}\theta} = \frac{2\pi\alpha^2}{2m_e^2 + s(1+\cos\theta)} \tag{20.95}$$

this leads to

$$\sigma_{\mathrm{tot}} = \frac{2\pi\alpha^2}{s}\log\frac{s}{m_e^2} \tag{20.96}$$

This $\log m_e$ is the same as the $\log m_\gamma$ terms we were seeing. At very high energies, your knowledge of the energy of the initial state electron would be limited. In that case, you would have to allow that the electron is off-shell by an energy E and include also $e^- \to e^-\gamma$. This would replace the $\log m_e$ by a $\log E$.

Another way to see that this total cross section formula is of limited validity is to consider radiative corrections. They would add terms of the form

$$\sigma_{\mathrm{tot}} = \frac{2\pi\alpha^2}{s}\left(\log\frac{s}{m_e^2} + \alpha\log^2\frac{s}{m_e^2} + \cdots\right) \tag{20.97}$$

If $s \gg m_e$ gets large enough than these subleading terms dominate your original estimate. Thus the processes with $\gamma\gamma e^- \to \gamma e^-$ will contribute. In that case you *must* allow for a finite resolution of the electron energy if you want to trust your answer. Then $\frac{2\pi\alpha^2}{s}\log\frac{s}{E^2}$ is the more useful formula.

Chapter 21
The Lamb Shift

21.1 Hydrogen atom spectrum – summary

At leading order, the energy of the states of the hydrogen atom only depend on the principle quantum number n

$$E_n = -m_e \frac{Z^2\alpha^2}{2n^2} \sim 3 \times 10^9 \,\text{MHz} \tag{21.1}$$

There are corrections to this due to the non-relativistic expansion of the Dirac equation. The Dirac Hamiltonian can be expanded and written as an effective Schrodinger-like equation

$$\left[\frac{p^2}{2m} - \frac{p^4}{8m^3} + \frac{Z\alpha}{4r^3m^2}\vec{B}\cdot\vec{\sigma} + \frac{Z\alpha\pi}{2m^2}\delta(r) \right]\psi = E\psi \tag{21.2}$$

The p^4 term is a relativistic corrections, as is the $\delta(r)$ term (the *Darwin* term). The $\vec{L}\cdot\vec{\sigma}$ term is the spin-orbit coupling. This changes the energy levels by

$$\Delta E = -m_e \frac{Z^4\alpha^4}{2n^3}\left(\frac{1}{j+\frac{1}{2}} - \frac{3}{4n} \right) \sim 4 \times 10^4 \,\text{MHz} \tag{21.3}$$

where j is the total angular momentum quantum number. This energy shift, to $\mathcal{O}(\alpha^4)$ is known as the *fine* structure.

The easiest way to see that the energies can only depend on the total angular momentum is by observing that the total angular momentum operator $\vec{J} = \vec{L} + \frac{1}{2}\vec{\sigma}$ commutes with the Dirac Hamiltonian. So the fine structure does not split states of the same j but different l and s. For example, there are two states with $n = 2$ and $j = 1/2$: the $2S_{1/2}$ and $2P_{1/2}$ states. These are degenerate as far as fine structure is concerned.

There is also a higher order correction, $\mathcal{O}(\alpha^6)$, which comes from the next order non-relativistic expansion of the Dirac equation. These will be effects of order

$$\Delta E \sim m_e \frac{Z^6\alpha^6}{2n^3} \sim 2 \,\text{MHz} \tag{21.4}$$

These can be important for large Z, but otherwise are negligible, and extremely difficult to measure experimentally. In any case, they will not split the states of the same j.

21.2 Bethe's Lamb shift calculation

The most quantitatively important part of the Lamb shift is a large logarithm, due to the infrared divergent vertex correction loop. This calculation was done first by Hans Bethe in 1947. Although Bethe did not understand the infrared divergences in our modern language, he was able to make physical arguments to arrive at the correct answer.

Consider the infrared and UV divergent form factor F_1, which corrects the photon vertex. We computed it already with an electron and photon mass. We are interested mainly in the momentum region relevant to correcting the energy levels of hydrogen. Since the electron Compton radius is much less the the Bohr radius: $\frac{1}{m_e} \ll a_0$, the physical effect should be dominated by momenta $p^2 \ll m_e^2$. The calculation is done by splitting up momenta into $p > E_B$ and $p < E_B$, where E_B is the binding energy of the orbit $E_B \sim 1\,\mathrm{Ry} \sim \alpha^2 m_e$.

Let us start with $E_B \ll p^2 \ll m_e$. Recalling that the renormalization conditions set

$$F_1(0) = 1 \tag{21.5}$$

exactly, we have to work to next-to-leading order in $\frac{p^2}{m_e^2}$. We then find

$$F_1(p^2) = 1 + \frac{\alpha}{6\pi}\frac{p^2}{m_e^2}\left(\log\frac{m_\gamma^2}{m_e^2} + \cdots\right) \tag{21.6}$$

The $\cdots$ may include things like the $\frac{2}{5}$ from the Uehling potential, which is really a vacuum polarization loop, or a $\frac{3}{4}$ which is spin-orbit coupling effect. So we will just concentrate on the log.

To get a correction to the energy from this, observe that F_1 corrects the photon vertex, so we get an interaction like

$$\Delta\mathcal{L} = F_1(p^2)\bar{\psi}\not{A}\psi \tag{21.7}$$

Since interaction terms in the Lagrangian are the same as energy density terms in the Hamiltonian, up to a sign, we can turn this into an effective energy. In the non-relativistic limit, only the electric field (A_0 term) is relevant, so we get

$$\Delta E = -\frac{\alpha}{6\pi m_e^2}\left(\log\frac{m_\gamma^2}{m_e^2} + \cdots\right)\psi^\dagger(q_1)p^2 A_0(p)\psi(q_2) \tag{21.8}$$

where I am leaving some integrals and Fourier transforms implicit.

Now, $A_0 = \frac{\alpha}{r} = \frac{\alpha}{p^2}$ in the Hydrogen atom, so $p^2 A_0 = Z\alpha\delta(x)$. Thus we end up evaluating $\bar{\psi}\psi$ at $x = 0$. But only the $l = 0$ modes have support at $x = 0$

$$\psi(0) = 2\left(\frac{Z\alpha m_e}{n}\right)^{3/2}\delta_{l,0} \tag{21.9}$$

That is, we pick up a shift for the $l = 0$ orbitals, like $2S_{1/2}$ but not the $l = 1$ orbitals, like $2P_{1/2}$. Thus this will contribute to a split of these levels. We find

$$\Delta E = -\frac{\alpha}{6\pi}\log\frac{m_\gamma^2}{m_e^2}\left(4\frac{Z^3\alpha^3 m_e^3}{8}\right)Z\alpha\delta_{l,0} \tag{21.10}$$

$$= -\frac{Z^4\alpha^5 m_e}{12\pi}\log\frac{m_\gamma^2}{m_e^2}\delta_{l,0} \tag{21.11}$$

This is still infrared divergent.

The infrared divergence is canceled by the really small momentum region $p^2 \ll E_B \sim \alpha^2 m_e$. In this region, the photon energy is not bigger than the energy of the photons coming out of the nucleus which are responsible for binding the electron in the first place. So we must work to all orders in the background Coulomb potential. That is, we need to consider diagrams with multiple insertions of the external current. The calculation is a real pain. The result is what you would have guessed by dimensional analysis. The low energy region contributes an amount

$$\Delta E = -\frac{Z^4\alpha^5 m_e}{12\pi}\log\frac{E_B^2}{m_\gamma^2}\delta_{l,0} \tag{21.12}$$

So that the combined value is

$$\Delta E = -\frac{Z^4\alpha^5 m_e}{12\pi}\log\frac{E_B^2}{m_e^2}\delta_{l,0} \tag{21.13}$$

$$\sim -\frac{Z^4\alpha^5 m_e}{12\pi}\log(\alpha^4 Z^4)\delta_{l,0} \tag{21.14}$$

$$\sim 1300\,\mathrm{MHz} \tag{21.15}$$

If you are interested in a less hand wavy derivation, I think clearest treatment can be found in Landau and Lifshitz volume 4, "Quantum Electrodynamics." Weinberg does the calculation also, in Section 14.3 of volume one, but it's a little messier.

21.3 Lamb shift phenomenology

The splitting between the $2S_{1/2}$ and $2P_{1/2}$ levels is a very small effect. It was not measured until 1947 by Willis Lamb, who was able to use microwave technology developed during WWII to measure the shift accurately. Lamb found

$$E(2S_{1/2}) - E(2P_{1/2}) \approx 1000\,\mathrm{MHz} \tag{21.16}$$

That is, the $2S_{1/2}$ state is lifted, by 1000 MHz. The understanding of this result was one of the great triumphs of QED. It is known as *hyperfine* structure, and do to radiative corrections. They are down by a loop factor α from the fine structure splittings. There are a number of contributions.

First, there is the finite piece that has support near $r = 0$.

$$V(r) = -\frac{\alpha}{r}\left[1 + \frac{4\alpha}{15\pi^2 m^2}\delta(r)\right] \tag{21.17}$$

This effects only the $l = 0$ modes and lowers their energy, to

$$\Delta E_{nlj} = m\frac{4Z^4\alpha^5}{3\pi n^3}\left[-\frac{1}{5}\right]\delta_{l0} \quad \Rightarrow \quad \Delta E(2S_{1/2}) = -27\,\mathrm{MHz} \tag{21.18}$$

This is known as the Uehling term, it was known since the mid 1930's. It shifts the $2S_{1/2}$ state by $-27\,\mathrm{MHz}$, which doesn't help at all to explain Lamb's result. The correction is fairly unambiguous and was derived from vacuum polarization by simply dropping all the logs. Renormalization was not at all understood by the mid 30s.

21.3.1 Bethe's log

Next, there is the radiative correction with the logarithm in it. We discussed this above. It was calculated first by Hans Bethe in 1947. He did the calculation using old-fashioned perturbation theory, but only the low momentum region. He argued that the energy of a free electron and a bound electron in a Hydrogen atom can have different energies, and worked out the radiative correction. He found that his answer was UV divergent $\Delta E \sim \log\frac{E_B}{\Lambda}$, independent of the electron mass. This is the same UV divergence that Oppenheimer had found in 1931, for which Oppenheimer concluded that QED was a failure. The difference is: Bethe new about Lamb's measurement, and his estimate agreed.

Bethe imposed that $\Lambda = m_e$ is a natural cutoff justified on physical grounds. This led to

$$\Delta E(2S_{1/2}) = m\frac{4Z^4\alpha^5}{3\pi n^3}\log\frac{m_e}{\bar{E}} \sim 1000\,\mathrm{MHz} \tag{21.19}$$

where $\bar{E}$ is some average excitation energy, not exactly the binding energy. For example, taking $\bar{E} = \frac{1}{4}13.6\,\mathrm{eV}$ gives $\log\frac{m_e}{\bar{E}} = 10$ which leads to a shift of around 1600 MHz. What Bethe used for $\bar{E}$ was a little more complicated, and he got 1000 MHz, in close agreement to Lamb's measurement.

Even though this shift accounts for the majority of the Lamb shift, but Bethe's derivation was somewhat *ad hoc*. Still, it was the first example of something coming out of a divergent loop and agreeing with experiment and was extremely important historically.

21.3.2 spin-orbit coupling

Then there is a spin-orbit coupling term, which only effects the modes with $l \neq 0$. This was first calculated by Schwinger, who got $\Delta E(2P_{1/2}) = 3/8$. Actually, Schwinger was able to argue on general grounds that this should be 3/8, but when he calculated it, he found 1/8. This was due to his use of a non gauge-invariant regulator.

The full effect was calculated correctly for the first time by Tomonaga in 1948. He found

$$\Delta E_{nlj} = m\frac{4Z^4\alpha^5}{3\pi n^3}\left[\frac{3}{8}\frac{j(j+1)-l(l+1)-\frac{3}{4}}{l(l+1)(2l+1)}\right](1-\delta_{l0}) \sim -16.875\,\mathrm{MHz} \tag{21.20}$$

This completed the calculation.

The total shift, first presented by Tomonaga in 1949 (Feynman and Schwinger calculated this formula correctly too, and independently of Tomonaga) is

$$\Delta E_{nlj} = m\frac{4Z^4\alpha^5}{3\pi n^3}\left[L_{nl} + \frac{19}{30}\delta_{l0} + \frac{3}{8}\frac{j(j+1)-l(l+1)-\frac{3}{4}}{l(l+1)(2l+1)}(1-\delta_{l0})\right] \tag{21.21}$$

where the L_{nl} come from the log piece, calculated properly. They are integrals over the hydrogen atom wavefunctions, and the Fourier transform of the proper combination of the loop graphs. For example, $L_{n=2,l=0} = 7.03$ and $L_{n=2,l=1}$ is 0.030. So this becomes

$$E(2S_{1/2}) - E(2P_{1/2}) = 0.411 m\alpha^5 = 1051\,\mathrm{MHz} \tag{21.22}$$

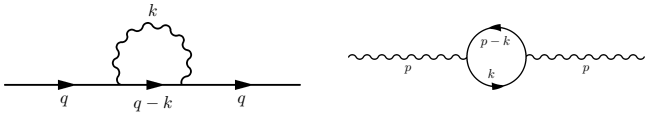
The current best measurement of this shift is 1054 MHz.

Chapter 22
Renormalizability of QED

22.1 Introduction

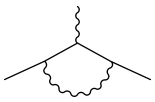
We have shown how to calculate 2, 3 and 4 point functions in QED. We found that the 1-loop corrections to these correlation functions often have UV divergences, but that these UV divergences were artifacts of not computing something physical. For physical quantities, the UV divergences always canceled, leaving a finite measurable result which can be compared to experiment. The next question we will address is, will this always be the case?

At 1-loop, it's not hard to see that in QED we can always cancel the divergences. We already found three UV divergent loops in QED. The first two corrected 2-point functions



(22.1)

and the third corrected a 3-point function



(22.2)

We were able to subtract off the UV divergent parts of these graphs with counterterms δ_1, δ_2, δ_3 and δ_m. These counterterms were introduced as a quick way to compare observables – they drop out in differences between correlations functions at different scales, which is often the quantity that we are really observing.

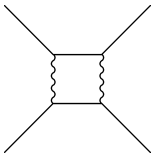
The UV divergences are the same whether or not the the external legs are on-shell; they come from regions of loop momenta with $k \gg p_i$ for any external momentum p_i. In particular, the same counterterms will cancel the UV divergences of graphs even when the graphs are sub-graphs in more complicated higher-order correlation functions. We saw this explicitly, and in some detail, for 4-point functions, like $e^+e^- \to \mu^+\mu^-$. The graphs above were all present and the counterterms we derived from 2 and 3-point functions nicely removed the UV divergences. This will hold quite generally because the counterterms work off-shell. Thus, we only have to make sure that we can remove all the UV divergences of graphs like these which are one-particle irreducible (1PI).

