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 most 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 the subject.

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, worth emphasizing, so let me stop and emphasize it. (Later on we’ll understand in what sense it’s a necessary consequence of Lorentz symmetry. The converse is false: particle production can happen even without relativity.)

Single-particle QM. In classes with the title ‘Quantum Mechanics’, we generally study quantum systems where the Hilbert space \( \mathcal{H}_1 \) holds states of a single particle (or sometimes a fixed small number of them).

The observables of such a system are represented by hermitian operators acting on \( \mathcal{H}_1 \). For example, the particle has a position \( \vec{x} \) and a momentum \( \vec{p} \) each of which is a \( d \)-vector of operators (for a particle in \( d \) space dimensions). The particle could be an electron (in which case it also has an inherent two-valuedness called spin) or a photon (in which case it also has an inherent two-valuedness called polarization).

Time evolution is generated by a Hamiltonian \( \mathbf{H} \) which is made from the position and momentum (and whatever internal degrees of freedom it has), \( \mathbf{i}\hbar \partial_t |\psi\rangle = \mathbf{H} |\psi\rangle \).

Finally, the fourth (most ersatz) axiom regards measurement: when measuring an observable \( \mathbf{A} \) in a state \( |\psi\rangle \in \mathcal{H} \), we should decompose the state the eigenbasis \( \mathbf{A} |a\rangle = a |a\rangle \), \( |\psi\rangle = \sum_a \langle a|\psi\rangle |a\rangle \); the probability to get the answer \( a \) is \( |\langle a|\psi\rangle|^2 \).
By the way: The components of the state vector in the position basis \( \langle \vec{x}|\psi \rangle = \psi(\vec{x}) \) is a function of space, the wavefunction. This looks like a field. It is not what we mean by a field in QFT. Meaningless phrases like ‘second quantization’ may conspire to try to confuse you about this.

Now suppose you want to describe quantumly the emission of a photon from an excited electron in an atom. Surely this is something for which we need QM. How do you do it?

In the first section of this course we’ll follow an organic route to discovering an answer to this question. This will have the advantage of making it manifest that the four axioms of QM just reviewed are still true in QFT. It will de-emphasize the role of Lorentz symmetry; in fact it will explicitly break it. It will emerge on its own!

‘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: space can go on forever (the infrared (IR)), and, 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.

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*

Zee, *Quantum Field Theory*

Le Bellac, *Quantum Physics*.

Many other bits of wisdom come from the Berkeley QFT course of Prof. Lawrence Hall.
0.2 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:

$$ds^2 = +dt^2 - d\vec{x} \cdot d\vec{x} = \eta_{\mu\nu} dx^\mu dx^\nu \quad \text{with} \quad \eta^{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}_{\mu\nu}.$$ 

(spacelike is negative). We will also write $\partial_\mu \equiv \partial_t, \vec{\nabla}_x$, and $\partial^\mu \equiv \eta^{\mu\nu}\partial_\nu$.

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

A useful generalization of the shorthand $\hbar \equiv \frac{\hbar}{2\pi}$ is

$$d\vec{k} \equiv \frac{d\vec{k}}{2\pi}.$$ 

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

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

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.

We work in units where $\hbar$ and the speed of light, $c$, are equal to one unless otherwise noted.

I reserve the right to add to this page as the notes evolve.

Please tell me if you find typos or errors or violations of the rules above.
1 From particles to fields to particles again

Here is a way to discover QFT starting with some prosaic ingredients. Besides the advantages mentioned above, it will allows us to check that we are on the right track with simple experiments.

1.1 Quantum sound: Phonons

Let’s think about a crystalline solid. The specific heat of solids (how much do you have to heat it up to change its internal energy by a given amount) was a mystery before QM. The first decent (QM) model was due to Einstein, where he supposed that each atom is a (independent) quantum harmonic oscillator with frequency $\omega$. This correctly predicts that the specific heat decreases as the temperature is lowered, but is very crude. Obviously the atoms interact: that’s why they make a nice crystal pattern, and that’s why there are sound waves, as we will see. By treating the elasticity of the solid quantum mechanically, we are going to discover quantum field theory. One immediate benefit of this will be a framework for quantum mechanics where particles can be created and annihilated.

As a more accurate toy model of a one-dimensional crystalline solid, let’s consider a linear chain of particles of mass $m$, each connected to its neighbors by springs with spring constant $\kappa$. 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. Now set $\lambda = 0$. (It will be a little while before we turn back on the interactions
resulting from nonzero $\lambda$; bear with me.) This hamiltonian above describes a collection of coupled harmonic oscillators\(^1\), with a matrix of spring constants $V = k_{ab}q_aq_b$. If we diagonalize the matrix of spring constants, we will have a description in terms of decoupled oscillators, called normal modes.

Since our system has (discrete) translation invariance, these modes are labelled by a wavenumber $k^2$:

\[
q_k = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{i k x_n} q_n, \quad p_k = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{i k x_n} p_n,
\]

(Notice that in the previous expressions I didn’t use boldface; that’s because this step is really just classical physics. Note the awkward (but in field theory, inevitable) fact that

\(^1\)In case you are rusty, or forget the numerical factors like I do, here is a review of the operator solution of the SHO:

\[
H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 q^2 = \frac{\hbar \omega}{2} (p^2 + Q^2) = \hbar \omega \left( a^\dagger a + \frac{1}{2} \right)
\]

with

\[
a = \frac{1}{\sqrt{2}} (Q + iP), \quad a^\dagger = \frac{1}{\sqrt{2}} (Q - iP).
\]

Here I’ve defined these new operators to hide the annoying factors:

\[
[Q, P] = i\hbar \mathbb{1}, \quad [a, a^\dagger] = \mathbb{1}.
\]

The number operator $N \equiv a^\dagger a$ satisfies

\[
[N, a] = -a, \quad [N, a^\dagger] = +a^\dagger.
\]

So $a$ and $a^\dagger$ are lowering and raising operators for the number operator. The eigenvalues of the number operator have to be positive, since

\[
0 \leq \| a | n \rangle \|^2 = \langle n | a^\dagger a | n \rangle = \langle n | N | n \rangle = n \langle n | n \rangle
\]

which means that for $n = 0$ we have $a | n = 0 \rangle = 0$. If it isn’t zero (i.e. if $n \geq 1$), $a | n \rangle$ is also an eigenvector of $N$ with eigenvalue $n - 1$. It has to stop somewhere! So the eigenstates of $N$ (and hence of $H = \hbar \omega \left( N + \frac{1}{2} \right)$ are

\[
|0\rangle, \quad |1\rangle \equiv a^\dagger |0\rangle, \quad ... , |n\rangle = c_n \left( a^\dagger \right)^n |0\rangle \ldots
\]

where we must choose $c_n$ to normalize these states. The answer which gives $\langle n | n \rangle = 1$ is $c_n = \frac{1}{\sqrt{n!}}$.

\(^2\)The inverse transformation is:

\[
q_n = \frac{1}{\sqrt{N}} \sum_{k>0}^{2\pi/a} e^{-ikx_n} q_k, \quad p_n = \frac{1}{\sqrt{N}} \sum_{k>0}^{2\pi/a} e^{-ikx_n} p_k. \tag{1.2}
\]
we'll have (field) momentum operators $p_k$ labelled by a wavenumber aka momentum.)
The nice thing about the fourier kernel is that it diagonalizes the translation operator:

$$Te^{ikx} \equiv e^{ik(x+a)} = e^{ika}e^{ikx}.$$  

**Regulators:** Because $N$ is finite, $k$ takes discrete values ($1 = e^{ikNa}$); this is a long-wavelength “IR” property. Because of the lattice structure, $k$ is periodic (only $e^{ikan}$, $n \in \mathbb{Z}$ appears): $k \equiv k + 2\pi/a$; this is a short-distance “UV” property. The range of $k$ can be taken to be

$$0 \leq k \leq \frac{2\pi(N-1)}{Na}.$$  

Because of the periodicity in $k$, we can equivalently label the set of wavenumbers by$^3$:

$$0 < k \leq \frac{2\pi}{a} \quad \text{or} \quad -\frac{\pi}{a} < k \leq \frac{\pi}{a}.$$  

(1.3)

Summary: Because the system is in a box (periodic), $k$-space is discrete. Because the system is on a lattice, $k$-space is periodic. There are $N$ oscillator modes altogether.

