## Contents

0.1 Introductory remarks ........................................... 4  
0.2 Sources and acknowledgement ................................ 5  
0.3 Conventions ...................................................... 6

1 The path integral makes some things easy .................. 7  
1.1 From particles to fields to particles again ................. 7  
1.2 Fields mediate forces ......................................... 13  
1.3 Euclidean path integral and Wick rotation ................. 17  
1.4 Feynman diagrams from the path integral ................... 21  
1.5 Lagrangian field theory ....................................... 29

2 From correlation functions to the $S$ matrix ............... 35

3 QED  ................................................................. 49  
3.1 Quantum light: Photons ...................................... 49  
3.2 More on vector fields ........................................... 51  
3.3 On the non-perturbative proof of the Ward identity ....... 57  
3.4 Feynman rules for QED ......................................... 58  
3.5 QED processes at leading order ............................... 63

4 To infinity and beyond .......................................... 73  
4.1 Casimir effect: vacuum energy is real ....................... 74  
4.2 A parable from quantum mechanics on the breaking of scale invariance ........................................... 77  
4.3 A simple example of perturbative renormalization in QFT .... 85  
4.4 Radiative corrections to the Mott formula .................. 88  
4.5 Electron self-energy in QED .................................... 90  
4.6 Big picture interlude .......................................... 96  
4.7 Vertex correction in QED ...................................... 104  
4.8 Vacuum polarization ........................................... 116

5 Consequences of unitarity ....................................... 125  
5.1 Spectral density ................................................. 125  
5.2 Cutting rules and optical theorem ............................ 132  
5.3 How to study hadrons with perturbative QCD ............... 138

6 Gauge theory ....................................................... 141  
6.1 Massive vector fields as gauge fields ....................... 141  
6.2 Festival of gauge invariance ................................... 143  
6.3 Interlude on differential forms and algebraic topology .... 146
0.1 Introductory remarks

Quantum field theory (QFT) is the quantum mechanics of extensive degrees of freedom. What I mean by this is that at each point of space, there’s some stuff that can wiggle.

It’s not surprising that QFT is so useful, since this situation happens all over the place. Some examples of ‘stuff’ are: the atoms in a solid, or the electrons in those atoms, or the spins of those electrons. A less obvious, but more visible, example is the electromagnetic field, even in vacuum. More examples are provided by other excitations of the vacuum, and it will be our job here to understand those very electrons and atoms that make up a solid in these terms. The vacuum has other less-long-lasting excitations which are described by the Standard Model of particle physics.

Some examples of QFT are Lorentz invariant (‘relativistic’). That’s a nice simplification when it happens. Indeed this seems to happen in particle physics. We’re going to focus on this case for much of this quarter. Still I would like to emphasize: though some of the most successful applications of QFT are in the domain of high energy particle physics, this is not a class on that subject, and I will look for opportunities to emphasize the universality of QFT.

A consequence of relativity is that the number of particles isn’t fixed. That is: there are processes where the number of particles changes in time. This is a crucial point of departure for QFT. It’s a necessary consequence of Lorentz symmetry, but the converse is false: particle production can happen without relativity.

‘Divergences’. Another intrinsic and famous feature of QFT discernible from the definition I gave above is its flirtation with infinity. I said that there is ‘stuff at each point of space’; how much stuff is that? Well, there are two senses in which ‘the number of points of space’ is infinite: (1) space can go on forever (the infrared (IR)), and (2) in the continuum, in between any two points of space are more points (the ultraviolet (UV)). The former may be familiar from statistical mechanics, where it is associated with the thermodynamic limit, which is where interesting things happen. For our own safety, we’ll begin our discussion in a padded room, protected on both sides from the terrors of the infinite.

Prof. Jenkins tells me that 215A ended just as Feynman diagrams were being drawn for the first time. I think therefore that it will be useful to retreat a bit and rederive the diagrammatic expansion from another (in many ways simpler) point of view, namely the path integral.
0.2 Sources and acknowledgement

The material in these notes is collected from many places, among which I should mention in particular the following:

- Peskin and Schroeder, *An introduction to quantum field theory* (Wiley)
- Banks, *Modern Quantum Field Theory: A Concise Introduction* (Cambridge)
- Schwartz, *Quantum field theory and the standard model* (Cambridge)
- David Tong’s lecture notes

Many other bits of wisdom come from the Berkeley QFT courses of Prof. L. Hall and Prof. M. Halpern.
0.3 Conventions

Following most QFT books, I am going to use the $+−−−$ signature convention for the Minkowski metric. I am used to the other convention, where time is the weird one, so I’ll need your help checking my signs. More explicitly, denoting a small spacetime displacement as $dx^\mu \equiv (dt, d\vec{x})^\mu$, the Lorentz-invariant distance is:

$$ds^2 = +dt^2 - d\vec{x} \cdot d\vec{x} = \eta_{\mu\nu} dx^\mu dx^\nu$$

with $\eta^{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}_\mu^\nu$.

(spacelike is negative). We will also write $\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \left(\partial_t, \vec{\nabla}_x\right)_\mu$, and $\partial^\mu \equiv \eta^{\mu\nu} \partial_\nu$. I’ll use $\mu, \nu$, ... for Lorentz indices, and $i, k, ...$ for spatial indices.

The convention that repeated indices are summed is always in effect unless otherwise indicated.

$D$ is the number of spacetime dimensions, $d$ is the number of space dimensions.

A consequence of the fact that english and math are written from left to right is that time goes to the left.

A useful generalization of the shorthand $\hbar \equiv \frac{\hbar}{2\pi}$ is $d\delta \equiv \frac{dk}{2\pi}$. I will also write $\delta^d(q) \equiv (2\pi)^d \delta^{(d)}(q)$. I will try to be consistent about writing Fourier transforms as

$$\int \frac{d^d k}{(2\pi)^d} e^{ikx} \tilde{f}(k) \equiv \int d^d k \ e^{ikx} \tilde{f}(k) \equiv f(x).$$

IFF $\equiv$ if and only if.

RHS $\equiv$ right-hand side. LHS $\equiv$ left-hand side. BHS $\equiv$ both-hand side.

IBP $\equiv$ integration by parts. WLOG $\equiv$ without loss of generality.

$+\mathcal{O}(x^n) \equiv$ plus terms which go like $x^n$ (and higher powers) when $x$ is small.

$+h.c. \equiv$ plus hermitian conjugate.

$L \supset \mathcal{O}$ means the object $L$ contains the term $\mathcal{O}$.

We work in units where $\hbar$ and the speed of light, $c$, are equal to one unless otherwise noted. When I say ‘Peskin’ I usually mean ‘Peskin & Schroeder’.

Please tell me if you find typos or errors or violations of the rules above.
1 The path integral makes some things easy

1.1 From particles to fields to particles again

Here is a way to discover QFT starting with some prosaic ingredients.

Consider a linear chain of particles of mass $m$, each connected to its neighbors by springs with spring constant $\kappa$. This is a model of a (one-dimensional) crystalline solid. When in equilibrium, the masses form a regular one-dimensional crystal lattice (equally spaced mass points). Now let $q_n$ denote the displacement of the $n$th mass from its equilibrium position $x_n$ and let $p_n$ be the corresponding momentum. Assume there are $N$ masses and (for simplicity) impose periodic boundary conditions: $q_{n+N} = q_n$. The equilibrium positions themselves are

$$x_n = na, n = 1, 2...N$$

where $a$ is the lattice spacing. The Hamiltonian for the collection of particles is:

$$H = \sum_{n=1}^{N} \left( \frac{p_n^2}{2m} + \frac{1}{2} \kappa (q_n - q_{n-1})^2 \right) + \lambda q^4.$$  \hspace{1cm} (1.1)

Notice that this system is an ordinary QM system, made of particles. In particular, the whole story below will take place within the fixed Hilbert space of the positions of the $N$ particles.

I’ve included a token anharmonic term $\lambda q^4$ to remind us that we are leaving stuff out; for example we might worry whether we could use this model to describe melting.

Set $\lambda = 0$ for a while. With $\lambda = 0$, the hamiltonian above describes a collection of coupled harmonic oscillators, with a matrix of spring constants $V = k_{ab} q_a q_b$. If we diagonalize the matrix of spring constants, we will have a description in terms of decoupled oscillators, called normal modes. Because the chain is translation invariant, the normal modes are labelled by a wavenumber $k$, and the eigenvalues are $\omega_k^2 = \frac{2\kappa}{m} \sin^2 \frac{ka}{2}$. Then we can use our knowledge of the spectrum of a single SHO $H = \hbar \omega (a^\dagger a + \frac{1}{2})$ to construct the whole spectrum of excitations of the chain,

$$H = \sum_k \hbar \omega_k \left( a_k^\dagger a_k + \frac{1}{2} \right) + \frac{p_0^2}{2m}.$$ (Here $p_0$ is the center-of-mass momentum of the chain.) The groundstate is $|0\rangle$, the state annihilated by all the annihilation operators $a_k |0\rangle = 0$, and excited states are built like $|k_1, k_2\rangle = a_{k_1}^\dagger a_{k_2}^\dagger |0\rangle$. In the context of an elastic solid, these excitations are called phonons.
Instead, let’s use the path integral.

---

**Path integral reminder in a box.**

Let’s remind ourselves how the path integral formulation of QM works for a particle in one dimension with $H = \frac{p^2}{2m} + V(q)$. The basic statement is the following formula for the propagator — the amplitude to propagate from position eigenstate $|q_0\rangle$ to position eigenstate $|q\rangle$ during a time interval $t$ is

$$
\langle q | e^{-iHt} | q_0 \rangle = \int_{q(0)=q_0}^{q(t)=q} [dq] e^{\int_{t_0}^t dt \left( \frac{1}{2} \dot{q}^2 - V(q) \right)}.
$$

Here $[dq] \equiv N \prod_{l=1}^{M_t} dq(t_l)$ — the path integral measure is defined by a limiting procedure ($M_t \equiv \frac{1}{\Delta t} \to \infty, \Delta t \to 0, t$ fixed), and $N$ is a normalization factor that always drops out of physical quantities so I don’t need to tell you what it is.

Recall that the key step in the derivation of this statement is the evaluation of the propagator for an infinitesimal time step:

$$
\langle q_2 | e^{-iH\Delta t} | q_1 \rangle = \langle q_2 | e^{-i\Delta t \frac{p^2}{2m}} e^{-i\Delta t V(q)} | q_2 \rangle + O(\Delta t^2).
$$

An integral expression for this can be obtained by inserting resolutions of the identity $\mathbb{1} = \mathbb{1}^2 = \left( \int dp |p\rangle \langle p| \right) \left( \int dq |q\rangle \langle q| \right)$ in between the two exponentials. For a more extensive reminder, please see §2.4 of this document.

---

**Scalar field theory in one dimension.** [Zee §1.3] The path integral for our collection of oscillators is

$$
Z = \int [dq_1 \cdots dq_N] e^{iS[q]} e^{iS[q]}
$$

with $S[q] = \int dt \left( \sum_n \frac{1}{2} m \dot{q}_n^2 - V(q) \right) \equiv \int dt L(q, \dot{q})$. The potential is $V(q) = \sum_n \frac{1}{2} \kappa (q_{n+1} - q_n)^2$. Now suppose we have poor eyesight and can’t resolve the individual atoms in the chain; rather we’re only interested in the long-wavelength (small-wavenumber) physics. So let’s try to take the continuum limit $a \to 0, N \to \infty$. Basically the only thing we need is to think of $q_n = q(x = na)$ as defining a smooth function: $\phi(x)$ instead of $q(x)$ for some reason. At least the letters $q(x)$ and $\phi(x)$ look similar.]
We’ll use
\[(q_n - q_{n-1})^2 \simeq a^2 (\partial_x q)^2 |_{x=na}, \quad a \sum_n f(q_n) \simeq \int dx f(q(x)).\]

The path integral becomes:
\[Z = \int [dq] e^{iS[q]}\]
with \([dq]\) now representing an integral over all configurations \(q(t, x)\) (defined by this limit) and
\[S[q] = \int dt \int dx \frac{1}{2} \left( \mu (\partial_t q)^2 - \mu v_s^2 (\partial_x q)^2 - aq^2 - uq^4 - \cdots \right) \equiv \int dt \int dx \mathcal{L}\]
where I’ve introduced some parameters \(\mu, v_s, r, u\) determined from \(m, \kappa, a\) in some ways that we needn’t worry about, except to say that they are finite in the continuum limit. The \(\cdots\) includes terms like \(a^4 (\partial_x q)^4\) which are small when \(k \ll \frac{1}{a}\), so we ignore them. \(\mathcal{L}\) is the Lagrangian density whose integral over space is the Lagrangian \(L = \int dx \mathcal{L}\).

The equation of motion is the stationary phase condition,
\[0 = \frac{\delta S}{\delta q(x, t)} = -\mu \ddot{q} + \mu v_s^2 \partial_x^2 q - r\dot{q} - 2uq^3 - \cdots\]
In this expression I have written a functional derivative; with our lattice regulator, it is simply a(n extremely useful) shorthand notation for the collection of partial derivatives \(\frac{\partial}{\partial q_n}\).

1 Functional derivatives will be very useful to us. The definition is
\[\frac{\delta \phi(x)}{\delta \phi(y)} = \delta(x - y) \quad \text{(1.2)}\]
plus the Liebniz properties (linearity, product rule). More prosaically, they are just partial derivatives, if we define a collection of values of the independent variable \(\{x_n\}\) to regard as grid points, and let
\[\phi_n \equiv \phi(x_n)\]
so that (1.2) is just
\[\frac{\partial \phi_n}{\partial \phi_m} = \delta_{nm}.\]

If you are not yet comfortable with the machinery of functional derivatives, please work through pages 2-28 through 2-30 of this document now.
that *translates* the whole crystal. It guarantees low-energy phonons near \( k = 0 \) because it means \( q(x) \) can only appear in \( S \) via its derivatives. (This is a general property of *Goldstone modes*; more on this later.)

We can construct a hamiltonian from this action by defining a canonical field-momentum density \( \pi(x) = \frac{\partial L}{\partial \dot{q}} = \mu \partial_t q \) and doing the Legendre transformation:

\[
H = \sum_n (p_n \dot{q}_n - L_n) = \int dx (\pi(x) \dot{q}(x) - \mathcal{L}) = \int dx \left( \frac{\pi(x)^2}{2\mu} + \mu v_s^2 (\partial_x q(x))^2 + r q^2 + u q^4 + \ldots \right).
\]  

(1.3)

Note that I suppress the dependence of all the fields on \( t \) just so it doesn’t get ugly, not because it isn’t there. Also, I emphasize that the position along the chain \( x \) here is just a *label* on the fields, not a degree of freedom or a quantum operator.

The field \( q \) is called a *scalar field* because it doesn’t have any indices decorating it. This is to be distinguished from *e.g.* the Maxwell field, which is a vector field, and which we’ll discuss soon. (Note that vibrations of a crystal in *three* dimensions actually do involve vector indices! We omit this complication.)

The lattice spacing \( a \) and the size of the box \( Na \) in the discussion above are playing very specific roles in *regularizing* our 1-dimensional scalar field theory. The lattice spacing \( a \) implies a maximum wavenumber or shortest wavelength and so is called an “ultraviolet (UV) cutoff”, because the UV is the short-wavelength end of the visible light spectrum. The size of the box \( Na \) implies a maximum wavelength mode which fits in the box and so is called an “infrared (IR) cutoff”.

If (in addition to the continuum limit) we also take the infinite volume limit, then the sums over \( k \) become integrals. In this limit we can make the replacement

\[
\frac{1}{L^d} \sum_k \rightarrow \int \! d^d k, \quad L^d \delta_{kk'} \rightarrow (2\pi)^d \delta^{(d)}(k - k').
\]

A check of the normalization factors comes from combining these two rules

\[
1 = \sum_k \delta_{kk'} = \int \! d^d k (2\pi)^d \delta^{(d)}(k - k').
\]

**Continuum (free) scalar field theory in \( d + 1 \) dimensions.** These continuum expressions are easy to generalize to scalar field theory in any number of dimensions. Let’s do this directly in infinite volume and set \( \mu = 1 \) by rescaling fields. The action is

\[
S[\phi] = \int \! d^d x dt \left( \frac{1}{2} \dot{\phi}^2 - \frac{1}{2} v_s^2 \nabla^2 \phi - \nabla \phi - V(\phi) \right).
\]  

(1.4)
This is almost what we would have found for the long-wavelength \((ka \ll 1)\) description of a \(d\)-dimensional lattice of masses on springs, like a mattress (except that there would have been one \(\phi\) for each direction in which the atoms can wiggle). The equation of motion is
\[
0 = \frac{\delta S[\phi]}{\delta \phi(x)} = -\partial^2_t \phi + v_s^2 \nabla^2 \phi - V'(\phi). \tag{1.5}
\]
For the harmonic case \(V(\phi) = \frac{1}{2} m^2 \phi^2\) we know what we’re doing, and (1.5) is called the Klein-Gordon equation,
\[
0 = \left( \partial_\mu \partial^\mu + m^2 \right) \phi. \tag{1.6}
\]
(Notice that I’ve set \(v_s = c = 1\) here, and this is where we have committed to a choice of signature convention; take a look at the conventions page §0.3.). In relativistic notation, the Lagrangian density is just \(L = \frac{1}{2} \left( \partial_\mu \phi \partial^\mu \phi - m^2 \phi^2 \right)\). This describes free continuum real massive relativistic scalar quantum field theory. (Match the adjectives to the associated features of the lagrangian; collect them all!)

The canonical momentum is \(\pi = \frac{\partial L}{\partial \dot{\phi}} = \dot{\phi}\) and the Hamiltonian (which we can instantly promote to a quantum operator by using boldface symbols) is then
\[
H = \int d^d x \left( \frac{\pi(x)^2}{2} + \frac{1}{2} v_s^2 \left( \vec{\nabla} \phi \cdot \vec{\nabla} \phi \right) + \frac{1}{2} m^2 \phi^2 \right).
\]
Note that all these terms are positive.

A translation invariant linear problem is solved by Fourier transforms: \(\phi(x) = \int d^d k \ e^{-i \vec{k} \cdot \vec{x}} \phi_k\), and \(\pi(x) = \int d^d k \ e^{-i \vec{k} \cdot \vec{x}} \pi_k\), this is
\[
H = \int d^d k \left( \frac{1}{2} \pi_k \pi_{-k} + \frac{1}{2} \left( v_s^2 k^2 + m^2 \right) \phi_k \phi_{-k} \right)
\]
where \(k^2 = (-i \vec{k}) \cdot (i \vec{k}) = \vec{k} \cdot \vec{k}\). This is merely a sum of decoupled oscillators, except for the coupling between wavenumbers \(k\) and \(-k\). We can read off the normal mode frequencies, aka the dispersion relation:
\[
\omega_k^2 = v_s^2 k^2 + m^2.
\]
Notice that this is also the condition for a Fourier mode \(e^{i \vec{k} \cdot \vec{x} - \omega t}\) to solve the Klein-Gordon equation (1.6).

We can decouple the modes with wavenumber \(k\) and \(-k\) by introducing the ladder operators\(^2\)
\[
\phi_k \equiv \sqrt{\frac{\hbar}{2 \omega_k}} \left( a_k + a_{-k}^\dagger \right), \quad \pi_k \equiv \frac{1}{i} \sqrt{\frac{\hbar \omega_k}{2}} \left( a_k - a_{-k}^\dagger \right), \quad [a_k, a_{k'}^\dagger] = (2\pi)^d \delta^{(d)}(k - k').
\]
\(^2\)Beware that the mode operators \(a_k\) defined here differ by powers of \(2\pi/L\) from the finite-volume objects in the previous discussion. These agree with Peskin’s conventions.
Their commutator follows from \([\phi(x), \pi(y)] = i\hbar \delta^{(d)}(x - y)\). In terms of the ladder operators,

\[
H = \int d^d k \, \hbar \omega(k) \left( a_k^\dagger a_k + \frac{1}{2} \right).
\]

The field operators

\[
\phi(\vec{x}) = \int d^d k \sqrt{\frac{\hbar}{2\omega_k}} \left( e^{ik \cdot \vec{x}} a_k + e^{-ik \cdot \vec{x}} a_k^\dagger \right),
\]

\[
\pi(\vec{x}) = \frac{1}{i} \int d^d k \sqrt{\frac{\hbar \omega_k}{2}} \left( e^{ik \cdot \vec{x}} a_k - e^{-ik \cdot \vec{x}} a_k^\dagger \right), \quad (1.7)
\]

satisfy the canonical commutation relation

\[
[\phi(\vec{x}), \pi(\vec{x}')] = i\hbar \delta^d(\vec{x} - \vec{x}').
\]

I emphasize that this is really the same equation as our starting point for each ball on springs:

\[
[q_n, p_{n'}] = i\hbar \delta_{nn'}.
\]

The mode expansions (1.7) contain a great deal of information. First notice that \(\phi\) is manifestly hermitian. Next, notice that from \(\phi(\vec{x}) \equiv \phi(\vec{x}, 0)\) by itself we cannot disentangle \(a_k\) and \(a_k^\dagger\), since only the combination \(a_k + a_k^\dagger\) multiplies \(e^{ik \cdot \vec{x}}\). The momentum \(\pi\) contains the other linear combination. However, if we evolve the field operator in time using the Heisenberg equation we find

\[
\phi(\vec{x}, t) \equiv e^{iHt} \phi(\vec{x}) e^{-iHt} = \int d^d k \sqrt{\frac{\hbar}{2\omega_k}} \left( e^{i(\vec{k} \cdot \vec{x} - \omega_k t)} a_k + e^{-i(\vec{k} \cdot \vec{x} + \omega_k t)} a_k^\dagger \right). \quad (1.8)
\]

Indeed we can check that the relation \(\pi = \dot{\phi}\) holds.