Recall that we introduced the notion of one-particle irreducibility when trying to deal with the electron self-energy. It justified using the exact renormalized propagator, with a pole at the physical mass instead of the bare propagator. Now we see that we only need to look at 1PI graphs when trying to figure out what UV divergences are present. Let us use as a working definition of 1PI

- One-particle irreducible (1PI): all internal lines have some loop momentum going through them

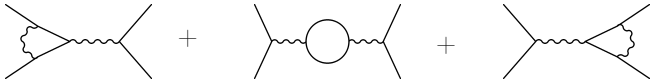
1PI graphs are also sometimes known as *proper* graphs.

An example of a 1PI 4-point graph is



(22.3)

The graphs



(22.4)

are not 1PI. These are amputated graphs, and very important to actually calculate in working out the loop-corrections to a particular process. But their UV divergences are canceled by counterterms which can be worked out for lower-order correlation functions, for which they are 1PI. So, keep in mind that for physics we need to compute all amputated graphs, but for studying general properties of renormalizability it is enough to study only the 1PI graphs.

22.2 Renormalizability of QED

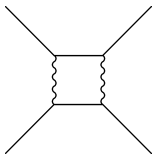
The graphs in Eqs (1) and (2) are the only non-vanishing 2 and 3-point functions in QED.

22.2.1 4-point functions

For 4-point functions, there is the function with 4-fermions

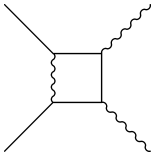
$$\langle 0|T\{\bar{\psi}\psi\bar{\psi}\psi\}|0\rangle \tag{22.5}$$

We have already looked at this, for $e^+e^- \to \mu^+\mu^-$, where the only relevant 1-loop graphs were the 3 above. If we took Moller scattering ($e^-e^- \to e^-e^-$) at 1-loop, there would be another one:

$$\langle 0|T\{\bar{\psi}\psi\bar{\psi}\psi\}|0\rangle \sim$$  $$\sim \int \frac{d^4k}{(2\pi)^4}\frac{1}{k^2}\frac{1}{k^2}\frac{1}{\not{k}}\frac{1}{\not{k}} \tag{22.6}$$

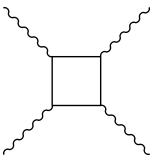
This amplitude goes like $\int \frac{d^4k}{k^6}$ at large k, and is therefore not UV divergent.

Similarly,

$$\langle 0|T\{\bar{\psi}\psi A_\mu A_\nu\}|0\rangle \sim$$  $$\sim \int \frac{d^4k}{(2\pi)^4}\frac{1}{k^2}\frac{1}{\not{k}}\frac{1}{\not{k}}\frac{1}{\not{k}} \tag{22.7}$$

This would contribute to $e^+e^- \to \gamma\gamma$ at 1-loop. It is also not UV divergent.

Finally, there is the 4-point photon for light-by-light scattering:

$$M_{\mu\nu\rho\sigma} = \langle 0|T\{A_\mu A_\nu A_\rho A_\sigma\}|0\rangle \sim$$  $$\sim \int \frac{d^4k}{(2\pi)^4}\frac{1}{\not{k}}\frac{1}{\not{k}}\frac{1}{\not{k}}\frac{1}{\not{k}} \tag{22.8}$$

This one appears logarithmically divergent:

$$M_{\mu\nu\rho\sigma} = c \log \Lambda^2 (\eta_{\mu\nu}\eta_{\rho\sigma} + \eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho}) + \text{finite} \tag{22.9}$$

For some constant c. The $\eta's$ have been symmetrized. However, we know by the Ward identity that this must vanish when one of the photons is replaced by it's momentum. Say A_μ has momentum q_μ. Then

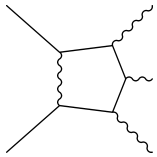
$$0 = q_\mu M_{\mu\nu\rho\sigma} = c \log \Lambda^2 (q_\nu \eta_{\rho\sigma} + q_\rho \eta_{\nu\sigma} + q_\sigma \eta_{\nu\rho}) + q_\mu \cdot \text{finite} \tag{22.10}$$

This must hold for all q^μ, which is impossible unless $c = 0$. Thus gauge invariance makes this amplitude finite.

22.2.2 5,6,... point functions

For 1-loop contributions to 1PI Green's functions with more than 4 legs, we get things like pentagon diagrams

$$\langle 0|T\{\bar{\psi}\psi A_\nu A_\rho A_\sigma\}|0\rangle \sim$$



$$\sim \int \frac{d^4k}{(2\pi)^4} \frac{1}{\not{k}} \frac{1}{\not{k}} \frac{1}{\not{k}} \frac{1}{\not{k}} \frac{1}{k^2} \tag{22.11}$$

This will all have at least 5 propagators, with 5 factors of loop momentum, so they will be finite. It doesn't matter if the propagators are fermions or photons anymore, we will always have more than 4 powers of k in the denominator.

So, the 4 counterterms δ_1, δ_2, δ_3 and δ_m suffice to cancel all the divergences in any Green's function of QED at 1-loop. The general definition of renormalizable is:

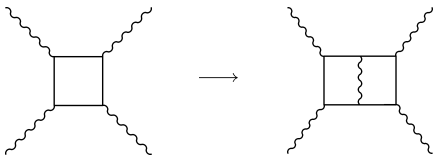
Renormalizable: all UV divergences are canceled with a finite number of counterterms

Thus, we have shown that at QED is *renormalizable* at 1-loop.

22.2.3 renormalizability to all orders

What about 2-loop and higher loop 1PI graphs? We can show they are finite by induction.

When we add a loop, we can add either a photon propagator or a fermion propagator. If we add an internal photon propagator,

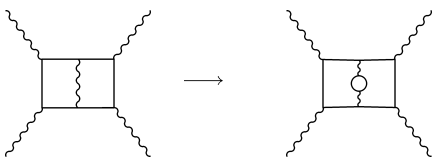


(22.12)

it must split two fermion lines, so the new loop has two additional fermion propagators as well. By the definition of 1PI, loop momenta go through all internal lines. So for $k \gg p_i$, where p_i is any external momentum, each internal line will get $\frac{1}{\not{k}}$ or $\frac{1}{k^2}$. So the matrix element gets modified to

$$\mathcal{M} \to \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^2} \frac{1}{\not{k}} \frac{1}{\not{k}} \mathcal{M} \tag{22.13}$$

If we add a fermion loop, we need to cut a photon line,



(22.14)

giving $\frac{1}{k^2}$ and by Lorentz invariance, we need to add at least 2 fermion propagators $\frac{1}{\not k}\frac{1}{\not k}$, so a fermion insertion also doesn't change the power counting.

This assumes that dominant divergence comes from when all the loop momenta go to infinity together. Clearly if the graph is not divergent when all the momenta are blowing up, it will not be divergent when only some of the momenta are blowing up.

There is another special case, when the two loop momenta come in as $k_1 - k_2$ in the denominator. Then the degree of divergence depends on precisely how we take the momenta to infinity together. But if we hold $p^\mu = k_1^\mu - k_2^\mu$ fixed in the limit, then we are effectively integrating over only one fewer loop momenta. Thus this is equivalent to a graph of one fewer loops. The actual proof that these special cases do not screw things up is fairly involved. For QED it involves gauge invariance in an essential way (for example, as we saw in the light-by-light scattering 4-point function above).

The result that there are only a finite number of divergent loops is known as the BPHZ theorem, after Bogoliubov, Parasiuk, Hepp, and Zimmermann. We have only sketched the proof here. Since we can cancel all the 1-loop divergences already, we can now cancel all the 2 and higher loop divergences by adding pieces to the same counterterms to higher order in α.

We can conclude that

QED is renormalizable

This means that that all the UV divergences in QED are canceled by the same 4 counterterms we've introduces at 1-loop. These are fit by 2 numbers: the physical value of the electric charge e_R measured in Coulomb's law at $p = 0$ and the physical value of the electron mass. The other two counterterms are fixed by canonically normalizing the electron and photon fields. (The physical observable for these is that we see 1-particle electron and photon states, which sounds like a rather coarse observable, but it is an observable nonetheless.)

This a is pretty amazing conclusion actually. With just 2 free parameters fit to data, we can make an infinite number of predictions in QED.

Renormalizability played a very important role in the historical development of quantum field theory and gauge theories. Nowadays, we have a larger context to place renormalizable theories. As it turns out, renormalizability is not obviously a quality we need in a theory, and it may even hurt us. For example, QED has a big problem: it has a Landau pole. So **we cannot predict *every* observable just because we can cancel all the UV divergences**. For example, Coulomb scattering above $E = 10^{300}$ GeV is a completely mystery in QED. Moreover, precisely because QED is renormalizable, we have no sensitivity at low energy to the UV structure of the theory, so we have no way of probing the mysterious high-energy regime without building a 10^{300} GeV collider. Other renormalizable theories are unpredictive in much more relevant regimes: the standard electroweak model (without the Higgs) is non-perturbative at $\sim 10^3$ GeV, as we will discuss in a moment, and QCD does not make perturbative predictions below ~ 1 GeV. Both of these theories are renormalizable.

We can understand these issues in more detail by asking a very simple question: what to non-renormalizable theories look like?

22.3 Non-renormalizable field theories

We saw that in QED, there are finite number of divergent one-particle irreducible contributions to Green's functions. The *superficial degree of divergence* of a Green's function is the overall number of powers of k in the loop momenta, assuming all $k \gg p_i$ for any external momentum p_i. Call the superficial degree of divergence of a Green's function with n external fermions and m external photons $D_{n,m}$. Then we found

$$\langle 0|T\{AA\}|0\rangle \sim \int \frac{d^4k}{(2\pi)^4}\frac{1}{\not k}\frac{1}{\not k} \quad \Rightarrow \quad D_{2,0} = 2 \tag{22.15}$$

$$\langle 0|T\{\bar\psi\psi\}|0\rangle \sim \int \frac{d^4k}{(2\pi)^4}\frac{1}{k^2}\frac{1}{\not k} \quad \Rightarrow \quad D_{2,0} = 1 \tag{22.16}$$

$$\langle 0|T\{\bar{\psi}\psi A\}|0\rangle \sim \int \frac{d^4k}{(2\pi)^4}\frac{1}{k^2}\frac{1}{\not{k}}\frac{1}{\not{k}} \quad \Rightarrow \quad D_{2,0}=0 \tag{22.17}$$

$$\langle 0|T\{A_\mu A_\nu A_\rho A_\sigma\}|0\rangle \sim \int \frac{d^4k}{(2\pi)^4}\frac{1}{\not{k}}\frac{1}{\not{k}}\frac{1}{\not{k}}\frac{1}{\not{k}} \quad \Rightarrow \quad D_{0,4}=0 \tag{22.18}$$

$$\langle 0|T\{\bar{\psi}\psi A_\mu A_\nu\}|0\rangle \sim \int \frac{d^4k}{(2\pi)^4}\frac{1}{k^2}\frac{1}{\not{k}}\frac{1}{\not{k}}\frac{1}{\not{k}} \quad \Rightarrow \quad D_{2,2}=-1 \tag{22.19}$$

$$\langle 0|T\{\bar{\psi}\psi\bar{\psi}\psi\}|0\rangle \sim \int \frac{d^4k}{(2\pi)^4}\frac{1}{k^2}\frac{1}{k^2}\frac{1}{\not{k}}\frac{1}{\not{k}} \quad \Rightarrow \quad D_{4,0}=-2 \tag{22.20}$$

And that anything with more than 4 fields has $D<0$. It is not hard to work out that

$$D_{n,m}=4-\frac{3}{2}n-m \tag{22.21}$$

The scaling comes simply from the dimension of the Green's function: fermions have dimension $\frac{3}{2}$ and photons dimension 1. With scalar external states, this would be

$$D_{n,m,s}=4-\frac{3}{2}n-m-s \tag{22.22}$$

where s is the number of scalars. Only the Green's functions with $D>0$ can possibly be divergent.

QED is a special theory because it only has a single interaction vertex

$$\mathcal{L}_{\text{QED}}=\mathcal{L}_{\text{kin}}+e\bar{\psi}A_\mu\gamma^\mu\psi \tag{22.23}$$

the coefficient of this interaction is the dimensionless charge e. More generally, we might have a theory with couplings of arbitrary dimension. For example,

$$\mathcal{L}=-\frac{1}{2}\phi(\Box+m^2)\phi+g_1\phi^3+g_2\phi^2\Box\phi^3+\cdots \tag{22.24}$$

This screws up the power counting.

Call the dimension of the coefficient of the ith interaction Δ_i. For example, the g_1 term above has $\Delta_1=1$ and g_2 has $\Delta_2=-5$. Now consider a loop contribution to a Green's function with n_i insertions of the vertices with dimension Δ_i. For $k\gg p_i$, the only scales that can appear are $k's$ and $\Delta's$. So by dimensional analysis, the superficial degree of divergence of the same integral changes as

$$\int k^D \to g_i^{n_i}\int k^{D-n_i\Delta_i} \tag{22.25}$$

Thus

$$D_{n,m}=4-\frac{3}{2}n-m-s-n_i\Delta_i \tag{22.26}$$

So if there are interactions with $\Delta_i<0$ then there can be an infinite number values of n and m, and therefore an infinite number of 1PI loops with $D>0$. Thus we will need an infinite number of counterterms to cancel all the infinities. Such theories are called non-renormalizable.

Non-renormalizable is equivalent to there being interactions with mass dimensions $\Delta_i<0$. On the other hand if all the interactions have mass dimension $\Delta_i>0$, then the theory is called *super-renormalizable*. ϕ^3 theory is a super-renormalizable theory.

We generalized this terminology also to describe individual interactions

- renormalizable interaction: coupling constant has dimension $\geqslant 0$
- non-renormalizable interaction: coupling constant has dimension <0

Also people sometimes describe interactions of dimension 0 as *marginal*, dimension > 0 as *relevant*, and dimension < 0 as *irrelevant*. These labels come from the renormalization group.

22.3.1 non-renormalizable theories are finite too

Non-renormalizable theories have an infinite number of superficially divergent integrals. This is due to the presence of couplings g_i with negative mass dimension. For example, $g = \frac{1}{M}$ for some scale M.