The whole hamiltonian is a bunch of decoupled oscillators, labelled by these funny wave numbers:

$$H = \sum_k \left( \frac{p_k p_{-k}}{2m} + \frac{1}{2} m\omega_k^2 q_k q_{-k} \right)$$  

(1.4)

where the frequency of the mode labelled $k$ is

$$\omega_k \equiv 2\sqrt{\frac{\kappa}{m}} \sin \frac{|k|a}{2}.$$  

(1.5)

Why might we care about this frequency? For one thing, consider the Heisenberg equation of motion for the deviation of one spring:

$$i\partial_t q_n = [q_n, H] = \frac{p_n}{m}, \quad i\partial_t p_n = [p_n, H]$$

Combining these gives:

$$m\ddot{q}_n = -\kappa ((q_n - q_{n-1}) - (q_n - q_{n+1})) = -\kappa (2q_n - q_{n-1} - q_{n+1}).$$

In terms of the fourier-mode operators:

$$m\ddot{q}_k = -\kappa (2 - 2 \cos ka) q_k.$$ 

$^3$This range of independent values of the wavenumber in a lattice model is called the Brillouin zone. There is some convention for choosing a fundamental domain which prefers the last one but I haven’t found a reason to care about this.
Plugging in a Fourier ansatz in time \( q_k(t) = \sum_\omega e^{-i\omega t} q_{k,\omega} \) turns this into an algebraic equation which says \( \omega^2_k = (2m)^2 \sin^2 \frac{|k|a}{2} \) for the allowed modes. We see that (the classical version of) this system describes waves:

\[
0 = (\omega^2 - \omega_k^2) q_{k,\omega} \quad \text{for} \quad k \ll a^{-1} \approx (\omega^2 - v_s^2 k^2) q_{k,\omega}.
\]

The result for small \( k \) is the Fourier transform of the wave equation:

\[
(\partial_t^2 - v_s^2 \partial_x^2) q(x,t) = 0.
\] (1.6)

\( v_s \) is the speed of propagation of the waves, in this case the speed of sound. Comparing to the dispersion relation (1.5), we have found

\[
v_s = \frac{\partial \omega_k}{\partial k} \bigg|_{k \to 0} = a \sqrt{\frac{\kappa}{m}}.
\]

The wave looks something like this:

So the story I am telling is a quantization of sound waves. Soon we will quantize electromagnetic (EM) waves, too.

So far the fact that quantumly \( [q_n, p_{n'}] = i\hbar \delta_{nn'} \| \) hasn’t really mattered in our analysis (go back and check – we could have derived the wave equation classically). For the Fourier modes, this implies the commutator

\[
[q_k, p_{k'}] = \sum_{n,n'} U_{kn} U_{k'n'} [q_n, p_{n'}'] = i\hbar \| \sum_n U_{kn} U_{k'n} = i\hbar \delta_{k,-k'} \|
\]

(In the previous expression I called \( U_{kn} = \frac{1}{\sqrt{N}} e^{ikx_n} \) the unitary matrix realizing the discrete Fourier kernel.)
To make the final step to decouple the modes with \( k \) and \(-k\), introduce the annihilation and creation operators\(^4\)

\[
\text{For } k \neq 0: \quad q_k = \sqrt{\frac{\hbar}{2m\omega_k}} (a_k + a_k^\dagger), \quad p_k = \frac{1}{i} \sqrt{\frac{\hbar m\omega_k}{2}} (a_k - a_k^\dagger).
\]

They satisfy

\[
[a_k, a_{k'}^\dagger] = \delta_{kk'}\mathbb{I}.
\]

In terms of these, the Hamiltonian is

\[
H_0 = \sum_k \hbar \omega_k \left( a_k^\dagger a_k + \frac{1}{2} \right) + \frac{p_0^2}{2m},
\]

- it is a sum of decoupled oscillators, and a free particle describing the center-of-mass.

**The discovery of Fock space (aka particle phononemology\(^5\)).** The ground state satisfies \( a_k |0\rangle = 0 \) for all \( k \) (and has eigenvalue zero for the center-of-mass momentum, \( p_0 = 0 \)). The first excitation above the ground state

\[
a_k^\dagger |0\rangle \propto |\text{one phonon with momentum } \hbar k\rangle
\]

has energy \( \hbar \omega_k \). It is called a phonon with momentum \( \hbar k \).\(^6\) This is what in undergrad QM we would have called “\( |k\rangle \)”; we can make a state with one phonon in a position eigenstate by taking superpositions:

\[
|\text{one phonon at position } x\rangle = \sum_k e^{ikx} |\text{one phonon with momentum } \hbar k\rangle \sim \sum_k e^{ikx} a_k^\dagger |0\rangle.
\]

The number operator (of the SHO with label \( k \)) \( N_k \equiv a_k^\dagger a_k \) counts the number of phonons with momentum \( k \). The ground state is the state with no phonons. We can also make a state with two phonons:

\[
|k, k'\rangle = a_k^\dagger a_{k'}^\dagger |0\rangle
\]

whose energy is \( E = \omega_k + \omega_{k'} \). Note that all these states have non-negative energy.

So this construction allows us to describe situations where the number of particles \( N = \sum_k N_k \) can vary! That is, we can now describe dynamical processes in which the

\(^4\)You might notice that when \( k = 0, \omega_k = 0 \). This last step applies to the modes with \( \omega_k \neq 0 \), hence \( k \neq 0 \). The ‘zero-mode’ must be treated specially. It is neglected in many treatments of this topic but actually as a lot of physics in it. If you are curious see this discussion, page 11.

\(^5\)This little gem is due to Sami Ortoleva.

\(^6\)I put ‘proportional to’ rather than ‘equal’ in (1.7) because there can be a \( k \)-dependent normalization factor. We’ll see soon that Lorentz symmetry prefers a particular normalization here which we will adopt.
number of particles changes. Let me emphasize: In QM, we would describe the Hilbert space of two (distinguishable) particles as a tensor product of the Hilbert space of each. How can we act with an operator which enlarges the Hilbert space? We just figured out how to do it.

We can specify basis states for this Hilbert space
\[ \left( \hat{a}_{k_1}^\dagger \right)^{n_{k_1}} \left( \hat{a}_{k_2}^\dagger \right)^{n_{k_2}} \cdots |0\rangle = |\{n_{k_1}, n_{k_2}, \ldots\}\rangle \]
by a collection of occupation numbers \( n_k \), eigenvalues of the number operator for each normal mode (and the center-of-mass momentum \( p_0 \)).

Notice that in this description it is manifest that a phonon has no identity. We only keep track of how many of them there are and what is their momentum. They cannot be distinguished. Also notice that we can have as many we want in the same mode – \( n_k \) can be any non-negative integer. These are an example of bosons.

Notice that there are some energies where there aren’t any phonon states. In particular, the function (1.5) has a maximum. More generally, in a system with discrete translation invariance, there are bands of allowed energies. In the continuum limit, to which we devolve soon, this maximum goes off to the sky.

**Heat capacity of (insulating) solids: phonons are real.** The simplest demonstration that phonons are real is the dramatic decrease at low temperatures of the heat capacity of insulating solids. At high temperatures, the equipartition theorem of classical thermodynamics correctly predicts that the energy of the solid from the lattice vibrations should be \( T \) times the number of atoms, so the capacity, \( C_V = \partial_T E \) should be independent of \( T \).

At low temperatures \( T < \Theta_D \), this is wrong. \( \Theta_D \) is the temperature scale associated with the frequencies of the lattice vibrations (say the maximum of the curve \( \omega_k \) above). The resolution lies in the thermal energy of a quantum harmonic oscillator for \( T < \omega \), the energy goes to a constant \( \frac{1}{2} \hbar \omega \):

So the heat capacity (the slope of this curve) goes to zero as \( T \to 0 \).

**The Mössbauer effect: phonons are real.** Here is another dramatic consequence of the quantization of the lattice vibrations of solids, known as the Mössbauer effect, first described in words. The nuclei of the atoms in a solid have various energy levels; when hit with a \( \gamma \)-ray photon, these nuclei can experience transitions from the groundstate to some excited energy level. If an excited nucleus somewhere in the lattice gets hit by a very energetic photon (a \( \gamma \)-ray) of some very specific energy
\[ E_\gamma = \Delta E \equiv E_{\text{excited}} - E_0, \] the nucleus can absorb and re-emit that photon. The resulting sharp resonant absorption lines at \( E_\gamma = \Delta E \) are indeed observed.

This sounds simple, but here is a mystery about this: Consider a nucleus alone in space in the excited state, after it gets hit by a photon. The photon carried a momentum \( p_\gamma = E_\gamma/c \). Momentum is conserved, and it must be made up by some recoil of the absorbing nucleus. When it emits a photon again, it needn’t do so in the same direction. This means that the nucleus remains in motion with momentum \( \Delta \vec{p} = \vec{p}_1 - \vec{p}_2 \). But if some of its energy \( \Delta E = E_{\text{excited}} - E_0 \) goes to kinetic energy of recoil, not all of that energy can go to the final photon, and the emitted photon energy will be less than \( E_\gamma \) by \( E_{\text{recoil}} = \frac{\Delta p^2}{2M} \). This can be as big as \( E_{\text{recoil}}^{\text{max}} = \frac{(2\vec{p}_\gamma^2)}{2M} = \frac{(2E_\gamma/c)^2}{2M} \) (in the case of scattering by angle \( \pi \)). So instead of a sharp absorption line, it seems that we should see a broad bump of width \( \frac{(E_\gamma/c)^2}{M} \). But we do see a sharp line!

The solution of the puzzle is phonons: for a nucleus in a lattice, its recoil means that the springs are stretched – it must excite a lattice vibration, it must create some phonons. But there is a nonzero probability for it to create zero phonons. In this case, the momentum conservation is made up by an acceleration of the whole solid, which is very massive, and therefore does not recoil very much at all (it loses only energy \( \frac{p^2_\gamma}{2NM} \)).

This allows for very sharp resonance lines. In turn, this effect has allowed for some very high-precision measurements. The different widths in these cartoon absorption spectra don’t do justice to the relative factor of \( N \). An essentially similar effect makes it possible to get precise peaks from scattering of X-rays off of a solid (Bragg scattering) – there is a finite amplitude for the scattering to occur without exciting any phonons.