Notice that the dependence on spacetime is via a sum of terms of the form:

\[
e^{i\vec{k} \cdot \vec{x} - i\omega_k t} = e^{i\vec{k} \cdot x^\mu} |_{k^0 = \omega_k}
\]

and their complex conjugates. These are precisely all the solutions to the wave equation (1.6). For each \(\vec{k}\), there are two solutions, one with positive frequency and one with negative frequency. You might have worried that solutions with both signs of the frequency mean that the world might explode or something (like it would if we tried to replace the Schrödinger equation for the wavefunction with a Klein-Gordon equation). This danger is evaded in a beautiful way: the coefficient of the positive frequency solution with wavenumber \(\vec{k}\) is the destruction operator for the mode; the associated negative frequency term comes with the creation operator for the same mode, as a consequence of reality of the field.
1.2 Fields mediate forces

[Zee §1.3] Consider again our chain of balls on springs. Suppose a giant hand reaches in and pushes the atom at position $x_n$ a little bit. This can be described by adding to the hamiltonian a term

$$\delta V(q) = -J_n(t)q_n(t)$$

which applies a force $J_n(t)$ to the $n$th atom. We can ask, in the presence of such a force, what is the amplitude to go from state $I$ to state $F$ in time $T$:

$$\langle F | e^{-\int_0^T dt H(t)} | I \rangle = \int [D\phi] e^{i\int dt d^3x \left( \frac{1}{2} (\partial \phi)^2 - V(\phi) + J(x)\phi(x) \right)}.$$  

As you see, this is a quantity for which we have a path integral representation. Here’s a reason we might care about this quantity: take the initial and final states to be the groundstate:

$$\langle 0 | e^{-\int_0^T dt H(t)} | 0 \rangle \approx e^{-\int_0^T dt E_{gs}(J)}.$$  

If the time-dependence is slow enough, the answer is obtained by the adiabatic approximation: just add up the instantaneous groundstate energy at each time step.

[End of Lecture 1]

Let’s retreat to the case where the action is quadratic in $\phi$, so that we can actually do the path integral:

$$\mathcal{L}(\phi) = \frac{1}{2} \left( \partial_\mu \phi \partial^\mu \phi - m^2 \phi^2 \right) \equiv -\frac{1}{2} \phi \left( \partial^2 + m^2 \right) \phi + \text{total derivative}. \quad (1.9)$$

Going back to the lattice to make the integrals slightly less scary, we have

$$e^{iW[J]} \equiv \int [D\phi] e^{i\int (L + J\phi)} = \int_{-\infty}^{\infty} \prod_{n,t} dq_{n,t} e^{\frac{i}{2} q_x A_{x y} q_y + i J_x q_x} = \sqrt{\frac{(2\pi i)^{N M_t}}{\det A}} e^{-\frac{1}{2} J_x A_{x y}^{-1} J_y}.$$  

Here repeated indices are summed as usual: $q_x A_{x y} q_y = \int dxdy \phi(x) A_{x y} \phi(y)$, etc... So you can see that the matrix $A$ multiplying the quadratic term in this gaussian integral is $A_{x y} = -\delta^{d+1}(x-y) (\partial^2 + m^2)$. It is an $NM_t \times NM_t$ matrix. Its inverse $A^{-1}$ satisfies by definition $A_{x z} A_{y z}^{-1} = \delta_{x y}$, which is the differential equation

$$-(\partial^2 + m^2) D(x-y) = \delta(x-y). \quad (1.10)$$

This equation says that $D$ is a Green’s function for the operator $-(\partial^2 + m^2)$. The fact that there is no special point in spacetime says $A_{x y}^{-1} = D(x-y)$ only depends on the difference of its arguments.

Does this integral actually converge? On the homework you saw an integral of the form $\int_\mathbb{R} dq e^{-\frac{1}{2}q^2 A q}$, which surely converges if $A$ is a positive matrix. Actually, this is
overkill – it is enough to replace $m^2 \to m^2 - i\epsilon$ to make all the integrals converge. Here $\epsilon$ is an infinitesimal, which means $\epsilon^2 = 0$ and $c\epsilon = \epsilon$ for any positive $c$. Then each $\int dq_{nt}$ will have a factor of $e^{-\epsilon \int q_{nt}^2}$ which suppresses the integrand in the dangerous large-field region$^3$.

The equation (1.10) is translation-invariant and linear so you should not be surprised that it is solved by going to Fourier space (in space and time):

$$D(x) = \int d^{d+1}k \ e^{ik\mu x^\mu} D_k, \quad \delta^{d+1}(x) = \int d^{d+1}k \ e^{ik\mu x^\mu}. $$

in terms of which (1.10) becomes the algebraic equation $1 = (k^2 - m^2 + i\epsilon)D_k$. Hence

$$D(x) = \int d^{d+1}k \frac{e^{ikx}}{k^2 - m^2 + i\epsilon}. $$

Notice that the shift by $\epsilon$ saves the day here: it keeps the integration contour from running right over the pole at $k^2 = m^2$, by moving slightly in the imaginary direction. More explicitly,

$$k^2 - m^2 + i\epsilon = \omega^2 - \vec{k}^2 - m^2 + i\epsilon$$

is zero when

$$\omega = \pm \sqrt{\vec{k}^2 + m^2 - i\epsilon} \quad \text{Taylor} \quad \omega_k \equiv \sqrt{\vec{k}^2 + m^2}. $$

In the second step I Taylor expanded $\sqrt{\omega_k^2 - i\epsilon} = \sqrt{\omega_k^2} - \frac{i\epsilon}{\omega_k} + O(\epsilon)^2$ and used the facts that $\omega_k > 0$, and that anything positive times an infinitesimal is an infinitesimal.

We can then do the $\omega$ integral by contours$^4$: if $t > 0 \ (t < 0)$, we can close the contour in the UHP (LHP) since the integrand goes like $e^{-\text{Im} \omega t}$, and the integral equals the residue of the pole at $\omega = \omega_k \mp i\epsilon$ (times $2\pi i$):

$$D(x) = -i \int d^d k \left( \theta(t) \frac{e^{-i(\omega_k t - \vec{k} \cdot \vec{x})}}{2\omega_k} + \theta(-t) \frac{e^{i(\omega_k t - \vec{k} \cdot \vec{x})}}{2\omega_k} \right). \quad (1.11)$$

We’ll learn to call this time-ordered in a moment.

---

$^3$Here I have shown you one way to make the integral well-defined. You might worry that there could be others (there are). Another thing you might be bothered by is the boundary conditions on the fields and their relation to the initial and final states. These issues are closely related! In the next subsection, we’ll say more.

$^4$We are using the Cauchy residue theorem $\oint_C dz f(z) = 2\pi i \sum_{z_j} \text{Res}_{z=z_j} f$ where $z_j$ are the poles of $f$. To remember the sign, consider a small circle $C_0$ counterclockwise around the origin and $f(z) = 1/z$, so $\oint_{C_0} \frac{dz}{z} = i \int_0^{2\pi} d\theta = 2\pi i$. 

14
The propagator. Who is $D(x)$, besides some quantity in terms of which we did a Gaussian integral? Recall from the homework that the inverse matrix can be extracted via a two-point correlation function:

$$A^{-1} = \int dq^2 e^{-\frac{i}{2} qAq}/Z.$$ 

Putting back all the labels, the same manipulations show that

$$D(x-y) \equiv \langle 0| \phi(x) \phi(y)|0 \rangle$$

– the amplitude to propagate an excitation created from the vacuum by $\phi(x)$ to be annihilated by $\phi(y)$. The propagator, for short.

(Notice that if the system is Lorentz invariant (which starting from (1.9) it is) then since $D(x)$ is a scalar quantity, it can only depend on $x$ through Lorentz invariants made from $x^\mu$, namely the proper distance $x^2 = t^2 - \vec{x}^2$, and the sign of $t$.)

Why the ‘?’ in (1.12)? For one thing, $\phi(x)$ and $\phi(y)$ are operators – the order matters. How do I know which order in which to write them? To reproduce (1.11) the thing to do is to time-order them:

$$\langle 0| T\phi(x)\phi(y)|0 \rangle \equiv \theta(x^0 - y^0) \langle 0| \phi(x) \phi(y)|0 \rangle + \theta(y^0 - x^0) \langle 0| \phi(y) \phi(x)|0 \rangle.$$ 

To see this, plug in the mode expansion (1.7) to see e.g.

$$\langle 0| \phi(x) \phi(y)|0 \rangle = \int \frac{dk}{2\omega_k} \frac{d^d q}{2\sqrt{\omega_k \omega_q}} e^{-ikx+iqy} \langle 0| a_k a_q^\dagger|0 \rangle = \int \frac{d^d k}{2\omega_k} e^{-ik(x-y)}$$

(where $k^0 = \omega_k$,$q^0 = \omega_q$ to satisfy the KG equation), which reproduces the first term in (1.11).

Now why should we care about the propagator? Look again at $W[J]$. We’ve learned that (up to terms independent of $J$),

$$W[J] = -\frac{1}{2} \int d^{d+1}x d^{d+1}y J(x) D(x-y) D(y) = -\frac{1}{2} \int d^{d+1}k \frac{1}{k^2 - m^2 + i\epsilon} J_k$$

Here $J(x) = \int d^{d+1}k e^{ikx} J_k$, $J_k^* = J_{-k}$ (since $J(x)$ is real).

We get to pick $J(x)$. Let’s choose $J = J_1 + J_2$ to describe (in Zee’s words) two lumps sitting still on the mattress: $J_a(x) = \delta^3(x - x_a), a = 1, 2$. Then $J_k =$

\[\text{The other ways of making the path integral well-defined correspond to other ways of ordering the } \phi \text{s, and other initial and final states.}\]

\[\text{In comparing to (1.11), it helps to notice that we can redefine the } \vec{k} \text{ integration variable to reverse the sign of the exponent of the spatial part, } \int d^{d} k f(\vec{k}^2) e^{i \vec{k} \cdot \vec{x}} = \int d^{d} k f(\vec{k}^2) e^{-i \vec{k} \cdot \vec{x}}. \text{ (Thanks to Hung-Hwa Lin for help during lecture.)}\]
\[ \int dx^0 e^{-ik^0x^0} \left( e^{i\vec{k}\cdot\vec{x}_1} + e^{i\vec{k}\cdot\vec{x}_2} \right) \] The interaction between the two lumps mediated by the mattress field \( \phi \) will then be described by the \( J_1 J_2 \) cross-terms in \( W[J] \):

\[
W[J] = -\frac{2}{2} \int dx^0 \int dy^0 \int dk^0 e^{ik^0(x^0-y^0)} \int d^3k \frac{e^{i\vec{k}\cdot(\vec{x}_1-\vec{x}_2)}}{k^2 - m^2 + i\epsilon + ...} \quad (1.13)
\]

\[
= -\int dx^0 \left( \int dk^0 2\pi \delta(k^0) \right) \int d^3k \frac{e^{i\vec{k}\cdot(\vec{x}_1-\vec{x}_2)}}{k^2 - m^2 + i\epsilon + ...} \quad (1.14)
\]

\[
= +\int dx^0 \int d^3k \frac{e^{i\vec{k}\cdot(\vec{x}_1-\vec{x}_2)}}{k^2 + m^2 - i\epsilon + ...} \quad (1.15)
\]

(The \( \ldots \) indicate terms which don’t depend on \( x_1, x_2 \), so let’s ignore them.)

For this choice of \( J \), the Hamiltonian is time-independent, and \( e^{iW} = \langle 0 | e^{-iHT} | 0 \rangle = e^{-iE_{gs}(J)T} \), so \( W = -E_{gs}(J)T \). We learn that

\[
E_{gs}(J) = -\int d^d k \frac{e^{i\vec{k}\cdot\vec{x}_{12}}}{k^2 + m^2} + \text{const}.
\]

Notice that we can drop the \( i\epsilon \) now, because this integrand is nonsingular for real \( \vec{k} \).

In \( d = 1 \), there are poles at \( k = \pm im \), and we can close the contour in the UHP for free to get\(^7\)

\[
E_{gs}(J) = -\frac{2\pi i}{2\pi} \frac{e^{-mx}}{2i} = -\frac{e^{-mx}}{2m}.
\]

Since \( x \) is the separation between the lumps, this means that our field has produced an attractive force between the lumps

\[
F = -\partial_x E_{gs}(J) = +\frac{1}{2} e^{-mx}
\]

which falls off exponentially with the separation between the lumps. The range of the potential goes inversely with the mass of the ‘force carrier’ \( \phi \). The 3d version of this potential \( e^{-mx/r} \) (see footnote 7) is called the Yukawa potential.

\(^7\)For convenience, here’s the integral in 3d:

\[
\int d^3k \frac{e^{i\vec{k}\cdot\vec{x}}}{k^2 + M^2} = \frac{1}{(2\pi)^2} \int_0^\infty \frac{k^2dk}{k^2 + M^2} \int_{-1}^1 dy e^{iky} = \frac{1}{(2\pi)^2} \int_{-\infty}^\infty \frac{dk}{k^2 + M^2} e^{ikr} \\
= \frac{1}{2\pi^2} \left( \frac{1}{2i} \int_{-\infty}^\infty \frac{dk}{k^2 + M^2} + h.c. \right) \quad \text{close contour in UHP for free} \\
= \frac{1}{2\pi^2} \frac{1}{2i} 2\pi iM e^{i(M)r} = e^{-Mr} \frac{1}{4\pi r}.
\]
1.3 Euclidean path integral and Wick rotation

Here is a route to defining the path integral (actually the same as the replacement $m^2 \rightarrow m^2 - i\epsilon$) which makes clearer what is going on with the initial and final states.

The whole point here can be made for a single mode of the field – a single harmonic oscillator – with action

$$S[q] = \frac{1}{2} \int dt \left( (\partial_t q)^2 - \Omega^2 q^2 \right) - \int Jq$$

(where $\Omega^2 = \vec{k}^2 + m^2$ if you like). Consider the replacement $\tau = it$ in the action:

$$S[q] = -\frac{1}{2} i \int d\tau \left( -(\partial_\tau q)^2 - \Omega^2 q^2 \right) + i \int d\tau Jq = i \int d\tau \left( \frac{1}{2} ((\partial_\tau q)^2 + \Omega^2 q^2) + Jq \right).$$

With this replacement, the path integral becomes

$$\int [Dq] e^{-\int d\tau \left( \frac{1}{2} ((\partial_\tau q)^2 + \Omega^2 q^2) + Jq \right)} \equiv \int [Dq] e^{-S_E[q]}.$$

This integrand suppresses configurations with large $q$, and large $\partial_\tau q$, and the integral is therefore totally well-defined. The euclidean action is

$$S_E[q] = \int d\tau \left( \frac{1}{2} ((\partial_\tau q)^2 + \Omega^2 q^2) + Jq \right) = \int d\tau \left( \frac{1}{2} q \left( -\partial_\tau^2 + \Omega^2 \right) q + Jq \right)$$

where $(-\partial_\tau^2 + \Omega^2)$ is a positive operator (meaning all of its eigenvalues are positive). Call its inverse $G$, which then, by definition, satisfies

$$(-\partial_\tau^2 + \Omega^2) G(\sigma, \tau) = \delta(\sigma - \tau)$$

The fact that our system is time-translation invariant means $G(\sigma, \tau) = G(\sigma - \tau)$. We can solve this equation in fourier space: $G(s) = \int d\omega e^{i\omega s} G_\omega$ makes it algebraic:

$$G_\omega = \frac{1}{\omega^2 + \Omega^2}$$

and we have

$$G(\tau) = \int d\omega \frac{e^{i\omega \tau}}{\omega^2 + \Omega^2} = e^{-\Omega |\tau|} \frac{1}{2\Omega}. \quad (1.16)$$

(Do it by residues: the integrand has poles at $\omega = \pm i\Omega$ (see the figure 1 below). The absolute value of $|\tau|$ is crucial, and comes from the fact that the contour at infinity converges in the upper (lower) half plane for $\tau < 0$ ($\tau > 0$).)

---

8It is called euclidean because the $(\partial_\tau q)^2$ has the same sign as the spatial derivatives $(\partial_x q)^2$, so this is the action we get in euclidean spacetime with metric $\delta_{\mu\nu}$, rather than $\eta_{\mu\nu}$. Exercise: put back the spatial derivative terms and check that this is the case.
I claim that the real-time calculation which keeps the oscillator in its groundstate is the analytic continuation of the one we did above, where we replace

$$\omega_{\text{Mink}} = e^{-i(\pi/2-\epsilon)}\omega_{\text{above}}$$  \hspace{1cm} (1.17)

where $\epsilon$ is (a familiar) infinitesimal. In the picture of the euclidean frequency plane in Fig. 1, this is a rotation by nearly 90 degrees. We don’t want to go all the way to 90 degrees, because then we would hit the poles at $\pm i\Omega$. The replacement (1.17) just means that if we integrate over real $\omega_{\text{Mink}}$, we rotate the contour in the integral over $\omega$ as follows:

as a result we pick up the same poles at $\omega_{\text{above}} = \pm i\Omega$ as in the euclidean calculation. Notice that we had better also rotate the argument of the function, $\tau$, at the same time to maintain convergence, that is:

$$\omega_{\text{eucl}} = -i\omega_{\text{Mink}}, \quad \omega_{\text{eucl}}t_{\text{eucl}} = \omega_{\text{Mink}}t_{\text{Mink}}, \quad t_{\text{eucl}} = +it_{\text{Mink}}.$$  \hspace{1cm} (1.18)

So this is giving us a contour prescription – a prescription for negotiating the poles – for the real-frequency integral. The result is the Feynman contour, and it is the same as
what we got from $m^2 \rightarrow m^2 - i\epsilon$: depending on the sign of the (real) time separation of the two operators (recall that $t$ is the difference), we close the contour around one pole or the other, giving the time-ordered propagator. For the case of a free scalar field, the replacement $m^2 \rightarrow m^2 - i\epsilon$ had the same effect of rotating the real-frequency contour away from the poles. It is also the same as shifting the frequency by $\Omega \rightarrow \Omega - i\epsilon$, as indicated in the right part of Fig. 2. This prescription works in a case where there is no $m^2$ term.

Notice for future reference that the euclidean action and real-time action are related by

$$S_{\text{eucl}}[Q] = \int dt_{\text{eucl}} \frac{1}{2} \left( \left( \frac{\partial Q}{\partial t_{\text{eucl}}} \right)^2 + \Omega^2 Q^2 \right) = -i S_{\text{Mink}}[Q] = -i \int dt_{\text{Mink}} \frac{1}{2} \left( \frac{\partial Q}{\partial t_{\text{Mink}}} \right)^2 - \Omega^2 Q^2 \right) .$$

because of (1.18). This means the path integrand is $e^{-S_{\text{eucl}}} = e^{iS_{\text{Mink}}}$. 

---

**Euclidean evolution.** Now, why does the contour coming from the euclidean path integral put the oscillator into its groundstate? The point in life of the euclidean time evolution to prepare the groundstate from an arbitrary state:

$$e^{-\beta H} |\text{any}\rangle = \sum_n e^{-\beta E_n} |n\rangle \langle n|\text{any}\rangle \propto |\text{gs}\rangle + \mathcal{O} \left( e^{-\beta (E_1 - E_{\text{gs}})} \right) \quad (1.19)$$

– the euclidean-time propagator $e^{-\beta H}$ beats down the amplitude of any excited state relative to the groundstate, for large enough $\beta$.

And the euclidean path integral gives a formula for this euclidean propagation amplitude. Recall that the path integral representation for the real-time propagation
amplitude is
\[ \langle f | e^{-iHt} | i \rangle = \int_{f \leftarrow i} [dq] e^{i \int_0^t dL} \]

On the RHS here, we sum over all paths between \( i \) and \( f \) in time \( t \) \( (i.e. q(0) = q_i, q(t) = q_f) \), weighted by a phase \( e^{i \int f \, dL} \). But that means you also know a representation for
\[ \sum_f \langle f | e^{-\beta H} | f \rangle \equiv \text{tr} e^{-\beta H} \]

- namely, you sum over all periodic paths \( q_i = q_f \) in imaginary time \( t = -i\beta \). So:
\[ Z(\beta) = \text{tr} e^{-\beta H} = \oint [dq] e^{-\int_0^\beta d\tau L} = \oint [dq] e^{-S_{\text{eucl}}[q]} \]

The LHS is the partition function in quantum statistical mechanics. The RHS is the euclidean functional integral we’ve been using. [For more on this, see Zee §V.2]

The period of imaginary time, \( \beta \equiv 1/T \), is the inverse temperature. We’ve been studying the limit as \( \beta \to \infty \). Taking \( \beta \to \infty \) means \( T \to 0 \), and you’ll agree that at \( T = 0 \) we project onto the groundstate (if there’s more than one groundstate we have to think more).

**Time-ordering.** To summarize the previous discussion: in real time, we must choose a state, and this means that there are many Green’s functions, not just one: \( \langle \psi | q(t)q(s) | \psi \rangle \) depends on \( |\psi\rangle \), unsurprisingly.

But we found a special one which arises by analytic continuation from the euclidean Green’s function, which is unique\(^9\). It is
\[ G(s, t) = \langle \mathcal{T} | q(s)q(t) \rangle, \]

the time-ordered, or Feynman, Green’s function, and I write the time-ordering symbol \( \mathcal{T} \) to emphasize this. I emphasize that from our starting point above, the time ordering arose because we have to close the contour in the UHP (LHP) for \( t < 0 \) \( (t > 0) \).