The divergent integrals can always be written as sums of terms of the form

$$\mathcal{I}_{\rm div} = (p_\mu \cdots p_\nu) g_1 \cdots g_n \int \frac{dk}{k^j} \tag{22.27}$$

for some number m of the various external momenta p_μ. This is in the region of loop momentum for which $k \gg p$ for all external momenta p. These integrals can produce logarithms of the regularization scale Λ, or powers of Λ

$$\mathcal{I}_{\rm div} = \sum g^n (p_\mu \cdots p_\nu)[c_0 \log\Lambda + c_1\Lambda + c_2\Lambda^2 + c_3\Lambda^3 + \cdots] \tag{22.28}$$

It is very important that there can never be terms like $\log p^2$ in the *divergent* part of the integral, that is, nothing like $\Lambda^2 \log p^2$ can appear. This is simply because we didn't have an $\log p^2$ terms to begin with and we can go to the divergent region of the integral by taking $k \gg p$ before integrating over anything that might give a log. It is not just $\Lambda^2 \log p^2$ terms that will never appear, but anything that is not analytic in the external momenta are forbidden as well.

Another way to see that the divergences are a polynomial in the external momenta comes from using the derivative trick to evaluate the integrals. Weinberg takes this approach in section 12.2. A general divergent integral will have various momenta factors in it, such as

$$\mathcal{I}(p) = \int_0^\infty \frac{k dk}{k + p} \tag{22.29}$$

There will always be at least one denominator with a factor of p. If we differentiate the integral with respect to p enough times, the integral becomes convergent

$$\mathcal{I}''(p) = \int_0^\infty \frac{2k dk}{(k+p)^3} = \frac{1}{p} \tag{22.30}$$

Thus we can then integrate over p to produce a polynomial, up to constants of integration

$$\mathcal{I}(p) = p \log \frac{p}{\Lambda_1} - p + \Lambda_2 = p \log p - p(\log \Lambda_1 + 1) + \Lambda_2 \tag{22.31}$$

So we do get non-analytic terms in the integral, but only polynomials in p can ever multiply the divergences.

The point is that polynomials in external momenta are exactly what we get at tree level from terms in the Lagrangian. That is, we can have counterterms for them. For example, suppose this integral $\mathcal{I}(p)$ were contribution to something like a ϕ^6 correlation function:

$$\langle 0 | T\{\phi(p_1) \cdots \phi(p)\} | 0 \rangle = -p \log \Lambda + \Lambda + \text{finite} \tag{22.32}$$

Then we could add terms

$$\mathcal{L}_{\rm new} = Z_5 (\partial\phi)\phi^5 + Z_6 \phi^6 = (\partial\phi)\phi^5 + \phi^6 + \delta_5 (\partial\phi)\phi^5 + \delta_6 \phi^6 \tag{22.33}$$

These new terms contribute to the 6 point function as

$$\langle 0 | T\{\phi(p_1) \cdots \phi(p_6)\} | 0 \rangle = p + 1 + p\delta_5 + \delta_6 \tag{22.34}$$

The counterterms can now be chosen to cancel the infinities: $\delta_5 = \log\Lambda$ and $\delta_6 = -\Lambda$. Then the Green's function is made finite. This can be done order by order in perturbation theory to cancel all of the infinities in the theory.

In order to have a counterterm, we need the corresponding term to actually be in our Lagrangian. So the easiest thing to do is just to add every possible term we can think of. But do we need terms like $\partial_x \phi^4$, which break Lorentz invariance? Obviously not, as we will never get Lorentz violating divergences in a Lorentz invariant theory. Clearly, we only need to add terms that preserve the symmetries of our theory. But, every possible term must be included, otherwise we cannot guarantee that all the infinities can be canceled. This can be subtle as symmetries are sometimes violated in the quantum theory or in the path integral measure, even though they are symmetries of the classical Lagrangian. Such violations are called *anomalies*, and we will get to them eventually.

In summary,

- non-renormalizable theories must include every term not forbidden by symmetries
- non-renormalizable theories need an infinite number of counterterms
- non-renormalizable theories are just as finite as renormalizable theories
- non-renormalizable theories can be renormalized. Their physical predictions are just as finite as in renormalizable theories.

The difference is that the infinities in renormalizable theories can be canceled with a finite number of counterterms, while an infinite number of counterterms is often necessary in non-renormalizable theories. However, non-renormalizable theories are still very predictive, as we will now see.

Chapter 23

Non-Renormalizable and Super-Renormalizable Field Theories

23.1 Introduction

Almost every quantum field theory we ever look at is non-renormalizable. This is simply because effective field theories are non-renormalizable, and every theory we know of is either an effective theory or can be treated as an effective theory for practical purposes. So you already know that these theories are wonderful for doing physics. We will look at 4 examples, corresponding to the 4 forces of nature: the Schrodinger equation (electromagnetism), the 4-Fermi theory (the weak force), the theory of mesons (the strong force), and quantum gravity. In each case we will see how the non-renormalizable theory is predictive despite the need for an infinite number of counterterms.

After studying some examples of non-renormalizable theories, we will look at the superrenormalizable theories. These theories all have couplings with positive mass dimension and have serious problems in making physical predictions..

23.2 Schrodinger Equation

Consider the Schrodinger equation

$$i\partial_t\psi = H\psi = \left[\frac{p^2}{2m} + V(r)\right]\psi \tag{23.1}$$

This is a non-renormalizable effective field theory. The parameter with negative mass-dimension is simply the coefficient of p^2: $\frac{1}{m}$.

Why are there not more terms of higher order in p in this equation? Well, there are, as we know from taking the non-relativistic limit of the Dirac equation

$$H = \frac{p^2}{2m}\left[1 + a_1\frac{p^2}{m^2} + a_2\frac{p^4}{m^4} + \cdots\right] + V(r) \tag{23.2}$$

All we are doing here is writing down the most general Hamiltonian consistent with rotational symmetry. The factors of m are just dimensional analysis. It turns out that $a_1 = -\frac{1}{4}$ for the Dirac equation, and the others as are also calculable.

The Schrodinger equation is useful even if we don't know about the Dirac equation or that $a_1 = -\frac{1}{4}$. In the non-relativistic limit $p \ll m$, the higher order terms have a very small effect. But they are measurable! a_1 contributes to the fine structure of the Hydrogen atom, and a_2 contributes to the hyperfine structure. It's nice that we can calculate a_1 and a_2 from the Dirac equation. But even if we didn't have the Dirac equation to calculate a_1 and a_2, we could have measured them form they Hydrogen atom and predicted things about Helium, or H_2 or lots of other things. So this non-renormalizable theory is very predictive indeed.

Also note that the Schrodinger equation is *not* predictive for energies $p \gtrsim m$. Then all of the higher order terms are equally important. So, the Schrodinger equation is predictive at low energy, but also indicates the scale at which perturbation theory breaks down. If we can find a theory which reduces to the Schrodinger equation at low energy, but for which perturbation theory still works at high energy, it is called *UV completion*. Thus, the Dirac equation is a UV completion of the Schrodinger equation. The Dirac equation (and QED) are predictive to much higher energies (but not at all energies, because of the Landau pole, not to mention the rest of the standard model and gravity). The Klein-Gordon equation is a different UV completion of the Schrodinger equation.

I want to emphasize that the Schrodinger equation is an extremely predictive *quantum* theory, independent of the Dirac equation. It made quantum predictions many years before the Dirac equation was discovered. So you already have seen how predictive a non-renormalizable theory can be, as long as we stay at low enough energy.

Let us now look at some other non-renormalizable quantum field theories and their UV completions.

23.3 The 4-Fermi theory

Weak decays were first modeled by Enrico Fermi. He observed that a proton can decay to a neutron, a positron and a neutrino, so he added an interaction to the free Lagrangian for these fields of the form

$$\mathcal{L}_{\text{fermi}} = G_F \bar{\psi}_p \psi_n \bar{\psi}_e \psi_\nu \tag{23.3}$$

with maybe some $\gamma'_\mu s$ or γ_5's thrown in. This is known as a 4-Fermi interaction, both because there are 4 fermions in it and also because Fermi used this term as a very successful model of radioactive decay

One prediction from this interaction is that the rate for β-decay $p^+ \to n e^+ \nu$ will be related to the rate for $n \to p^+ e^- \bar{\nu}$. Both of these reactions occur in nuclei. It also makes a prediction for the angular dependence and energy distribution of the decay products. In addition, the 4-Fermi theory can also be used to study parity violation, say, by comparing the predictions of this interaction to that of an interaction like $\bar{\psi}\gamma_5\gamma^\mu\psi\bar{\psi}\gamma_5\gamma^\mu\psi$.

The Fermi constant is

$$G_F = 1.166 \times 10^{-5} \text{GeV}^{-2} \sim \left(\frac{1}{300\,\text{GeV}}\right)^2 \tag{23.4}$$

Since G_F has mass dimension -2, this is a non-renormalizable interaction.

Since it is non-renormalizable, we should add all terms consistent with symmetry

$$\mathcal{L} = G_F \bar{\psi}\psi\bar{\psi}\psi + a_1 G_F^2 \bar{\psi}\psi\Box\bar{\psi}\psi + a_2 G_F^3 \bar{\psi}\psi\Box^2\bar{\psi}\psi + \cdots \tag{23.5}$$

We know the first term has coefficient G_F. Keep in mind, this is the renormalized value of the coupling, because it was extracted from a physical experiment. The $a'_i s$ are arbitrary, and the $G'_F s$ that accompany them have been added by dimensional analysis.

Despite these additional terms, the 4-Fermi theory is very predictive. For example, these higher order terms will effect the β-decay rate by factors of $(G_F E^2)^j$ where E is some energy in the process. Since the masses of the particles and energies involved in β-decay are much less than $G_F^{-1/2}$, these higher order terms will do practically nothing

Or, consider a low energy $\psi\psi \to \psi\psi$ scattering. The matrix element will get tree-level contributions from all of these terms, and we would find, in the s channel

$$\mathcal{M}(s) = G_F + a_1 G_F s + a_2 G_F s^2 + \cdots \tag{23.6}$$

No matter what a_1 is, for $s \ll (a_1 G_F)^{-1}$ this second term will be negligible. If it turns out that $a_1 \sim a_2 \sim 1$ then for $\sqrt{s} \ll 300$ GeV, the entire distribution is given practically by the first term, which is from the original 4-Fermi coupling (that's what happens in nature). On the other hand, if we go measure this scatting cross section as a function of s, and we find that it does depend on s, we can fit a_1 and a_2 to data. Then we have a more advanced theory, which is still very predictive. We only need to measure the cross section at 3 energies to fit G_F, a_1 and a_2, then we can say a lot about the detailed energy dependence, or other crossed (e.g. t-channel) scatterings, or decays. It is true we would need an infinite number of experiments to fit all the $a_j s$, but from a practical point of view, at low energies, this higher order $a_j s$ are totally irrelevant, just as the p^{10} term in the Schrodinger equation is irrelevant. All our discussion of renormalizability is secondary to the fact that non-renormalizable theories are extremely predictive at tree level, that is as *classical* field theories.

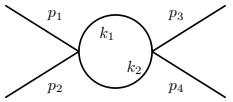
But what happens if $s \gtrsim \frac{1}{G_F}$? Then all orders in perturbation theory will contribute. Thus perturbation theory breaks down. So we need a UV completion to calculate scattering for $s > \frac{1}{G_F}$.

23.3.0.1 quantum predictions of the Fermi theory

Perhaps counter-intuitively, the 4-Fermi theory is predictive even at the quantum level. Many people often forget this fact!

Consider the loop correction to the 4-point function:

$$i\mathcal{M}_2(s) = \quad \sim G_F^2 \int \frac{d^4k}{(2\pi)^4} \frac{1}{\not{k}} \frac{1}{\not{k}} \sim G_F^2 (b_1 s + b_2 s \log\frac{\Lambda^2}{s} + b_3 \Lambda^2) \tag{23.7}$$



This graph is actually very similar to the vacuum polarization graph. We have just parametrized the possible form the result could take with 3 constants b_1, b_2, and b_3. Without any symmetry argument, like gauge invariance, there is no reason to expect that any of these constants should vanish.

The quadratically divergent part $b_3\Lambda^2$ is an (infinite) number, which gives an (infinite) correction to G_F. We can renormalize by writing

$$\mathcal{L} = Z_F \bar{\psi}\psi\bar{\psi}\psi \tag{23.8}$$

$$= G_F \bar{\psi}\psi\bar{\psi}\psi + \delta_F \bar{\psi}\psi\bar{\psi}\psi \tag{23.9}$$

and using the new counterterm δ_F to absorb the quadratically divergent correction.

$$\delta_F = -G_F^2 b_3 \Lambda^2 \tag{23.10}$$

In the same way the $b_2 \log\Lambda^2$ term would come from the renormalization of the a_1 term in the Lagrangian

$$\mathcal{L} = Z_{a_1} \bar{\psi}\psi \Box \bar{\psi}\psi \tag{23.11}$$

Thus we want $Z_{a_1} = a_1 - b_2 G_F^2 \log\frac{\Lambda^2}{\mu^2} + \cdots$ to cancel the $b_2 G_F^2 s \log\Lambda^2$ term from the loop. Then we get for the 4-point function

$$\mathcal{M}_{\rm tot}(s) = Z_F + Z_{a_1} s + \mathcal{M}_2(s) = G_F + G_F^2 \left(b_1 s + a_1 s + b_2 s \log\frac{\mu^2}{s} \right) \tag{23.12}$$

So we have absorbed the infinities into the renormalized values of G_F and a_1. Note that we needed the a_1 term in the Lagrangian to do this, which agrees with our general argument that all operators not forbidden by symmetry must be included.

The quantum prediction comes from the log term. This term is non-analytic in s, so no term in the Lagrangian could give a tree-level contribution of this form. The coefficient of the log, b_2 is a calculable number. Thus, for example, we can predict that

$$\frac{\mathcal{M}(s_1) - G_F}{G_F^2 s_1} - \frac{\mathcal{M}(s_2) - G_F}{G_F^2 s_2} = b_2 \log\frac{s_1}{s_2} \tag{23.13}$$

The left-hand side of this is measurable, so the right hand side is a genuine measurable prediction of the quantum theory. It is not sensitive to higher-order terms in the Lagrangian, since they would be suppressed parametrically by additional factors of $G_F s$.

23.3.0.2 UV-completing the Fermi theory

Today, we know that the 4-Fermi interaction comes from the the low-energy limit of a renormalizable theory with massive vector bosons. The actual 4-Fermi interaction has $\gamma^\mu s$ in it and looks like

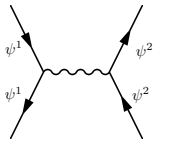
$$\mathcal{L} = G_F \bar{\psi}\gamma^\mu\psi\bar{\psi}\gamma_\mu\psi + a_1 G_F^2 \bar{\psi}\gamma^\mu\psi\Box\bar{\psi}\gamma_\mu\psi + a_2 G_F^3 \bar{\psi}\gamma^\mu\psi\Box^2\bar{\psi}\gamma_\mu\psi + \cdots \tag{23.14}$$

There are additional terms with $\gamma_5 s$, which are phenomenologically important, but irrelevant for the following discussion.