This is actually a remarkable thing: although solids seem ordinary to us because we encounter them frequently, the rigidity of solids is a quantum mechanical emergent phenomenon. You can elastically scatter photons off of a solid only because the atoms making up the solid participate in this collective behavior wherein the whole solid acts like a single quantum object!

**Towards scalar field theory.** It is worthwhile to put together the final relation between the ‘position operator’ and the phonon annihilation and creation operators:

\[ q_n = \sqrt{\frac{\hbar}{2\mu}} \sum_k \frac{1}{\sqrt{\omega_k}} \left( e^{ikx_n}a_k + e^{-ikx_n}a_k^\dagger \right) + \frac{1}{\sqrt{N}} q_0 \]  

(1.8)
and the corresponding relation for its canonical conjugate momentum

$$p_n = \frac{1}{i} \sqrt{\frac{\hbar \mu}{2}} \sum_k \sqrt{\omega_k} \left( e^{ikx_n} a_k - e^{-ikx_n} a_k^\dagger \right) + \frac{1}{\sqrt{N}} p_0.$$ 

The items in red are the ways in which $p$ and $q$ differ; they can all be understood from the relation $p = \mu \dot{q}$ as you will see on the homework. Notice that these expressions are formally identical to the formulae in a QFT textbook expressing a scalar field in terms of creation and annihilation operators (such as Peskin eqns. (2.25) and (2.26)). The stray factors of $\mu$ arise because we didn’t ‘canonically normalize’ our fields and absorb the $m$s into the field, e.g. defining $\phi \equiv \sqrt{\mu} q$ would get rid of them. The other difference is because we still have an IR regulator in place.

Path integral reminder in a box. At this point I will use the path-integral description. Let’s remind ourselves how this 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

$$\langle q \mid e^{-iHt} \mid q_0 \rangle = \int_{q(0)=q_0}^{q(t)=q} [dq] e^{\frac{1}{2} \int_0^t dt \left( \frac{1}{2} \dot{q}^2 - V(q) \right)}.$$

Here $[dq] \equiv N \prod_{l=1}^{M_t} dq(t_l)$ – the path integral measure is defined by a limiting procedure, 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:

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

An integral expression for this can be obtained by inserting resolutions of the identity

$$\mathbb{I} = \mathbb{I}^2 = \left( \int dp \langle p \mid \langle p \rangle \right) \left( \int dq \langle q \mid \langle q \rangle \right)$$

in between the two exponentials. For a more extensive reminder, please see §2.4 of this document.

1.2 Scalar field theory

Scalar field theory in one dimension. Notice that if we use the path integral description, some of these things (in particular the continuum, sound-wave limit) are more obvious-seeming. The path integral for our collection of oscillators is

$$Z = \int [dq_1, \ldots dq_N] e^{iS[q]}$$
with \( S[q] = \int dt \left( \sum_n \frac{1}{2} m_n \dot{q}_n^2 - V(\{q\}) \right) \equiv \int dt L(q, \dot{q}) \). The potential is \( V(\{q\}) = \sum_n \frac{1}{2} \kappa (q_{n+1} - q_n)^2 \). Now let’s try to take the continuum limit \( a \to 0, N \to \infty \) (now \( N \) is the number of points in space, not in time like in the last chapter). Basically the only thing we need is to think of \( q_n = q(x = na) \) as defining a smooth function:

![Continuum Limit q(x)](image)

[Note that the continuum field is often called \( \phi(x) \) instead of \( q(x) \) for some reason. At least the letters \( q(x) \) and \( \phi(x) \) look similar.]

We now have

\[
(q_n - q_{n-1})^2 \simeq a^2 (\partial_x q)^2 \big|_{x=na}
\]

Now 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

\[
S[q] = \int dt \int dx \frac{1}{2} \left( \mu (\partial_t q)^2 - \mu v_s^2 (\partial_x q)^2 - rq^2 - uq^4 - \ldots \right) \equiv \int dt \int dx L
\]

where I’ve introduced some parameters \( \mu, v_s, r, u \) determined from \( m, \kappa \) ... in some ways that we needn’t worry about. \( L \) is the Lagrangian density whose integral over space is the Lagrangian \( L = \int dx L \).

The equation of motion is the stationary phase condition,

\[
0 = \frac{\delta S}{\delta q(x, t)} = -\mu \ddot{q} - \mu v_s^2 \partial_x^2 q - rq - 2uq^3 - \ldots
\]

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

\[
\frac{\partial}{\partial q_n}.
\]

Functional derivatives will be very useful to us. The definition is

\[
\frac{\delta \phi(x)}{\delta \phi(y)} = \delta(x - y)
\]

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 (1.9) 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.
From the phonon problem, we automatically found \( r = u = 0 \), and the equation of motion is just the wave equation (1.6). This happened because of the symmetry
\[ q_n \to q_n + \epsilon. \]
This is the operation 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.)

The following will be quite useful for our subsequent discussion of quantum light. We can construct a hamiltonian from this action by defining a canonical field-momentum \( \pi(x) = \frac{\partial L}{\partial \dot{q}} = \mu \frac{\partial}{\partial t} q \) and doing the Legendre transformation:
\[
H = \sum_n (p_n \dot{q}_n - L_n) = \int dx \left( \pi(x) \dot{q}(x) - \mathcal{L} \right) = \int dx \left( \frac{\pi(x)^2}{2\mu} + \mu v_s^2 (\partial_x q(x))^2 + rq^2 + uq^4 + \ldots \right). \tag{1.10}
\]
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 the Maxwell field, which is a vector field, and which is our next subject. (Note that vibrations of a crystal in three dimensions actually do involve vector indices! We will omit this complication from our discussion.)

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 we also take the infinite volume limit, then the sums over \( k \) become integrals. In this limit we can make the replacement
\[
\frac{1}{L^d} \sum_k \to \int d^dk, \quad L^d \delta_{kk'} \to (2\pi)^d \delta^{(d)}(k - k').
\]
A check of the normalization factors comes from combining these two rules 1 = \( \sum_k \delta_{k,k'} = \int d^dk (2\pi)^d \delta^{(d)}(k - k') \).

**Continuum (free) scalar field theory in \( d + 1 \) dimensions.** Notice that these continuum expressions are easy to generalize to scalar field theory in any number of dimensions. Let’s do them directly in infinite volume and set \( \mu = 1 \) by rescaling fields. The action is
\[
S[\phi] = \int d^dx dt \left( \frac{1}{2} \ddot{\phi}^2 - \frac{1}{2} v_s^2 \nabla \cdot \nabla \phi - \nabla \phi - V(\phi) \right) \tag{1.11}
\]
This is 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. The equation of motion is

\[
0 = \frac{\delta S[\phi]}{\delta \phi(x)} = -\partial_t^2 \phi + v_s^2 \nabla^2 \phi - V'(\phi). \tag{1.12}
\]

For the harmonic case \(V(\phi) = \frac{1}{2}m^2\phi^2\) we know what we’re doing, and (1.12) is called the Klein-Gordon equation,

\[
0 = (\partial_\mu \partial^\mu + m^2) \phi. \tag{1.13}
\]

(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 §0.2.). 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

\[
H = \int d^d x \left( \frac{\pi(x)^2}{2} + \frac{1}{2} v_s^2 \nabla^2 \phi + \frac{1}{2} m^2 \phi^2 \right).
\]

Note that all these terms are positive.

A translation invariant problem is solved by Fourier transforms\(^8\): \(\phi(x) = \int d^d k \ e^{-i\vec{k} \cdot \vec{x}} \phi_k\), 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 \(k^2 = (-i\vec{k}) \cdot (i\vec{k}) = \vec{k} \cdot \vec{k}\). Just as in (1.4), this is merely a sum of decoupled oscillators, except for the coupling between wavenumbers \(k\) and \(-k\). Comparing with (1.4), 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 \(e^{i\vec{k} \cdot \vec{x} - i\omega t}\) to solve the Klein-Gordon equation (1.13).

We can decouple the modes with wavenumber \(k\) and \(-k\) as above by introducing the ladder operators

\[
\phi_k \equiv \sqrt{\frac{\hbar}{2\omega_k}} \left( a_k + a_{-k}^\dagger \right), \quad \pi_k \equiv \frac{1}{i} \sqrt{\frac{\hbar \omega_k^2}{2}} \left( a_k - a_{-k}^\dagger \right), \quad [a_k, a_{k'}^\dagger] = (2\pi)^d \delta^{(d)}(k - k').
\]

\(^8\)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 \([\phi(x), \pi(y)] = i\delta^{(d)}(x - y)\). 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

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

\[
\pi(\vec{x}) = \frac{1}{i} \int d^d k \; \sqrt{\frac{\hbar \omega_k}{2}} \left( e^{i\vec{k} \cdot \vec{x}} a^\dagger_k - e^{-i\vec{k} \cdot \vec{x}} a^\dagger_k^\dagger \right), \tag{1.14}
\]

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 (1.14) 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 \(e^{i\vec{k} \cdot \vec{x}}\). The momentum \(\pi\) contains the other linear combination. However, if we evolve the field operator in time using the Heisenberg equation (as you did on the HW), we find

\[
\phi(\vec{x}, t) \equiv e^{iHt} \phi(\vec{x}) e^{-iHt} = \int d^d k \; \sqrt{\frac{\hbar}{2\omega_k}} \left( e^{i\vec{k} \cdot \vec{x} - i\omega_k t} a_k + e^{-i\vec{k} \cdot \vec{x} + i\omega_k t} a_k^\dagger \right). \tag{1.15}
\]

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 (1.13). 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. (Some words about antimatter would be appropriate here, but it will be clearer later when we talk about an example of particles which are not their own antiparticles.)