\(^9\) Another important perspective on the uniqueness of the euclidean Green’s function and the non-uniqueness in real time: in euclidean time, we are inverting an operator of the form \(-\partial_x^2 + \Omega^2\) which is **positive** \( (\equiv \) all its eigenvalues are positive) \( – \) recall that \(-\partial_x^2 = \hat{p}^2\) is the square of a hermitian operator. If all the eigenvalues are positive, the operator has no kernel, so it is completely and unambiguously invertible. This is why there are no poles on the axis of the (euclidean) \( \omega \) integral in (1.16). In real time, in contrast, we are inverting something like \(+\partial_t^2 + \Omega^2\) which annihilates modes with \( \partial_t = i\Omega \) (if we were doing QFT in \( d > 0 + 1 \) this equation would be the familiar \( p^2 - m^2 = 0 \)). These are called ‘on-shell states’, they are actual states in the spectrum of the Hamiltonian of the system. So the operator we are trying to invert has a kernel and this is the source of the ambiguity. In frequency space, this is reflected in the presence of poles of the integrand on the contour of integration; the choice of how to negotiate them encodes the choice of Green’s function.
Let’s pursue this one more step. The same argument tells us that the generating functional for real-time, time-ordered correlation functions of $q$ is

$$Z[J] = \left< T e^{i \int dq(t) J(t)} \right> = \langle 0 | T e^{i \int J q} | 0 \rangle ,$$

(1.20)
in the sense that

$$\langle T q(t_1) q(t_2) \ldots \rangle = \frac{1}{Z} \frac{\delta}{\delta J(t_1)} \frac{\delta}{\delta J(t_2)} \ldots Z[J]|_{J=0} .$$

In the second step of (1.20) I just emphasized that the real time expectation value here is really a vacuum expectation value, as long as we use the $i \epsilon$ prescription above to define the integrals. This quantity has the picturesque interpretation as the vacuum persistence amplitude, in the presence of the source $J$.\(^{10}\)

So we see that in general, the correlation functions that are computed by this “$i \epsilon$ prescription” of Wick rotating from Euclidean spacetime are time-ordered vacuum expectation values:

$$\frac{1}{Z} \int [D\phi] e^{iS_{m^2 \rightarrow m'^2 - \epsilon} f(\phi)} = \langle 0 | T f(\phi) | 0 \rangle .$$

Causality. In other treatments of this subject, you will see the Feynman contour motivated by ideas about causality. This was not the logic of our discussion here but it is reassuring that we end up in the same place. Note that even in 0+1 dimensions there is a useful notion of causality: effects should come after their causes.

### 1.4 Feynman diagrams from the path integral

Subsection §1.3 was a sophisticated discussion of QFT in 0+1 dimensions (i.e. ordinary quantum mechanics of a single particle), since we focussed on a single mode. To attempt to demystify some more of the structure we’ll discover in QFT, let’s regress even further, and consider the case of QFT in 0+0 dimensions. By the path-integral representation, this means ordinary integrals. If everything is positive, this is probability theory.

Suppose we want to do the integral

$$Z(J) = \int_{-\infty}^{\infty} dq \ e^{-\frac{1}{2} m^2 q^2 - \frac{g}{4!} q^4 + J q} \equiv \int dq \ e^{-S(q)} .$$

(1.21)

\(^{10}\)Actually, more useful is the generating function of connected correlation functions:

$$\langle T q(t_1) q(t_2) \ldots \rangle_c = \frac{\delta}{\delta J(t_1)} \frac{\delta}{\delta J(t_2)} \ldots \log Z[J]|_{J=0} .$$

where $\langle q_1 q_2 \rangle_c \equiv \langle q_1 q_2 \rangle - \langle q_1 \rangle \langle q_2 \rangle$. Higher-point connected correlation functions are defined by subtracting the gaussian answer. Connected correlation functions are well-named because they are computed by connected Feynman diagrams, as we’ll discuss more next.
It is the path integral for $\phi^4$ theory with fewer labels. For $g = 0$, this is a gaussian integral which we know how to do. For $g \neq 0$ it’s not an elementary function of its arguments. We can develop a (non-convergent!) series expansion in $g$ by writing it as

$$Z(J) = \int_{-\infty}^{\infty} dq \ e^{-\frac{1}{2}m^2q^2 + Jq} \left( 1 - \frac{g}{4!} q^4 + \frac{1}{2!} \left( -\frac{g}{4!} q^4 \right)^2 + \cdots \right)$$

and integrating term by term. And the term with $q^{4n}$ (that is, the coefficient of $\frac{1}{n!} \left( -\frac{g}{4!} \right)^n$) is

$$\int_{-\infty}^{\infty} dq \ e^{-\frac{1}{2}m^2q^2 + Jq} q^{4n} = \left( \frac{\partial}{\partial J} \right)^{4n} \int_{-\infty}^{\infty} dq \ e^{-\frac{1}{2}m^2q^2 + Jq} \left( \frac{\partial}{\partial J} \right)^{4n} e^{\frac{1}{2}J \frac{1}{m^2} J} \sqrt{\frac{2\pi}{m^2}}.$$  

So:

$$Z(J) = \sqrt{\frac{2\pi}{m^2}} e^{-\frac{g}{4!} \left( \frac{\partial}{\partial J} \right)^{4} e^{\frac{1}{2}J \frac{1}{m^2} J}.}$$

This is a double expansion in powers of $J$ and powers of $g$. The process of computing the coefficient of $J^n g^m$ can be described usefully in terms of diagrams. There is a factor of $1/m^2$ for each line (the propagator), and a factor of $(-g)$ for each 4-point vertex (the coupling), and a factor of $J$ for each external line (the source). For example, the coefficient of $gJ^4$ comes from:

\[
\sim \left( \frac{1}{m^2} \right)^4 gJ^4.
\]

\[
\begin{array}{ccc}
\quad & J & J \\
J & q & J \\
\quad & J & J
\end{array} + \begin{array}{c}
\quad \\
J
\end{array} + \begin{array}{c}
\quad \\
J
\end{array} + \begin{array}{c}
\quad \\
J
\end{array} + \begin{array}{c}
\quad \\
J
\end{array} < \left( \frac{1}{m^2} \right)^4 J^4.
\]

There is a symmetry factor which comes from expanding the exponential: if the diagram has some symmetry preserving the external labels, the multiplicity of diagrams does not completely cancel the $1/n!$.

As another example, consider the analog of the two-point function:

$$G \equiv \langle q^2 \rangle \mid_{J=0} = \frac{\int dq \ q^2 \ e^{-S(q)}}{\int dq \ e^{-S(q)}} = -2 \frac{\partial}{\partial m^2} \log Z(J = 0). \quad (1.22)$$

In perturbation theory this is:

\[
G \simeq \quad + \quad + \quad + \quad + \quad + \mathcal{O}(g^3)
\]

22
\begin{align*}
= m^{-2} \left( 1 - \frac{1}{2}gm^{-4} + \frac{2}{3}g^2m^{-8} + \mathcal{O}(g^3) \right) \tag{1.23}
\end{align*}

To get the numerical coefficients note that Wick’s theorem for this simple case is

\[ \langle q^k \rangle_0 = \begin{cases} 0, & k \text{ odd} \\ (k - 1)!!, & k \text{ even} \end{cases} \tag{1.24} \]

– the number of ways of pairing \( k \) objects. Here \((k - 1)!! \equiv (k - 1)(k - 3)(k - 5) \cdots 3 \cdot 1\).

This is because there are \( k - 1 \) choices of partner for the first \( q \), after which there are \( k - 3 \) choices of partner for the next one, etc.

Some important structural comments: A diagram contributing to \( G \) which has any part not connected to the external legs is cancelled by the expansion of the denominator \( Z = \int dq e^{-S(q)} \). The contributions to \( Z \) are called ‘vacuum diagrams’ (since they have no external lines, so they are like an amplitude for nothing to turn back into nothing).

\( Z \) is a sum over all diagrams with no external lines, including disconnected ones. As you saw in 215A, this sum exponentiates: \( Z = e^{\sum \text{(connected diagrams)}} \).

**Some labels.** Some of these points are clearer if we put back some of the labels.

So consider the slightly more complicated case

\[ Z = \int \prod_{i=1}^N dq_i e^{-S(q)} , \quad S(q) = \frac{1}{2}q_iA_{ij}q_j + \frac{g}{4!} \sum_i q_i^4 \equiv S_0 + \frac{g}{4!} \sum_i q_i^4. \]

(Think of \( i \) as like a position index, and \( A \) as a difference operator, so this is a discretization of \( \phi^4 \) theory.) Then we can develop a perturbative expansion by writing

\[ \langle q_1 \cdots q_k \rangle = \frac{\int \prod_{i=1}^N dq_i e^{-S(q)} q_1 \cdots q_k}{Z} \tag{1.25} \]

\[ = \frac{\int \prod_{i=1}^N dq_i e^{-S_0(q)} e^{-\frac{g}{4!} \sum_i q_i^4} q_1 \cdots q_k}{\int \prod_{i=1}^N dq_i e^{-S_0(q)} e^{-\frac{g}{4!} \sum_i q_i^4}} \tag{1.26} \]

\[ \approx \frac{\sum_{n=0} (-\frac{g}{4!})^n /n! \int dq e^{-S_0} \sum_{i_1, \ldots, i_n} \prod_{i_a} q_{i_a}^4 q_1 \cdots q_k}{\sum_{n=0} (-\frac{g}{4!})^n /n! \int dq e^{-S_0} \sum_{i_1, \ldots, i_n} \prod_{i_a} q_{i_a}^4}. \tag{1.27} \]

(Note the step with the \( \approx \) is where we exchange the sum over \( n \) with the integral over \( q \).)

Now the general statement of Wick’s theorem is:

\[ \langle q_1 \cdots q_k \rangle_0 = \begin{cases} 0, & k \text{ odd} \\ \sum \text{contractions} A_{i_1 i_2}^{-1} \cdots A_{i_{k/2-1} i_{k/2}}^{-1}, & k \text{ even} \end{cases} \]
Here the sum is over all ways of pairing up the \( k \) fields. (Note that this reduces to (1.24) if we remove all the labels.)

Let’s think about a particular term, e.g. an \( \mathcal{O}(g^2) \) contribution to \( \langle q_i q_j \rangle \):

\[
\langle q_i q_j \rangle \approx \frac{1}{6} g^2 \sum_{k,l} A^{-1}_{ik} A^{-1}_{kj} A^{-3}_{kl}
\]

The factor of 1/6 is called a symmetry factor. Instead of by explicit combinatorics, we could have gotten this number by dividing by the order of the automorphism group of the diagram. An automorphism of the diagram is a map from the diagram to itself which preserves the external lines and the connectivity. In this diagram, we can permute the three internal lines amongst themselves, giving \( |S_3| = 3! = 6 \) in the denominator. Don’t get hung up on the symmetry factors.

As a final example for now, here is the expansion of the four-point function:

\[
\langle q_1 q_2 q_3 q_4 \rangle \approx A^{-1}_{14} A^{-1}_{23} + A^{-1}_{12} A^{-1}_{34} + A^{-1}_{13} A^{-1}_{24} - g \sum_i \left( A^{-1}_{i1} A^{-1}_{i2} A^{-1}_{i3} A^{-1}_{i4} \right) \left( A^{-1}_{i1} A^{-1}_{i4} A^{-1}_{i2} A^{-1}_{i3} + \cdots \right)
\]

\[
+ g^2 \sum_{ij} \left( \frac{1}{2} A^{-1}_{i1} A^{-1}_{i4} A^{-2}_{i2} A^{-1}_{i3} A^{-1}_{j4} + \cdots \right) \left( \frac{1}{2} A^{-1}_{i1} A^{-1}_{i2} A^{-1}_{i3} A^{-1}_{i4} A^{-1}_{j1} + \cdots \right) + \cdots
\]

All the labels. The Feynman diagrams we’re going to draw all the time are the same but with more labels. Notice that each of the \( q_i \) in our integral could come with a label, \( q \rightarrow q_a \). Then each line in our diagram would be associated with a matrix \( (m^{-2})_{ab} \) which is the inverse of the quadratic term \( q_a m^2_{ab} q_b \) in the action. If our diagrams have loops we get free sums over the label. If that label is conserved by the interactions, the vertices will have some delta functions. In the case of translation-invariant field theories we can label lines by the conserved momentum \( k \). Each comes with a factor of the free propagator \( \frac{1}{k^2 + m^2 + i\epsilon} \), each vertex conserves momentum, so
comes with $i g \delta^D (\sum k) (2\pi)^D$, and we must integrate over momenta on internal lines $\int d^D k$.

**Brief comments about large orders of perturbation theory.**

- The perturbation series about $g = 0$ does not converge. How do I know? One way to see this is to notice that if I made $g$ even infinitesimally negative, the integral itself would not converge (the potential would be unbounded below), and $Z_{g = -|\epsilon|}$ is not defined. Therefore $Z_g$ as a function of $g$ cannot be analytic in a neighborhood of $g = 0$. This argument is due to Dyson, and applies also in most QFTs. This means there is more to QFT than perturbation theory: the perturbation series does not define the field theory amplitudes.

- The expansion of the exponential in the integrand is clearly convergent for each $q$. The place where we went wrong is exchanging the order of integration over $q$ and summation over $n$.

- In this case, the perturbation expansion can be given a closed form expression:

$$Z(0) \simeq \left( \frac{2\pi}{m^2} \right) \sum_n \frac{(-1)^n}{n!} \frac{2^{n+\frac{1}{2}}}{(4!)^n} \Gamma \left( 2n + \frac{1}{2} \right) \left( \frac{g}{m^2} \right)^n \tag{1.28}$$

- The expansion for $G$ is of the form

$$G \simeq m^{-2} \sum_{n=0}^{\infty} c_n \left( \frac{g}{m^2} \right)^n.$$

When $n$ is large, the coefficients satisfy $c_{n+1} \geq \frac{2}{3} n c_n$ (you can see this by looking at the coefficients in (1.28)) so that $|c_n| \sim n!$. This factorial growth of the number of diagrams is general in QFT and is another way to see that the series does not converge.

- The fact that the coefficients $c_n$ grow means that there is a best number of orders to keep. The errors start getting bigger when $c_{n+1} \left( \frac{g}{m^2} \right) \sim c_n$, that is, at order $n \sim \frac{3m^4}{2g}$. So if you want to evaluate $G$ at this value of the coupling, you should stop at that order of $n$. An estimate of the error by the next term left out gives something that goes like $e^{-\#/g}$.

- I said above that the fact that the perturbation series doesn’t converge means that it doesn’t define the field theory amplitudes. What does it miss? To answer
this, consider trying to do the integral (1.21) by saddle point (at $J = 0$ for simplicity):

$$0 = S'(q_*) = m^2 q_* + \frac{g}{3!} q_*^3.$$  

(Note the resemblance to the equations of motion.) This has three solutions:

$$q_* = 0, \quad q_* = \pm i \sqrt{\frac{3! m^2}{g}}.$$  

The expansion about the ‘trivial’ saddle at $q_*$ (where the action is $S(q_* = 0) = 0$) reproduces the perturbation series. At the other saddles,

$$S \left( q_* = \pm i \sqrt{\frac{3! m^2}{g}} \right) = -\frac{3m^4}{2g}, \quad (1.29)$$

which means their contribution would go like $e^{+\frac{3m^4}{2g}}$, which actually would blow up at weak coupling, $g \to 0$. These saddles are not on the contour and don’t contribute for small positive $g$, but more generally (as for example when $m^2 < 0$), there will be effects that go like $e^{-\frac{3m^4}{2g}}$. This is a function whose series expansion in $g$ at $g = 0$ is identically zero. You can never find it by doing perturbation theory in $g$ about $g = 0$.

- A technique called Borel resummation can sometimes produce a well-defined function of $g$ from an asymptotic series whose coefficients diverge like $n!$. The idea is to make a new series

$$B(z) \equiv \sum_{n=0}^{\infty} \frac{c_n}{n!} z^n$$

whose coefficients are ensmallened by $n!$. Then to get back $Z(g)$ we use the identity

$$1 = \frac{1}{n!} \int_0^\infty dz e^{-z} z^n$$

and do the Laplace transform of $B(z)$:

$$\int_0^\infty \! dz B(z) e^{-z/g} = \sum_{m=0}^{\infty} \int_0^\infty \! dz e^{-\frac{z}{g}} z^m \frac{c_m}{m!} = g \sum_{m=0}^{\infty} \frac{c_m g^m}{m!} = g Z(g).$$

This procedure requires both that the series in $B(z)$ converges and that the Laplace transform can be done. In fact this procedure works in this case.

The fact that the number of diagrams at large order grows like $n!$ is correlated with the existence of saddle-point contributions to $Z(g)$ which go like $e^{-a/g}$.  

26
This is because they are associated with singularities of $B(z)$ at $z = a$; such a singularity means the sum of $\frac{c_n}{n!} z^n$ must diverge at $z = a$. (More generally, non-perturbative effects which go like $e^{-a/g^{1/p}}$ (larger if $p > 1$) are associated with (faster) growth like $(pn)!$. In string theory, $p = 2$. See this classic work.)

- In fact in this case, we know the whole function. The integral actually does have a name – it’s a Bessel function:

$$Z(J = 0) = \frac{2}{\sqrt{m^2}} \sqrt{\rho} e^{\rho} K_{\frac{1}{4}}(\rho), \quad \rho \equiv \frac{3m^4}{4g}$$

(for $\text{Re} \sqrt{\rho} > 0$), as Mathematica will tell you. Because we know about Bessel functions, in this case we can actually figure out what happens at strong coupling, when $g \gg m^4$, using the asymptotics of the Bessel function.

- The functions $G(g)$ and $Z(g)$ can be analytically continued in $g$ away from the real axis, and can in fact be defined on the whole complex $g$ plane. It has a branch cut on the negative real axis, across which its discontinuity is related to its imaginary part. The imaginary part goes like $e^{-a |g|}$ near the origin and can be computed by a tunneling calculation like (1.29).

How did we know $Z$ has a branch cut? One way is from the asymptotics of the Bessel function. But, better, why does $Z$ satisfy the Bessel differential equation as a function of the couplings? The answer, as you’ll check on the homework, is that the Bessel equation is a Schwinger-Dyson equation,

$$0 = \int_{-\infty}^{\infty} \frac{\partial}{\partial q} \left( \text{something } e^{-S(q)} \right)$$

(1.30)

which results from demanding that we can change integration variables in the path integral.

For a bit more about this, you might look at sections 3 and 4 of this recent paper from which I got some of the details here. See also the giant book by Zinn-Justin. There is a deep connection between the large-order behavior of the perturbation series about the trivial saddle point and the contributions of non-trivial saddle points. The keywords for this connection are resurgence and trans-series and a starting reference is here.

The class of equations (1.30) is very important: it shows that the equations of motion are true in Green’s functions, up to contact terms, the contributions where the
\frac{\partial}{\partial \eta} \text{ hits the 'something'. For example, in scalar field theory}

\[ 0 = \int [D\phi] \frac{\delta}{\delta \phi(y)} (\phi(x)e^{iS[\phi]}) = \left\langle \phi(x)i \frac{\delta S}{\delta \phi(y)} \right\rangle + \delta^{d+1}(x - y). \]

In the special case where \( S \) is quadratic, \( S = \int A\phi \), this shows that the two-point function is a Green’s function for the quadratic operator \( A \).
1.5 Lagrangian field theory

[Here we fill in the bits of Peskin §2.2 that we missed above.] Let’s consider a classical field theory in the Lagrangian description. This means that the degrees of freedom are a set of fields $\phi_r(x)$, where $r$ is a discrete index (for maybe spin or polarization or flavor), and we specify the dynamics by the classical action. If the world is kind to us (in this class we assume this), the action is an integral over space and time of a Lagrangian density

$$S[\phi] \equiv \int d^{d+1}x L(\phi, \partial^\mu \phi).$$

This important assumption is an implementation of locality.

This central object encodes the field equations, the canonical structure on the phase space, the Hamiltonian, the symmetries of the theory.

I’ve sneakily implied that we are going to assume Lorentz invariance, so that $L$ depends on the 4-vector $\partial^\mu \phi$, and not its components separately. I am also going to assume that the action $S$ is real.

Two examples to keep in mind are the Klein-Gordon Lagrangian:

$$L_{KG} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2$$

and the Maxwell Lagrangian:

$$L_{EM} = -\frac{1}{4e^2} F_{\mu\nu} F^{\mu\nu} = \frac{1}{4e^2} (E^2 - B^2)$$

with $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$ and $A_\mu$ regarded as the independent degrees of freedom.

A word about units: in units with $\hbar = c = 1$, everything has units of mass to some power, called its mass dimension. Energy and momentum $p_\mu = \hbar k_\mu$ have mass dimension $+1$. The space and time coordinates $x^\mu$ have mass dimension $-1$. The action goes in the exponential of the path integral measure $\int [D\phi] e^{iS/\hbar}$ and so must be dimensionless. So the Lagrangian density has mass dimension $d+1$. This means that the KG field has mass dimension $d + 1/2$ (and the mass $m$ has mass dimension 1 (yay!)). In $d + 1 = 3 + 1$ dimensions, $E \sim \dot{A}$, $B \sim \nabla A$ have mass dimension 2 and $A$ has mass dimension one (and $e$ is dimensionless). This is nice because then the covariant derivative $\partial_\mu + A_\mu$ has mass dimension one. Notice that $E^2 + B^2$ has dimension 4 which is good for an energy per unit volume.

The equation of motion is

$$0 = \frac{\delta S}{\delta \phi_r(x)}.$$

<table>
<thead>
<tr>
<th>object</th>
<th>mass dim.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_\mu$</td>
<td>1</td>
</tr>
<tr>
<td>$x^\mu$</td>
<td>-1</td>
</tr>
<tr>
<td>$S$</td>
<td>0</td>
</tr>
<tr>
<td>$L$</td>
<td>$d + 1$</td>
</tr>
<tr>
<td>$\phi$</td>
<td>$\frac{d + 1}{2}$</td>
</tr>
<tr>
<td>$A_\mu$</td>
<td>1</td>
</tr>
<tr>
<td>$E, B, F_{\mu\nu}$</td>
<td>2</td>
</tr>
</tbody>
</table>
Note the functional derivative. You can check that in the case when $L$ depends only on $\phi$ and $\partial_\mu \phi$, this is the same as the Lagrange EOM

$$0 = \frac{\partial L}{\partial \phi_r} - \partial_\mu \frac{\partial L}{\partial (\partial_\mu \phi_r)}$$

(for each $r$) which I can’t remember. Note that since we are interested here in the bulk equations of motion, we ignore boundary terms unless we are interested in field theory on a space with boundary. That is a worthy subject but an unnecessary complication for now.

By redefining the field by e.g. $\phi \equiv \frac{1}{D} (\chi - B/C)$, we can make the KG theory uglier

$$L = A + B\chi + \frac{1}{2} C\chi^2 + \frac{1}{2} D\partial_\mu \chi \partial_\mu \chi.$$
places along our never-ending chains of ‘why?’ questions. For example, one answer to the question “Why QFT?” is (certainly this is the one given in Weinberg’s text, but just as certainly it is not the only one): quantum mechanics plus Poincaré symmetry.