Consider the Lagrangian for a fermion interacting with a massive vector

$$\mathcal{L}_{\text{EW}} = -\frac{1}{4}F_{\mu\nu}^2 + \frac{1}{2}M^2 A_\mu^2 + \bar{\psi}\left(i\partial\!\!\!/ + g_w A_\mu\right)\psi \tag{23.15}$$

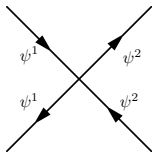
The vector is the Z or W boson, and the fermions are the various quarks and leptons. The tree level $\psi\psi \to \psi\psi$ matrix element in the s channel is given by

$i\mathcal{M} =$ 

$= (i g_w)^2 \bar{v}_1\gamma^\mu u_1 \dfrac{-i\left(g^{\mu\nu} + \frac{p^\mu p^\nu}{M^2}\right)}{s - M^2}\bar{u}_2\gamma^\mu v_2$ (23.16)

At low energy, $s \ll M$, this is well approximated by

$$\mathcal{M} = \frac{g_w^2}{M^2}\bar{v}_1\gamma^\mu u_1 \bar{u}_2\gamma^\mu v_2 = \tag{23.17}$$



This is the same matrix element as we would get from the 4-Fermi interaction if $G_F = \frac{g_w^2}{M^2}$. The actual expression for the Fermi constant in terms of the weak coupling constant and the W mass is

$$G_F = \frac{\sqrt{2}}{8}\frac{g_w^2}{m_W^2} \tag{23.18}$$

where $m_W = 80.4$ GeV and $g_w = 0.65$.

Another way to derive the 4-Fermi interaction from the electroweak Lagrangian is by integrating out the heavy vector. At energies much less than m_W, pair creation of W bosons is a very small effect. Then only tree-level diagrams involving the W are relevant and the W can be treated classically. The $W's$ equation of motion is

$$A^\mu = \frac{g_w}{\Box + M^2}\bar{\psi}\gamma^\mu\psi = \frac{g_w}{M^2}\left(1 - \frac{\Box}{M^2} + \cdots\right)\bar{\psi}\gamma^\mu\psi \tag{23.19}$$

Plugging back into the Lagrangian we find

$$\mathcal{L}_{\text{EW}} = -\frac{1}{2}A_\mu(\Box + M^2)A_\mu + g_w A_\mu\bar{\psi}\gamma^\mu\psi \tag{23.20}$$

$$= \frac{1}{2}\bar{\psi}\gamma^\mu\psi\frac{g_w^2}{\Box + M^2}\bar{\psi}\gamma^\mu\psi \tag{23.21}$$

$$= \frac{g_w^2}{M^2}\bar{\psi}\gamma^\mu\psi\bar{\psi}\gamma^\mu\psi - \frac{g_w}{M^2}\bar{\psi}\gamma^\mu\psi\frac{\Box}{M^2}\bar{\psi}\gamma^\mu\psi + \cdots \tag{23.22}$$

which includes the Fermi interaction. It also predicts that the a_1 term above is just $1/g_w^2$. So the theory with a massive vector boson is a UV completion of the 4-Fermi theory. Since we already observed that we can measure a_1, we might have used this to guess the UV completion were it not already known.

In conclusion, for non-renormalizable field theories

- Dependence on powers of external momenta can be fit to data and give hints about the UV completion.
- Non-analytic corrections are genuine predictions of the quantum theory
- The dimensionful coupling indicates a break down of perturbation theory at the scale of the coupling

By the way, the theory with the massive vector boson is renormalizable, but it has its own problems. Due to the massive vector propagator

$$\mathcal{M} \sim p^\nu \frac{i(g^{\mu\nu} - \frac{p^\mu p^\nu}{M^2})}{s - M^2} p^\mu \sim g^2 \frac{s - \frac{s^2}{M^2}}{s - M^2} \tag{23.23}$$

For large $s \gg M^2$, this goes like $g^2 s/M^2$ (this is due to exchange of the longitudinal polarization of the massive vector $\varepsilon_\mu \sim \frac{1}{M} k_\mu$). Loop corrections may go like $g^4 \frac{s^2}{M^4}$ and so on. Thus, even though the theory is renormalizable, perturbation theory breaks down at energies $\sqrt{s} \sim \frac{M}{g}$. For the W boson, this scale is ~ 1 TeV. We have no idea what will happen above that scale, and so we are building the LHC.

23.4 Theory of Mesons

The first field theoretic model of cuclear structure was conceived by Hideki Yukawa in 1934 (Nobel prize 1949). He noted that nuclear interactions seem to be confined within the nucleus, so the are very short range. Keep in mind, he was trying to explain why neutrons and protons stuck together, not anything have to do with the structure of the neutron or proton themselves, which were still thought to be point particles. The confusion in the 30s was whether the thing binding the neutrons and protons had anything to do with the thing that caused radioactive decay (the weak force). Yukawa was the first person to speculate that they are different. Actually, the more profound and lasting insight that he made was the connection between forces and virtual particle exchange. In 1934 people were still using old-fashioned perturbation theory, and nobody thought of virtual particles as actually existing.

We already know that the exchange of a massive particle leads not to a Coulomb potential but to a Yukawa potential

$$V(r) = \frac{1}{r} e^{-mr} \tag{23.24}$$

Yukawa saw at $m \sim 100$ MeV was the appropriate scale for nuclear interactions, and therefore postulated that there should be particles of mass intermediate between the nucleons (~ 1 GeV) and the electron ($\sim$ 1 MeV) and he called them *mesons*. We know define mesons as bosonic quark anti-quark bound states. The mesons responsible for the nuclear interactions are called pions.

Incidentally, the first meson was disovered in 1936 by Carl Anderson (Nobel 1936 for positron discovery in 1932). It had a mass of 100 MeV, very nearly what Yukawa predicted. However, this was the muon and not the pion. It took another 10 years, until 1947 for the pion to be discovered (Cecil Powell, Nobel 1950). Pions are strongly interacting and shorter lived than muons so they are harder to see. The confusion of the relationship between the cosmic ray that Anderson found and Yukawas theoretical prediction led to the rapid advancement of quantum field theory in the 30s and helped people to start taking virtual particles seriously.

The pions provide an effective description of nuclear forces. These days, we don't think pions are really responsible for nuclear forces. QCD is the real theory of the strong force. But it is very difficult to use QCD to study nuclear physics. Even the simple explanation of why the strong force is short range had to wait until asymptotic freedom was understood in the 1970's, 40 years after Yukawa's phenomenological explanation. The low energy theory of pions is known as the *chiral Lagrangian*, and it is a very powerful non-renormalizable field theory still being actively explored today.

The nucleons, more generally known as hadrons, include the proton and neutron, as well as the meson. There are three π mesons labelled by their electric charges: π^0, π^+ and π^-. We can guess that Lagrangian for the pions has normal kinetic terms and photon interactions from scalar QED:

$$\mathcal{L}_{\rm kin} = -\frac{1}{2}\pi^0(\Box + m_0^2)\pi^0 + (D_\mu\pi^+)^2 - \frac{1}{2}m_+^2(\pi^+)^2 + (D_\mu\pi^-)^2 - \frac{1}{2}m_-^2(\pi^-)^2 \tag{23.25}$$

The masses are of order $\sim 140\,\text{MeV}$. The pions should also interact with each other, with terms like

$$\mathcal{L}_{\rm int} = \frac{1}{f_\pi^2}\pi^0\pi^0\partial_\mu\pi^+\partial_\mu\pi^- + \frac{1}{f_\pi^4}(\pi^-\pi^+)^2\partial_\mu\pi^0\partial_\mu\pi^0 + \cdots \tag{23.26}$$

where f_π is some constant with dimensions of mass, making these interactions non-renormalizable. This is as far as we can get just using Lorentz invariance and gauge invariance to constrain the Lagrangian.

It turns out the interactions are strongly constrained by an additional symmetry, called chiral symmetry, which is related to the isospin symmetry which relates the neutron and the proton. We now know that this symmetry is ultimately due to the fact that all these particles are made up of up and down quarks u and d, which are almost massless, and the symmetry is an SU(2) symmetry which rotates them into each other. (There are actually 2 SU(2) symmetries, one for the left handed quarks and one for the right-handed quarks, and the chiral symmetry is one linear combination. But lets not worry about the details now.)

The point of this symmetry on the pions is that all the interaction can be expressed in terms of a single matrix

$$U = \exp\left[\frac{i}{f_\pi}\begin{pmatrix} \pi_0 & \pi_+ + i\pi_- \\ \pi_+ - i\pi_- & -\pi_0 \end{pmatrix}\right] = \exp\left[\frac{i}{f_\pi}\tau^a\pi^a\right] \tag{23.27}$$

where τ^a are the Pauli matrices. Then the simplest term we can write down involving U is

$$\mathcal{L} = \frac{f_\pi^2}{4}\text{Tr}\big[(D_\mu U)\big(D_\mu U^\dagger\big)\big] + \cdots \tag{23.28}$$

This is known as the *Chiral Lagrangian*.

This leading order term contains interactions, such as the ones in $\mathcal{L}_{\rm int}$ above. As well as being SU(2) invariant, all the interactions coming from this term must have 2 derivatives. So the interactions have a very special form. That form has been checked to great accuracy by low energy pion scattering experiments.

Since theory is non-renormalizable, we should also add more terms. These terms must always have derivatives acting on the $U's$, since $U^\dagger U = 1$. In fact, there are only 3 terms you can write down with 4-derivatives

$$\mathcal{L}_4 = L_1\text{Tr}[(D_\mu U)(D_\mu U^\dagger)]^2 + L_2\text{Tr}[(D_\mu U)(D_\nu U^\dagger)]^2 + L_3\text{Tr}\big[(D_\mu U)(D_\mu U^\dagger)(D_\nu U)(D_\nu U^\dagger)\big]$$

The coefficients of these terms can be fit from low energy pion scattering experiments. It has been found that $L_1 = 0.65$, $L_2 = 1.89$ and $L_3 = -3.06$. From the Lagrangian with these interactions we can say practically everything we would want about pion scattering. Additional interactions are suppressed by powers of momentum divided by the parameter $f_\pi \sim 92$. Actually if you are careful about factors of π you will find that this is a predictive theory for $E < 4\pi f_\pi \sim 1200$ GeV. So we are doing an expansion in $\frac{E}{4\pi f_\pi}$.

Moreover, the quantum effects are calculable and measurable as well. We are talking about the non-analytic logarithmic corrections. These have been worked out in great deal for the chiral Lagrangian and agree wonderfully with experiment.

As with any non-renormalizable theory, the chiral Lagrangian indicates that it becomes non-perturbative at a scale $\Lambda \sim 4\pi f_\pi$. Above this scale, all the higher order interactions become relevant and the theory is not predictive. The UV completion of the chiral Lagrangian is Quantum Chromodynamics, the theory of quarks and gluons. This is a completely different type of UV completion than the electroweak theory that UV completed the Fermi theory or the Dirac equation completing the Schrodinger equation. For both of these theories, the fermions in the low energy theory were present in the UV completion, but with different interactions. The theory of QCD does not have pions in it at all! Thus one cannot ask about pion scattering at high energy in QCD. Instead, one must try to match the two theories more indirectly, for example, through correlation functions of external currents. Experimentally, we do this by scattering photons or electrons off pions. Theoretically, we need a non-perturbative description of QCD, such as the lattice. Lattice QCD can show the existence of pions as poles in correlation functions. It can also give a pretty good estimate of f_π and other parameters of the chiral Lagrangian.

QCD is renormalizable. It is a UV completion of the chiral Lagrangian in the sense that it is well-defined and perturbative up to arbitrarily high energies. However, it cannot answer perturbative the questions that the chiral Lagrangian couldn't answer: what does $\pi\pi$ scattering look like for $s \gg f_\pi^2$? In fact, QCD doesn't even have asymptotic states. We cannot talk about an S-matrix of quarks and gluons, since quarks and gluons are confined. Thus we have to rephrase the questions we ask. Moreover, if we actually want to scatter pions, the chiral Lagrangian is much more useful than QCD: we can calculate things perturbatively.

23.5 Quantum Gravity

The final non-renormalizable field theory I want to discuss is Quantum Gravity. This is the effective description of a massless spin-2 particle. As we saw with Weinberg's soft graviton theorems, an interacting theory of a massless spin-2 field must have an associated conservation law. This is the conservation of momentum. The symmetry it corresponds to is symmetry under spacetime translations

$$x^\alpha \to x^\alpha + \xi^\alpha(x) \tag{23.29}$$

This generates the group of *general coordinate transformations*. The Noether current for this symmetry is the energy-momentum tensor T^α_μ.

The spin two field $h_{\mu\nu}$ is a gauge field for this symmetry. For infinitesimal transformation, it transforms as

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu + (\partial_\mu \xi^\alpha) h_{\alpha\nu} + (\partial_\nu \xi^\alpha) h_{\mu\alpha} + \xi^\alpha \partial_\alpha h_{\mu\nu} \tag{23.30}$$

The first two terms are the gauge part, they are the analog of

$$A_\mu \to A_\mu + \partial_\mu \alpha \tag{23.31}$$

but there are 4 $\alpha' s$ now, called ξ^α. The last three terms are just the transformation properties of a tensor representation of the Poincare group under translations. For example, a scalar transforms as

$$\phi(x) \to \phi(x+\xi) = \phi(x) + \xi^\alpha \partial_\alpha \phi(x) + \mathcal{O}(\xi^2) \tag{23.32}$$

Vectors and tensors get additional pieces because of their Lorentz properties. Requiring that $\partial_\mu V_\mu$ be general coordinate invariant implies that

$$V_\mu(x) \to \partial_\mu(x^\alpha + \xi^\alpha) V_\alpha(x+\xi) = V_\mu(x) + \xi^\alpha \partial_\alpha V_\mu(x) + (\partial_\mu \xi^\alpha) V_\alpha(x) \tag{23.33}$$

and so on. Note that transformation on h is non-linear: it has terms proportional to h itself.

I am particularly fond of this language. By Taylor expanding the transformation we never have to talk about x changing, just fields changing. x always refers to a space-time point, which is independent of any coordinate representation we might have for it.