Finally, in a relativistic system, anything we can say about time should also be true of space, up to some signs. So the fact that we were able to generate the time dependence by conjugation with the unitary operator $e^{iHt}$ (as in (1.15)) says that we should be able to generate the space dependence by conjugating by a unitary operator of the form $e^{-iP \cdot x}$. Here $\vec{P}$ is the last in a long list of object with a claim to the name ‘momentum’. It is the conserved charge associated with spatial translation symmetry, the generator of spatial translations. Its form in terms of the fields will be revealed below when we speak about Noether’s theorem. For now, let me emphasize that is distinct from the objects $p_n, \pi(x)$ (which were ‘momenta’ in the sense of canonical momenta of various excitations) and also from the wavenumbers $\vec{k}$ of various modes, which (when multiplied by $\hbar$) are indeed spatial momenta of single particles. (This statement gives us an expectation for what is the total momentum of a state of a collection of particles which we will check below in §1.4.) In terms of the momentum operator, then, we can write

$$\phi(x^\mu) = e^{iP_\mu x^\mu} \phi(0) e^{-iP_\mu x^\mu}$$

with $P_\mu \equiv (H, \vec{P})_\mu$.

## 1.3 Quantum light: Photons

The quantization of the Maxwell field is logically very similar to the preceding discussion. 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 are:

$$\tilde{\nabla} \cdot \vec{B} = 0, \quad \tilde{\nabla} \times \vec{E} = -\partial_t \vec{B}, \quad (1.16)$$

$$\tilde{\nabla} \cdot \vec{E} = 4\pi \rho, \quad \nabla \times \vec{B} = \partial_t \vec{E} + \frac{4\pi}{c} \vec{j}. \quad (1.17)$$

The first two equations (1.16) 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 which determine the fields by taking derivatives
and which automatically solve the constraints (1.16):
\[ \vec{E} = -\vec{\nabla}\Phi - \partial_t \vec{A}, \quad \vec{B} = \vec{\nabla} \times \vec{A}. \]

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

This wave equation is different from our scalar wave equation (1.6) 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.

Recall that the energy density of a configuration of Maxwell fields is
\[ u = \frac{\epsilon_0}{2} \left( \vec{E}^2 + \vec{B}^2 \right). \]

So the quantum Hamlitonian is
\[ H = \frac{\epsilon_0}{2} \int d^3 r \left( \vec{E}^2 + c^2 \vec{B}^2 \right). \tag{1.18} \]

Here \( \vec{E} = -\partial_t \vec{A} \) plays the role of field momentum \( \pi(x) \) in (1.10), 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
\[ [\phi(x), \pi(x')] = i\hbar \delta(x - x') \quad \leadsto \quad [A_i(\vec{r}), E_j(\vec{r}')] = -i\hbar \delta^3(\vec{r} - \vec{r}') \delta_{ij} \]
where \(i, j = 1, 3\) are spatial indices. 9 So we can immediately write down an expression for the quantum Maxwell field in terms of independent creation and annihilation operators:

\[
\vec{A}(\vec{r}) = \sum_{\vec{k}} \sqrt{\frac{\hbar}{2\epsilon_0 \omega_k L^3}} \sum_{s=1,2} \left( a_{\vec{k},s} \vec{e}_s(\hat{k}) e^{i\vec{k} \cdot \vec{r}} + a_{\vec{k},s}^\dagger \vec{e}_s^*(\hat{k}) e^{-i\vec{k} \cdot \vec{r}} \right)
\]

The field momentum is \(\vec{E} = -\partial_t \vec{A}\):

\[
\vec{E}(\vec{r}) = i \sum_{\vec{k}} \sqrt{\frac{\hbar \omega_k}{2\epsilon_0 L^3}} \sum_{s=1,2} \left( a_{\vec{k},s} \vec{e}_s(\hat{k}) e^{i\vec{k} \cdot \vec{r}} - a_{\vec{k},s}^\dagger \vec{e}_s^*(\hat{k}) e^{-i\vec{k} \cdot \vec{r}} \right)
\]

10 Also, the magnetic field operator is

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

the magnetic field is analogous to \(\vec{\nabla} \phi\) in the scalar field theory. Plugging these expressions into the Hamiltonian (1.18), 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:

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

Notice that the vacuum energy is

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

---

9 As a check, note that using this Hamiltonian and the canonical commutator, we can reproduce Maxwell’s equations using Ehrenfest’s theorem:

\[
\langle \partial_t^2 \vec{A} \rangle = \partial_t \langle \vec{E} \rangle = -\frac{i}{\hbar} \langle [\vec{H}, \vec{E}] \rangle = \langle c^2 \vec{\nabla}^2 \vec{A} \rangle.
\]

10 I am short-changing you a little bit here on an explanation of the polarization vectors, \(\vec{e}_s\). They conspire to make it so that there are only two independent states for each \(\vec{k}\) and they are transverse \(\vec{k} \cdot \vec{e}_s(\hat{k}) = 0\), so \(s = 1, 2\). The bit that I’m leaving out is the completeness relation satisfied by the polarization vectors of a given \(\vec{k}\):

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

This says that they span the plane perpendicular to \(\hat{k}\).
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 fact that $\sum_k$ is no longer a finite sum might be something to worry about. We will see below in §1.5 that this vacuum energy has physical consequences.
Consolidation of understanding

So far in this chapter, we have studied systems of increasing complexity: the simple harmonic oscillator, a non-interacting scalar field, and the EM field in vacuum (i.e. in the absence of charge). All these free field theories have the same structure, in the following sense.

In the following, Here Re $\mathbf{A} \equiv \frac{1}{2} (\mathbf{A} + \mathbf{A}^\dagger)$ as usual. The normalization constant is $\mathcal{N} = \frac{1}{2} \sqrt{\frac{\hbar}{2m\omega}}$.

$$H_{\text{SHO}} = \frac{1}{2m} p^2 + \frac{1}{2} m \omega^2 q^2 = \hbar \omega \left( a^\dagger a + \frac{1}{2} \right)$$

$$[q,p] = i\hbar \implies [a,a^\dagger] = \mathbb{1}.$$ $q = \text{Re} \mathcal{N} a$, $p = m \text{Im} \omega \mathcal{N} a.$

$$H_{\text{scalar}} = \int \! dx \left( \frac{1}{2\mu} \pi^2 + \frac{1}{2} \mu c^2 (\partial_x \phi)^2 \right) = \sum_k \hbar \omega_k \left( a_k^\dagger a_k + \frac{1}{2} \right)$$

$$\phi(x), \pi(x') = i\hbar \delta(x - x') \implies [a_k, a_{k'}^\dagger] = i\hbar \delta_{kk'}.$$ $\phi(x) = \text{Re} \left( \sum_k \mathcal{N}_k e^{ikx} a_k \right)$, $\pi(x) = \mu \text{Im} \left( \sum_k \omega_k \mathcal{N}_k e^{ikx} a_k \right).$

$$H_{\text{EM}} = \int \! d^3 x \left( \frac{\epsilon_0}{2} \mathbf{E}^2 + \frac{\epsilon_0 c^2}{2} \mathbf{B}^2 \right) = \sum_{k,s=1,2} \hbar \omega_k \left( a_{ks}^\dagger a_{ks} + \frac{1}{2} \right)$$

$$[\mathbf{A}_i(x), \mathbf{E}_j(x')] = i\hbar \delta^3(x - x')\delta_{ij} \implies [a_{ks}, a_{ks'}^\dagger] = \hbar \delta_{kk'} \delta_{ss'}.$$ $\mathbf{A}(x) = \text{Re} \left( \sum_k \mathcal{N}_k e^{ikx} a_{ks} \right)$, $\mathbf{E}(x) = \mu \text{Im} \left( \sum_k \omega_k \mathcal{N}_k e^{ikx} a_{ks} \right)$.

Note that $\mathbf{E}$ is the canonical momentum of $\mathbf{A}$ since (in Coulomb gauge) $\mathbf{E} = -\partial_t \mathbf{A}$.
1.4 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

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

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 $\mathcal{L}$ 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.

We’ve seen basically two examples so far

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

and (you get to do the Legendre transformation from the Hamiltonian of §1.3 on the homework)

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

with $A_\mu$ 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 $p_\mu = \vec{k}_\mu$ 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 $\int [D\phi] e^{i\frac{S}{\hbar}}$ and so must be dimensionless. So the Lagrangian density has mass dimension $d+1$. This means that the KG field has mass dimension $\frac{d-1}{2}$ (and the mass $m$ has mass dimension 1 (yay!)). In $d+1 = 3+1$ dimensions, $E \sim \vec{A}$, $B \sim \vec{\nabla} A$ 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.