They are also helpful for solving physical systems: Continuous symmetries are associated with conserved currents. Suppose the action is invariant under a continuous transformation of the fields \( \phi, \phi(x) \mapsto \phi'(x) \). (The invariance of the action is what makes the transformation a symmetry.) ‘continuous’ here means we can do the transformation just a little bit, so that \( \phi(x) \mapsto \phi(x) + \epsilon \Delta \phi(x) \) where \( \epsilon \) is an infinitesimal parameter.

If the transformation with constant \( \epsilon \) (independent of space and time) is a symmetry, then the variation of the action with \( \epsilon = \epsilon(x,t) \) must be proportional to \( \partial_\mu \epsilon \) (at least assuming some smoothness properties of the action), and so that it vanishes \( \forall \phi \) when \( \epsilon \) is constant:

\[
S[\phi + \epsilon(x) \Delta \phi] - S[\phi] = \int d^d x d t \partial_\mu \epsilon(x) j^\mu \text{IBP} = - \int d^d x d t \epsilon(x) \partial_\mu j^\mu.
\]

But if the equations of motion are obeyed, then the action is invariant under any variation of \( \phi \), including this one, for arbitrary \( \epsilon(x) \). But this means that \( \partial_\mu j^\mu = 0 \), the current is conserved. These words are an accurate description of the equation because they mean that the charge

\[
Q_R \equiv \int_R d^d x \int R d^d x j^0
\]

in some region of space \( R \) can only change by leaving the region (assume the definition of \( R \) is independent of time):

\[
\partial_t Q_R = \int_R d^d x \partial_t j^0 = - \int_R d^d x \nabla \cdot j = - \int_{\partial R} d^{d-1} x \hat{n} \cdot j
\]

where in the last step we used Stokes’ theorem.

This trick with pretending the parameter depends on space is called the Noether method. More prosaically, the condition that the action is invariant means that the Lagrangian density changes by a total derivative (we assume boundary terms in the action can be ignored):

\[
\mathcal{L}(\phi', \partial_\mu \phi') \overset{\text{symmetry}}{=} \mathcal{L}(\phi, \partial_\mu \phi) + \epsilon \partial_\mu \mathcal{J}^\mu
\]

but on the other hand, by Taylor expansion,

\[
\mathcal{L}(\phi', \partial_\mu \phi') \overset{\text{calculus}}{=} \mathcal{L}(\phi, \partial_\mu \phi) + \epsilon \left( \frac{\partial \mathcal{L}}{\partial \phi} \Delta \phi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \partial_\mu \Delta \phi \right)
\]
\[
\mathcal{L}(\phi, \partial \mu \phi) + \epsilon \left( \frac{\partial \mathcal{L}}{\partial \phi} - \partial \mu \frac{\partial \mathcal{L}}{\partial (\partial \mu \phi)} \right) \Delta \phi + \epsilon \partial \mu \left( \frac{\partial \mathcal{L}}{\partial (\partial \mu \phi)} \Delta \phi \right)
\]

By combining the previous two equations for \( \mathcal{L}(\phi') \), we see that on configurations which satisfy the EOM, \( 0 = \partial \mu j^\mu \) with

\[
 j^\mu = \frac{\partial \mathcal{L}}{\partial (\partial \mu \phi)} \Delta \phi - \mathcal{J}^\mu.
\]

Notice that I stuck the index back in at the last step.

There is a converse to the Noether theorem, which is easier to discuss directly in quantum mechanics. Given a conserved charge \( Q \), that is, a hermitian operator with \([H, Q] = 0\), we can make a symmetry transformation of the fields \( \phi \) by

\[
\delta \phi \equiv i \epsilon [Q, \phi].
\]

We’ll say that \( Q \) *generates* the symmetry, for the following reason. \((1.32)\) is the infinitesimal version of the finite transformation

\[
\phi \rightarrow \phi' \equiv e^{i\epsilon Q} \phi e^{-i\epsilon Q}.
\]

The object \( U \equiv e^{i\epsilon Q} \) is a unitary operator (since \( Q \) is hermitian) which represents the action of the symmetry on the Hilbert space of the QFT. It is a symmetry in the sense that it commutes with the time evolution operator \( e^{-iHt} \).

Some examples will be useful:

- For example, suppose \( S[\phi] \) only depends on \( \phi \) through its derivatives, for example, \( S[\phi] = \int \frac{1}{2} \partial \mu \phi \partial \mu \phi \). Then there is a shift symmetry \( \phi \rightarrow \phi' \equiv \phi + \epsilon \). Letting \( \epsilon \) depend on spacetime, the variation of the action is \( \delta S[\phi + \epsilon(x)] - S[\phi] = -\int \epsilon \partial \mu \partial^\mu \phi \), so the current is \( j^\mu = \partial \mu \phi \). Let’s check the converse: Indeed, the charge \( Q = \int_{\text{space}} j_0 \) generates the symmetry in the sense that for small \( \epsilon \), the variation in the field is

\[
\delta \phi \equiv \phi' - \phi = \epsilon = i \epsilon [Q, \phi]
\]

(if we were doing classical mechanics, we should replace \( i [Q, \phi] \) with the Poisson bracket). Using our expression for the current this is

\[
\delta \phi = i \epsilon \left[ \int d^d y \frac{\dot{\phi}(y)}{\pi(y)}, \phi(x) \right] = \epsilon
\]

which is indeed true by the canonical commutation relations. In this case the finite transformation is again \( \phi \rightarrow \phi + \epsilon \).
• Consider a complex scalar $\Phi$, and suppose $S[\Phi, \Phi^*]$ is invariant under $\Phi \rightarrow e^{i\epsilon}\Phi = \Phi + i\epsilon\Phi + \mathcal{O}(\epsilon^2)$, such as $S = \int (\partial\Phi^* \partial\Phi - V(\Phi^*\Phi))$. This $U(1)$ phase transformation can be rewritten in terms of the real and imaginary parts as an $SO(2)$ rotation. The charge can be written as

$$Q = \int d^d x j^0 = \int d^d p (a^\dagger p a_p - b^\dagger p b_p)$$

where the two sets of creation and annihilation operators are associated with excitations of $\Phi$ and $\Phi^\dagger$ respectively. (That is, quantize $\phi_{1,2}$ as we did for a single real scalar field, in terms of mode operators $a_{1,2}$ respectively. Then let $a \equiv a_1 + ia_2, b \equiv a_1 - ia_2$, up to numerical prefactors.) So the particles created by $a$ and $b$ have opposite charge (this follows given the mode expansion $\Phi_k \sim a_k + b^\dagger_{-k}$) and can be interpreted as each others’ antiparticles: there can be symmetry-respecting processes where an $a$ particle and $b$ particle take each other out.

The previous two examples are related. Consider the case where $V(\Phi^*\Phi) = \lambda(\Phi^*\Phi - v^2)^2$ Changing variables to polar coordinates in field space, $\Phi = \rho e^{i\phi}$, the Lagrangian is

$$\mathcal{L} = \rho^2(\partial\phi)^2 + (\partial\rho)^2 - \lambda(\rho^2 - v^2)^2.$$ 

If $\lambda$ is big, the potential forces $\rho = v$, and its fluctuations are heavy, and we are left with $\mathcal{L} = v^2(\partial\phi)^2$, where $\phi \rightarrow \phi + \epsilon$ is a symmetry. Notice that $\phi \equiv \phi + 2\pi$ is periodic.

• Consider spacetime translations, $x^\mu \rightarrow x^\mu - a^\mu$. We can think of this as a transformation of the fields by

$$\phi(x) \mapsto \phi(x + a) = \phi(x) + a^\nu \partial_\nu \phi + \mathcal{O}(a^2).$$

Our transformation parameter is now itself a four-vector, so we’ll get a four-vector of currents $T^\mu_\nu$. This will be a symmetry as long as the lagrangian doesn’t depend explicitly on space and time (so $\partial_\mu \mathcal{L} = 0$) but rather depends on space and time only via the fields (so $0 \neq \frac{d}{dx^\mu} \mathcal{L}$ chain rule $= \partial_\mu \phi \partial_\nu \mathcal{L} + \partial_\nu \phi \partial_\mu \mathcal{L} \partial_{\mu,\nu} \mathcal{L}$). Let’s use the prosaic method for this one: the shift in the Lagrangian density also can be found by Taylor expansion

$$\mathcal{L} \mapsto \mathcal{L} + a^\mu \frac{d}{dx^\mu} \mathcal{L} = \mathcal{L} + a^\nu \partial_\mu (\delta^\mu_\nu \mathcal{L}).$$
So the formula (1.31) gives

\[ T^\mu_\nu = \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \partial_\nu \phi - \mathcal{L} \delta^\mu_\nu. \]

For the time translation, the conserved charge \( T^0_0 \) gives back the hamiltonian density \( \mathfrak{h} = \pi \dot{\phi} - \mathcal{L} \) obtained by Legendre transformation. The conserved quantity from spatial translations is the momentum carried by the field, which for the KG field is

\[ \mathbf{P}_i = \int d^d x \ T^0_i = - \int d^d x \ \pi \partial_i \phi. \]

For the Maxwell field, this gives the Poynting vector.

There is some ambiguity in the definition of the stress tensor (associated with the possibility of adding total derivatives to \( \mathcal{L} \)).

Let’s check that the expression above for the conserved momentum agrees with our expectations. In particular, in free field theory the total momentum of the state \( | \mathbf{k}_1, \cdots, \mathbf{k}_n \rangle \) should be just the sum of the momenta of the particles, \( \mathbf{P} = \sum_{\ell=1}^n \hbar \mathbf{k}_\ell \) (with interactions the story can be more complicated). Indeed

\[ \mathbf{P}_i = - \int d^d x \ \pi \partial_i \phi = \int d^d k k_i \mathbf{a}_k^\dagger \mathbf{a}_k \]

agrees with this. (Notice that I used rotation invariance of the vacuum to not worry about a possible constant term.)

- I have the impression that you learned all about the rest of the Poincaré group already in 215A.
2 From correlation functions to the $S$ matrix

We’ve derived an expression for correlation functions, such as (1.22), in terms of a sum of diagrams connected to the external lines, ordered by the number of powers of the coupling constant. Our next goal is to organize this sum.

First let’s make contact with the way the diagrammatic expansion was (I think) introduced in 215A. A time-ordered real-time Green’s function has path integral representation (with $i\epsilon$ prescription implicit, and $\phi_i \equiv \phi(x_i)$)

$$G^n \equiv \langle 0 | T \phi_1 \cdots \phi_n | 0 \rangle = Z^{-1} \int [D\phi] \phi_1 \cdots \phi_n e^{iS[\phi]}$$ \hspace{1cm} (2.1)

$$= \frac{\int [D\phi] \phi_1 \cdots \phi_n e^{-i \int V(\phi)e^{iS_0[\phi]}}}{\int [D\phi] e^{iS_0[\phi]} e^{-i \int V(\phi)}} \hspace{1cm} (2.2)$$

$$= \frac{\langle 0 | T \phi_1 \cdots \phi_n e^{-i \int V(\phi)} | 0 \rangle}{\langle 0 | T e^{-i \int V(\phi)} | 0 \rangle}. \hspace{1cm} (2.3)$$

Here we’ve written $S = S_0 - \int V$ where $S_0$ is gaussian. The last object here is a time-ordered expectation value in the free theory, which we know how to compute by Wick contraction.

Two comments about this formula: (1) It must be admitted that in (2.1) the variable ‘$\phi$’ is seriously overloaded: on the LHS it is used to represent a (Heisenberg-picture) operator, while on the RHS it is used to represent a (functional) integration variable. (2) This formula (or a related one) is sometimes called the Dyson formula for interaction-picture time evolution.

Taylor expanding the exponential $e^{-i \int V}$ in (2.3) reproduces the diagrammatic expansion. (Notice that in real-time quantities, the interaction vertex comes with a factor of $-ig$.) The denominator is the sum of bubble diagrams. In both numerator and denominator, the disconnected diagrams exponentiate, and therefore cancel. Here is a reminder of why this is true:

---

**The exponentiation of the disconnected diagrams.** [Peskin page 96] There are some patterns in these sums of diagrams to which it behooves us to attend. (The
following discussion transcends the $\phi^4$ example.) The general diagram has the form:

Only some of the components are attached to the external legs; for a given diagram $A$, call the factor associated with these components $A_c$ (note that $A_c$ need not be fully connected). The rest of the diagram is made of a pile of ‘bubbles’ of various types $V_i$ (each one internally connected, but disconnected from the external lines) and multiplicities $n_i$ (e.g. $V_1$ could be a figure eight, and there could be $n_1 = 2$ of them. These bubbles (or ‘vacuum bubbles’) would be there even if we didn’t have any external lines, and they would have the same value; they are describing the fluctuations intrinsic to the vacuum. The amplitude associated with the general diagram is then

$$M_A = M_{A_c} \cdot \frac{V_1^{n_1}}{n_1!} \cdot \frac{V_2^{n_2}}{n_2!} \cdots \frac{V_\alpha^{n_\alpha}}{n_\alpha!}$$

where the $n_i!$ factors are the most important appearance of symmetry factors: they count the number of ways to permute the identical copies of $V_i$ amongst themselves.

The numerator of $G^{(n)}$ is then

$$G^{(n)}_{\text{numerator}} = \langle 0 | T \left( \phi_1 \cdots \phi_n e^{-i \int V} \right) | 0 \rangle = \sum_A M_A = \sum_{A_c} M_{A_c} \sum_{\{n_i=0\}} V_1^{n_1} \cdot \frac{V_2^{n_2}}{n_2!} \cdots \frac{V_\alpha^{n_\alpha}}{n_\alpha!}$$

$$= \sum_{A_c} M_{A_c} \cdot e^{\sum_i V_i}$$

$$= \sum_{A_c} M_{A_c} e^{\sum_i V_i} \quad (2.4)$$

– the bubbles always exponentiate to give the same factor of $e^{\sum_i V_i}$, independent of the external data in $G$. In particular, consider the case of $n = 0$, where there are no external lines and hence no $A_c$:

$$G^{(0)}_{\text{numerator}} = \langle 0 | T e^{-i \int V} | 0 \rangle = 1 \cdot e^{\sum_i V_i}$$

But we care about this because it is the denominator of the actual Green’s function:

$$G^{(n)} = \frac{\langle 0 | T \left( \phi_1 \cdots \phi_n e^{-i \int V} \right) | 0 \rangle}{\langle 0 | T e^{-i \int V} | 0 \rangle} = \frac{G^{(n)}_{\text{numerator}}}{G^{(0)}_{\text{numerator}}} = \sum_{A_c} M_{A_c} . \quad (2.5)$$
And with that we can forget all about the bubbles. So for example,

\[
G^{(2)} = \phantom{\text{Diagram}} + \frac{8}{1 + 8 + 88 + \cdots} = \frac{G^{(4)}}{1 + 8 + 88 + \cdots} + \frac{\delta^{d+1}(q_0 - p)}{q_0^2 - m^2 + \epsilon}. 
\]

Notice that in this manipulation (2.5) we are adding terms of many orders in perturbation theory in the coupling \( g \). If we want an answer to a fixed order in \( g \), we can regard anything of higher order as zero, so for example, it makes perfect sense to write

\[
G^{(2)} = \frac{\delta^{d+1}(q_0 - p)}{q_0^2 - m^2 + \epsilon} + \mathcal{O}(g) = \frac{\delta^{d+1}(q_0 - p)}{q_0^2 - m^2 + \epsilon} + \mathcal{O}(g) = \frac{\delta^{d+1}(q_0 - p)}{q_0^2 - m^2 + \epsilon}. 
\]

(I only drew one kind of bubble in the previous expression since that one was easy to type.)

**Momentum space Green’s functions from Feynman diagrams.** In translation-invariant problems, things are usually a little nicer in momentum space. In \( \phi^4 \) theory in \( d + 1 \) dimensions, let’s think about

\[
\tilde{G}^{(n)}(p_1 \cdots p_n) \equiv \prod_{i=1}^{n} \int d^{d+1} p_i e^{-i p_i x_i} G^{(n)}(x_1 \cdots x_n). 
\]

This an off-shell Green’s function, a function of general \( p \), not necessarily \( p^2 = m^2 \). It will, however, vanish unless \( \sum_{i} p_i^0 = 0 \) by translation invariance. Consider a fully-connected contribution to it, at order \( g^N \). (We’ll get the others by multiplying these bits.)

In \( \phi^4 \) theory, we need to make a diagram by connecting \( n \) external position vertices \( x_i \) to \( N \) 4-valent vertices \( z_a \) using Feynman propagators \( \Delta_F(y_A - y_B) = \int d^{d+1} q e^{-i (y_A - y_B) q} \frac{1}{q^2 - m^2 + \epsilon} \), where \( \{y_A\} = \{z_a, x_i\} \). All of the position dependence is in these exponentials.

Since each propagator has two ends, the number of lines (by the fully-connected assumption) is

\[
N_I = \frac{\# \text{ of ends of lines}}{2} = \frac{n + 4N}{2} = \frac{n}{2} + 2N.
\]

This is the number of \( q \) integrals, a priori.

The integral over the external positions \( x_i \) (in the Fourier transform) gives \( \int d^{d+1} x_i e^{i \sum (q_r - p_i)} = \delta^{d+1}(q_i - p_i) \) and so we can label the external lines by \( p_i \) (and we lose \( n q \) integrals).

The integral over the position of each internal vertex is of the form \( \int d^{d+1} z e^{i \sum (\Sigma_r q_r)} \) where \( q_r \) are the momenta associated to the lines coming into the vertex. So each
internal vertex decreases the number of $q$ integrals by 1. One combination of the momenta is fixed by overall momentum conservation so we have left

$$N_I - n - (N - 1) = N - \frac{n}{2} + 1 \equiv N_L$$

momentum integrals. This number is $\geq 0$ for fully connected diagrams, and it is the number of loops in the diagram. (This counting is the same as in a Kirchoff’s law resistor network problem.)\(^{11}\)

For example, consider a particular contribution to $G^{(4)}$ ($n = 4$ external legs) and $N = 2$ interaction vertices

which has $N_L = \frac{4 + 2 \cdot 4}{2} = 6$.

In the example, $N_L = 2 - 2 + 1 = 1$ which agrees with one undetermined momentum integral. This gives the amplitude

$$\mathcal{M}_{FC}(p_1 \ldots p_n) = (-ig)^N \cdot s(FC) \phi^{(d+1)} \left( \sum p_i \right) \int \prod_{\text{loops, } \alpha=1}^{N_L} d^{d+1} k_\alpha \prod_{\text{lines, } r=1}^{N} \frac{i}{q_i^2 - m^2 + i\epsilon}$$

(You might notice that the integral over $k$ is in fact formally infinite, since at large $k$ it goes like $\int \Lambda^d \frac{d^4 k}{k^2} \sim \log(\Lambda)$. Try to postpone that worry.) The propagators for the external lines just factor out, and can be brought outside the momentum integrals. Notice that here $p$ is general, and this function has poles when the external particles go on-shell, $p_i^2 = m^2$.

So here are the momentum space Feynman rules for Green’s function in $\phi^4$ theory:

- Every line gives a factor of $\sqrt{\frac{i}{p^2 - m^2 + i\epsilon}} = \tilde{\Delta}_F(p)$. Notice that since $\Delta_F(x - y) = \Delta_F(y - x)$, the choice of how we orient the momenta is not so fateful.

- An internal vertex gives $\sim (-ig) \int d^{d+1} z e^{-is} \sum_i p_i z = (-ig) \phi^{(d+1)}(\sum_i p_i)$, momentum conservation at each vertex. So, set $\sum_i p_i = 0$ at each vertex (I’ve assumed the arrows are all pointing toward the vertex). After imposing momentum

\(^{11}\) Here’s a proof that (2.6) is the number of loops in the diagram: place the $N + n$ internal and external vertices on the page. Add the propagators one at a time. You must add $N + n - 1$ just to make the diagram fully connected. After that, each line you add makes a new loop.
conservation, the remaining consequence of the vertex is

\[ = -ig. \]

- Integrate over the loop momenta \( \prod_{\alpha=1}^{N_L} d^{d+1} q_\alpha \) for each undetermined momentum variable. There is one for each loop in the diagram. You should think of these integrals as just like the Feynman path integral: if there is more than one way to get from here to there, we should sum over the amplitudes.

- Multiply by the wretched symmetry factor \( s(A) \).

- For \( \tilde{G}(p) \), multiply by an overall \( \phi^{d+1}(\sum p) \) in each diagram.

- An external vertex at fixed position, \( e^{-ipx} \). (Such vertices would arise if we wanted to compute \( G(x) \) using momomentum-space feynman rules.) More generally, external vertices are associated with the wavefunctions of the states we are inserting; here they are plane waves.

Here is another perspective on the exponentiation of the vacuum bubbles. Consider the diagram:

\[ (-ig)^2 \prod_{i=1}^{4} \int d^{d+1}p_i \phi^{d+1}(p_1 + p_2) \phi^{d+1}(p_1 + p_2) \cdots \]

The two delta functions come from the integrals over \( z_{1,2} \), and we can restore sense by remembering this:

\[ \left( \phi^{d+1}(p_1 + p_2) \right)^2 = \phi^{d+1}(p_1 + p_2) \int d^{d+1}z_2 = \phi^{d+1}(p_1 + p_2)VT \]

where \( VT \) is the volume of spacetime. This factor arises because this process can happen anywhere, anytime. There is one such factor for each connected component of a collection of vacuum bubbles, so for example the diagram \( \tilde{G}(0) \) is proportional to \( (VT)^2 \). But the free energy \( \propto \log Z = \log G(0) \) should be extensive, \( \propto VT \). Therefore, the vacuum bubbles must exponentiate.

The whole two point function in momentum space is then (through order \( g^2 \)):

\[ \tilde{G}^{(2)} = \sum_{N_L = 0}^{N_L = 2} \mathcal{O}(g^3) \] (2.7)
I draw the blue dots to emphasize the external propagators. Notice that for the two-point function, the number of loops is $N_L = N - \frac{n}{2} + 1 = N$, the same as the number of powers of $g$. More generally, for $n \neq 2$, there is an additive shift: $N_L = \text{constant plus number of powers of } g$.