In the same way that $F_{\mu\nu}^2$ with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the unique kinetic term we can write down for A_μ, there is a unique kinetic term we can write down for $h_{\mu\nu}$. It is

$$\mathcal{L}_{\rm kin} = \frac{1}{4}(\partial_\alpha h_{\mu\nu})(\partial_\alpha h_{\mu\nu}) - \frac{1}{2}(\partial_\nu h_{\beta\nu})(\partial_\mu h_{\beta\mu}) - \frac{1}{2}(\partial_\nu h_{\mu\nu})(\partial_\mu h_{\beta\beta}) + \frac{1}{2}(\partial_\mu h_{\alpha\alpha})(\partial_\mu h_{\beta\beta}) \tag{23.34}$$

You can check that this is invariant under the gauge symmetry, up to terms cubic in h. Those cubic terms can then be determined by demanding invariance under the full translation. We get a bunch of terms like

$$\mathcal{L}_{\rm int} = a_1 \frac{1}{M_P}(\partial_\mu h_{\alpha\beta})(\partial_\nu h_{\beta\alpha})h_{\mu\nu} + a_2 \frac{1}{M_P}(\partial_\alpha h_{\mu\nu})(\partial_\beta h_{\mu\alpha})h_{\nu\beta} + \cdots \tag{23.35}$$

Where we have added a scale M_P by dimensional analysis. You can continue, and work out what those cubic terms are, and then the quartic terms, and so on. This approach is done beautifully in Feynman's Lectures on Gravitation.

It turns out you can sum the whole series and write it in a compact form

$$\mathcal{L}_{\rm kin} + \mathcal{L}_{\rm int} = M_P^2 \sqrt{\det(g)} R_{\mu\mu} \tag{23.36}$$

where $g_{\mu\nu} = \eta_{\mu\nu} + \frac{1}{M_P} h_{\mu\nu}$, with $\eta = \mathrm{diag}(1,-1,-1,-1)$ the Minkowski metric. $R_{\mu\nu}$ is the Ricci curvature tensor for a Riemannian manifold based on the metric $g_{\mu\nu}$. This is the Einstein Lagrangian.

23.5.0.3 Predictions of gravity as a non-renormalizable field theory

In some sense the Riemann curvature tensor looks like

$$R_{\mu\nu\alpha\beta} \sim \partial_\mu \partial_\nu \exp\left(\frac{1}{M_P} h_{\alpha\beta}\right), \quad R_{\mu\nu} \sim R_{\mu\nu\alpha\alpha} \tag{23.37}$$

This is *very* heuristic, but shows that all the terms in the expansion of the curvature have two derivatives and lots of $h's$. So then $R_{\mu\mu} = \mathrm{Tr}[R_{\mu\nu}]$ becomes very similar to the form of the chiral Lagrangian $\mathrm{Tr}[D_\mu U D_\mu U^\dagger]$. Just like the chiral Lagrangian, each term has two derivatives, and the interactions are strongly constrained by symmetries. We could add higher order terms

$$\mathcal{L} = M_P^2 R_{\mu\mu} + L_1 R_{\alpha\alpha} R_{\beta\beta} + L_2 R_{\alpha\beta} R_{\alpha\beta} + L_3 R_{\alpha\beta\gamma\delta} R_{\alpha\beta\gamma\delta} \tag{23.38}$$

In this case, there are 3 terms, just like in the chiral Lagrangian. (Actually one is a total derivative, called the Gauss-Bonnet term, so there are really only 2). Since R has 2 derivatives, these all have 4 derivatives, so the leading interactions go like

$$\mathcal{L} = \cdots + L_i \frac{1}{M_P^3} \Box^2 h^3 \tag{23.39}$$

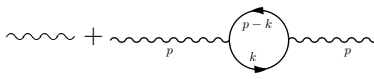
The reason gravity is predictive is because $M_P = 10^{19}$ GeV, so $E \ll M_P$ for any reasonable E. In fact, it is very difficult even to test the terms in the Lagrangian cubic in h with 2 derivatives. That is terms like $\sim \frac{a_1}{M_P} h^2 \Box h$ coming from $\sqrt{g}R$. For these, we don't get additional E/M_P suppression, but since the interaction has an extra factor of h/M_P, would need large field values $h \gtrsim M_P$. This happens, for example, from the gravitational field around the sun. There,

$$h \sim \phi_{\rm newton} \sim \frac{M_{\rm sun}}{M_P} \frac{1}{r} \sim 10^{38} \frac{1}{r} \tag{23.40}$$

By the way, $M_P \sim G_N^{-1/2}$, so $h \sim M_P$ for $r \sim G_N M_{\rm sun} \sim 100\,m$, which is still much less than the radius of the sun $r_{\rm sun} \sim 10^6$ km. Nevertheless, these cubic terms have been measured in the perihelion shift of Mercury. Mercury is about 10^7 km from the sun, so the effect is only one part in 10^8. Still it has been calculated precisely and measured and the agreement is perfect.

To see the effect of the higher order terms, like $L_{1,2,3}$, we would need to use interactions like $\frac{L_1}{M_P^3}\Box^2 h^3$. These have additional factors of $\frac{E}{M_P} \sim \frac{1}{M_P r}$ suppression. For mercury, this factor is 10^{-45}, so we would need to measure the perihelion shift to 1 part in 10^{-98} to see these terms. Needless to say, that is never going to happen. Thus the higher order terms in the Lagrangian are totally irrelevant for astrophysics.

So we can't measure the higher order terms. But there are still additional predictions we can make from the quantum theory. We can calculate a loop correction to graviton exchange. This is a correction to the graviton propagator, which has the same general form as the vacuum polarization graph



$$\sim \frac{1}{p^2}+\frac{1}{p^2}\left[\frac{p^4}{M_P^2}\left(c_1+c_2\log\frac{\Lambda^2}{p^2}\right)+c_3\frac{\Lambda^2 p^2}{M_P^2}\right]\frac{1}{p^2} \tag{23.41}$$

where the p^4/M_P^4 comes from the two vertices which have factors of $\frac{p^2}{M_P}$ in them. The rest is dimensional analysis. To be concrete

$$c_2=\frac{141}{30\pi^2} \tag{23.42}$$

You might have also expected a term in brackets going like Λ^4, but that would produce a double pole at $p=0$, and when the 1PI graphs are resummed it would shift the graviton mass. The mass is protected by gauge invariance, so this can't happen, just like for the photon. So,

$$\Pi_{\rm grav}=\frac{1}{p^2}\left[1+\frac{p^2}{M_P^2}(c_1+c_2\log\frac{\Lambda^2}{p^2})+c_3\frac{\Lambda^2}{M_P^2}\right] \tag{23.43}$$

This is exactly the same form that we got from the radiative correction in the Fermi theory above. The quadratic divergence c_3 is absorbed into the renormalization of M_P, The c_1 term is absorbed into counterterms for $L_{1,2,3}$. That leaves the logarithm with it's calculable coefficient.

If we Fourier transform back into position space, and put it all together we find the Newton potential becomes

$$h(r)=\frac{M_{\rm sun}}{M_P}\frac{1}{r}\left[1+\frac{c_2}{M_P^2 r^2}+\frac{L_1}{M_P^3 r^3}+\frac{M_{\rm sun}}{M_P}\left(\frac{a_1}{M_P r}+\frac{L_1}{M_P^3 r^3}\right)\right] \tag{23.44}$$

The first three terms come from $h\Box h$, the $\log p^2$ quantum correction, and the $L_1 h\Box^2 h$ term in Eq. (23.38). The final two are the $\Box h^3$ term from $\sqrt{g}R$ and the $L_1\Box^2 h^3$ term from Eq. (23.38). L_1 is a free parameter, but both c_2 and a_1 are known. Thus the radiative correction, the c_2 term, gives a calculable effect this is parametrically more important than the L_1 term. This term has not yet been observed, but it is a genuine prediction of quantum gravity. In fact, it is **the only prediction of quantum gravity** I know of. Don't forget: this prediction is entirely independent on the UV completion of the Einstein Lagrangian.

In any case, this discussion should make clear that

Gravity and quantum mechanics are completely compatible

There is **nothing** inconsistent about general relativity and quantum mechanics. Gravity is the only consistent theory of an interacting spin 2 particle. It is a quantum theory, just as solid and calculable as the 4-Fermi theory. It is just non-renormalizable, and therefore non-perturbative for energies $E\gtrsim M_P$, but it is not inconsistent. At distances $r\sim\frac{1}{M_P}\sim 10^{-33}$ cm, all of the quantum corrections, and all of the higher order terms in the Lagrangian become important. So if we want to use gravity at very short distances we need a UV completion.

String theory is one such theory. It is capable of calculating the L_i terms in Eq (23.38). It does this very indirectly. If we could measure the L_i, then we could test string theory, but as we noted above, we would need to improve our measurements by 90 orders of magnitude. In the same way that QCD is not a perturbative description of pions, string theory is not a perturbative theory of gravitons. However, we can prove in string theory that there is a massless spin-2 excitation, as we can (almost) prove that there are bound states in QCD. And since the only consistent theory of spin-2 particles is the Einstein Lagrangian plus higher order terms, since string theory is consistent, it must match on to Einstein gravity at low energy. However, unlike QCD, string theory does not (yet) have a non-perturbative formulation so we don't know how to define it, or even if it exists!

23.6 Summary of non-renormalizable theories

We have looked at 4 important non-renormalizable theories.

- The Schrodinger equation is perturbative for $E < m_e$. It's UV completion is the Dirac equation and QED, which is perturbative up to its Landau pole, $E \sim 10^{100}$ GeV.
- The Fermi theory of weak interactions is perturbative for $E < G_F^{-1/2} \sim 300\,\text{GeV}$. It's UV completion is the electroweak theory with massive vector bosons W and Z. The electroweak theory is renormalizable, but only perturbative up to $E \sim 1$ TeV. One possible UV completion of the electroweak theory is to add a Higgs boson.
- The chiral Lagrangian is the low energy theory of pions. It is perturbative and very predictive for $E < 4\pi f_\pi \sim 1200$ MeV. It's UV completion is QCD. QCD is predictive at high energies, but it has no asymptotic states. At low energy, QCD must be studied non-perturbatively on the lattice. So, at low energy the chiral Lagrangian which is perturbative, is much more useful.
- Einstein quantum gravity is the low energy theory of gravity. It is perturbative for $E < M_P \sim 10^{19}$ GeV. It is extremely predictive at low energies, including predictive quantum corrections. One possible UV completion is string theory.

Incidentally, these 4 examples correspond to the 4 forces of nature: the electromagnetic force, the weak force, the strong force, and gravity. Notice that the UV completions are all qualitatively very different. In some cases, certainly for many physical applications, the non-renormalizable theory is more useful than the renormalizable one. Renormalizable just means there are a finite number of counterterms, it does *not* mean that you can calculate every observable perturbatively.

I want to emphasize again: non-renormalizable theories are very predictive, not just at tree-level (classically) but also through quantum effects. The quantum effects are calculable and non-analytic in momentum space. In fact, non-renormalizable theories are so predictive that it is often better to make a non-renormalizable approximation to a renormalizable theory than to use the full renormalizable theory itself. Examples include Einstein gravity instead of string theory, Heavy Quark Effective Theory (HQET) instead of QCD, Ginsburg-Landau theories in condensed matter, the Schrodinger equation instead of the Dirac equation, and chemistry instead of physics.

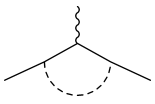
23.7 Quantum predictions which are not logs

We have seen that non-renormalizable field theories are predictive at the classical level, but also at the quantum level through terms like $\log p^2$. Is it possible that we could also compute corrections which are analytic in p^2?

We already know of one example. Recall that in QED we were able to calculate a finite correction to the magnetic moment of the electron:

$$\mu_e = \frac{e}{2m_e}\Big[2 + \frac{\alpha}{\pi} + \mathcal{O}(\alpha^2)\Big] \tag{23.45}$$

What happens to this prediction if we embed QED in a larger theory, such as quantum gravity? The contribution of gravity to the electron magnetic moment is divergent:



$$\sim \frac{m_e}{M_P^2}\log\frac{\Lambda^2}{m_e^2} + b_\sigma \frac{m_e}{M_P^2} \tag{23.46}$$

where b_σ is finite and Λ is the UV cutoff. (You don't get a linearly divergent piece because $\int \frac{k^\mu}{k^4} d^4k = 0$. Also, for a magnetic moment, you always a mass insertion because $\bar{\psi}\sigma^{\mu\nu}\psi$ couples left and right handed spinors.)

So we need another term in the Lagrangian to absorb the infinity:

$$\mathcal{L}_\sigma = Z_\sigma \bar{\psi}\sigma^{\mu\nu}\psi F_{\mu\nu} \tag{23.47}$$

As usual, we split Z_σ into a finite part c_σ and counterterm: $Z_\sigma = c_\sigma + \delta_\sigma$ with $\delta_\sigma = -\frac{m_e}{M_P^2}\log\frac{\Lambda}{m_e}$. Then

$$\mu_e = \frac{e}{2m_e}\Big[2+\frac{\alpha}{\pi}\Big] + c_\sigma + b_\sigma\frac{m_e}{M_P^2} \tag{23.48}$$

Now we need a renormalization condition to fix c_σ. If we demand that the measured value of the magnetic moment is $\mu_e^{\text{experiment}}$, then

$$c_\sigma = \mu_e^{\text{experiment}} - \frac{e}{2m}\Big[2+\frac{\alpha}{\pi}\Big] - b_\sigma\frac{m_e}{M_P^2} \tag{23.49}$$

$$\Rightarrow \mu_e = \mu_e^{\text{experiment}} \tag{23.50}$$

But this implies that the $\frac{\alpha}{\pi}$ part we started with is totally unobservable. But that doesn't make any sense, since we have observed it! So it's a paradox!

23.7.1 from physics to philosophy

The resolution of this paradox lies outside the realm of physics. I'm going to start italicizing words that are more related to philosophy then experiment, and I will try to be very clear about distinguishing the two.

One way to think about it is to suppose that the *ultimate theory of nature* were finite. Then we don't need the counterterm, and can just compute the loop in gravity. We would find

$$\mu_e = \frac{e}{2m_e}\Big[2+\frac{\alpha}{\pi}\Big] + \frac{m_e}{M_P^2}\log\frac{\Lambda^2}{m_e^2} + b_\sigma\frac{m_e}{M_P^2} \tag{23.51}$$

where now Λ is some physical cutoff scale. We need at least $\Lambda > M_P$. For example, in string theory, $\Lambda \sim \frac{1}{g_{\text{string}}}M_P$. b_σ is calculable, and probably not much larger than 10. Thus as long as

$$\Lambda < m_e\exp(\frac{M_P^2}{m_e^2}) \sim 10^{10^{44}}\,\text{eV} \tag{23.52}$$

the effect of gravity on g-2 is completely negligible.

What if the *ultimate theory of nature* is simply renormalizable? For example, imagine that the graviton is a bound state in some exotic theory (like string theory), as the pions are bound states of QCD. Then there are no coefficients of positive mass dimension and g-2 will come out finite. Thus the effective value of Λ would be around the compositeness scale of the graviton, which would *likely* be around M_P. For $\Lambda \sim M_P$, the contribution to g-2 of the graviton is still exponentially small.