<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>$\mathcal{L}$</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>
</tbody>
</table>

[End of Lecture 3]
The equation of motion is

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

Note the functional derivative. You can check that in the case when \( \mathcal{L} \) depends only on \( \phi \) and \( \partial_\mu \phi \), this is the same as the Lagrange EOM

\[ 0 = \frac{\partial \mathcal{L}}{\partial \phi_r} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_r)} \]

(for each \( r \)) which I can’t remember.

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

\[ \mathcal{L} = A + B\chi + \frac{1}{2} C\chi^2 + \frac{1}{2} D\partial_\mu \chi \partial_\mu \chi + \ldots. \]

From the path integral point of view, the field is just an integration variable. Sometimes, its normalization is meaningful, like in the phonon example where it began its life as the displacement of the atoms from their equilibrium. So you see that we are not losing generality except in our neglect of interactions, and in our neglect of terms with more derivatives. The former neglect we will repair little by little in this course, by doing perturbation theory. The latter is justified well by the renormalization group philosophy, which is a subject for Physics 215C, i.e. the Spring Quarter.

**Canonical field momentum and Hamiltonian.** The Hamiltonian viewpoint in field theory has the great virtue of bringing out the physical degrees of freedom. It has the great shortcoming that it picks out the time coordinate as special and obscures Lorentz symmetry.

The canonical field momentum is defined to be

\[ \pi(x) = \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi(x))}. \]

Notice that this expression assumes a local Lagrangian density. \( \pi \) is actually a ‘field momentum density’ in the sense that the literal canonical momentum is \( \frac{\partial}{\partial \phi(x)} L = d^4 x \pi(x) \) (as opposed to \( \mathcal{L} \)). I will often forget to say ‘density’ here.

The hamiltonian is then

\[ H = \sum_n p_n q_n - L = \int d^4 x \left( \pi(x) \dot{\phi}(x) - \mathcal{L} \right) \equiv \int d^4 x \mathfrak{h}. \]

**Noether’s theorem and the Noether method.** Yay, symmetries. Why do physicists love symmetries so much? One reason is that they offer possible resting places along our never-ending chains of ‘why?’ questions. For example, one answer
(certainly the one given in Weinberg’s text, but just as certainly not the only one) to the question “Why QFT?” is: 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:

\[
S[\phi + \epsilon(x)] - S[\phi] = \int d^d x dt \partial_\mu \epsilon(x) j_\mu \quad \text{IBP} \equiv - \int d^d x dt \epsilon(x) \partial_\mu j_\mu .
\]

But if the equations of motion are obeyed, then the action is invariant under any variation, 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

\[
Q_R \equiv \int_R d^d x \ j_0
\]

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):

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

but on the other hand, by Taylor expansion,

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

\[
\quad \text{IBP} \quad \mathcal{L}(\phi, \partial_\mu \phi) + \epsilon \left( \frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \partial_\mu \right) \Delta \phi + \epsilon \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \Delta \phi \right).
\]
By combining the previous two equations for $L(\phi')$, we see that on configurations which satisfy the EOM, $0 = \partial_\mu j^\mu$ with

$$j^\mu = \frac{\partial L}{\partial (\partial_\mu \phi_r)} \Delta \phi_r - J^\mu. \quad (1.19)$$

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

$$\delta \phi \equiv i\epsilon [Q, \phi]. \quad (1.20)$$

We’ll say that $Q$ generates the symmetry, for the following reason. (1.20) is the infinitesimal version of the finite transformation

$$\phi \to \phi' \equiv e^{i\epsilon Q} \phi e^{-i\epsilon Q}.$$

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 its action 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 $\phi \to \phi' \equiv \phi + \epsilon$. Letting $\epsilon$ depend on spacetime, the variation of the action is $S[\phi + \epsilon(x)] = \int \epsilon \partial_\mu \partial^{\mu} \phi$, 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

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

which is indeed true. In this case the finite transformation is again $\phi \to \phi + \epsilon$.

- On the homework you’re studying a complex scalar $\Phi$, with $S[\Phi, \Phi^\dagger]$ is invariant under $\Phi \to e^{i\epsilon} \Phi = \Phi + i\epsilon \Phi + \mathcal{O}(\epsilon^2)$. This $U(1)$ phase transformation can be...
rewritten in terms of the real and imaginary parts as an SO(2) rotation. You’ll find that the charge can be written as
\[ Q = \int d^d x j^0 = \int d^dp \left( a_p^\dagger a_p - b_p^\dagger b_p \right) \]

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 which I tried to get right in the posted solutions.) So the particles created by \( a \) and \( b \) have opposite charge (this is essential given the mode expansion \( \Phi_k \sim a_k + b_k^\dagger \)) 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.

- Consider spatial translations, \( x^\mu \to x^\mu - a^\mu \). We can think of this as a transformation of the fields by

\[ \phi(x) \mapsto \phi(x + a) = \phi(x) + a^\nu \partial_\nu \phi + O(a^2). \]

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 \( \partial_\nu L = 0 \) but rather depends on space and time only via the fields (so \( \delta \mu L = \partial_\nu \phi \delta \mu \phi + \partial_\mu \partial_\nu \phi \delta \mu \phi \)). Let’s use the prosaic method for this one: the shift in the Lagrangian density also can be found by Taylor expansion

\[ L \mapsto L + a^\mu \frac{d}{dx^\mu} L = L + a^\nu \partial_\nu \left( \delta \mu L \right). \]

So the formula (1.19) gives

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

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

\[ P_i = \int d^d x T^0_i = - \int d^d x \pi \partial_i \phi. \]

For the Maxwell field, this gives the Poynting vector.
There is some ambiguity in the definition of the stress tensor which you’ll study on the homework.

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

\[
P_i = - \int d^d x \, \pi \partial_i \phi = \int d^d k k_i a_k^0 a_k
\]

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

- We’ll say more about the rest of the Poincaré group, i.e. rotations and boosts, later on.

1.5 Casimir effect: vacuum energy is real

[A. Zee, Quantum Field Theory in a Nutshell, §I.9] This subsection has two purposes. One is to show that the \(\frac{1}{2} \hbar \omega\) energy of the vacuum of the quantum harmonic oscillator is real. Sometimes we can get rid of it by choosing the zero of energy (which doesn’t matter unless we are studying dynamical gravity). But it is meaningful if we can vary \(\omega\) (or the collection of \(\omega\)s if we have many oscillators as for the radiation field) and compare the difference.

The other purpose is to give an object lesson in asking the right questions. In physics, the right question is often a question which can be answered by an experiment, at least in principle. The answers to such questions are less sensitive to our silly theoretical prejudices, e.g. about what happens to physics at very short distances.

In the context of the bunch of oscillators making up the radiation field, we can change the spectrum of frequencies of these oscillators \(\{\omega_k\}\) by putting it in a box and varying the size of the box. In particular, two parallel conducting plates separated by some distance \(d\) experience an attractive force from the change in the vacuum energy of the EM field resulting from their presence. The plates put boundary conditions on the field, and therefore on which normal modes are present.

To avoid some complications of E&M which are not essential for our point here, we’re going to make two simplifications:

- we’re going to solve the problem in 1+1 dimensions
- and we’re going to solve it for a scalar field.
To avoid the problem of changing the boundary conditions outside the plates we use the following device with three plates:

\[ | ← d → | ←− L − d −→ | \]

(We will consider \( L \gg d \), so we don’t really care about the far right plate.) The ‘perfectly conducting’ plates impose the boundary condition that our scalar field \( q(x) \) vanishes there. The normal modes of the scalar field \( q(x) \) in the left cavity are then

\[ q_j = \sin \left( j\pi x/d \right), \quad j = 1, 2, \ldots \]

with frequencies \( \omega_j = \frac{\pi |j|}{d} c \). There is a similar expression for the modes in the right cavity which we won’t need. We’re going to add up all the \( \frac{1}{2}\hbar \omega_j \) for all the modes in both cavities to get the vacuum energy \( E_0(d) \); the force on the middle plate is then \(-\partial_d E_0\).

The vacuum energy in the whole region of interest between the outer plates is the sum of the vacuum energies of the two cavities

\[ E_0(d) = f(d) + f(L - d) \]

where

\[ f(d) = \frac{1}{2} \hbar \sum_{j=1}^{\infty} \omega_j = \hbar c \frac{\pi}{2d} \sum_{j=1}^{\infty} j \equiv \infty. \]

We have done something wrong. What? Cliffhanger! [End of Lecture 4]

Our crime is hubris: we assumed that we knew what the modes of arbitrarily large mode number \( k \) (arbitrarily short wavelength, arbitrarily high frequency) are doing, and in particular we assumed that they cared about our silly plates. In fact, no metal in existence can put boundary conditions on the modes of large enough frequency – those modes don’t care about \( d \). The reason a conductor puts boundary conditions on the EM field is that the electrons move around to compensate for an applied field, but there is a limit on how fast the electrons can move (e.g. the speed of light). The resulting cutoff frequency is called the plasma frequency but we don’t actually need to know about all these details. To parametrize our ignorance of what the high-frequency modes do, we must cut off (or regularize) the contribution of the high-frequency modes. Let’s call modes with \( \omega_j \gg \pi/a \) high frequency where \( a \) is some short time\(^\text{11}\). Replace

\[ f(d) \sim f(a, d) = \hbar c \frac{\pi}{2d} \sum_{j=1}^{\infty} e^{-\omega_j/\pi j} \]

\(^{11}\text{You can think of } a \text{ as the time it takes the waves to move by one lattice spacing. If we work in units where the velocity is } c = 1, \text{ this is just the lattice spacing. I will do so for the rest of this discussion.} \)
\[ F = -\partial_d E_0 = -(f'(d) - f'(L - d)) \]
\[ = -\hbar \left( \left( \frac{\pi}{2a^2} + \frac{\pi}{24d^2} + O(a^2) \right) - \left( \frac{\pi}{2a^2} + \frac{\pi}{24(L - d)^2} + O(a^2) \right) \right) \]
\[ \xrightarrow{a \to 0} \left( \frac{\pi\hbar}{24} \left( 1 \frac{L}{d^2} - \frac{1}{(L - d)^2} \right) \right) \]
\[ \geq \frac{\pi\hbar c}{24d^2} \left( 1 + O(d/L) \right) . \]
force (which would lead the solid to want to shrink) is that it is hard to distinguish
the ‘phonon vacuum energy’ from the rest of the energy of formation of the solid,
that is, the energy difference between the crystalline configuration of the atoms and
the configuration when they are all infinitely separated. Certainly the latter is not
well-described in the harmonic approximation ($\lambda = 0$ in (1.1)).