**Organizing the propagator.** We would like to unpack the physics contained in the correlation functions which we’ve learned to compute in perturbation theory. The first interesting one is the two-point function aka the propagator. Let’s factor out the overall delta function by writing:

$$\tilde{G}^{(2)}(p_1, p_2) \equiv \delta^{d+1}(p_1 + p_2)\tilde{G}^{(2)}(p_1).$$

It will be useful to re-organize this sum, in the following way:

Here’s the pattern: we define a diagram to be one-particle irreducible (1PI) if it cannot be disconnected by cutting through a single internal propagator. So for example, is 1PI, but is not; rather, the latter contributes to the bit with two 1PI insertions. Then

$$\tilde{G}^{(2)}(p) = \quad \text{1PI} \quad + \quad \text{1PI} \quad + \quad \text{1PI} \quad + \quad \text{1PI} \quad + \ldots$$

So that we may write equations without pictures, let

$$-i\Sigma(p) \equiv \text{1PI}$$

denote the 1PI two-point function. $\Sigma$ being 1PI means that the external lines sticking out of it are ‘nubbins,’ placeholders where propagators may be attached. That’s why there are no blue dots at the ends.
Now suppose we know $\Sigma$. It is known as the self-energy, for reasons we will see next. Then we can write

$$\tilde{G}^{(2)}(p) = \frac{i}{p^2 - m_0^2} + \frac{i}{p^2 - m_0^2} (-i\Sigma(p)) \frac{i}{p^2 - m_0^2} + \frac{i}{p^2 - m_0^2} (-i\Sigma(p)) \frac{i}{p^2 - m_0^2} (\Sigma(p)) \frac{i}{p^2 - m_0^2} + \cdots$$

$$= \frac{i}{p^2 - m_0^2} \left( 1 + \frac{\Sigma}{p^2 - m_0^2} + \left( \frac{\Sigma}{p^2 - m_0^2} \right)^2 + \cdots \right)$$

$$= \frac{i}{p^2 - m_0^2} \left( 1 - \frac{\Sigma}{p^2 - m_0^2} \right) = \frac{i}{p^2 - m_0^2 - \Sigma(p)}. \tag{2.8}$$

We see that the self-energy shifts the $m^2$ of the particle – it moves the location of the pole in the propagator. In the interacting theory, $m_0^2 + \Sigma(p)_{|\text{pole}}$ is the physical mass, while $m_0$ (what we’ve been calling $m$ until just now) is deprecatingly called the ‘bare mass’. For $p^2 \sim m^2$, we will write

$$\tilde{G}^{(2)}(p) \equiv \left( \frac{iZ}{p^2 - m^2} + \text{regular bits} \right) \tag{2.9}$$

This equation defines the residue $Z$ which is called the ‘wavefunction renormalization factor’. It is 1 in the free theory, and represents the amplitude for the field to create a particle, and the other terms, which are not singular at $p^2 = m^2$, represent the amplitude for the field to do something else (such as create multiparticle states), and are not present in the free theory. Later we will see that unitarity requires $Z \leq 1$. Notice that if we know $\Sigma$ only to some order in perturbation theory, then (2.8) is still true, up to corrections at higher order.

The notion of 1PI extends to diagrams for $\tilde{G}^{(n>2)}(p_1 \cdots p_n)$. Let

$$\tilde{G}_{1PI}^{(n)}(p_1 \cdots p_n) \equiv \begin{array}{c}
\text{1PI} \\
\text{1PI}
\end{array}$$

where the blob indicates the sum over all 1PI diagrams with $n$ external nubbins (notice that these do not have the blue circles that were present before). This means $G_{1PI}$ does not include diagrams like:
Notice that 1PI diagrams are *amputated* – their external limbs have been cut off.

**LSZ reduction formula.** This is almost what we need to make $S$-matrix elements. If we multiply the $n$-point function by $\prod_{i=1}^{n} p_i^2 - m^2 \sqrt{Z}$ we cancel out the propagators from the external legs, near the mass shell. This object is naturally called the *amputated* $n$-point function. (It differs from the 1PI $n$-point Green’s function because of diagrams like this one which is amputated but not 1PI.) If we then take

$p_i^2 \to m^2$, we keep only the part of $\tilde{G}$ which is singular on the mass-shell. And here’s why we care about that:

Claim (the LSZ reduction formula):

$$ S_{fi} \equiv \langle \bar{p}_1 \cdots \bar{p}_n | S | k_1 \cdots k_m \rangle = \prod_{a=1}^{n+m} \left( \lim_{P_a \to E} \frac{P_a^2 - m^2}{i \sqrt{Z}} \right) \tilde{G}^{(n+m)}(k_1 \cdots k_m, -p_1 \cdots -p_n) $$

(2.10)

where $P_a \in \{p_i, k_i\}$. In words: the $S$-matrix elements are obtained from Green’s functions by amputating the external legs, and putting the momenta on-shell. Notice that choosing all the final momenta $p_i$ different from all the initial momenta $k_i$ goes a long way towards eliminating diagrams which are not fully connected.

This formula provides the bridge from time-ordered Green’s functions (which we know how to compute in perturbation theory now) and the $S$-matrix, which collects probability amplitudes for things to happen to particles, in terms of which we may compute cross sections and lifetimes. Let us spend just another moment inspecting the construction of this fine conveyance.

---

**Why is LSZ true?** Here’s the argument I’ve found which best combines concision and truthiness. [It is mainly from the nice book by Maggiore §5.2; I also like Schwartz’ chapter 6; Peskin’s argument is in section 4.6.] The argument has several steps. The field operators in this discussion are all in Heisenberg picture.

1. First, for a free field, the mode expansion implies that we can extract the ladder operators by:

$$ \sqrt{2 \omega_k} a_k = i \int d^d x \ e^{ikx} (-i \omega_k + \partial_0) \phi_{\text{free}}(x) $$
\[ \sqrt{2\omega_k} a_k^\dagger = -i \int d^d x \ e^{-ikx} (i \omega_k + \partial_0) \phi_{\text{free}}(x) \quad (2.11) \]

Notice that the LHS is independent of time, but the integrand of the RHS is not.

2. Now let’s pretend that we can turn the interactions off at \( t = \pm \infty \), so that the asymptotic states we are scattering are free particles.\(^\text{12}\) This allows us to write the field in terms of some pretend free fields of mass \( m \) (not \( m_0 \! \) !)

\[ \phi(x) \left\{ \begin{array}{ll} t \to -\infty & Z^{\frac{1}{2}} \phi_{\text{in}}(x) \\ t \to +\infty & Z^{\frac{1}{2}} \phi_{\text{out}}(x) \end{array} \right. \]

The factors of \( \sqrt{Z} \) are required to get the correct two point functions (2.9) near the mass shell. The mode operators for \( \phi_{\text{in}} \) are called \( a_{\text{(in)}} \) etc. \( \phi_{\text{in}}, \phi_{\text{out}} \) are free fields: their full hamiltonian is \( H_0 \). They are in Heisenberg picture, and the reference time for \( \phi_{\text{in}}, \phi_{\text{out}} \) is \( \pm \infty \) respectively. Since they are free fields, we can use (2.11) to write

\[ \sqrt{2\omega_k} a_{\text{(in)}}^\dagger = -i \int d^d x \ e^{-ikx} (i \omega_k + \partial_0) \sqrt{Z}^{-1/2} \int d^d x \ e^{-ikx} (i \omega_k + \partial_0) \phi(x) \left| t \to -\infty \right. \]

where in the second step we used the independence on time in (2.11), even though \( \phi(x) \) is not a free field. An expression for \( a_{\text{(out)}}^\dagger \) obtains if we take \( t \to +\infty \) instead.

3. Now make this expression manifestly covariant using the fundamental theorem of calculus:

\[ \sqrt{2\omega_k} (a_{\text{(in)}}^\dagger - a_{\text{(out)}}^\dagger) = i Z^{-1/2} \int_{-\infty}^{\infty} dt \partial_t \left( \int d^d x \ e^{-ikx} (i \omega_k + \partial_0) \phi(x) \right) \]

IBP in time \( i Z^{-1/2} \int d^{d+1} x \ e^{-ikx} \left( \partial_0^2 \phi - \phi \cdot \partial_0^2 e^{-ikx} \right) \)

IBP in space \( i Z^{-1/2} \int d^{d+1} x e^{-ikx} (\Box + m^2) \phi(x) \quad (2.12) \]

In the last step we made a promise to only use wavepackets for external states, so that we can do IBP in space.

4. Now, here’s where the \( S \)-matrix enters. Assume none of the incoming momenta \( k_i \) is the same as any outgoing momentum \( p_j \).

\[ \langle p_1 \cdots p_n | S | k_1 \cdots k_m \rangle \]

\(^{12}\)Here’s why this is really bad: nearly everything we might scatter is a boundstate. For example: atoms, nuclei, nucleons etc... But if there are no interactions there are no boundstates.
\[ \begin{align*}
&= \prod_{p,k} \sqrt{2\omega} \langle \Omega | \hat{a}_p^\text{out} S \prod_k a_k^\text{inf} | \Omega \rangle \\
&= \prod_{p,k} \sqrt{2\omega} \langle \Omega | T \left( \prod_p \hat{a}_p^\text{out} S \prod_k a_k^\text{inf} \right) | \Omega \rangle \\
&= \prod_{p,k} \sqrt{2\omega} \langle \Omega | T \left( \prod_p \hat{a}_p^\text{out} S \left( a_{k_1}^\text{inf} - a_{k_1}^\text{out} \right) \prod_m \frac{m}{2} a_k^\text{inf} \right) | \Omega \rangle \\
&= \prod_{p,k} \sqrt{2\omega} \langle \Omega | T \left( \prod_p \sqrt{\omega_p} \hat{a}_p^\text{out} S \phi(x_1) \prod_m \frac{m}{2} \sqrt{\omega_k} a_k^\text{inf} \right) | \Omega \rangle + X
\end{align*} \]

\[ (2.12) \]

In the last step, \( X \) comes from where the \( \Box x_1 \) hits the time ordering symbol. This gives terms which will not matter when we take \( k^2 \to m^2 \), I promise.

5. Now do this for every particle to get

\[ \langle p_1 \cdots p_n | S | k_1 \cdots k_m \rangle = \prod_{j=1}^m \int d^{d+1}y_j \ e^{ip_j y_j} iZ^{-1/2} (\Box_j + m^2) \prod_{i=1}^n \int d^{d+1}x_i \ e^{-ik_i x_i} iZ^{-1/2} (\Box_i + m^2) \langle \Omega | T \phi(x_1) \cdots \phi(y_j) S | \Omega \rangle + X \]

The \( x \) and \( y \) integrals are just Fourier transforms, and this says that near the mass shell,

\[ \tilde{G}^{(n+m)}(k_1 \cdots k_m, -p_1 \cdots -p_n) = \prod_a \frac{i\sqrt{Z}}{P_a^2 - m^2} \langle p_1 \cdots p_n | S | k_1 \cdots k_m \rangle + \text{regular} \]

(where \( P_a \in \{p_j, k_i\} \)) which is the same as \( (2.10) \).

---

Comment: In our discussion of QFT, a special role has been played by fields called \( \phi \). Suppose we have some other (say hermitian) local operator \( \mathcal{O} \) such that

\[ \langle p | \mathcal{O}(x) | \Omega \rangle = Z_{\mathcal{O}} e^{ipx} \]

where \( \langle p \rangle \) is a one-particle state made by our friend \( \phi \) (we could put some labels, \( e.g. \) for spin or polarization or flavor, on both the operator and the state, but let’s not). Such an \( \mathcal{O} \) is called an ‘interpolating field’ or ‘interpolating operator’. And suppose we have information about the correlation functions of \( \mathcal{O} \):

\[ G^{(n)}_{\mathcal{O}}(1 \cdots n) \equiv \langle \Omega | T (\mathcal{O}_1(x_1) \cdots \mathcal{O}_n(x_n)) | \Omega \rangle . \]

In this case, there is a more general statement of LSZ:

\[ \prod_{a \in i} \left( Z_a^{-1/2} i \int d^{d+1}x_a e^{-ip_a x_a} (\Box_a + m_a^2) \right) \]
\[
\prod_{b \in f} \left( Z_b^{-1/2} i \int d^{d+1}x_b e^{i p_b x_b} (\Box_b + m_b^2) \right) \mathcal{O}_\mathcal{O}^{(n)}(1 \cdots n) = \langle \{p_f\} | S | \{p_a\} \rangle \tag{2.13}
\]

This more general statement follows as above if we can write \( \mathcal{O}_a \to -\infty \sqrt{Z_a} \phi_m \). This more general formula allows us to scatter particles that are not 'elementary' in the sense that they are made by the fields in terms of which we write our Lagrangian.

Here is a summary of the long logical route connecting Feynman diagrams to measurable quantities in particle physics:

The final step was covered in 215A.
**S-matrix from Feynman diagrams.** The end result of the previous discussion is a prescription to compute $S$-matrix elements from Feynman diagrams. In a translation-invariant system, the $S$ matrix always has a delta function outside of it. Also we are not so interested in the diagonal elements of the $S$ matrix where nothing happens. So more useful than the $S$ matrix itself are the scattering amplitudes $M$ defined by

$$
\langle f | (S - \mathbb{1}) | i \rangle \equiv (2\pi)^{d+1}\delta^{(d+1)} \left( \sum_f p_f - \sum_i p_i \right) i M_{fi}.
$$

(2.14)

(The object $i M \delta^{d+1}(\sum p)$ is sometimes called the transfer matrix. The $i$ is a convention.)

The rules for the Feynman diagram calculation of $M$ (for $\phi^4$ theory, as a representative example) are:

1. Draw all amputated diagrams with appropriate external nubbins for the initial and final states. For a diagram with $N_L$ loops think of $N_L$ letters that are like $k$ or $q$ or $p$ to call the undetermined loop momenta.

2. For each vertex, impose momentum conservation and multiply by the coupling $(-i\lambda)$.

3. For each *internal* line, put a propagator.

4. For each loop, integrate over the associated momentum $\int d^{d+1}k$.

A comment about rule 1: For tree-level diagrams (diagrams with no loops), ‘amputate’ just means leave off the propagators for the external lines. More generally, it means leave off the resummed propagator (2.8). In particular, a diagram like is already included by using the correct $Z$ and the correct $m$.

---

I skipped the example in lecture. Please read through it and make sure you are happy about it. Please ask questions if you are not. We will do some examples in QED soon.

**Example: nucleon scattering.** [Here we follow Tong §3.5 very closely] Let’s consider an example with a complex scalar field $\Phi$ interacting with a real scalar field $\phi$ with Lagrangian

$$
\mathcal{L} = \frac{1}{2} \partial_\mu \Phi^* \partial^\mu \Phi - \frac{1}{2} m^2 \Phi^* \Phi + \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} M^2 \phi^2 + \mathcal{L}_I
$$

(2.15)

with $\mathcal{L}_I = -g \Phi^* \Phi \phi$. 

46
In specifying initial states below, I will need names for the mode operators of the two fields:

\[ \phi = \int \frac{d^dp}{\sqrt{2\omega_p}} (a_p e^{-ipx} + a^+_p e^{ipx}) |p^0=\omega_p \]

\[ \Phi = \int \frac{d^dp}{\sqrt{2E_p}} (b_p e^{-ipx} + c^+_p e^{ipx}) |p^0=E_p \]

where I’ve written \( \omega_p \equiv \sqrt{M^2 + p^2}, E_q \equiv \sqrt{m^2 + q^2} \). Notice that the \( \Phi \to e^{-i \alpha \Phi} \) symmetry is conserved; the charge is

\[ q = N_c - N_b. \]

But the \( \phi \) particles are not conserved.\(^\text{13}\) Relative to \( \phi^4 \) theory, the differences in the Feynman rules are: we have two kinds of propagators, one of which is oriented (to keep track of the flow of the conserved \( \Phi \) number through the diagrams), and instead of a 4-point vertex which costs \(-ig\), we have a 3-point vertex for \( \phi \Phi \Phi \) which costs \(-ig\).

Let’s consider \( 2 \to 2 \) scattering of \( \Phi \) particles, so

\[ |i\rangle = |\vec{p}_1, \vec{p}_2\rangle, |f\rangle = |\vec{p}_3, \vec{p}_4\rangle \quad \text{with} \quad |\vec{p}_i, \vec{p}_j\rangle \equiv \sqrt{2E_{\vec{p}_i}} \sqrt{2E_{\vec{p}_j}} b^\dagger_{\vec{p}_i} b^\dagger_{\vec{p}_j} |0\rangle. \]

(To appreciate some of the beauty of the diagram technique, see Tong §3.3.3 for the artisanal version of this calculation.) The Feynman rules above give, to leading nonzero order,

\[ iM = (-ig)^2 \left( \frac{i}{(p_1 - p_3)^2 - M^2 + i\epsilon} + \frac{i}{(p_1 - p_4)^2 - M^2 + i\epsilon} \right). \]

(2.16)

The diagrams depict two ‘snucleons’ \( \Phi \) (solid lines with arrows indicating snucleons versus antinucleons) exchanging a meson \( \phi \) (double gray line, with no arrow) with momentum \( k \equiv p_1 - p_3 = p_4 - p_2 \) (first term) or \( k \equiv p_4 - p_1 = p_2 - p_3 \) (second term). Time goes to the left as always. Notice that here I am being careful about using arrows on the lines to indicate flow of particle number through the diagram, while the extra (light blue) arrows indicate momentum flow.

\(^{13}\)You might notice a possible problem with this theory: what happens to the quadratic term for \( \Phi \) when \( \phi \) is very negative? Let’s not take it too seriously.
The meson in these diagrams is virtual, or off-shell, in the sense that it does not satisfy its equation of motion $k^2 \neq M^2$. In fact, each of these diagrams is actually the sum of retarded and advanced exchange of real on-shell particles. (For more on this statement, see Schwartz chapter 4 or §4.5 of the notes here.) The two diagrams included in (2.16) make the amplitude symmetric under interchanging the two particles in the initial or final state, as it must be because they are indistinguishable bosons.

Two more examples with the same ingredients are useful for comparison. If we instead scatter a snucleon and an anti-snucleon, so $|i\rangle = \sqrt{\frac{2E_{p_1}}{2E_{p_2}}} b^\dagger_{p_1} c^\dagger_{p_2} |0\rangle$, then the leading diagrams are

\[
iM = (-ig)^2 \left( \frac{i}{(p_1 + p_2)^2 - M^2 + i\epsilon} + \frac{i}{(p_1 - p_3)^2 - M^2 + i\epsilon} \right). \tag{2.17}
\]

This one has a new ingredient: in the first diagram, the meson momentum is $k = p_1 + p_2$, which can be on-shell, and the $i\epsilon$ matters. This will produce a big bump, a resonance, in the answer as a function of the incoming center-of-mass energy $\sqrt{s} \equiv \sqrt{(p_1 + p_2)^2}$.

Finally, we can scatter a meson and a snucleon:

\[
iM = (-ig)^2 \left( \frac{i}{(p + k)^2 - m^2 + i\epsilon} + \frac{i}{(p' - k')^2 - m^2 + i\epsilon} \right). \tag{2.18}
\]

Now the intermediate state is a snucleon.

There is a common notation for the Lorentz-invariant combinations of the momenta appearing in these various processes, called Mandelstam variables, of which $s$ is one. A concise summary appears in §3.5.1 of Tong’s notes.
3 QED

3.1 Quantum light: Photons

I skipped this subsection in lecture. Please read through it and make sure you are happy about it. Please ask questions if you are not.

The quantization of the Maxwell field is logically very similar to the case of a harmonic chain. There are just a few complications from its several polarizations, and from the fact that quantum mechanics means that the vector potential is real and necessary (whereas classically it is just a convenience). This is a quick-and-dirty version of the story. I mention it here to emphasize that the machinery we are developing applies to a system you have already thought a lot about!