Another way to think about the paradox is to suppose that we actually had measured some deviation of g-2 from it's value in QED (or in the renormalizable standard model). For example, there is a little room for this at order α^3 and a better measurement could conceivable reveal a deviation. If you are real positivist you can then just add $\mathcal{L}_\sigma$ and tune the renormalized value of g-2 to match experiment. From a practical point of view that would mean that $c_\sigma \sim \alpha^3\frac{1}{m_e} \sim \frac{1}{1300\,\text{GeV}}$. So we've introduced a new scale in our theory, and made it non-renormalizable. It is unlikely, but not impossible, that this effect comes from quantum gravity. For that to happen, we would have needed $\Lambda \sim 10^{10^{44}}$ eV. Moreover, if $\Lambda \sim 10^{10^{45}}$ eV the effect would have been so enormous that the whole universe would have been paramagnetic.

So let's add this operator to account for the $\sim\alpha^3$ deviation. That works fine for g-2, but it screws up everything else. That is, this new operator would contribute to other observables, and we would need more higher dimension operators, etc. Now we can keep choosing the coefficients of those operators to cancel the contribution from this one and agree with the predictions of the standard model for everything. So nothing is guaranteed to happen from this term. (Actually, there will be calculable quantum effects of this term. But since these *must* show up, they don't tell us anything new.) It makes more sense to *expect* deviations to show up elsewhere. These deviations are *motivated* by the g-2 observation, but they are not guaranteed to show up. By looking for those deviations, we can fit more higher-dimension operators, and this may help us guess the UV completion.

In summary, looking for deviations in finite predictions of the standard model is a clue to new physics. Finding such a such a deviation *probably* indicates a new scale where we might find other new effects, but nothing is guaranteed. In the context of a non-renormalizable theory, there are infinite contributions to the same quantity. That formally makes the theory unpredictive for this quantity. However in practice, with any reasonable finite cutoff, or through an embedding into a renormalizable theory, the infinite correction from the non-renormalizable terms becomes calculable and small. The modern *viewpoint* is to think of the higher dimension operators in the non-renormalizable theory as encoding information about the UV completion.

23.8 Super-renormalizable field theories

Now let us turn to theories with coupling constants of positive mass dimension. In 4 dimensions there aren't many options.

23.8.1 tadpoles

One possibility is that we could have a term linear in a scalar field. For example,

$$\mathcal{L} = -\frac{1}{2}\phi(\Box + m^2)\phi + \Lambda^3\phi \tag{23.53}$$

for some number Λ^3 with mass-dimension 3, this would generate $\langle 0|\phi|0\rangle = \Lambda^3 \neq 0$. But also contribute to things like $\langle 0|T\{\phi^4\}|0\rangle$, and everything else. The way to deal with this term is to redefine $\phi \to \phi + \phi_0$, where $\phi_0 = \frac{\Lambda^3}{m^2}$. Then we remove the tadpole. So the existence of a term like $\Lambda^3\phi$ in the Lagrangian means we have been working with the wrong degrees of freedom. That's a serious problem.

The only possible term of mass dimension 4 a constant

$$\mathcal{L} = \cdots + \rho \tag{23.54}$$

This constant has a name: ρ is the *cosmological constant*. By itself, this term does nothing. It couples to nothing and in fact just be pulled out of the path integral. The reason it is dangerous is because when one couples to gravity, it becomes

$$\rho \to \sqrt{g}\,\rho = \frac{1}{2}h_{\mu\mu}\rho + h^2_{\mu\nu}\rho + \cdots \tag{23.55}$$

The first term generates a tadpole diagram, so $\langle 0|h_{\mu\mu}|0\rangle = \frac{1}{2}\rho \neq 0$. This indicates that we are expanding around the wrong vacuum. By redefining $h_{\mu\nu} \to h_{\mu\nu} + h^0_{\mu\nu}$ for some constant but x-dependent field $h^0_{\mu\nu}(x)$, we can remove this term. We know it has to be x-dependent because all the terms in the Einstein Lagrangian have derivatives, so they will kill any spacetime-independent $h^0_{\mu\nu}$. That has huge consequences. For example, it violates Lorentz invariance. In the end, it preserves some other symmetry group, the de-Sitter group or anti de-Sitter group, depending on the sign of ρ. Still, if this term is around, we have to go back and rethink all the results we had in quantum field theory that depend on Lorentz invariance.

So far, these are pretty awful things. Other dimensionful couplings include mass terms, like $m^2\phi^2$, $m^2A^2_\mu$ or $m\bar{\psi}\psi$, or perhaps a cubic coupling in a scalar field theory, $g\phi^3$. That actually exhausts the possibilities in 4 dimensions. There are only a handful of super-renormalizable terms.

23.8.2 relevant interactions

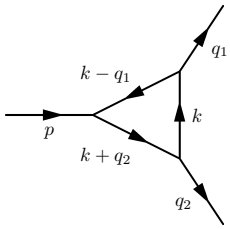
Let's look at ϕ^3 theory first, then the mass terms. We are interested in radiative corrections, so let's couple ϕ to something else, say a fermion ψ. We will set the mass of ϕ to zero for simplicity. The Lagrangian is

$$\mathcal{L} = -\frac{1}{2}\phi\Box\phi + \frac{g}{3!}\phi^3 + \lambda\phi\bar{\psi}\psi + \bar{\psi}\,(i\not{\partial} - M)\,\psi \tag{23.56}$$

Now consider the renormalization of g. We can compute this by evaluating the 3-point function

$$\langle 0|T\{\phi(p)\phi(q_1)\phi(q_2)\}|0\rangle = g + \cdots \tag{23.57}$$

There is a radiative correction from the loop of ϕ. This is UV finite, proportional to g^3 and not particularly interesting. A more important radiative correction comes from the loop of the fermion.

$\mathcal{M}_2 =$ 

$$= \lambda^3 \int \frac{d^4k}{(2\pi)^4} \mathrm{Tr}\left[\frac{1}{\not{k} - M}\frac{1}{\not{k} - \not{q}_1 - M}\frac{1}{\not{k} + \not{q}_2 - M}\right] \tag{23.58}$$

$$= \frac{\lambda^3}{8\pi^2} M \int_0^1 dz \int_0^{1-z} dy \left[\frac{M^2 + p^2(y + z - 3yz - \frac{1}{2})}{M^2 - p^2 yz} + 3\log\frac{\Lambda^2}{M^2 - p^2 yz}\right] \tag{23.59}$$

Taking $p^2 = 0$, this gives

$$\langle 0|T\{\phi(p)\phi(q_1)\phi(q_2)\}|0\rangle = g + \frac{\lambda^3}{16\pi^2} M\left(1 + 3\log\frac{\Lambda^2}{M^2}\right) \tag{23.60}$$

So we find a (divergent) shift in g proportional to the mass of the fermion M. This is fine if $M \sim g$, but if parametrically it's really weird. As M gets larger, the correction grows. That means that the theory is sensitive to scales much larger than the scale associated with g. So,

- Super-renormalizable theories are sensitive to UV physics

Is this a problem? Not in the technical sense, because the correction can be absorbed in a renormalization of g. We just add a counterterm by writing

$$g = g_R - \frac{\lambda^3}{16\pi^2} M\left(1 + 3\log\frac{\Lambda^2}{M^2}\right) \tag{23.61}$$

This will cancel the radiative correction exactly (at $p = 0$). Again, this is totally reasonable if $g_R \sim M$, and theoretically and experimentally consistent for all g, but if $g_R \ll M$ it *looks strange*.

Will there be problems due to the quantum corrections? We can look at the Green's function at a different value of p^2. This would contribute to something like the scalar Coulomb potential. We find that of, course, the divergence cancels, and the result at small p is something like

$$\langle 0|T\{\phi(p)\phi(q_1)\phi(q_2)\}|0\rangle \approx g_R + \frac{\lambda^3}{16\pi^2}\left[\frac{p^2}{M} + 3M\log\left(1 - \frac{p^2}{M^2}\right) + \cdots\right] \tag{23.62}$$

$$\approx g_R + \frac{\lambda^3}{16\pi^2}\frac{p^2}{M} \tag{23.63}$$

So for $p \ll M$ these corrections are tiny and do in fact decouple as $M \to \infty$.

As for the non-renormalizable case, we enjoy speculating that the *ultimate theory of nature* is finite. Then Λ is physical. If we suppose that there is some bare g which is fixed, then we need a very special value of Λ to get the measured coupling to be $g_R \ll M$. Of course, there is some solution for Λ, for any g_0 and g_R. But if we change Λ by just a little bit, then we get

$$g_R \to g_R + \frac{\lambda^3}{16\pi^2} M \log\frac{\Lambda^2_{\rm old}}{\Lambda^2_{\rm new}} \tag{23.64}$$

which is a huge correction if $M \gg g_R$.

This means that generically the theory *wants* to have $g_R \sim M$. It is *unnatural* to have a relevant coupling in a quantum field theory.

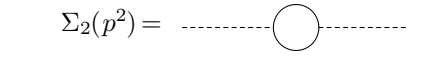
23.8.3 scalar masses

The only other possible terms of positive mass dimension in a Lagrangian are masses. First, consider a scalar mass. Let's go back to our Yukawa theory and drop that ridiculous ϕ^3 term.

$$\mathcal{L} = -\frac{1}{2}\phi(\Box + m^2)\phi + \lambda\phi\bar{\psi}\psi + \bar{\psi}\,(i\not{\partial} - M)\psi \tag{23.65}$$

Now consider the radiative corrections to m.

We need to compute the self-energy graph for the scalar (see P&S 10.33)

$\Sigma_2(p^2) =$  $= 4\frac{\lambda^2(d-1)}{(4\pi)^{d/2}}\Gamma(1-\frac{d}{2})\mu^{4-\varepsilon}\int_0^1 dx(M^2 - x(1-x)p^2)^{d/2-1}$ (23.66)

This is quadratically divergent, as evidenced by the pole at $d = 2$. If we expand for small p, we find

$$\Sigma_2(p^2) = \frac{\lambda^2}{4\pi^2}\left[\frac{p^2 - 6M^2}{\varepsilon} - \frac{1}{2}(p^2 - 6M^2)\log\frac{M^2}{\mu^2} + p^2 - 5M^2 + \frac{p^4}{20M^2} + \cdots\right] \tag{23.67}$$

It is somewhat nicer to use a regulator which makes the UV divergences explicit. Then we would have

$$\Sigma_2(p^2) = \frac{\lambda^2}{4\pi^2}\left[\Lambda^2 + (2p^2 + 3M^2)\log\frac{\Lambda^2}{M^2} + 4p^2 + 5M^2 + 6\frac{p^4}{M^2} + \cdots\right] \tag{23.68}$$

where I have just made up the coefficients of these terms.

The divergences are canceled with counterterms from the field strength and mass renormalizations giving

 $= -p^2\delta_\phi - \delta_m$ (23.69)

We can now use physical renormalization conditions to set the pole of the propagator at the renormalized mass, with residue 1. Then,

$$\Sigma_2(m^2) = m^2\delta_\phi + \delta_m \tag{23.70}$$

$$\Sigma_2'(m^2) = \delta_\phi \tag{23.71}$$

Explicitly,

$$\delta_\phi = \frac{\lambda^2}{4\pi^2}\left[(2m^2)\log\frac{\Lambda^2}{M^2} + 4 + 12\frac{m^2}{M^2} + \cdots\right] \tag{23.72}$$

$$\delta_m = \frac{\lambda^2}{4\pi^2}\left[\Lambda^2 + 3M^2\log\frac{\Lambda^2}{M^2} + 5M^2 + 6\frac{m^4 - 2m^2M^2}{M^2} + \cdots\right] \tag{23.73}$$

$$\Rightarrow \Sigma(p^2) = \Sigma_2(p^2) + m^2\delta_\phi + \delta_m = \frac{\lambda^2}{4\pi^2}\left[\frac{1}{20}\frac{(p^2 - m^2)^2}{M^2}\right] \tag{23.74}$$

Note that δ_m is the shift in the mass. It is getting not only a quadratically divergent correction, but also a correction proportional to M^2, which can be huge if $M \gg m$.

Suppose we used minimal subtraction instead. Call the renormalized mass m_R, to distinguish from the pole mass m above.

$$\delta_\phi = \frac{\lambda^2}{4\pi^2}\left[2\log\frac{\Lambda^2}{M^2}\right] \tag{23.75}$$

$$\delta_m = \frac{\lambda^2}{4\pi^2}\left[\Lambda^2 + 3M^2\log\frac{\Lambda^2}{M^2}\right] \tag{23.76}$$

$$\Rightarrow \Sigma(p^2) = \Sigma_2(p^2) + m_R^2\delta_\phi + \delta_m = \frac{\lambda^2}{4\pi^2}\left[4p^2 + 5M^2 + 6\frac{p^4}{M^2} + \cdots\right] \tag{23.77}$$

Now consider the location of the pole in this case. The propagator is

$$\Pi = \frac{i}{p^2 - m_R^2 - \Sigma(p^2)} \tag{23.78}$$

So the pole is at

$$0 = m_P^2 - m_R^2 - \Sigma(m_P^2) \tag{23.79}$$

$$\Rightarrow m_R^2 = m_P^2 - \frac{\lambda^2}{4\pi^2}\left[4m_P^2 + 5M^2 + 6\frac{m_P^4}{M^2} + \cdots\right] \tag{23.80}$$

If $M \sim m_P$ then $m_R \sim m_P$. But if $M \gg m_P$ then there is an enormous difference between the masses defined in these two subtraction schemes. In either scheme, the large corrections are absorbed into the physical mass, defined as either the pole mass or the $\overline{\text{MS}}$ mass. So this difference is not observable. Yet it is *disturbing*.

Just as for the ϕ^3 theory, we can see that the mass here is UV sensitive. Suppose the theory really is finite. Then to get $m \ll M$ we need

$$\delta_m = \frac{\lambda^2}{4\pi^2}\left[\Lambda^2 + 3M^2 \log\frac{\Lambda^2}{M^2}\right] \sim m^2 \ll M^2 \tag{23.81}$$

This requires a careful selection of Λ with respect to M, to a part in $\frac{M^2}{m^2}$. Or suppose we add a new fermion to the theory of mass $M_{\text{new}} \sim M$. This one will also contribute, and so the special value of Λ that allows for this cancellation will have to change. Either way something *fishy* is going on in the UV theory that allows $m \ll M$.