A few comments about the 3+1 dimensional case of E&M. Assume the size
of the plates is much larger than their separation $L$. Dimensional analysis shows that
the force per unit area from vacuum fluctuations must be of the form

$$P = A \frac{\hbar c}{L^4}$$

where $A$ is a numerical number. $A$ is not zero!

Use periodic boundary conditions in the xy planes (along the plates). The allowed
wave vectors are then

$$\vec{k} = \left( \frac{2\pi n_x}{L_x}, \frac{2\pi n_y}{L_y} \right)$$

with $n_x, n_y$ integers.

We have to do a bit of E&M here. Assume the plates are perfect conductors
(this where the hubris about the high-frequency modes enters). This means that the
transverse component of the electric field must vanish at the surface. Instead of plane
waves in $z$, we get standing waves: $\phi(z) \propto \sin(n\pi z/L)$.

The frequencies of the associated standing waves are then

$$\omega_n(\vec{k}) = c\sqrt{\frac{\pi^2 n^2}{L^2} + \vec{k}^2}, n = 0, 1, 2$$

Also, there is only one polarization state for $n = 0$.

So the zero-point energy is

$$E_0(L) = \frac{\hbar}{2} \left( 2 \sum_{n,\vec{k}}' \omega_n(\vec{k}) \right)$$

where it’s useful to define

$$\sum'_{n,\vec{k}} \equiv \frac{1}{2} \sum_{n=0,\vec{k}} + \sum_{n\geq 1,\vec{k}}$$

Now you can imagine introducing a regulator like the one we used above, and replacing

$$\sum'_{n,\vec{k}} \cdot \sum_{n,\vec{k}} e^{-\omega_n(\vec{k})/\pi}.$$
and doing the sums and integrals and extracting the small-\(a\) behavior.

### 1.6 Lessons

Starting from a collection of particles, we chained them together, and made a field; treating this system quantumly, we found a new set of particles. The new particles (the normal modes) are collective excitations: their properties can be very different from those of the constituent particles. (For example, the constituent particles are distinguishable by their locations, but phonons are indistinguishable from each other.)

Some lessons from all this hard work:

**Identical particles.** Every photon is the same as every other photon, except for their position (or momentum) and polarization state. For photons this is an immediate consequence of how we discovered them by quantizing the Maxwell field: the state with \(n\) photons of the same momentum and polarization is

\[
| n \text{ photons with } \vec{k}, \alpha \rangle = \left( \frac{a_{\vec{k},\alpha}^\dagger}{\sqrt{n!}} \right)^n |0\rangle.
\]

The same is true of all the other kinds of particles we know about, including electrons (for which we haven’t seen a similar classical field description).

This means that we can write the state of \(N\) such indistinguishable particles merely by specifying a collection of positions and of spin states – we don’t need to say which is which (and in fact, we cannot).

A (momentum-space) wavefunction for \(N\) such particles is of the form

\[
\Psi(k_1, \alpha_1; \ldots; k_N, \alpha_N) \equiv \langle k_1\alpha_1; \ldots; k_N, \alpha_N | \Psi \rangle = \langle 0| a_{k_1\alpha_1} a_{k_2\alpha_2} \cdots a_{k_N\alpha_N} | \Psi \rangle.
\]

But the same state is described if we switch the labels of any two of the particles:

\[
\Psi(k_2, \alpha_2; k_1, \alpha_1; \ldots) = a \Psi(k_1, \alpha_1; k_2, \alpha_2; \ldots)
\]

where \(a\) is some phase (recall: multiplying the whole wavefunction by a phase does not change the state). Switching them back gives back the first state:

\[
\Psi(k_1, \alpha_1; k_2, \alpha_2; \ldots) = a^2 \Psi(k_1, \alpha_1; k_2, \alpha_2; \ldots)
\]

so \(a^2 = 1\). There are two solutions: \(a = +1\) and \(a = -1\) and the two classes of particles associated with these two choices are called respectively bosons and fermions.

Note that the Hilbert space of \(N\) indistinguishable particles (bosons or fermions) is therefore not quite a tensor product of the Hilbert spaces of the individual particles.
An immediate consequence of the minus sign under exchange of fermion labels is the *Pauli exclusion principle*:

\[ \Psi_{\text{Fermions}}(k_1, \alpha_1; k_1, \alpha_1; ...) = 0. \]

No two fermions can occupy the same single-particle state. The ground state of a collection of (non-interacting) fermions is therefore quite interesting, since we must find a different single-particle state in which to put each of our fermions. This has many dramatic consequences, including the periodic table of elements, and the distinction between metals and insulators.

A puzzle to consider: the bosonic statistics of phonons and photons was an immediate consequence of the ladder operator algebra \([a, a^\dagger] = 1\). How could we possibly get a different answer?

**Lorentz invariance can emerge.** The dispersion relation for the sound mode and for the light mode was \(\omega^2 = v^2 k^2\). This is the Fourier transform of \(\partial_\mu \partial^\mu \phi(x) = 0\), a wave-equation which has Lorentz symmetry (if \(v\) is the speed appearing in the Minkowski metric). In the case of sound, we had to ignore the \(O(a^4 k^4)\) terms in the long-wavelength expansion of the dispersion relation, \(\cos(ka)\). The lattice breaks Lorentz symmetry, but its effects go away for \(ka \ll 1\). This point might make you think that the Lorentz symmetry which is so precious to us in particle physics could emerge in a similar way, but with a much smaller \(a\) than the lattice spacing in solids. There are strong constraints on how small this can be (*e.g.* this well-appreciated paper) so it is very useful to treat it as a fundamental principle.

## 2 Lorentz invariance and causality

[Peskin §2.2, 2.3, 2.4] Now we take Lorentz invariance to be an exact symmetry and see what its consequences are for QFT.

**Relativistic normalization of 1-particle states.** Fock space is spanned by the states \(|p_1, \cdots, p_n\rangle \propto a^\dagger_{p_1} \cdots a^\dagger_{p_n} |0\rangle\) where \(a_p |0\rangle = 0\). Now it is time to turn that \(\propto\) into an \(=\). Fock space is a direct sum of sectors labelled by the number of particles: \(\sum_k N_k = 0, 1, 2, \ldots\). (Without interactions, the Hamiltonian is block diagonal in this decomposition.) In the no-particle sector, it is clear what we should do: \(\langle 0 | 0 \rangle = 1\).

A one-particle state is \(|p\rangle \equiv c_p a^\dagger_p |0\rangle\). How best to choose \(c_p\)?

(This discussion is shaded because it contains equations which will not be true in the normalization we’ll use below. In this regard, beware the section of Peskin called
“how not to quantize the Dirac equation.” Suppose we choose $c_p \equiv 1$. Then

$$\left\langle \vec{k} | \vec{p} \right\rangle \equiv \langle 0 | a_k a_p^\dagger | 0 \rangle a_k | 0 \rangle = \langle 0 | [a_k, a_p^\dagger] | 0 \rangle = (2\pi)^d \delta^{(d)}(\vec{k} - \vec{p}) \equiv \delta(\vec{k} - \vec{p}).$$

Suppose the previous equation is true in my rest frame $F$. Since $1 = \int d^d\vec{p} \delta(\vec{p} - \vec{k})$, we see that $\int d^d\vec{p} \delta(\vec{p} - \vec{k})$ is Lorentz invariant. More precisely,

$$\delta(f(x)) = \sum_{\text{zeros } x_0 \text{ of } f} \frac{\delta(x - x_0)}{|f'(x_0)|}.$$

If another $F'$ is obtained by a boost in the $x$ direction, $p'_\mu = \Lambda^\mu_{\nu} p_\nu$,

$$\begin{pmatrix} E' \\ p'_x \\ p'_y \\ p'_z \end{pmatrix} = \begin{pmatrix} \gamma & -\beta \gamma & 0 & 0 \\ -\beta \gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} E \\ p_x \\ p_y \\ p_z \end{pmatrix} \implies \frac{dp'_x}{dp_x} = \gamma \left( 1 - \beta \frac{dE}{dp_x} \right) \quad \Rightarrow \quad \frac{dp'_x}{dp_x} = \gamma \left( 1 - \beta \frac{E}{p_x} \right) = \gamma \left( E - \beta p_x \right) = \frac{E'}{E}$$

where we used $E^2 = \vec{p}^2 + m^2 = p_x^2 + p_\perp^2 + m^2$ and $2E \frac{dE}{dp_x} = 2p_x$ and $\frac{dE}{dp_x} = \frac{p_x}{E}$.