Maxwell’s equations (with $c = 1$) are:

\[
\begin{align*}
\epsilon_{\mu
\nu\rho\sigma} \partial_{\nu} F_{\rho\sigma} &= 0 \\
\partial_{\mu} F_{\mu\nu} &= 4\pi j_{\nu}
\end{align*}
\]

(3.1)

\[
\begin{align*}
\vec{\nabla} \cdot \vec{B} &= 0, & \vec{\nabla} \times \vec{E} &= -\partial_t \vec{B} \\
\vec{\nabla} \cdot \vec{E} &= 4\pi \rho, & \vec{\nabla} \times \vec{B} &= \partial_t \vec{E} + 4\pi \vec{j}
\end{align*}
\]

(3.2)

(where the familiar electric and magnetic fields are $E^{i} = -F^{0i}$ and $\epsilon^{ijk} B^{j} = -F^{ik}$). The first two equations (3.1) are constraints on $\vec{E}$ and $\vec{B}$ which mean that their components are not independent. This is annoying for trying to treat them quantumly. To get around this we introduce potentials $A_{\mu} = (\Phi, \vec{A})\mu$ which determine the fields by taking derivatives and which automatically solve the constraints (3.1):

\[
F_{\mu\nu} = \partial_{\mu} A_{\nu} - \partial_{\nu} A_{\mu},
\]

aka

\[
\vec{E} = -\vec{\nabla} \Phi - \partial_t \vec{A}, \quad \vec{B} = \vec{\nabla} \times \vec{A}.
\]

Potentials related by a gauge transformation

\[
\vec{A} \rightarrow \vec{A}^{\lambda} = \vec{A} - \vec{\nabla} \lambda, \quad \Phi \rightarrow \Phi^{\lambda} = \Phi + \partial_t \lambda
\]

for any function $\lambda(\vec{r}, t)$, give the same $\vec{E}, \vec{B}$. The Bohm-Aharonov effect is proof that (some of the information in) the potential is real and useful, despite this redundancy. We can partially remove this redundancy be choosing our potentials to satisfy Coulomb gauge

\[
\vec{\nabla} \cdot \vec{A} = 0.
\]

In the absence of sources $\rho = 0 = \vec{j}$, we can also set $\Phi = 0$. In this gauge, Ampère’s law becomes

\[
c^2 \vec{\nabla} \times \left( \vec{\nabla} \times \vec{A} \right) = c^2 \vec{\nabla} \cdot \left( \vec{\nabla} \cdot \vec{A} \right) - c^2 \nabla^2 \vec{A} = -\partial_t^2 \vec{A} \quad \text{i.e.} \quad \left[ \nabla^2 - c^2 \right] \vec{A} = 0.
\]

This wave equation is different from the scalar wave equation $\Box \phi = 0$ in three ways:
• we’re in three spatial dimensions,
• the speed of sound \( v_s \) has been replaced by the speed of light \( c \),
• the field \( \vec{A} \) is a vector field obeying the constraint \( \vec{\nabla} \cdot \vec{A} = 0 \). In fourier space \( \vec{A}(x) = \sum_k e^{i\vec{k} \cdot \vec{x}} \vec{A}(k) \) this condition is
  
  \[ 0 = \vec{k} \cdot \vec{A}(k) \]

  – the vector field is *transverse.*

An action which gives rise to Maxwell’s equations is

\[
S[A] = \int d^4x \left( -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \right) = \int d^4x L_{\text{Maxwell}}. \quad L_{\text{Maxwell}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} = \frac{1}{2} (E^2 - B^2).
\]

Note that we must regard \( A \) as the dynamical variable to obtain (3.2) by

\[
0 = \frac{\delta S}{\delta A_\mu(x)}.
\]

The canonical momentum of \( A \) is then \( \Pi_{Ai} = \partial L_{\text{Maxwell}} / \partial \dot{A}_i = E_i \). So the Hamiltonian is

\[
H = \frac{1}{2} \int d^3x \left( \vec{E}^2 + c^2 \vec{B}^2 \right).
\]

(3.3)

Here \( \vec{E} = -\partial_t \vec{A} \) plays the role of field momentum \( \pi(x) \) in (1.3), and \( \vec{B} = \vec{\nabla} \times \vec{A} \) plays the role of the spatial derivative \( \partial_x q \). We immediately see that we can quantize this system just like for the scalar case, with the canonical commutator

\[
[\phi(x), \pi(x')] = i\hbar \delta(x - x') \quad \rightarrow \quad [A_i(\vec{r}), E_j(\vec{r}')] = -i\hbar \delta^3(\vec{r} - \vec{r}') \delta_{ij}
\]

where \( i, j = 1..3 \) are spatial indices. So we can immediately write down an expression for the quantum Maxwell field in terms of independent creation and annihilation operators:

\[
\hat{\vec{A}}(\vec{r}) = \int d^3k \frac{1}{\sqrt{2\omega_k}} \sum_{s=1,2} \left( a_{k,s} \hat{e}^*_s(\vec{k}) e^{i\vec{k} \cdot \vec{r}} + a_{k,s}^\dagger \hat{e}^*_s(\vec{k}) e^{-i\vec{k} \cdot \vec{r}} \right)
\]

\[\text{You may also recall that the energy density of a configuration of Maxwell fields is } u = \frac{1}{2} \left( \vec{E}^2 + \vec{B}^2 \right). \text{ This result can be obtained either by Legendre transformation of } L_{\text{Maxwell}}, \text{ or from } T_{00}, \text{ the energy momentum tensor.}\]

\[\text{As a check, note that using this Hamiltonian and the canonical commutator, we can reproduce Maxwell’s equations using Ehrenfest’s theorem:}\]

\[
\langle \partial_t^2 \hat{A} \rangle = \partial_t \langle \hat{E} \rangle = -\frac{i}{\hbar} \langle [H, \hat{E}] \rangle = \langle c^2 \vec{\nabla}^2 \hat{A} \rangle.
\]
The field momentum is \( \vec{E} = -\partial_t \vec{A} \):

\[
\vec{E}(\vec{r}) = i \sqrt{\frac{\omega_k}{2}} \sum_{s=1,2} \left( a_{k,s} \hat{e}_s(\hat{k}) e^{i \vec{k} \cdot \vec{r}} - a_{k,s}^\dagger \hat{e}_s^*(\hat{k}) e^{-i \vec{k} \cdot \vec{r}} \right)
\]

Also, the magnetic field operator is

\[
\vec{B} = \vec{\nabla} \times \vec{A} = \sum_{\vec{k}} \sum_s \sqrt{\frac{\hbar}{2 \epsilon_0 \omega_k L^3}} i \vec{k} \times \left( a_{k,s} \hat{e}_s(\hat{k}) e^{i \vec{k} \cdot \vec{r}} - a_{k,s}^\dagger \hat{e}_s^*(\hat{k}) e^{-i \vec{k} \cdot \vec{r}} \right); 
\]

the magnetic field is analogous to \( \vec{\nabla} \phi \) in the scalar field theory\(^\text{16}\). Plugging these expressions into the Hamiltonian (3.3), we can write it in terms of these oscillator modes (which create and annihilate photons). As for the scalar field, the definitions of these modes were designed to make this simple: It is:

\[
H = \sum_{\vec{k},s} \hbar \omega_k \left( a_{k,s}^\dagger a_{k,s} + \frac{1}{2} \right).
\]

Notice that in this case we began our story in the continuum, rather than with microscopic particles connected by springs. (However, if you read Maxwell’s papers you’ll see that he had in mind a particular UV completion involving gears and cogs. I actually don’t understand it; if you do please explain it to me.)

The vacuum energy is

\[
E_0 = \frac{1}{2} \sum_{\vec{k},s} \hbar \omega_k = \frac{L^3}{(2\pi)^3} \int d^3k \hbar c \kappa.
\]

The fact that \( \sum_k \) is no longer a finite sum might be something to worry about. This vacuum energy has physical consequences, since it can depend on boundary conditions placed on the field by conducting objects, as we’ll discuss in \( \S 4.1 \).

### 3.2 More on vector fields

A few things we did not do yet for vector fields: study the propagator, figure out the data on external states, and understand the relation of between the masslessness of the photon and gauge invariance.

\(^\text{16}\)I should say a little more about the polarization vectors, \( \hat{e}_s \). They conspire to make it so that there are only two independent states for each \( \hat{k} \) and they are transverse \( \hat{k} \cdot \hat{e}_s(\hat{k}) = 0 \), so \( s = 1, 2 \). The polarization vectors of a given \( \vec{k} \) can be chosen to satisfy the following completeness relation:

\[
\sum_s \hat{e}_{si}(\hat{k}) \hat{e}_{sj}^*(\hat{k}) = \delta_{ij} - \hat{k}_i \hat{k}_j .
\]

This says that they span the plane perpendicular to \( \hat{k} \).
Consider the following Lagrangian for a vector field $A_\mu$ (which I claim is the most general quadratic Poincaré-invariant Lagrangian with at most two derivatives):

$$\mathcal{L} = -\frac{1}{2} \left( \partial_\mu A^\nu \partial_\nu A_\mu + a \partial_\mu A_\nu \partial_\nu A^\mu + b A_\mu A^\mu + c \epsilon^{\mu\nu\rho\sigma} \partial_\mu A_\nu \partial_\rho A_\sigma \right).$$

The sign is chosen so that spatial derivatives are suppressed, and the normalization of the first term is fixed by rescaling $A$. (Another possible-seeming term, $\partial_\mu A^\nu \partial_\nu A_\mu$, is related to the second term by two IBPs.) The last term is a total derivative, $\epsilon^{\mu\nu\rho\sigma} \partial_\mu A_\nu \partial_\rho A_\sigma \propto \partial_\mu (\epsilon^{\mu\nu\rho\sigma} A_\nu \partial_\rho A_\sigma)$, and will not affect the EoM or anything at all in perturbation theory; it is called a $\theta$ term.

The EoM are

$$0 = \frac{\delta}{\delta A^\mu(x)} \int L = -\partial^2 A_\mu - a \partial_\mu (\partial \cdot A) + b A_\mu$$

which (like any translation-invariant linear equation) is solved by Fourier transforms $A_\mu(x) = \epsilon_\mu e^{-i k x}$, if

$$k^2 \epsilon_\mu + a k_\mu (k \cdot \epsilon) + b \epsilon_\mu = 0.$$

There are two kinds of solutions: longitudinal ones with $\epsilon_\mu \propto k_\mu$ (for which the dispersion relation is $k^2 = -\frac{b}{1+a}$), and transverse solutions $\epsilon \cdot k = 0$ with dispersion $k^2 = -b$. The longitudinal mode may be removed by taking $b \neq 0$ and $a \to -1$, which we will do from now on. This gives the Proca Lagrangian:

$$\mathcal{L}_{a=-1,b=-\mu^2} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} \mu^2 A_\mu A^\mu,$$

where as usual $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$. Note that the EOM (Proca equation) $0 = \partial \cdot F_{\nu} + \mu^2 A_\nu$ implies $0 = \partial^2 A_\nu$ by $0 = \partial^\mu \partial^\nu F_{\mu\nu}$. So each component of $A_\mu$ satisfies (by (3.5)) the KG equation, $k^2 = \mu^2$, and the transverse condition $\epsilon \cdot k = 0$. In the rest frame, $k^\mu = (k^0, \vec{0})^\mu$, we can choose a basis of plane-wave transverse solutions which are eigenstates of the vector rotation generator

$$J^z = i \begin{pmatrix} +1 \\ -1 \end{pmatrix}, \text{ namely, } \epsilon^{(\pm)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \mp i \end{pmatrix}, \epsilon^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

They are normalized so that $\epsilon^{(r)} \cdot \epsilon^{(s)} = +\delta^{rs}$ and $\sum_{r=\pm 1,0} \epsilon^{(r)}_\mu \epsilon^{(r)*}_\nu = -\eta_{\mu\nu} + \frac{k_\mu k_\nu}{\mu^2}$ so that they project out $\epsilon \propto k$. Notice that in the massless case, only two of these three polarization states will be transverse to $k^\mu$. If $\vec{k} \propto \vec{z}$ (for example in the massless case with $k^\mu = (E, 0, 0, E)^\mu$) then these $\epsilon$ are also all helicity eigenstates: $h = \vec{J} \cdot \vec{k} = J^z$. 52
Canonical stuff: The canonical momenta are $\pi^i = \frac{\partial L}{\partial \dot{A}^i} = -F^0_i = E^i$ (as for electrodynamics in §3.1) and $\pi^0 = \frac{\partial L}{\partial \dot{A}^0} = 0$. This last bit is a little awkward, but it just means we can solve the equations of motion for $A_0$ algebraically in terms of the other (real) dofs:

$$0 = \frac{\delta S}{\delta A_0} = \vec{\nabla} \cdot \vec{E} - \mu^2 A_0 = (-\nabla^2 + \mu^2) A_0 + \vec{\nabla} \cdot \vec{A} \implies A_0(\vec{x}) = \int d^3 y e^{-\mu |\vec{x} - \vec{y}|} \frac{(-\vec{\nabla} \cdot \vec{A})}{4\pi |\vec{x} - \vec{y}|}. \tag{3.6}$$

So at each moment $A_0$ is determined by $A_i$. (Notice that this is still true for $\mu \to 0$.) The hamiltonian density is (after using $\pi^i = F^0_i$, integration by parts, and the equations of motion for $A_0$)

$$h = \frac{1}{2} \left( F_{0i}^2 + \frac{1}{2} F_{ij}^2 + \mu^2 A_i^2 + \mu^2 A_0^2 \right) = \frac{1}{2} \left( \vec{E}^2 + \vec{B}^2 + \mu^2 \vec{A}^2 + \mu^2 A_0^2 \right) \geq 0,$$

where positivity follows from the fact that it is a sum of squares of real things.

The canonical equal time commutators are then

$$[A_i(t, \vec{x}), F^j_0(t, \vec{y})] = i \delta^j_i \delta^{(3)}(\vec{x} - \vec{y})$$

which if we add up the plane wave solutions as

$$A_\mu(x) = \sum_{r=1,2,3} \int \frac{d^3 k}{\sqrt{2\omega_k}} \left( e^{-ikx} a^r_\mu \epsilon^{(r)}_\mu + e^{+ikx} a^r_\mu \epsilon^{(r)}_{\mu} \right)$$

give the bosonic ladder algebra for each mode

$$[a^r_\mu, a^s_\nu] = \delta^{(3)}(\vec{k} - \vec{p}) \delta^{rs}.$$

The normal-ordered hamiltonian is

$$:H:= \sum_r \int d^3 k \omega_k a^{\dagger}_{\mu} a^\mu_{\nu}.$$

Using the mode expansion above, the propagator for the $A_\mu(x)$ field is found to be

$$\langle 0 | \mathcal{T} A_\mu(x) A_\nu(y) | 0 \rangle = \int d^4 k e^{-ik(x-y)} \left[ \frac{-i(\eta_{\mu\nu} - k_\mu k_\nu / \mu^2)}{k^2 - \mu^2 + i\epsilon} \right]. \tag{3.7}$$

Notice that like in the spinor case the polarization sum $\sum_r \epsilon^{r*}_\mu \epsilon^{(r)}_\nu = -(\eta_{\mu\nu} - k_\mu k_\nu / \mu^2)$ appears in the numerator of the propagator. (Note that there are 3 orthonormal polarizations, so this is a rank-3 matrix; its kernel is the longitudinal direction, $k^\mu$.) The quantity in square brackets is then the momentum-space propagator. Since $\langle 0 | A_\mu(x) | k, r \rangle =$
ε^r_μ(k)e^{-i k \cdot x}$, a vector in the initial state produces a factor of $ε^r_μ(k)$, and in the final state gives $ε^*$. 

**Massless case.** In the limit $\mu \to 0$ some weird stuff happens. If we couple $A_\mu$ to some object $j^\mu$ made of other matter, by adding $\Delta L = j^\mu A_\mu$, then we learn that $\partial_\mu A^\mu = \mu^{-2} \partial_\mu j^\mu$. This means that in order to take $\mu \to 0$, it will be best if the current is conserved $\partial_\mu j^\mu$.

One example is the QED coupling, $j^\mu = \bar{\Psi} \gamma^\mu \Psi$. Here $j^\mu$ is the Noether current for the symmetry $\Psi \to e^{i \alpha} \Psi$ of the Dirac Lagrangian. This coupling $A_\mu j^\mu$ arises from the ‘minimal coupling’ prescription of replacing $\partial_\mu \to D_\mu = \partial_\mu + ieq A_\mu$ in the Dirac Lagrangian. In this case, the model, with Lagrangian

$$L = \bar{\Psi} (i \slashed{D} - m) \Psi - \frac{1}{4} F^\mu_\nu F^{\mu \nu} - \frac{\mu^2}{2} A_\mu A^\mu |_{\mu^2 = 0},$$

has a local invariance under $A_\mu \to A_\mu + \partial_\mu \lambda(x)/e, \Psi(x) \to e^{iq\lambda(x)} \Psi(x)$. For $\lambda$ non-constant (and going to zero far away), this is a redundancy of our description rather than a symmetry (for example, they have the same configuration of $\vec{E}, \vec{B}, \oint A$). That is, configurations related by this gauge transformation should be regarded as equivalent.

[End of Lecture 6]

Another example can be obtained by taking a complex scalar and doing the same replacement: $L = D_\mu \Phi^* D^\mu \Phi + ...$ Notice that in this case the vertex involves a derivative, so it comes with a factor of $\ = -ieq(p_\Phi + p_{\Phi^*})^\mu$. Also, there is a $A_\mu A_\nu \Phi^* \Phi$ coupling, which gives a vertex $\ = -ie^2 q^2 \eta_{\mu \nu}$.

How do I know that configurations related by a gauge transformation should be regarded as equivalent? If not, the kinetic operator for the massless vector field $(\eta_{\mu \nu} (\partial_\mu \partial_\nu) - \partial_\mu \partial_\nu) A^\nu = 0$ is not invertible (even in Euclidean section!), since it annihilates $A_\nu = \partial_\nu \lambda$.

What’s the propagator for a massless vector field, then? One strategy is to simply ignore the gauge equivalence and use the same propagator (3.7) that we found in the massive case with $\mu \to 0$. Since the dynamics are gauge invariant, it will never make gauge-variant stuff, and the longitudinal bits $\propto k_\mu k_\nu$ in (3.7) (which depend on $\mu$) will just drop out, and we can take $\mu \to 0$ in the denominator at the end. This actually works. The guarantee that it works is the QED Ward identity: any amplitude with an
external vector $\epsilon(k)_\mu$ is of the form

$$\mathcal{M} = iM = iM^\mu(k)\epsilon_\mu(k)$$

and if all external fermion lines are on-shell then

$$\mathcal{M}^\mu(k)k_\mu = 0.$$ 

There is a complicated diagrammatic proof of this statement in Peskin; Schwartz §8.4 argues that it is a necessary condition for Lorentz invariance of $\mathcal{M} = \epsilon_\mu \cdot M^\mu$; and we will see some illustrations of it below (I also recommend Zee §II.7). But it is basically a statement of current conservation: such an amplitude is made (by LSZ and the photon Schwinger-Dyson equation) from a correlation function involving an insertion of the electromagnetic current $j^\mu(k) = \int d^4x e^{-ikx} j^\mu(x)$, in the form, $\mathcal{M}^\mu \sim \langle \Omega | ... j^\mu(k) ... | \Omega \rangle$, and $k_\mu j^\mu(k) = 0$ is current conservation $^{17}$. 

This property guarantees that we will not emit either of the unphysical polarizations of massless photons, since the amplitude to do so is either $\mathcal{A}(\text{emit } \epsilon_\lambda \propto k_\lambda) = \epsilon_\mu \cdot M^\mu \propto k_\mu M^\mu \overset{\text{Ward}}{=} 0$, or the $\mu \to 0$ limit of

$$\mathcal{A}\left(\text{emit } \epsilon^\lambda_\alpha = \frac{1}{\mu}(k,0,0,-\omega)_\lambda \right) \propto \epsilon^L_\mu \cdot M^\mu = \frac{1}{\mu} (k_\mu M^0_\mu - \omega M^3_0) = \frac{1}{\mu} \left( k_\mu M^0_\mu - \sqrt{k^2 + \mu^2} M^3_0 \right) = k_\mu \frac{\mu}{2k} M^2_0 + \mathcal{O}(\mu^3) \overset{\mu \to 0}{\to} 0.$$ 

For the same reason, in summing over photon final-states (for example in computing a cross section), we’ll have

$$\sum_\epsilon |\mathcal{M}|^2 = \sum_\epsilon \epsilon(\epsilon)\epsilon(\epsilon)^* \cdot M^\mu(k) \cdot M^\mu(k)^*.$$ 

$^{17}$ Current conservation $\partial_\mu j^\mu$ is a statement which requires the equations of motion (recall the proof of Noether’s theorem). Recall that equations of motion are true in correlation functions, up to contact terms, using the independence of the path integral on choice of integration variables. By contact terms, I mean terms which are only nonzero when two operators are at the same point. So you can worry about the contact terms in the argument for the Ward identity. The reason they do not contribute is that all the operators in the correlation function (using the LSZ formula) correspond to external states. A collision between the operators creating the external particles would lead to a disconnected amplitude, which could only contribute for degenerate kinematical configurations, and we can ignore them. If you would like to read more words about this, look at Schwartz §14.8, or §3.3.
This sum is only over the two physical polarizations of the massless photon. If we choose a frame with \( k^\mu = (k, 0, 0, k)^\mu \), the Ward identity says \( 0 = k^\mu M^\mu = k (M^0 - M^3) \), so

\[
\sum_\epsilon \epsilon (k)_\mu \epsilon (k^\star)_\nu M^\mu (k) M^\nu (k)^\star = |M^1|^2 + |M^2|^2 = |M^1|^2 + |M^2|^2 + |M^3|^2 - |M^0|^2 = -\eta_{\mu\nu} M^\mu (k) M^\nu (k)^\star,
\]

that is, just like in the numerator of the propagator, we can replace

\[
\sum_\epsilon \epsilon (k)_\mu \epsilon (k^\star)_\nu \to -\eta_{\mu\nu}
\]

since they differ by stuff proportional to \( k^\mu \) which vanishes when contracted with the rest of the amplitude. In (3.8) we see explicitly that the crazy timelike polarization (which looks like negative probability) cancels the longitudinal polarization \( \vec{\epsilon} \propto \vec{k} \).

**Gauge fixing.** You might not be happy with the accounting procedure I’ve advocated above, where unphysical degrees of freedom are floating around in intermediate states and only drop out at the end by some formal trick. In that case, a whole zoo of formal tricks called *gauge fixing* has been prepared for you. Here’s a brief summary to hold you over until we really need it for the non-Abelian case.

At the price of Lorentz invariance, we can make manifest the physical dofs, by choosing *Coulomb gauge*. That means we restrict \( \partial_\mu A^\mu = 0 \) (so far, so Lorentz invariant) and also \( \vec{\nabla} \cdot \vec{A} = 0 \). Looking at (3.6), we see that this kills off the bit of \( A_0 \) that depended on \( \vec{A} \). We also lose the helicity-zero polarization \( \vec{\nabla} \cdot \vec{A} \propto \epsilon^{(0)} \). But the Coulomb interaction is instantaneous action at a distance.

To keep Lorentz invariance, we can instead merely discourage configurations with \( \partial \cdot A \neq 0 \) by adding a term to the action

\[
\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2\xi} (\partial \cdot A)^2
\]

for some arbitrary number \( \xi \). Physics should not depend on \( \xi \), and this is a check on calculations. The propagator is

\[
\langle T A_\mu (x) A_\nu (y) \rangle_\alpha = \int d^4 k \ e^{-ik(x-y)} \left[ \frac{-i(\eta_{\mu\nu} - (1 - \xi) k^\mu k^\nu / \mu^2)}{k^2 - \mu^2 + i\epsilon} \right]
\]

and again the bit with \( k^\mu k^\nu \) must drop out. \( \xi = 1 \) is called Feynman gauge and makes this explicit. \( \xi = 0 \) is called Landau gauge and makes the propagator into a projector onto \( k_\perp \).

It becomes much more important to be careful about this business in non-Abelian gauge theory.
3.3 On the non-perturbative proof of the Ward identity

[Schwartz §14.8] First, consider a Green’s function from which we might make an S-matrix element by LSZ,

\[ G \equiv \langle \Omega | T \mathcal{O}_1(x_1) \cdots \mathcal{O}_n(x_n) | \Omega \rangle = \int D\Psi e^{iS} \mathcal{O}_1(x_1) \cdots \mathcal{O}_n(x_n) \]

where the operators \( \mathcal{O}_1(x_1) \mapsto e^{-iQ_1(x_1)} \) have charge \( Q_i \) under a global \( \mathbb{U}(1) \) symmetry. For example the \( \mathcal{O}(x) \) could be just the elementary field \( \Psi(x) \).