An important example of an unnatural scalar mass is the Higgs boson mass in the standard model. We have direct and indirect bounds on the Higgs boson implying $114\,\text{GeV} < m_h < 196\,\text{GeV}$. However, the top quark is around 170 GeV, so it gives a large radiative correction to the Higgs mass. As only the renormalized Higgs pole mass is actually observable, this radiative correction is not in itself physical. But if one imagines a finite UV completion for the standard model at very high energies, say $\Lambda \sim M_P$, then parameters have to be tuned to one part in $\frac{m_{\text{higgs}}^2}{M_P^2} \sim 10^{-40}$ for the Higgs mass to come out right. Again, this tuning is not a problem in the standard model itself, only when the standard model is embedded in a larger theory. Since we *believe* that there should be physics beyond the standard model, in particular there's gravity, then getting a light Higgs mass becomes an extraordinary practical challenge because of this fine-tuning problem.

23.8.4 fermion and gauge boson masses

What about fermions and gauge boson masses? Why doesn't the electron pick up a large mass due to other couplings in the theory?

We already calculated the self-energy graph of the electron. We found for the counterterm with the pole-mass renormalization scheme

$$\delta m = \frac{\alpha}{4\pi} m_e \int_0^1 dx (4 - 2x) \log\frac{x\Lambda^2}{(1-x)^2 m_e^2 + x\, m_\gamma^2} \tag{23.82}$$

The $\overline{\text{MS}}$ bass correction is also proportional to m_e.

First of all, note that the correction is not quadratically or linearly divergent. That is a non-trivial fact. In non-relativistic quantum mechanics, you do get a linearly divergent shift. This can be seen from a simple integral over the classical electron-self energy. In the non-relativistic limit, the energy density of electromagnetic field is $\rho = |\vec{E}|^2 + |\vec{B}|^2$. So

$$\Delta m \sim \int d^3r\, |E|^2 = \int d^3r \left(\frac{e}{r^2}\right)^2 \sim \alpha \int d^3r \frac{e^2}{r^4} \sim \frac{\alpha}{r} \sim \alpha\Lambda \tag{23.83}$$

Thus the quantum theory is canceling something.

More importantly, the shift is proportional to the electron mass, not the other mass going around the loop. In this case, the a fictitious photon mass, but if the photon were a real heavy gauge boson, like the Z, the correction would still be proportional to m_e not m_Z. What if we throw some more fermions in, or a scalar or two, and look at 6 loops. It turns out that the mass shift will still be proportional to m_e. Thus heavy fields do decouple for the electron self-energy, and its mass is not UV sensitive.

The reason this happens is because the electron mass is protected by a custodial chiral symmetry. The chiral symmetry is a global symmetry acting as

$$\psi \to e^{i\alpha\gamma_5}\psi \tag{23.84}$$

Thus,

$$\bar{\psi}\not{\partial}\psi \to \psi^\dagger e^{-i\alpha\gamma_5^\dagger}\gamma_0\not{\partial}e^{i\alpha\gamma_5}\psi = \bar{\psi}\not{\partial}\psi \tag{23.85}$$

since $\gamma_5^\dagger = \gamma_5$ and $[\gamma_5, \gamma_0\gamma^\mu] = 0$. However,

$$m\bar{\psi}\psi \to m\bar{\psi}e^{2i\alpha\gamma_5}\psi \neq m\bar{\psi}\psi \tag{23.86}$$

Thus the mass term breaks the chiral symmetry. Another way to see this is that the symmetry rotates left and right handed fields into each other, but in the absence of a mass, the left and right-handed fields decouple.

The chiral symmetry is called custodial because it exists in the limit $m \to 0$. That means that if $m = 0$ then m will stay 0 to all orders in perturbations theory. Therefore, all violations of chiral symmetry will be proportional to the symmetry breaking parameter. So the chiral symmetry acts like a custodian and protects the mass. If we think of the mass term as an interaction rather than a kinetic term, every diagram that violates the custodial symmetry, including a correction to the mass itself, must be proportional to the mass. This is true whether or not you know about the symmetry. In practice, that's quite a useful thing because if you find something is canceling, or always proportional to something else, you can go back and look for the symmetry. But you don't need to know about the symmetry ahead of time to get the cancellation.

We have another way of describing this phenomenon. We say that the setting $m_e = 0$ is technically natural.

- Technically natural: corrections to a parameter are proportional to the parameter itself

This idea is usually attributed to Gerard 't Hooft.

The same thing holds for a photon mass. The term

$$\mathcal{L} = \cdots + m^2 A_\mu^2 \tag{23.87}$$

breaks gauge invariance. Gauge invariance is another custodial symmetry. Thus any contribution to the gauge boson mass will be proportional to the mass itself. This is only true if the *only* term that breaks gauge invariance is the mass term. If there are other interactions breaking gauge invariance, the mass correction can be proportional to them as well.

23.8.5 summary

We have seen that super-renormalizable interactions are not super at all. The cosmological constant causes a violation of Lorentz invariance (not to mention the collapse of the universe if $\Lambda \sim M_P$). A $g\phi^3$ interaction in a scalar field theory and a scalar mass $m^2\phi^2$ are both sensitive to UV physics. That is, if we change the UV theory for example by adding another massive fermion, then the bare (unphysical) parameters in our Lagrangian have to change in a very specific way. UV sensitivity means sensitivity to playing with the model, and is not physically observable directly. It violates our intuition that heavy fields should decouple.

The other super-renormalizable terms we could write down are fermion and vector boson masses, both of which are protected by custodial symmetries. Although they are super-renormalizable terms, they are not UV sensitive.

Theories with relevant parameters appear fine-tuned. The word *fine-tuned* comes from condensed matter. In condensed matter physics it is possible to construct super-renormalizable field theories by playing with external fields. For example, if you dial the magnetic field in such a way that a system becomes close to a phase transition, there can be long range correlations. These long range correlations are indicative of a light, almost massless pseudo-particle. It is much lighter than characteristic scales in the system; one normally expects a system to communicate only over atomic distances. This is exactly the situation in which $m \ll \frac{1}{a_0}$. You achieve this by carefully tuning the external magnetic field.

In nature, we might find a light scalar. The Higgs is one possibility. Perhaps $m_h \ll M_P$. If all we find is the Higgs, then we have no indication of any scale of new physics up to M_P. So the theory would be tuned, but we would never know who is tuning the dial of the magnetic field. And we would have no way to find out!

A fun way to think about the development of our understanding of the standard model is

- 4-Fermi theory. This theory is **Non-renormalizable** and breaks down at $E \sim 100\,\text{GeV}$
- Its UV completion is the standard model with W and Z bosons. This theory is **Renormalizable**, but still breaks down at $E \sim 1\,\text{TeV}$.
- Its UV completion is the standard model with a Higgs. This theory is **Super-renormalizable**. It does not break down, but is UV sensitive.

We don't know if there is really anything wrong with the standard model with a Higgs. But we haven't seen the Higgs. We haven't ever seen any *fundamental* scalar in fact. So, perhaps we're missing something and super-renormalizable theories aren't even consistent. In that case, we are back to the standard model with the W and Z bosons. So we must find either the Higgs or evidence of some other UV completion at the LHC.

Chapter 24
Implications of Unitarity

24.1 Introduction

In a quantum theory, we expect probabilities to add up to 1. That is, we expect to get out as much stuff as we put in. This means that the norm of a state $|s\rangle$ at time $t=0$ should be the same at a later time t.

$$\langle s, t=0|s, t=0\rangle = \langle s, t|s, t\rangle \tag{24.1}$$

This means the Hamiltonian should be Hermetian, $H^\dagger = H$, because

$$|s, t\rangle = e^{iHt}|s, 0\rangle \tag{24.2}$$

It also means the S-matrix should be unitary, since $S = e^{iHt}$. So

$$S^\dagger S = 1 \tag{24.3}$$

This equation has remarkable practical consequences in quantum field theory, as we will now explore.

24.2 The Optical Theorem

One of the most important implications of unitarity is the optical theorem. Write

$$S = \mathbb{1} + iT \tag{24.4}$$

Recall the matrix elements we have been calculating with Feynman graphs are matrix elements of this transfer matrix T with a factor of overall momentum conservation stripped off.

$$\langle f|T|i\rangle = (2\pi)^4\delta^4(p_i - p_f)\mathcal{M}(i \to f) \tag{24.5}$$

The matrix T is not Hermetian. In fact, unitarity implies

$$1 = S^\dagger S = (1 - iT^\dagger)(1 + iT) \tag{24.6}$$

$$\Rightarrow i(T^\dagger - T) = T^\dagger T \tag{24.7}$$

The left-hand side is

$$\big\langle f\big|i(T^\dagger - T)\big|i\big\rangle = i(2\pi)^4\delta^4(p_i - p_f)\big(\mathcal{M}^\dagger(i \to f) - \mathcal{M}(f \to i)\big) \tag{24.8}$$

Summing over intermediate states, the right-hand side is

$$\big\langle f|T^\dagger T|i\big\rangle = \sum_X d\Pi^{(X)}_{\rm LIPS}\big\langle f|T^\dagger|X\big\rangle\langle X|T|i\rangle \tag{24.9}$$

$$= \big[i(2\pi)^4\delta^4(p_f - p_X)\big]\big[i(2\pi)^4\delta^4(p_i - p_X)\big]\sum_X \int d\Pi^{(X)}_{\rm LIPS}\mathcal{M}(i \to X)\mathcal{M}^\dagger(X \to f) \tag{24.10}$$

Where $d\Pi_{\rm LIPS}$ is the phase space factor we need to insure that we've only inserted the operator $\mathbb{1}$. This is the same phase space factor that appears in differential cross-sections.

Then unitarity implies

$$\boxed{\mathcal{M}(i\to f)-\mathcal{M}^\dagger(f\to i)=-i\sum_X\int\Pi^{(X)}_{\rm LIPS}(2\pi)^4\delta^4(p_i-p_X)\mathcal{M}(i\to X)\mathcal{M}^\dagger(X\to f)} \tag{24.11}$$

This is the Generalized Optical Theorem.

Note that they must work order by order in perturbation theory. But while the left hand side is matrix elements, the right hand side is matrix elements squared. This means that at order λ^2 in some coupling the left hand side must be a loop to match a tree-level calculation on the right hand side. Thus, the imaginary parts of loop amplitudes are determined completely by tree-level amplitudes. In particular, we must have the loops – the classical theory by itself, without loops, violates unitarity.

The most important case is when $|i\rangle=|f\rangle=|A\rangle$ for some state A. Then

$$2i\,{\rm Im}\mathcal{M}(A\to A)=-i\sum_X\int\Pi^{(X)}_{\rm LIPS}(2\pi)^4\delta^4(p_i-p_X)|\mathcal{M}(A\to X)|^2 \tag{24.12}$$

In particular, when $|A\rangle$ is a 1-particle state, then the total decay rate is

$$\Gamma(A\to X)=\frac{1}{2m_A}\int\Pi^{(X)}_{\rm LIPS}|\mathcal{M}(A\to X)|^2 \tag{24.13}$$

So,

$$\boxed{{\rm Im}\mathcal{M}(A\to A)=m_A\sum_X\Gamma(\phi\to X)} \tag{24.14}$$

Here $\mathcal{M}(A\to A)$ is just the 2-point function, aka the propagator. So the imaginary part of the propagator is equal to the sum of the decay rates into every possible particle.

If $|A\rangle$ is a 2-particle state, then the cross section is

$$\sigma(A\to X)=\frac{1}{2E_{\rm CM}|\vec{p}_{\rm CM}|}\int\Pi^{(X)}_{\rm LIPS}|\mathcal{M}(A\to X)|^2 \tag{24.15}$$

So,

$$\boxed{{\rm Im}\,\mathcal{M}(A\to A)=2E_{\rm CM}|\vec{p}_{\rm CM}|\sum_X\sigma(A\to X)} \tag{24.16}$$

This says that the imaginary part of the forward scattering amplitude is proportional to the total scattering cross section, which is the optical theorem from optics.

Both of these examples have nice and interesting applications.

24.3 Decay rates

Let's start by thinking about the 1-particle case. Then $|i\rangle=|f\rangle=|\phi\rangle$. The tree-level contribution to $\mathcal{M}(\phi\to\phi)$ is just the propagator

$$i\mathcal{M}(\phi\to\phi)=\frac{i}{p^2-m^2+i\epsilon} \tag{24.17}$$

The imaginary part of a propagator is

$${\rm Im}\frac{1}{p^2-m^2+i\epsilon}=\frac{1}{2i}\left(\frac{1}{p^2-m^2+i\epsilon}-\frac{1}{p^2-m^2-i\epsilon}\right)=-\frac{\epsilon}{(p^2-m^2)^2+\epsilon^2} \tag{24.18}$$

This vanishes as $\epsilon\to0$, except near $p^2=m^2$. If we integrate over p^2, we find

$$\int_{-\infty}^{\infty}\frac{\epsilon}{(p^2-m^2)^2+\epsilon^2}dp^2=-\pi \tag{24.19}$$

Implying that

$$\boxed{\mathrm{Im}\frac{1}{p^2-m^2+i\epsilon}=-\pi\delta(p^2-m^2)} \tag{24.20}$$

this is a useful formula. It says that the propagator is real except for when the particle goes on-shell.

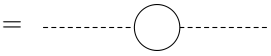
Let's check the optical theorem at 1-loop. To be specific, we'll take ϕ^3 theory but with two fields ϕ and χ with interaction

$$\mathcal{L}=-\frac{1}{2}\phi(\Box+M^2)\phi-\frac{1}{2}\chi(\Box+m^2)\chi+\frac{\lambda}{2}\phi\chi^2 \tag{24.21}$$

For $M>2m$ we should get a contribution to the imaginary part of the propagator corresponding to decays of $\phi\to 2\chi$, as given by the optical theorem.

The 1-loop amplitude is

$$i\,\mathcal{M}(p^2)= \tag{24.22}$$



$$\mathcal{M}=i\lambda^2\int\frac{d^4k}{(2\pi)^4}\frac{1}{[(k+p)^2-m^2+i\varepsilon][k^2-m^2+i\varepsilon]} \tag{24.23}$$

$$=\frac{g^2}{16\pi^2}\int_0^1 dx\log\frac{m^2-M^2x(1-x)}{\Lambda^2} \tag{24.24}$$

where we have set $p^2=M^2$ to put ϕ on shell – after all, we are computing S-matrix elements. Now, $x(1-x)<\frac{1}{4}$, so for $M<2m$ this expression is real, and therefore $\mathrm{Im}\mathcal{M}=0$. For $M>2m$, we use

$$\log(-x)=\log x+i\pi \tag{24.25}$$

Then,

$$\mathrm{Im}\mathcal{M}=\frac{g^2}{16\pi}\int_0^1 dx\theta(M^2x(1-x)-m^2) \tag{24.26}$$

$$=\frac{g^2}{16\pi}\sqrt{1-4\frac{m^2}{M^2}}\theta(M-2m) \tag{24.27}$$

The decay rate is

$$\Gamma=\frac{g^2}{16\pi M}\sqrt{1-4\frac{m^2}{M^2}}\theta(M-2m) \tag{24.28}$$

So,

$$\mathrm{Im}\mathcal{M}(\phi\to\phi)=M\Gamma(\phi\to X) \tag{24.29}$$

As desired. Thus the optical theorem works.