So

$$\delta^{(d)}(\vec{p}' - \vec{k}) = \frac{d^d p'}{d^d p} \delta^{(d)}(\vec{p}' - \vec{k}') = \frac{dp'_x}{dp_x} \delta^{(d)}(\vec{p}' - \vec{k}') = \frac{E'}{E} \delta^{(d)}(\vec{p}' - \vec{k}').$$

Which means that in $F'$ we would have

$$\left\langle \vec{k}' | \vec{p}' \right\rangle \equiv \frac{E'}{E} \delta^{(d)}(\vec{p}' - \vec{k}').$$

There is a special frame, it’s no good.

There is an easy fix:

$$|\vec{p}'\rangle \equiv \sqrt{2\omega_p} a_p^\dagger |0\rangle.$$

In that case the calculation in the shaded text is replaced by

$$\left\langle \vec{k} | \vec{p} \right\rangle = \sqrt{4\omega_k \omega_p} \delta^{(d)}(k - p) = 2\omega_p \delta^{(d)}(k - p)$$

while

$$\left\langle \vec{k}' | \vec{p}' \right\rangle = 2\omega_p \delta^{(d)}(k' - p').$$

So the expression is the same in any frame, yay.
Now you can ask why the factor of $\sqrt{2}$. We’d like to use these states to resolve the identity in the 1-particle sector, $\mathbb{1}_1 \equiv \sum_i \sum_i |i\rangle \langle i|$. I claim that the following expression does this and makes Lorentz symmetry manifest:

$$\mathbb{1}_1 \equiv \int d^{d+1}k \frac{\theta(k^0)}{2\pi} \frac{\delta(k^2 - m^2)}{k^0} |k\rangle \langle k|$$

$$= \int d^d k \int dk_0 \theta(k^0) \left( \frac{\delta \left( k^0 - \sqrt{k^2 - m^2} \right)}{2k^0} \right) |k\rangle \langle k| = \int \frac{d^d k}{2\omega_k} \left| \omega_k, \tilde{k} \right\rangle \left\langle \omega_k, \tilde{k} \right| .$$

We used the general fact $\delta(f(x)) = \sum_{x_0 \mid f(x_0) = 0} \frac{1}{\left| f'\right|} \delta(x - x_0)$.

So in retrospect, a quick way to check the normalization is to notice that the following combination is Lorentz invariant:

$$\frac{d^d k}{2\omega_k} = \int dk_0 \theta(k_0) \frac{d^d k}{2\omega_k} \delta(k^2 - m^2) = \frac{d^d k'}{2\omega_{k'}}.$$ 

Actually, this statement has a hidden assumption, that $m^2 > 0$. In that case, the 4-vector $k^\mu$ satisfying $k^2 = m^2 > 0$ is timelike, and no Lorentz transformation connected to the identity can change the sign of $k_0$, it can only move it around within the lightcone. So the $\theta(k_0)$ is Lorentz invariant.

Notice that our convenient choice of normalization doesn’t show that our Hamiltonian description of scalar field theory is actually Lorentz invariant. For example, we have

$$[\phi(\tilde{x}), \pi(\tilde{y})]_{ETCR} = i\delta^{(d)}(\tilde{x} - \tilde{y})$$

at equal times, in one frame. What about other frames?

A second reason to study commutators is ...

### 2.1 Causality and antiparticles

**Causality:** This is the very reasonable condition on our description of physics that *events should not precede their causes.*

It will be worthwhile to think about how to implement this condition in a QFT. (The following discussion is based on appendix D of this paper.) Suppose $B$ wants to send a message to $A$. How does he do this? He applies an operator\footnote{"Applying an operator" is more complicated than it seems. Actually it means ‘changing the Hamiltonian so that the time evolution operator is $\mathbf{B}$.’}, call it $\mathbf{B}$, localized near $B$, to the shared state of their quantum many body system $|\psi\rangle_{ABE}$.\footnote{\url{https://example.org}}
(Here $E$ is for ‘environment’, the rest of the world besides $A$ and $B$.) Then $A$ measures some observable $A$; let’s assume $1 = \text{tr}A = \sum (\text{eigenvalues of } A)$. To send a different message, he should apply a different operator, say $B'$.

Under the circumstances just stated, the expectation for $A$’s measurement of $A$ is

$$\langle A \rangle_B = \langle \psi | B^\dagger e^{iHt} A e^{-iHt} B | \psi \rangle = \langle A(t) \rangle + \langle B^\dagger [B, A(t)] \rangle .$$

Therefore, if $[B, A(t)] = 0$, the expectation doesn’t depend on what $B$ did. In fact, replacing $A$ with $A^\eta$ for any $\eta$ and using ($[B, A(t)] = 0$ $\Rightarrow$ $[B, A(t)^\eta] = 0$) shows that all the moments of the distribution for $A$’s measurement will also be independent of what $B$ did, so no message gets through$^{13}$.

**Causality in relativistic QFT.** In a Lorentz invariant system, ‘precede’ is sometimes a frame-dependent notion. If $A$ is in the future lightcone of $B$, i.e. $0 < (x_A - x_B)^2 = (t_A - t_B)^2 - (\vec{x}_A - \vec{x}_B)^2$ and $t_A > t_B$, then everyone agrees that $A$ is after $B$. This is the easy case. But if $A$ and $B$ are spacelike separated, $0 > (x_A - x_B)^2 = (t_A - t_B)^2 - (\vec{x}_A - \vec{x}_B)^2$, then there is a frame where they occur at the same time, and frames where they occur in either order. This is the dangerous case.

So: causality will follow if $[A(x_A), B(x_B)] = 0$ whenever $x_A$ and $x_B$ are spacelike separated, $0 > (x_A - x_B)^2$. Recall that spacelike separated means that there is a Lorentz frame where $A$ and $B$ are at the same time.

A general operator in a scalar QFT can be made from $\phi$s and $\partial_\mu \phi$s, so the general

$^{13}$The loophole-seeking reader (ideally, this is everyone) will worry that a distribution is not in general determined by its moments. (For example, there are perfectly normalizable distributions with finite averages but where the higher moments are all infinite, such as $p(x) = \sqrt{2a}/\pi a$ on the real line: $\langle x^2 \rangle = a^2$, but $\langle x^{2n} \rangle = \infty$ for $n > 1$.) What we would really like to show is that the conditional probability $p(a|B)$ is independent of $B$, in which case for sure $A$ couldn’t learn anything about what $B$ did. That is

$$p(a|B) = \langle \psi | B^\dagger e^{iHt} | a \rangle \langle a | e^{-iHt} B | \psi \rangle = \langle P_a(t) \rangle + \langle B^\dagger [B, P_a(t)] \rangle .$$

Does $[A, B] = 0$ imply that the projector onto a particular eigenvalue of $a$ commutes with $B$? In a finite dimensional Hilbert space, it does for sure, since $0 = [A^\eta, B] = \sum a^\eta [P_a, B]$ is true for all $\eta$, which gives infinitely many equations for $[P_a, B]$. In the case of infinite dimensional $\mathcal{H}$ I think there is some room for functional analysis horror. On the other hand, any measurement has finite resolution. Thanks to Sami Ortoleva and Chunccheong Lam for help with this point.
statement will follow from considering commutators of

\[\phi(x) = \int \frac{d^d p}{\sqrt{2\omega_p}} \left( a_p e^{-ip_\mu x_\mu} + a_p^\dagger e^{+ip_\mu x_\mu} \right) |p^\mu = \omega_p \rangle \equiv \phi^{(+)} + \phi^{(-)} .\]

Here we have decomposed the field into positive- and negative-frequency parts. Notice that since \(\phi^{(+)}\) (\(\phi^{(-)}\)) only involves annihilation (creation) operators, \([\phi^{(\pm)}(x), \phi^{(\pm)}(y)] = 0\) for any \(x, y\). Using the ladder algebra,

\[
[\phi(x), \phi(y)] = \int \frac{d^d p}{2\omega_p} \left( e^{-ip_\mu(x-y)_\mu} - e^{+ip_\mu(x-y)_\mu} \right) \\
= \pm \int d^{d+1} p \ (2\pi\delta(p^2 - m^2)\theta(p^0) \left( e^{-ip_\mu(x-y)_\mu} - e^{+ip_\mu(x-y)_\mu} \right) \ (2.1) \]

Here comes the slippery stuff: Suppose \(x - y\) is spacelike. Let’s choose a frame where they are at the same time, and let \(\Lambda\) be the Lorentz matrix that gets us there: \(\Lambda_\mu^\nu(x - y)_\nu = (0, \Delta x)_\mu \equiv \tilde{x}_\mu\). Then we can change integration variable to \(\tilde{p}^\mu = \Lambda_\mu^\nu p^\nu\).

\[\phi(x), \phi(y) \sim_0 \int d^{d+1} \tilde{p} \ 2\pi\delta(p^2 - m^2)\theta(p^0) \left( e^{-i\tilde{p}_\mu \Delta x_\mu} - e^{+i\tilde{p}_\mu \Delta x_\mu} \right) = 0. \ (2.2) \]

We conclude that \([\phi(x), \phi(y)] = 0\) if \((x - y)^2 < 0\), i.e. for spacelike separation.