Now change variables in the path integral so that \( \mathcal{O}_i(x_i) \mapsto e^{-iQ_i(x_i)} \mathcal{O}_i(x_i) \); the action will shift by \( S \mapsto S - \int \alpha \partial_{\mu} j^\mu \) where \( j^\mu \) is the Noether current. The path integral doesn’t change at all, so its infinitesimal variation is

\[ 0 = \delta G = \int D\Psi \left( -\int i\alpha \partial^\mu j_\mu e^{iS} \mathcal{O}_1 \cdots \mathcal{O}_n - i \sum_i Q_i \alpha(x_i) e^{iS} \mathcal{O}_1 \cdots \mathcal{O}_n \right) \]

(3.9)

\[ = \int d^D x \alpha(x) \left[ i\partial_\mu \langle j^\mu(x) \mathcal{O}_1 \cdots \mathcal{O}_n \rangle - \sum_i Q_i \delta^D (x-x_i) G \right]. \]

(3.10)

Since this is true for any \( \alpha(x) \), we learn that the thing in square brackets is zero: \( \partial_\mu j^\mu = 0 \) up to contact terms. This is called the Ward-Takahashi identity.

Now suppose we do this same manipulation in a gauge theory, like QED. The additional terms in \( S \) are \(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + iA_\mu \bar{\Psi} \gamma^\mu \Psi \), which are invariant under the transformation, so don’t change these statements. Notice that the transformation we’re doing here is not the gauge transformation, since \( A_\mu \) doesn’t transform – we’re only doing the gauge transformation on the matter fields here, so their kinetic terms actually shift and produce the \( \alpha \partial_\mu j_\mu \) term above. Photon field insertions in \( G \) don’t contribute, since they have charge zero here.

Next, think about the LSZ formula for an S-matrix element with (say) two external photons:

\[ \mathcal{M} = \langle \epsilon, ... \varepsilon_k | S | ... \rangle_{LSZ} = e^\mu e^\nu i^n \int d^4 x e^{ipx} \Box_{\mu\nu} \int d^4 x_1 e^{ip_1 x_1} \Box_{\rho\sigma}^{k} \int ... \langle A^\nu(x) ... A_\sigma(x_k) ... \rangle \]

(3.11)

where \( \Box_{\mu\nu} \) is shorthand for the photon kinetic operator \( \Box_{\mu\nu} \equiv \Box_{\mu\nu} - \partial_\mu \partial_\nu / \mu^2 \). The Schwinger-Dyson equation for \( A_\mu \) then implies that

\[ \Box_{\rho\sigma}^{k} \Box_{\mu\nu} \langle A^\nu(x) ... A_\sigma(x_k) ... \rangle = \Box_{\rho\sigma}^{k} \left( \langle j_\mu(x) ... A_\sigma(x_k) ... \rangle - i \delta^4(x-x_k) \eta_{\mu\sigma} \langle ... \rangle \right) \]

(3.12)

\[ = \langle j_\mu(x) ... j_\sigma(x_k) ... \rangle - i \delta^4(x-x_k) \eta_{\mu\sigma} \langle ... \rangle \]

(3.13)

\[^{18}\text{You’ll have to trust me for now that the path integral for fermionic fields exists. That’s the only information about it we’ll need here.}\]
First of all, this is why I said we could get the S-matrix elements with photons from correlators with currents. But notice that this is only true up to the contact terms. But those are disconnected amplitudes which we can ignore.

Finally, set the polarization of one of the photons equal to its momentum $\epsilon = p$. Then

$$p^\mu \mathcal{M}_\mu = e_\lambda^\mu e^{\lambda x_1} \int d^4 xe^{\lambda x_1} \int d^4 x_1 e^{\lambda x_1} \int dy e^{\lambda y} (i\partial_y + m_1) \ldots \langle -i\partial_\mu j^\mu \ldots j_\rho (x_1) \ldots \Psi(y) \rangle$$

$$= \left( q_1 - m_1 \right) \left( q_2 - m_2 \right) \cdots \sum_j Q_j \tilde{G}(\ldots, q_j \pm p, \ldots) \quad (3.15)$$

where the $\pm$ depends on whether particle $j$ is incoming or outgoing. At the last step we used the Fourier transform of (3.10).

Now here’s the punchline: The $\tilde{G}$ on the RHS of (3.15) has poles at $(q_j \pm p)^2 = m_j^2$, and not at $q_j^2 = m_j^2$. So when it’s multiplied by $q_j - m_j = \frac{q_j^2 - m_j^2}{\frac{q_j^2}{1} + m_j}$ it will vanish. End of story. Notice that no use of perturbation theory was made here.

### 3.4 Feynman rules for QED

First, Feynman rules for Dirac fermion fields, more generally. As always in these notes, time goes to the left, so I draw the initial state on the right (like the ket) and the final state on the left (like the bra).

1. An internal fermion line gives

$$= \frac{i}{\hat{k} - m_\Psi}$$

19Another good example of a QFT with interacting fermions is the Yukawa theory theory of a Dirac fermion field plus a scalar $\phi$ and an interaction

$$V = g\phi\bar{\Psi}\Psi \quad \Rightarrow \quad = -i\delta^{rr'} \cdot (3.16)$$

Notice that in 3 + 1 dimensions, $|g| = +4 - |\phi| - 2|\Psi| = 4 - 1 - 2\frac{3}{2} = 0$, the coupling is dimensionless. This describes more realistically the interactions between nucleons (which are fermions, as opposed to nucleons) and scalar pions, which hold together nuclei. It also is a crude sketch of the Higgs coupling to matter; notice that if $\phi$ is some nonzero constant $\langle \phi \rangle$, then there is a contribution to the mass of the fermions, $g \langle \phi \rangle$. 

58
which is a matrix on the spinor indices.

There are four possibilities for an external fermion line of definite momentum. Here \( u, v \) are respectively the positive- and negative-energy solutions of the Dirac equation,

\[
(\not{p} - m) \ u^r(k) = 0, \ (\not{p} + m) \ v^r(k) = 0. \quad \text{(For a reminder, see e.g. §5.4 of my 215A lecture notes.)}
\]

2. \[ \begin{array}{c}
\cdots \ddots \end{array} = \begin{array}{c}
\begin{array}{c}
\not{k} \\
\not{r}
\end{array}
\end{array} = \Psi[k, r] = u^r(k) \]

3. \[ \begin{array}{c}
\begin{array}{c}
\not{k} \\
\not{r}
\end{array} \ddots \end{array} = \begin{array}{c}
\begin{array}{c}
\not{k} \\
\not{r}
\end{array}
\end{array} = \langle k, r | \bar{\Psi} = \bar{u}^r(k) \]

4. \[ \begin{array}{c}
\begin{array}{c}
\not{k} \\
\not{r}
\end{array} \ddots \end{array} = \begin{array}{c}
\begin{array}{c}
\not{k} \\
\not{r}
\end{array}
\end{array} = \bar{\Psi}[k, r] = \bar{v}^r(k) \]

5. \[ \begin{array}{c}
\begin{array}{c}
\not{k} \\
\not{r}
\end{array} \ddots \end{array} = \begin{array}{c}
\begin{array}{c}
\not{k} \\
\not{r}
\end{array}
\end{array} = \langle k, r | \bar{\Psi} = v^r(k) \]

6. Some advice: When evaluating a Feynman diagram with spinor particles, always begin at the head of the particle-number arrows on the fermion lines, and keep going along the fermion line until you can’t anymore. This will keep the spinor indices in the form of matrix multiplication. Why: every Lagrangian you’ll ever encounter has fermion parity symmetry, under which every fermionic field gets a minus sign; this means fermion lines cannot end, except on external legs. The result is always of the form of a scalar function (not a matrix or a spinor) made by sandwiching gamma matrices between external spinors:

\[
\begin{array}{c}
\begin{array}{c}
\not{r'} \\
\not{p'}
\end{array} \ddots \not{r} \not{p} = \sum_{a, b \ldots = 1, 4} \bar{u}^{r'}(p')_a (\text{pile of gamma matrices})_{ab} u^r(p)_b
\end{array}
\]

Furthermore, in S-matrix elements the external spinors \( u(p), v(p) \) satisfy the equations of motion \( (\not{p} - m)u(p) = 0 \), a fact which can be used to our advantage to shrink the pile of gammas.

There can also be fermion lines which form internal loops (though not at tree level, by definition). In this case, the spinor indices form a trace,

\[
\sum_a (\text{pile of gamma matrices})_{aa} \equiv \text{tr} \ (\text{pile of gamma matrices})
\]

We’ll learn to compute such traces below (around (3.18)); in fact, traces appear even in the case with external fermions if we do not measure the spins.
7. Diagrams related by exchanging external fermions have a relative minus sign.

8. Diagrams with an odd number of fermion loops have an extra minus sign.

The last two rules are best understood by looking at an example in detail.

To understand rule 8 consider the following amplitude in the Yukawa theory with interaction (3.16):

\[ \sum \bar{\Psi}_a(x) \Psi_b(y) \bar{\Psi}_c(y) \Psi_d(x) = (-1) \text{tr} \Psi(x) \bar{\Psi}(y) \bar{\Psi}(x) \Psi(y) = (-1) \text{tr} S_F(x - y) S_F(x - y) \]

[Peskin page 119] To understand rule 7 consider \( \Psi \Psi \rightarrow \Psi \Psi \) (nucleon) scattering in the Yukawa theory:

The blob represents the matrix element

\[ 0 \langle p_3 r_3; p_4 r_4 | T e^{-i \int V d^4 z} | p_1 r_1; p_2 r_2 \rangle_0 \]

where the initial state is

\[ | p_1 r_1; p_2 r_2 \rangle_0 \propto a_{p_1}^r a_{p_2}^r | 0 \rangle \]

and the final state is

\[ 0 \langle p_3 r_3; p_4 r_4 | = (| p_3 r_3; p_4 r_4 \rangle_0 )^\dagger \propto \langle 0 | a_{p_4}^r a_{p_3}^r = - \langle 0 | a_{p_3}^r a_{p_4}^r \]

where note that the dagger reverses the order.

The leading contribution comes at second order in \( V \):

\[ 0 \langle p_3 r_3; p_4 r_4 | T \left( \frac{1}{2i} (ig)^2 \int d^4 z_1 \int d^4 z_2 (\bar{\Psi} \Psi \phi)_1 (\bar{\Psi} \Psi \phi)_2 \right) | p_1 r_1; p_2 r_2 \rangle_0 \]

To get something nonzero we must contract the \( \phi \)s with each other. The diagrams at right indicate best the possible ways to contract the fermions. Exchanging the roles of \( z_1 \) and \( z_2 \) interchanges two pairs of fermions so costs no signs and cancels the \( \frac{1}{2i} \).

The overall sign is annoying but can be fixed by demanding that the diagonal bit of the \( S \)-matrix give

\[ \langle p_3 p_4 | (1 + \ldots) | p_1 p_2 \rangle = + \delta(p_1 - p_3) \delta(p_2 - p_4) + \ldots \]

The relative sign is what we’re after, and it comes by comparing the locations of fermion
operators in the contractions in the two diagrams at right. In terms of the contractions, these \( t - \) and \( u - \) channel diagrams are related by leaving the annihilation operators alone and switching the contractions between the creation operators and the final state. Denoting by \( a_{1,2}^{\dagger} \) the fermion creation operators coming from the vertex at \( z_{1,2} \),

\[
\langle 0 | a_{p_4} a_{p_3} a_{p_2}^{\dagger} a_{p_1}^{\dagger} ... + \langle 0 | a_{p_4} a_{p_3} a_{p_2}^{\dagger} a_{p_1}^{\dagger} ... = \langle 0 | a_{p_4} a_{p_3} a_{p_2}^{\dagger} a_{p_1}^{\dagger} ... - \langle 0 | a_{p_4} a_{p_3} a_{p_2}^{\dagger} a_{p_1}^{\dagger} ...
\]

In the last expression the fermion operators to be contracted are all right next to each other and we see the relative minus sign.

\[ S_{fi} = -g^2 \int dz_1 dz_2 \int d^4q \frac{-e^{iq(z_1-z_2)}}{q^2 - m^2 + i\epsilon} (e^{-i\delta(p_1-p_3)}u_1^{*}u_2^{*}u_1^{*}(p_1) \cdot e^{-i\delta(p_2-p_4)}u_4^{*}u_2^{*}(p_2) - (3 \leftrightarrow 4)) \]

In the first \( (t-\) channel) term, the integrals over \( z_{1,2} \) gives \( \delta(p_1 - p_3 - q) \delta(p_2 - p_4 - q) \), and the \( q \) integral then gives \( \delta(p_1 + p_2 - p_3 - p_4) \), overall momentum conservation. In the second \( (u-\) channel) term, \( q = p_1 - p_4 = p_3 - p_2 \). Altogether,

\[ S_{fi} = \mathbb{1} + \delta^4(p_T)i\mathcal{M} \]

with, to leading order,

\[ i\mathcal{M} = -ig^2 \left( \frac{1}{t-m^2} (\bar{u}_3 u_1^{*}) (\bar{u}_4 u_2^{*}) - \frac{1}{u-m^2} (\bar{u}_4 u_1^{*}) (\bar{u}_3 u_2^{*}) \right) \quad (3.17) \]

with \( t \equiv (p_1 - p_3)^2, u \equiv (p_1 - p_4)^2 \). This minus sign implements Fermi statistics.

**Yukawa force revisited.** In the non-relativistic limit, we can again relate this amplitude to the force between particles, this time with the actual spin and statistics of nucleons. In the COM frame, \( p_1 = (m, \vec{p}) \), \( p_2 = (m, -\vec{p}) \) and \( p_3 = (m, \vec{p}') \), \( p_4 = (m, -\vec{p}') \).

In the non-relativistic limit, the spinors become \( u_p^{\dagger} = \frac{\sqrt{\sigma \cdot \vec{p} \xi^r}}{\sqrt{\sigma \cdot \vec{p} \xi^r}} \rightarrow \sqrt{m} \left( \frac{\xi'}{\xi^r} \right) \) so that \( \bar{u}_3 u_1 = \bar{u}(p_3)^r u(p_1)^r = 2m\xi_{r_3}^{\dagger} \xi_{r_1} = 2m\delta_{r_3 r_1} \). Let’s simplify our lives and take two *distinguishable* fermions (poetically, they could be proton and neutron, but let’s just add a label to our fermion fields; they could have different masses, for example, or different couplings to \( \phi \), call them \( g_1, g_2 \)). Then we only get the \( t-\) channel diagram. The intermediate scalar momentum is \( q = p_1 - p_3 = (0, \vec{p} - \vec{p}') \) so \( t = (p_1 - p_3)^2 = -q^2 = - (\vec{p} - \vec{p}')^2 \) and

\[ i\mathcal{M}_{NR,COM} = ig_1 g_2 \frac{1}{q^2 + m_\phi^2} 4m^2 \delta_{r_3 r_1} \delta_{r_3 r_1}. \]
Compare this to the NR Born approximation matrix element

\[ 2\pi\delta(E_p - E_{p'}) \left( -i\tilde{V}(\vec{q}) \right) = NR \langle \vec{p}' | S | \vec{p} \rangle_{NR} \]

\[ = \sum_{r4} \int d^3p_4 V \prod_{i=1}^{4} \frac{1}{\sqrt{2E_i}} S(34 \leftarrow 12) \]

\[ = 2\pi\delta(E_p - E_{p'}) \delta^{r1r3} \frac{i g_1 g_2}{q^2 + m_\phi^2} \]

where in the second line we summed over possible final states of the second (target) particle, and corrected the relativistic normalization, so that \( NR \langle \vec{p}' | \vec{p} \rangle_{NR} = \delta^3(p - p') \).

This is completely independent of the properties of the second particle. We infer that the scalar mediates a force with potential \( U(x) = -\frac{g_1 g_2 e^{-m_\phi r}}{4\pi r} \). It is attractive if \( g_1 g_2 > 0 \).

---

Back to QED. The new ingredients in QED are the propagating vectors, and the interaction hamiltonian \( V = e\bar{\Psi}\gamma^\mu\Psi A_\mu \). The rest of the Feynman rules are

9. The interaction vertex gets a

\[ = -ie\gamma^\mu \]

10. An external photon in the initial state gets a \( \epsilon^\mu(p) \), and in the final state gets a \( \epsilon^{\mu*}(p) \).

11. An internal photon line gets a

\[ \frac{i}{k^2 - m_\gamma^2} \left( -\eta^{\mu\nu} + (1 - \xi)k^\mu k^\nu/k^2 \right) \]

where \( m_\gamma = 0 \) (it’s sometimes useful to keep it in there for a while as an IR regulator) and the value of \( \xi \) is up to you (meaning that your answers for physical quantities should be independent of \( \xi \)).

---

Spinor trace ninjutsu.

The trace is cyclic:

\[ \text{tr} (AB \cdots C) = \text{tr} (CAB \cdots) . \] (3.18)

Our gamma matrices are \( 4 \times 4 \), so \( \text{tr} \mathbb{1} = 4 \).
\[ \text{tr} \gamma^\mu = \text{tr} (\gamma^5)^2 \gamma^\mu \overset{(3.18)}{=} \text{tr} \gamma^5 \gamma^\mu \gamma^5 \{ \gamma^5, \gamma^\mu \} = 0 - \text{tr} \gamma^\mu = 0. \] (3.19)

The same trick shows that the trace of any odd number of gammas vanishes. The idea is that an odd number of gammas is a map between the \( L \) and \( R \) subspaces, so it has only off-diagonal terms in the Weyl basis.

\[ \text{tr} \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma = 4 (\eta^\mu^\nu \eta^\rho^\sigma + \eta^\sigma^\mu \eta^\nu^\rho - \eta^\mu^\rho \eta^\nu^\sigma). \] (3.20)

Why is this? The completely antisymmetric bit vanishes because it is proportional to \( \gamma^5 \) which is traceless (by the same argument as (3.19)). If any pair of indices is the same then the other two must be too by (3.20). If adjacent pairs are the same they can just square to one and we get +1; if alternating pairs are the same (and different from each other) then we must move them through each other with the anticommutator. If they are all the same we get 4.

\[ \text{tr} \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \gamma^5 = -4i \epsilon^{\mu^\nu^\rho^\sigma}. \]

### 3.5 QED processes at leading order

Now we are ready to do lots of examples, nearly all of which (when pushed to the end) predict cross sections which are verified by experiments to about one part in 137.\(^{20}\) Here \( \frac{1}{137} \approx \alpha \equiv \frac{e^2}{4\pi} \) is the small number by which the next order corrections are suppressed.

Did I mention that the antiparticle of the electron, predicted by the quantum Dirac theory (i.e. by Dirac), is the positron? It has the same mass as the electron and the opposite electromagnetic charge, since the charge density is the 0 component of the electromagnetic current, \( j^\mu = \bar{\Psi} \gamma^\mu \Psi \), so the charge is

\[ \int \! d^3 x \, j^0 (x) = \int \! \bar{\Psi} \gamma^0 \Psi = \int \! \Psi^\dagger \Psi = \int \! d^3 p \sum_s (a^\dagger_{p,s} a_{p,s} - b^\dagger_{p,s} b_{p,s}). \]

So \( b^\dagger \) creates a positron.

\(^{20}\)I guess it is this overabundance of scientific victory in this area that leads to the intrusion of so many names of physicists in the following discussion.

\(^{21}\)This statement is true naively (in the sense that the next diagrams which are nonzero come with two more powers of \( e \)), and also true in fact, but in between naiveté and the truth is a long road of renormalization, which begins in the next section.
Perhaps the simplest to start with is scattering of electrons and positrons. We can make things even simpler (one diagram instead of two) by including also the muon, which is a heavy version of the electron\textsuperscript{22}, and asking about the process $\mu^+\mu^- \leftarrow e^+e^-$. At leading order in $e$, this comes from

\[
i\mathcal{M}_{\mu^+\mu^-\leftarrow e^+e^-} = (-i\bar{u}^{\text{muons}}s_3(p_3)\gamma^\mu v^{\text{muons}}s_4(p_4)) - \frac{i(\eta_{\mu\nu} - (1-\xi)k_{\mu}k_{\nu})}{k^2}(-i\bar{u}^{\text{electrons}}s_2(p_2)\gamma^\nu u^{\text{electrons}}s_1(p_1))\tag{3.22}
\]

with $k \equiv p_1 + p_2 = p_3 + p_4$ by momentum conservation at each vertex. I’ve labelled the spinors according to the particle types, since they depend on the mass.

**Ward identity in action.** What about the $k_{\mu}k_{\nu}$ term in the photon propagator? The spinors satisfy their equations of motion, $\slashed{p}_1 u_1 = m_e u_1$ (where $u_1 \equiv u_{\mu_1}^s$ for short) and $\bar{v}_2 \slashed{p}_2 = -m_e \bar{v}_2$. The $k_{\nu}$ appears in

\[k_{\nu} \bar{v}_2 \gamma^\nu u_1 = \bar{v}_2 (\slashed{p}_1 + \slashed{p}_2) u_1 = \bar{v}_2 \slashed{p}_1 u_1 + \bar{v}_2 \slashed{p}_2 u_1 = (m - m) \bar{v} u = 0.
\]

(The other factor is also zero, but one factor of zero is enough.) Therefore

\[\mathcal{M} = \frac{e^2}{s} \bar{u}_s \gamma_{\mu} v_4 \cdot \bar{v}_2 \gamma^\mu u_1
\]

where $s \equiv k^2 = (p_1 + p_2)^2 = E_{\text{CoM}}^2$ is the Mandelstam variable. And I am relying on you to remember which spinors refer to muons (3,4) and which to electrons (1,2).