24.4 Cutting Rules

We can understand this calculation a little more clearly by looking back at the original integral.

$$\mathcal{M}=i\lambda^2\int\frac{d^4k}{(2\pi)^4}\frac{1}{[(k+p)^2-m^2+i\varepsilon][k^2-m^2+i\varepsilon]} \tag{24.30}$$

Because of the i, we want the real part of the product of two propagators. That means either both real parts, or both imaginary parts.

We saw above that the propagators only have imaginary parts when they go on-shell. If both parts are real, then they are both off-shell, and so there are no poles. Then we can drop the $i\varepsilon$, and after a Wick-rotation, we will pick up a factor of i. So the result is both real and imaginary, meaning it's zero. Thus the contribution to the imaginary part when both propagators are zero vanishes. We conclude that both propagators must by imaginary.

Therefore we can evaluate this by replacing

$$\frac{1}{(k+p)^2-m^2+i\varepsilon} \to \mathrm{Im}[\frac{1}{(k+p)^2-m^2+i\varepsilon}] = -\pi\delta((k+p)^2-m^2) \tag{24.31}$$

And

$$\frac{1}{k^2-m^2+i\varepsilon} \to -\pi\delta(k^2-m^2) \tag{24.32}$$

So we get

$$\mathrm{Im}\mathcal{M} = \pi^2\lambda^2 \int \frac{d^4k}{(2\pi)^4}\delta((k+p)^2-m^2)\delta(k^2-m^2) \tag{24.33}$$

Now write

$$q_1 = \frac{p}{2}+k \tag{24.34}$$

$$q_2 = \frac{p}{2}-k \tag{24.35}$$

And this is

$$\mathrm{Im}\mathcal{M} = \frac{\lambda^2}{2}\int \frac{d^3q_1}{(2\pi)^3}\frac{1}{2m}\int \frac{d^3q_2}{(2\pi)^3}\frac{1}{2m}(2\pi)^4\delta^4(q_1+q_2-p) \tag{24.36}$$

$$= M\Gamma(\phi \to X) \tag{24.37}$$

Which agrees with the optical theorem again.

This result is very general:

- Imaginary parts of loop amplitudes arise when intermediate states go on-shell together

It's easiest to see this as an implication of the optical theorem, rather than using it to derive the optical theorem as we have done. This statement is almost the definition of the optical theorem: the decay rate comes from on-shell final states. These states also appear as virtual states in loop momentum. Each way of putting intermediate states on-shell is known as a *cut*, after Cutkosky, and this method for computing the imaginary part of scattering diagrams is known as the *cutting rules*.

24.5 Propagators and polarization sums

The cutting rules work for scalars, as we have shown, but also for fermions and vectors, etc.. They must work for any state simply as a consequence of unitarity.

For the decay rate, we don't use the propagators of the outgoing particles. For example, take Yukawa theory

$$\mathcal{L}_{\mathrm{higgs}} = -\frac{1}{2}\phi(\Box+m_h^2)\phi + \lambda\phi\bar{\psi}\psi \tag{24.38}$$

Then for the decay of ϕ into $\bar{\psi}\psi$, we would have

$$\Gamma(\phi(p) \to \bar{\psi}(q_1)\psi(q_2)) = \sum_s \lambda^2(\bar{v}_s u_s)^2 = \lambda^2\mathrm{Tr}[(\not{q}_1+m)(\not{q}_2-m)] \tag{24.39}$$

The loop is

$$\int \frac{d^4k}{(2\pi)^4}\frac{\mathrm{Tr}[(\frac{1}{2}\not{p}+\not{k}+m)(\frac{1}{2}\not{p}-\not{k}+m)]}{[(\frac{1}{2}p-k)^2-m^2+i\varepsilon][(\frac{1}{2}p+k)^2-m^2+i\varepsilon]} \tag{24.40}$$

For the imaginary part of the loop, we have to put the intermediate states on shell. The denominators give us the δ-functions, just like for the scalar case. The numerators give

$$\mathrm{Tr}[(\not{q}_1+m)(\not{q}_2-m)] \tag{24.41}$$

Which is exactly what we got for the spin sum. This works because the spin sum is the same as the numerator of the propagator. This must hold, and is a consequence of unitarity.

For a massive spin 1 field, we have

$$\sum_{i=1}^{3} \varepsilon_\mu^i \varepsilon_\nu^{i\star} = -\eta_{\mu\nu} + \frac{p^\mu p^\nu}{m^2} \tag{24.42}$$

And the propagator is

$$\Pi_{\mu\nu}(p^2) = \frac{-\eta_{\mu\nu} + \frac{p^\mu p^\nu}{m^2}}{p^2 + i\varepsilon} \tag{24.43}$$

So the numerator is again given by the sum over physical spin states. Thus the optical theorem holds.

What about a massless spin-1 field? There, the spin sum includes only transverse polarizations. There is no way to write the sum in a Lorentz invariant way, but we can write it as

$$\sum_{i=1}^{2} \varepsilon_\mu^i \varepsilon_\nu^{i\star} = -\eta_{\mu\nu} + \frac{1}{2E^2}(p^\mu \tilde{p}^\nu + \tilde{p}^\nu p^\mu) \tag{24.44}$$

in terms of the two unphysical polarizations

$$\varepsilon_f^\mu = p^\nu = (E, \vec{p}\,) \tag{24.45}$$

$$\varepsilon_b^\mu = \tilde{p}^\mu = (E, -\vec{p}\,) \tag{24.46}$$

The propagator is

$$\frac{-\eta_{\mu\nu} + (1-\xi)\frac{p^\mu p^\nu}{p^2}}{p^2 + i\varepsilon} \tag{24.47}$$

So the numerator of the propagator is *not* just the sum over physical polarizations.

However, because of gauge invariance (for the propagator) and the Ward identity (for the decay rate), all the p^μ terms drop out in physical calculations. Thus we see that gauge invariance and the Ward identity are tied together, and moreover, both are required for a unitary theory of a massless spin 1 particle. That is

- unitarity for massless spin 1 fields requires gauge invariance

The same analysis can be made for massive and massless spin 2. It is not terribly illuminating, so we will not bother with it here.

In general, we can therefore write

$$\Pi(p^2) = \frac{i\sum_s \varepsilon_s \varepsilon_s^\dagger}{p^2 - m^2} \tag{24.48}$$

where ε_s are a basis of polarizations for a particle of arbitrary spin.

24.6 Breit-Wigner

We showed that the propagator to all-orders can be written as

$$i\Pi(p^2) = \frac{1}{p^2 - m^2 + M(p^2)} \tag{24.49}$$

where the on-shell renormalization conditions imply $M(m^2) = M'(m^2) = 0$. Now that we see that $i\Pi(m^2)$ has a substantial imaginary part, proportional to the decay rate. Thus $M(p^2)$ cannot be purely real. The counterterms, however, were real, so we cannot completely absorb $M(m^2)$ into the mass and field strength renormalization.

There are a number of consequences of this. The most obvious one is that a pole mass is not well-defined for an unstable particle. That's ok, because unstable particles don't really appear as asymptotic states in the S-matrix anyway. We can instead make a convention for defining the *real pole mass* by

$$\mathrm{Re}[M(m^2)] = 0 \tag{24.50}$$

Then

$$i\Pi(p^2) = \frac{1}{p^2 - m^2 + \mathrm{Im}\, M(p^2)} \tag{24.51}$$

The optical theorem tells us that $\mathrm{Im}\, M(m^2) = im\Gamma$. So if the state is close to on-shell, $p^2 \sim m^2$ and $\Gamma \ll m$ this is well approximated by,

$$\Pi(p^2) = \frac{i}{p^2 - m^2 + im\Gamma}, \quad p^2 \sim m^2, \quad \Gamma \ll m \tag{24.52}$$

If we plop this propagator into an s-channel diagram, we get a Breit-Wigner resonance:

$$d\sigma \sim |\frac{i}{p^2 - m^2 + im\Gamma}|^2 = \frac{1}{(q^2 - m^2)^2 + (m\Gamma)^2} \tag{24.53}$$

If Γ is large, then the width is broad, and we will need the full momentum dependence of $\mathcal{M}(p^2)$, which we have from computing the full self-energy graph.

As $\Gamma \to 0$, this resonance is well approximated by a δ-function

$$|\frac{i}{p^2 - m^2 + im\Gamma}|^2 \to \frac{\pi}{M\Gamma}\delta(p^2 - m^2) \tag{24.54}$$

Say we have an s-channel process close to a resonance, then the intermediate state is close to on shell. So we can write

$$d\sigma = |J_1 \frac{i\sum_s \varepsilon_s \varepsilon_s^\dagger}{p^2 - m^2 + im\Gamma} J_2|^2 \tag{24.55}$$

where J_1 and J_2 are whatever produced the state and whatever happened to it. If we are within the width of the resonance, $p^2 - m^2 \ll m\Gamma$, then

$$d\sigma \sim \frac{\pi}{M\Gamma}\delta(p^2 - m^2) \sum_s |J_1 \epsilon_s|^2 |\epsilon_s J_2|^2 \tag{24.56}$$

$$\sim \frac{\pi}{M}\delta(p^2 - m^2) \sum_s |J_1 \epsilon_s|^2 \tag{24.57}$$

That is, the total cross section is given by the production cross section. This is known as the *narrow width approximation*. For resonance production, the cross-section for production is independent of how the particle decays.

For example, suppose the interactions between the currents and the resonance are proportional to some coupling e. Naively, we would have expected

$$d\sigma \sim |J_1 \frac{i\sum_s \varepsilon_s \varepsilon_s^\dagger}{p^2 - m^2 + im\Gamma} J_2|^2 \sim e^4 \tag{24.58}$$

But the narrow width approximation implies that near a resonance, the cross-section goes as e^2. Thus if we are trying to produce a particle, it is very helpful to work near resonance, since the cross-section will be tremendously enhanced.

24.7 Cross sections: The Froissart Bound

For a 2 particle initial state, $|A_2\rangle = |p_1 p_2\rangle$, the optical theorem becomes

$$\mathrm{Im}\, \mathcal{M}(A_2 \to A_2) = 2E_{\mathrm{CM}} |\vec{p}_{\mathrm{CM}}| \sum_X \sigma(A_2 \to X) \tag{24.59}$$

which says that the imaginary part of the matrix element for forward scattering is proportional to the cross section for $A \to$ anything.

As a rough estimate, the total cross section is at least as big as the cross section for forward scattering,. So,

$$\mathrm{Im}\,\mathcal{M}(A_2\to A_2)=2E^2\sigma_{\rm tot}>2E^2\frac{1}{64\pi^2E^2}|\mathcal{M}(A_2\to A_2)|^2 \tag{24.60}$$

So,

$$|\mathcal{M}(A_2\to A_2)|<64\pi^2 \tag{24.61}$$

Thus the matrix element for forward scattering is bounded.

Let's go even further and look at two body final states in general, so $X=X_2=|p_3p_4\rangle$. Then we can use our formula

$$\sigma(A_2\to X_2)=\frac{1}{64\pi^2E_{\rm CM}^2}\int d\Omega|\mathcal{M}(A_2\to X_2)|^2 \tag{24.62}$$

to get

$$\mathrm{Im}\,\mathcal{M}(A_2\to A_2)=2E_{\rm CM}|\vec{p}\,|\sigma_{\rm tot}\geqslant 2E_{\rm CM}^2\,\sigma(A_2\to X_2)=\frac{1}{32\pi^2}\int d\Omega|\mathcal{M}(A_2\to X_2)|^2 \tag{24.63}$$

Now let's write in general the partial wave decomposition

$$\mathcal{M}(A_2\to X_2)=16\pi\sum\,(2n+1)P_n(\cos\theta)a_n \tag{24.64}$$

Where P_n are the Legendre polynomials. They satisfy

$$P_n(1)=1 \tag{24.65}$$

$$\int_{-1}^{1}P_n(x)P_m(x)dx=\frac{2}{2n+1}\delta_{nm} \tag{24.66}$$

So,

$$\mathrm{Im}\mathcal{M}(A_2\to A_2)=8\pi\sum_n\,(2n+1)\mathrm{Im}\,a_n \tag{24.67}$$

$$\int d\Omega|\mathcal{M}(A_2\to X_2)|^2=(8\pi)^2 4\pi\sum_n\,(2n+1)|a_n|^2 \tag{24.68}$$

Thus,

$$|a_n|^2\leqslant\mathrm{Im}\,a_n\leqslant|a_n| \tag{24.69}$$

In particular

$$\boxed{|a_n|<1} \tag{24.70}$$

Thus the partial wave amplitudes are bounded. The matrix elements are bounded

$$\int d\cos\theta P_n(\cos\theta)\mathcal{M}(A_2\to X_2)=16\pi a_n<16\pi \tag{24.71}$$

This also means that

$$\boxed{\text{Cross sections cannot be arbitrarily large}}$$

The precise bound is that for any process, there exists a constant c such that

$$\sigma_{\rm tot}(s)<c\log^2 s \tag{24.72}$$

for all s. This is known as the Froissart bound, and is a consequence of unitarity.

The partial wave unitary bound, or the Froissart bound, provide extremely important limitations on the behavior of scattering amplitudes. For example, suppose we have a theory with higher-derivative interaction, like

$$\mathcal{L}=\phi\Box\phi+\frac{1}{\Lambda}\phi^2\Box\phi \tag{24.73}$$

The s channel exchange diagram gives

$$\mathcal{M}(\psi\psi\to\psi\psi)=$$  $$=\frac{p^2}{\Lambda}\frac{1}{p^2}\frac{p^2}{\Lambda}=\frac{s}{\Lambda^2} \tag{24.74}$$

This has no angular dependence, so for $s > 16\pi\Lambda$ it violates the partial unitarity bound. Thus this diagram cannot represent the physics of this process for $s > 16\pi\Lambda$. Of course, we already knew that because this is a non-renormalizable theory, and so loops become important around that scale too. But the bound also implies that the loops *must* be important. The perturbation expansion simply breaks down.

The same thing happens with W exchange around a TeV. Due to the $\frac{E}{m}$ dependence of the longitudinal polarizations of the W, the amplitude for W scattering violates the unitary bound at $\sim 1\,\text{TeV}$. That means the standard model does not describe physics perturbative at a TeV without a Higgs. The Higgs restores perturbative unitarity.

Keep in mind, this is not a statement that unitarity is violated in these theories. It says that unitary *would be violated*, if we could trust perturbation theory, which we can't. The standard resolution is to introduce new particles or to look for a UV completion of these models above that scale.