**Squaring the amplitude.** We need to find $\mathcal{M}^\dagger$ (the dagger here really just means complex conjugate, but let’s put dagger to remind ourselves to transpose and reverse the order of all the matrices). Recall the special role of $\gamma^0$ here:

\[\gamma^\dagger_0 = \gamma_0 \gamma^\dagger_0 = \gamma_0.
\]

This means that for any two Dirac spinors,

\[(\bar{\Psi}_1 \gamma^\mu \Psi_2)^\dagger = \bar{\Psi}_2 \gamma^\mu \Psi_1.
\]

\textsuperscript{22}Who ordered that? (I. I. Rabi’s reaction to learning about the muon.) I hope you don’t find it too jarring that the number of ‘elementary’ particles in our discussion increased by three in the last two paragraphs. People used to get really disgruntled about this kind of thing. But here we have, at last, uncovered the true purpose of the muon, which is to halve the number of Feynman diagrams in this calculation (compare (3.28)).
(This is the same manipulation that showed that the Dirac Lagrangian was hermitian.) So
\[ \mathcal{M}^\dagger = \frac{e^2}{s} (\bar{\psi}_4 \gamma^\mu u_3) (\bar{\psi}_1 \gamma_\mu v_2). \]
and therefore
\[ |\mathcal{M}_{\mu^+ \mu^-} e^+ e^-|^2 = \frac{e^4}{s^2} (\bar{\psi}_4 \gamma^\mu u_3) (\bar{\psi}_3 \gamma^\nu v_4) \cdot (\bar{\psi}_1 \gamma_\mu v_2) (\bar{\psi}_2 \gamma_\mu u_1). \] (3.23)

These objects in parentheses are just c-numbers, so we can move them around, no problem. I've grouped them into a bit depending only on the initial state (the electron stuff 1, 2) and a bit depending only on the final state (the muon stuff 3, 4).

**Average over initial, sum over final.** In the amplitude above, we have fixed the spin states of all the particles. Only very sophisticated experiments are able to discern this information. So suppose we wish to predict the outcome of an experiment which does not measure the spins of the fermions involved. We must sum over the final-state spins using
\[ \sum_{s_4} v^s_4(p_4) v^{s_4}_4(p_4) = \left( \frac{q_4 - m_\mu}{ab} \right) = \sum_{s_4} v^{s_4}_4(p_4) v^s_4(p_4) \]
(where I wrote the last expression to emphasize that these are just c-numbers) and
\[ \sum_{s_3} u^s_3(p_3) u^{s_3}_3(p_3) = \left( \frac{p_3 + m_\mu}{ab} \right). \]
Looking at just the ‘out’ factor of $|\mathcal{M}|^2$ in (3.23), we see that putting these together produces a spinor trace, as promised:
\[ \sum_{s_3, s_4} (\bar{\psi}_4^s_3 (p_3 a_b) \gamma^\mu v^s_4 (p_4 b_c)) (\bar{\psi}_3^s_3 (p_3 c_d) \gamma^\nu u^s_4 (p_4 d) = \gamma^\mu \gamma^\nu \gamma^\sigma \gamma^\rho \gamma^\sigma - m_\mu^2 \gamma^\mu \gamma^\nu) \] (3.20), (3.21)
\[ \equiv p_{34} \gamma^\mu \gamma^\nu \gamma^\sigma \gamma^\rho \gamma^\sigma - m_\mu^2 \gamma^\mu \gamma^\nu) \]

\[ \text{If also we don’t know the initial (electron) spins, then the outcome of our experiment is the average over the initial spins, of which there are four possibilities. Therefore, the relevant probability for unpolarized scattering is} \]
\[ \frac{1}{4} \sum_{s_1, s_2, s_3, s_4} |\mathcal{M}|^2 = \frac{e^4}{4s^2} \text{tr} \left( \gamma^\mu \left( p_1 - m_\mu \right) \gamma^\nu \left( p_3 + m_\mu \right) \right) \text{tr} \left( \gamma^\nu \left( p_2 - m_\mu \right) \gamma_\mu \left( p_1 + m_\mu \right) \right) \]
\[(3.24)\text{ twice} \quad \frac{8e^4}{s^2} (p_{13}p_{24} + p_{14}p_{23} + m_{\mu}^2 p_{12} + m_e^2 p_{34} + 2m_e^2 m_\mu^2) \]
\[\text{algebra} \quad \frac{2e^4}{s^2} \left( t^2 + u^2 + 4s(m_e^2 + m_\mu^2) - 2(m_e^2 + m_\mu^2)^2 \right) \quad (3.25)\]

In the second step of (3.25) the \(p_{12}p_{34}\) terms cancel. In the last step of (3.25) we used all the Mandelstam variables:

\[s \equiv (p_1 + p_2)^2 = (p_3 + p_4)^2 = E_{\text{CoM}}^2 = 4E^2\]
\[t \equiv (p_1 - p_3)^2 = (p_2 - p_4)^2 = m_e^2 + m_\mu^2 - 2E^2 + 2\vec{k} \cdot \vec{p}\]
\[u \equiv (p_1 - p_4)^2 = (p_2 - p_3)^2 = m_e^2 + m_\mu^2 - 2E^2 - 2\vec{k} \cdot \vec{p}\]

where the particular kinematic variables (in the rightmost equalities) are special to this problem, in the center of mass frame (CoM), and are defined in the figure at right. Really there are only two independent Lorentz-invariant kinematical variables, since \(s + t + u = \sum_i m_i^2\).

Now we can use the formula for a differential cross section with a two-body final state, in the CoM frame (for the derivation, see these notes, §4.7):

\[
\left( \frac{d\sigma}{d\Omega} \right)_{\text{CoM}} = \frac{1}{64\pi^2 E_{\text{CoM}}^2 |\vec{p}|^2} \left( \frac{1}{4} \sum_{\text{spins}} |\mathcal{M}|^2 \right)
= \frac{\alpha^2}{16E^6 |\vec{k}|} \left( E^4 + |\vec{k}|^2 |\vec{p}|^2 \cos^2 \theta + E^2 (m_e^2 + m_\mu^2) \right) \quad (3.26)
\]

where \(\alpha \equiv \frac{e^2}{\pi}\) is the fine structure constant. This can be boiled a bit with kinematical relations \(|\vec{k}| = \sqrt{E^2 - m_e^2}, |\vec{p}| = \sqrt{E^2 - m_\mu^2}\) to make manifest that it depends only on two independent kinematical variables, which we can take to be the CoM energy \(E\) and the scattering angle \(\theta\) in \(\vec{k} \cdot \vec{p} = |\vec{k}||\vec{p}| \cos \theta\) (best understood from the figure). It simplifies a bit if we take \(E \gg m_e\), and more if we take \(E \gg m_\mu \sim 200m_e\), to

\[
\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E_{\text{CoM}}^2} \left( 1 + \cos^2 \theta \right). \quad (3.27)
\]

In fact, the two terms here come respectively from spins transverse to the scattering plane and in the scattering plane; see Schwartz §5.3 for an explanation.

There is a lot more to say about what happens when we scatter an electron and a positron! Another thing that can happen is that the final state could be an electron and positron again (Bhabha scattering\footnote{See figure 3 here. Now remember that a person doesn’t have much control over their name. By the way, I totally believe the bit about non-perturbative strings = lint.}).
They are not necessarily the same $e^-$ and $e^+$, though (except in the sense that they are all the same), because another way to get there at tree level is the second, $t$-channel, diagram, at right. The intermediate photon in that diagram has $k_t = (p_1 - p_3)$, so that the denominator of the propagator is $t = k_t^2 = (p_1 - p_3)^2$ instead of $s$.

Squaring this amplitude gives

$$|M_s + M_t|^2 = |M_s|^2 + |M_t|^2 + 2\text{Re}(M_s M^*_t),$$

interference terms. Interference terms mean that you have to be careful about the overall sign or phase of the amplitudes.

You may be surprised that the cross section (3.27) decreases with energy. Mechanically this comes mainly from the $1/s^2$ from the photon propagator: as $s$ grows, the intermediate photon is more and more off-shell. But more deeply, it’s because above we’ve studied an exclusive cross-section, in the sense that we fixed the final state to be exactly a muon and an antimuon. At higher energies, nothing new happens, because the final state is fixed.

It has also been very valuable to think about inclusive cross-sections for $e^+e^-$ scattering, because in this way you can make anything that the $s$-channel photon couples to, if you put enough energy into it. The inclusive cross section for $(e^+e^-)$ goes to anything) does grow with energy, and jumps at energies which are thresholds for new particles in the final state. In this way, for example, we can also make quarks (more specifically quark-antiquark pairs) since they also carry electric charge. See Peskin pp 139-140 for a bit more about that, and in particular how this observable gives evidence that there are three colors of quarks.

\[ e^- e^- \leftrightarrow e^- e^- \]. What happens if instead we scatter two electrons (Möller scattering)? In that case, the leading order diagrams are the ones at right. Now the intermediate photons have $k_t = (p_1 - p_3)$ and $k_u = (p_1 - p_4)$ respectively, so that the denominator of the propagator is $t$ and $u$ in the two diagrams. The evaluation of these diagrams has a lot in common with the ones for $e^+e^- \rightarrow e^+e^-$, namely you just switch some of the legs between initial and final state.
The relation between such amplitudes is called crossing symmetry. Let’s illustrate it instead for $e^- \mu^- \leftarrow e^- \mu^-$, where again there is only one diagram, related by crossing to (3.30). The diagram is the one at right. (The muon is the thicker fermion line.)

$$iM = \left( -ie \bar{\nu}_3 \gamma^\mu u_1 \right)_{\text{electrons}} \frac{-i \left( \eta_{\mu\nu} - \frac{(1-\xi)k_\mu k_\nu}{k^2} \right)}{k^2} \left( -ie \bar{\nu}_2 \gamma^\nu u_4 \right)_{\text{muons}} (3.29)$$

with $k \equiv p_1 - p_3 = p_2 - p_4$. It differs from (3.30) by replacing the relevant $\nu$s with $\mu$s for the initial/final antiparticles that were moved into final/initial particles, and relabelling the momenta. After the spin sum,

$$\frac{1}{4} \sum_{s,t,u} |M|^2 = \frac{e^4}{4t^2} \left( \gamma^\mu \left( \frac{p_4 + m_\mu}{p} \right) \gamma^\nu \left( \frac{p_2 + m_\mu}{p} \right) \right) \left( \gamma^\nu \left( \frac{p_3 - m_\mu}{p} \right) \gamma^\mu \left( \frac{p_1 - m_\mu}{p} \right) \right)$$

this amounts to the replacement $(p_1, p_2, p_3, p_4) \rightarrow (p_1, -p_3, p_4, -p_2)$; on the Mandelstam variables, this is just the permutation $(s, t, u) \rightarrow (t, u, s)$.

**Crossing symmetry more generally.** If you look at a Feynman diagram on its side (for example because someone else fails to use the convention that time goes to the left) it is still a valid amplitude for some process. Similarly, dragging particles between the initial and final state also produces a valid amplitude. Making this relation precise can save us some work. The precise relation for dragging an incoming particle into the final state, so that it is an outgoing antiparticle, is:

$$iM_{f \rightarrow iA}(p_f, p_i, p_A) = iM_{A \rightarrow f}(p_f, k = -p_A; p_i) = \left( -1 \right)^{\text{number of fermions shuffled between in and out}}$$

(If you must, note that this is another sense in which an antiparticle is a particle going backwards in time.) If $A$ is a spinor particle, the sum relations for particles and antiparticles are different:

$$\sum_r u^r(p) \bar{u}^r(p) = \slashed{p} + m, \quad \sum_r v^r(k) \bar{v}^r(k) = \slashed{k} - m = -\left( \slashed{p} + m \right)$$

– after accounting for $k = -p_A$, they differ by an overall sign. Hence we must also append a fermion sign factor $\left( -1 \right)^{\text{number of fermions shuffled between in and out}}$ in the unpolarized scattering probability. We’ll study a well-motivated example in more detail next.
Mott formula. By studying scattering of an electron from a heavy charged fermion (a muon is convenient) we can reconstruct the cross section for scattering off a Coulomb potential (named after Mott). This example will be important next quarter, where we’ll figure out how it is corrected by other QED processes.

\[
\mu^+\mu^- \leftrightarrow e^+e^-.
\]

Consider again the process \(\mu^+\mu^- \leftrightarrow e^+e^-\). To try to keep things straight, I’ll call the electron momenta \(p, p'\) and the muon momenta \(k, k'\), since that won’t change under crossing. We found the amplitude

\[
iM_{\mu^+\mu^- \leftrightarrow e^+e^-} = \left( -ie\bar{u}(k)\gamma^\mu u'(k') \right)_{\text{muons}} \frac{-i}{q^2} \left( \eta_{\mu
u} - \frac{1-\xi}{q^2} q_\mu q_\nu \right) \left( -ie\bar{v}'(p')\gamma^\nu v'(p) \right)_{\text{electrons}} (3.30)
\]

(with \(q \equiv p + p' = k + k'\)) and the (unpolarized) scattering probability density

\[
\frac{1}{4} \sum_{\text{spins}} |M|^2 \text{ spinor traces} = \frac{1}{4} \frac{e^4}{s^2} E_{\mu\nu} M_{\mu\nu},
\]

where the tensor objects \(E_{\mu\nu}, M_{\mu\nu}\) come respectively from the electron and muon lines,

\[
\frac{1}{4} E_{\mu\nu} = p_\mu p'_\nu + p'_\mu p_\nu - \eta_{\mu\nu} (p \cdot p' + m_e^2)
\]

\[
\frac{1}{4} M_{\mu\nu} = k_\mu k'_\nu + k'_\mu k_\nu - \eta_{\mu\nu} (k \cdot k' + m_\mu^2),
\]

and they are contracted by the photon line, with \(s = q^2 = (p + p')^2\).

\[
e^-\mu^- \leftrightarrow e^-\mu^-.
\]

To get from this the amplitude (tree level, so far) for the process \(e^-\mu^- \leftrightarrow e^-\mu^-\), we must move the incoming positron line to an outgoing electron line, and move the outgoing antimuon line to an incoming muon line (hence the sign in \(\sigma\) will be \((-1)^{\text{number of fermions shuffled between in and out}} = (-1)^2 = 1\)). Relative to the amplitude for \(\mu^+\mu^- \leftrightarrow e^+e^-\) (3.30), we must replace the relevant \(v\) with \(u\) for the initial/final antiparticles that were moved into final/initial particles, and we must replace \(p' \rightarrow -p', k' \rightarrow -k'\):

\[
iM = \left( -ie\bar{u}(p')\gamma^\mu u(p) \right)_{\text{electrons}} \frac{-i}{q^2} \left( \eta_{\mu\nu} - \frac{1-\xi}{q^2} q_\mu q_\nu \right) \left( -ie\bar{u}(k)\gamma^\nu u(k') \right)_{\text{muons}} (3.31)
\]

\[\text{Relative to the notation I used earlier, } p_1 = p, p_2 = p', p_3 = k, p_4 = k'.\]
with \( q_t \equiv p - p' = k - k' \). After the spin sum,

\[
\frac{1}{4} \sum_{s,s',r,r'} |M|^2 = \frac{e^4}{k^2} \left( -p_\mu p'_\nu - p'_\mu p_\nu - \eta_{\mu\nu}(-p \cdot p' + m^2_\pi) \right) \\
\cdot \left( -k_\mu k'_\nu - k'_\mu k_\nu - \eta_{\mu\nu}(-k \cdot k' + m^2) \right)
\]

(3.32)

On the Mandelstam variables, this is just the permutation \((s, t, u) \to (t, u, s)\).

**Payoff: the Mott formula.** Recall other ways of figuring out the scattering cross section from a Coulomb formula from a point charge of charge \(ze\).

We think about scattering from a fixed electrostatic potential \(A_0 = \frac{ze}{r}\) and do classical mechanics. I can never remember how this goes. Instead, let’s just scatter an electron off a heavy charge, such as a muon. If the charge of the heavy object were \(z\) times that of the electron, we would multiply the amplitude by \(z\) and the cross section by \(z^2\).

‘Heavy’ here means that we can approximate the CoM frame by its rest frame, and its initial and final energy as \(k'_0 = m_\mu, k_0 = \sqrt{m^2_\mu + \vec{k}^2} = m_\mu + \frac{1}{2} \vec{k}^2/m_\mu + \cdots \simeq m_\mu\). Also, this means the collision is approximately elastic, \(E' \simeq E\). In the diagram of the kinematics at right, annoyingly, \(s \equiv \cos \theta, c \equiv \sin \theta\). (Sorry.)

This means that the muon-line tensor factor \(M_{\mu\nu}\) in (4.11) simplifies dramatically:

\[
-\frac{1}{4} M_{\mu\nu} \simeq \delta_{\mu0} \delta_{\nu0} 2m^2_\mu.
\]

In the electron line, we’ll need the ingredient.

\[
-p \cdot p' + m^2_\pi = -E^2 + \vec{p}^2 \cos \theta + m^2_\pi = -\vec{p}^2(1 - \cos \theta).
\]

(3.33)

So

\[
E_{\mu\nu} M_{\mu\nu} = 32m^2_\mu E^{00} = 32m^2_\mu (2E^2 + \eta^{00}(-p \cdot p' + m^2_\pi))
\]

\[
\overset{(3.33)}{=} 32m^2_\mu (2E^2 - \vec{p}^2(1 - \cos \theta))
\]

\[
\overset{\text{trig}}{=} 32m^2_\mu 2(E^2 - \vec{p}^2 \sin^2 \theta/2)^{\beta^2 = \vec{p}^2/E^2} 64m^2_\mu E^2(1 - \beta^2 \sin^2 \theta/2) \overset{32.4}{=}
\]

70
From the two-body phase space, the cross section is
\[ d\sigma = \frac{1}{v_{\text{rel}}^2} \frac{1}{2E} \frac{1}{2m_\mu} \frac{z^2 e^4}{t^2} - 64m_\mu^2 E^2 (1 - \beta^2 \sin^2 \theta/2) \frac{d\Omega}{16\pi^2 E_{\text{total}}} \]
\[ E_{\text{total}} \simeq m_\mu \]
\[ \beta \]
\[ = \]
\[ \frac{4E}{\beta} \]
\[ \frac{2\bar{p}^2 (1 - \cos \theta)}{t^2} d\Omega \]

Noting that \( t = (p - p')^2 = -2\bar{p}^2 (1 - \cos \theta) \), we get
\[ \frac{d\sigma}{d\Omega_{\text{Mott}}} = \frac{z^2 \alpha^2 (1 - \beta^2 \sin^2 \theta/2)}{4\beta^2 \bar{p}^2 \sin^4 \theta/2} . \]

If we take \( \beta \ll 1 \) in this formula we get the Rutherford formula. Notice that it blows up at \( \theta \to 0 \). This is a symptom of the long-range nature of the Coulomb potential, \( \text{i.e.} \) the masslessness of the photon.

**Electron-proton scattering.** The answer is basically the same if we think of the heavy particle in (4.10) as a proton (we have to flip the sign of the charge but this gets squared away since there is no interference in this case). \( ep \to ep \) is called Rutherford scattering, for good reason\(^{25}\).

**Electron-photon scattering.** In the case of the process \( e^- \gamma \leftrightarrow e^- \gamma \), \( ^{26} \) we meet a new ingredient, namely external photons:

\[ i\mathcal{M} = \left( \begin{array}{c} \gamma e \\ \bar{\Psi} A \Psi \\ \bar{\Psi} A \Psi \\ \end{array} \right) \]
\[ \equiv i\mathcal{M}_s + i\mathcal{M}_t \]

\[ = (-ie)^2 \epsilon_1^{\mu} \epsilon_4^{\nu} \bar{u}_3 \left( \gamma_{\nu} \frac{i\gamma_k + m}{s - m^2} \gamma_\mu + \gamma_\mu \frac{i\gamma_l + m}{t - m^2} \gamma_\nu \right) u_2 . \]

The two amplitudes have a relative plus since we only mucked with the photon contractions, they just differ by how the gamma matrices are attached. If you don’t believe me, draw the contractions on this:

\[ \langle \gamma e | (\bar{\Psi} A \Psi)_1 (\bar{\Psi} A \Psi)_2 | \gamma e \rangle \]

(\( I’m \) not going to TeX it, thank you).

\(^{25}\)If you don’t know why, you should go read *Inward Bound*, by Abraham Pais, as soon as possible.

\(^{26}\)which at high energy is called Compton scattering and at low energies is called Thomson scattering. Despite my previous curmudgeonly footnote chastising the innocent reader for an imagined incomplete knowledge of the history of science, I do have a hard time remembering which name goes where. Moreover, as much as I revere the contributions of many of these folks, I find that using their names makes me think about the people instead of the physics. No one owns the physics! It’s the same physics for lots of space aliens, too.
Now, if we don’t measure the photon polarizations, we need
\[ P = \frac{1}{4} \sum_{\text{polarizations, spins}} |\mathcal{M}|^2. \]

The key ingredient is the completeness relation
\[ \sum_{i=1,2} \epsilon^{i*}_\mu(k) \epsilon^i_\nu(k) = -\eta_{\mu\nu} + \text{something proportional to } k^\mu k^\nu. \]

We can do various incantations to find a definite coefficient, but it will not matter because of the Ward identity: anytime there is an external photon \( \epsilon(k)_\mu \), the amplitude is \( \mathcal{M} = M_\mu \epsilon^\mu(k) \) and satisfies \( k^\mu \mathcal{M}_\mu = 0 \). Therefore, we can ignore the term about which I was vague and we have
\[ \sum_{\text{polarizations}} |\mathcal{M}|^2 = \sum_i \epsilon^{i*}_\mu \mathcal{M}^{\mu*} \mathcal{M}^\nu \epsilon^i_\nu = -\eta_{\mu\nu} \mathcal{M}^{\mu*} \mathcal{M}^\nu + (\text{terms with } \mathcal{M}_\mu k^\mu) \]
\[ = -\mathcal{M}^{\mu*} \mathcal{M}_\mu. \]

Don’t be scared of the minus sign, it’s because of the mostly minus signature, and makes the thing positive. But notice the opportunity to get negative probabilities if the gauge bosons don’t behave!

A dramatic process related by crossing to Compton scattering is pair annihilation, \( \mathcal{M}_{\gamma\gamma \rightarrow e^+ e^-} \). See the end of Peskin §5, where he has a nice plot comparing to experimental data the result for \( \frac{d\sigma}{d\Omega} \) as a function of scattering angle.