The same argument works for \([\phi, \pi]\) and \([\pi, \pi]\). For \([\phi, \pi]\), the strict inequality is important.

So vanishing ETCR (for nonzero separation) plus Lorentz symmetry implies causality.

Notice that the argument fails if \((x - y)^2 > 0\), since then we can’t get rid of the time component of the exponents by a Lorentz transformation, and they don’t cancel.

Now let’s think more about the bit which is nonzero:

\[\phi(x), \phi(y) = \left[\phi^{(+)}(x), \phi^{(-)}(y) \right] + \left[\phi^{(-)}(x), \phi^{(+)}(y) \right] .
\]

Because \([a, a^\dagger] \propto \mathbb{1}, \hat{\Delta}^\pm\) is a c-number, independent of what state it acts on. So, for any normalized state \(|\psi\rangle\),

\[\Delta^+(x - y) = \langle \psi | \hat{\Delta}^+(x - y) | \psi \rangle = \langle 0 | \hat{\Delta}^+(x - y) | 0 \rangle = \langle 0 | \left[\phi^{(+)}(x), \phi^{(-)}(y) \right] | 0 \rangle = \langle 0 | \phi^{(+)}(x) \phi^{(-)}(y) | 0 \rangle + \langle 0 | \phi^{(-)}(y) \phi^{(+)}(x) | 0 \rangle = 0 ; \phi^+ \ni a, a^\dagger | 0 \rangle = 0 .\]
\[ \langle 0 | \phi(x) \phi(y) | 0 \rangle \equiv D(x - y) \]

where in the last step we again used \( \phi^+ |0\rangle = 0 \). This is the vacuum-to-vacuum amplitude, or propagator, in the sense that

\[ \phi(y) |0\rangle = |\text{particle created at } y\rangle \]

\[ \langle 0 | \phi(x) = |\text{particle destroyed at } x\rangle \]

That is,

\[ \Delta^+(x - y) = \sum_p (x \cdot y - y \cdot x) = \int \frac{d^d p}{2 \omega_p} e^{-i p \cdot (x-y)}. \]

This integral can be done in terms of functions with names\(^\text{14}\), but the most useful information is about its asymptotics in the very timelike (\( t \equiv |x^0 - y^0| \gg |\vec{x} - \vec{y}| \)) and very spacelike (\(|x^0 - y^0| \ll |\vec{x} - \vec{y}| \equiv r \)) limits. You can read about how to arrive at these expressions in Peskin (page 27):

\[ \Delta^+(x - y) \simeq \begin{cases} e^{-imt}, & |x^0 - y^0| \gg |\vec{x} - \vec{y}|, (x - y)^2 \equiv t^2 \\ e^{-mr}, & |x^0 - y^0| \ll |\vec{x} - \vec{y}|, (x - y)^2 \equiv -r^2 \end{cases} \]

Notice that this quantity \( \langle \phi(x) \phi(y) \rangle \) is not zero outside the lightcone. What gives?

Causality only requires the vanishing of commutators outside the lightcone, which we already showed in (2.2). I got mixed up in the book-keeping at this point in lecture. The purpose of the following discussion is to interpret the cancellation in (2.2) as destructive interference between particles and antiparticles. What happened here in terms of physics pictures? It’s clearer for the complex scalar field, where

\[ \Phi^+(x) = \int \frac{d^d p}{\sqrt{2 \omega_p}} e^{-ipx} a_p, \quad \Phi^-(x) = \int \frac{d^d p}{\sqrt{2 \omega_p}} e^{ipx} b_p^\dagger \]

(with the expressions for the + and − frequency components for \( \Phi^* \) following by taking hermitian conjugates). So consider the analogous

\[ D(x - y) \equiv \langle 0 | \Phi(x) \Phi^*(y) | 0 \rangle = \langle 0 | [\Phi^+(x), \Phi^-(y)] | 0 \rangle = \Delta_a^+(x - y) \]

\[ D^*(y - x) \equiv \langle 0 | \Phi^*(y) \Phi(x) | 0 \rangle = \langle 0 | [\Phi^+(y), \Phi^-(x)] | 0 \rangle = -\Delta_b^-(x - y) \]

\(^{14}\)Note from the definition that \( \Delta^\pm(x - y) = [\phi^{(\pm)}(x), \phi^{(\mp)}(y)] = -\Delta^\mp(y - x). \)

\(^{15}\)Specifically, in four spacetime, dimensions and spacelike separation, \((x - y)^2 \equiv -r^2, \Delta^+(x - y) = \frac{m}{2\pi^2} K_1(r). \)
So if we consider the commutator,
\[
\langle 0 | [\Phi(x), \Phi^*(y)] | 0 \rangle = D(x - y) - D^*(y - x)
\]
\[
= \sum_p \left( \begin{array}{c}
\begin{array}{c}
(x \rightarrow -p \rightarrow -y)
\end{array}
\end{array}
\right)_{\text{particle}}
- \left( \begin{array}{c}
\begin{array}{c}
(x \rightarrow -p \rightarrow -y)
\end{array}
\end{array}
\right)_{\text{antiparticle!}}
\]
then in the spacelike case, the antiparticle bit from the first term of the commutator cancels the particle bit of the second, as above in (2.2). Antimatter makes QFT causal.

### 2.2 Propagators, Green’s functions, contour integrals

I claim that the propagator for a real free scalar field can be represented as:

\[
\Delta(x) = \int_C d^{d+1}p \frac{\text{e}^{-ip \cdot x}}{p^2 - m^2} \frac{i}{p^2 - m^2}.
\]

To see that this is related to the object we discussed above, first note that the denominator is

\[
p^2 - m^2 = (p_0 - \omega_{\vec{p}})(p_0 + \omega_{\vec{p}}) = \sqrt{\vec{p} \cdot \vec{p} + m^2},
\]

so there are two poles, which seem to be on the real axis; this means that our integral is ambiguous and we need more information, indeed some physical input.

We can specify the contour \( C \) by taking linear combinations of \( C_{\pm} \) which are small circles going clockwise around the poles at \( \pm \omega_{\vec{p}} \).

These basic integrals are \(^{16}\):

\[
\int_{C_{\pm}} d^{d+1}p \frac{\text{e}^{-ip \cdot x}}{p^2 - m^2} = \int d^d p \frac{1}{2\omega_{\vec{p}}} e^{-ip \cdot x} \big|_{p_0 = \omega_{\vec{p}}} = \Delta^\pm(x).
\]

If we add these up, we get the full propagator:

\[
\Delta(x) = \Delta^+(x) + \Delta^-(x) = \int_{C = C_{\pm}} d^{d+1}p \frac{i}{p^2 - m^2} e^{-ip \cdot x}.
\]

\(^{16}\)We are going to use the Cauchy residue theorem \( \oint_C dz = f(z) = 2\pi i \sum \text{Res}_z f \) where \( z_j \) are the poles of \( f \). To remember the sign, consider a small circle \( C_0 \) counter-clockwise around the origin and \( f(z) = 1/z \), so \( \oint_{C_0} \frac{dz}{z} = 2\pi i \).
This is one particular choice of contour, and others are also interesting. Consider the \textit{retarded} propagator,

\[ D_R(x-y) \equiv \theta(x^0-y^0) \langle 0| [\phi(x), \phi(y)] |0 \rangle. \]

We can reproduce this by routing our contour to go above the poles in the complex \( p^0 \) plane: if \( x^0-y^0 > 0 \), then the factor \( e^{-ip^0(x^0-y^0)} \) decays when \( \text{Im} p^0 < 0 \), so we can close the contour for free in the LHP, and we pick up both poles; if \( x^0-y^0 > 0 \), we must close in the UHP and we pick up no poles and get zero. Notice that we could get the same result by replacing \( p^0 \to p^0 + i\epsilon \) in the denominator, where \( \epsilon \) is an \textit{infinitesimal} (this means that \( \epsilon^2 = 0 \) and \( \epsilon c = \epsilon \) for any positive quantity \( c \)).

Another interesting way to navigate the poles is by replacing \( p^2 - m^2 \) with \( p^2 - m^2 + i\epsilon \). This shifts the poles to

\[ p^0 = \pm \omega_p \sqrt{1 - i\epsilon/\omega_p} = \pm \omega_p (1 - i\epsilon). \]

The \textit{Feynman contour}, \( C_F \), and it seems rather ad hoc; if we had built our understanding of field theory using the path integral, as Zee does, this would have popped out as an inevitable consequence of making the path integral well-defined (see page 20 of the 2d edition). To salvage some self-respect here, consider the \textit{euclidean} propagator, where we get rid of the relative sign in the metric:

\[ I_E(\vec{x}) \equiv \int d^D p \frac{-i}{\sum_{i=1}^D p_i^2 + m^2} e^{-i \sum_{i=1}^D p_i x_i}. \]

Its poles are at \( p^D = \pm i \sqrt{\vec{p} \cdot \vec{p} + m^2} \), far from the integration contour, so there is no problem defining it. Now consider smoothly rotating the contour by varying \( \theta \) from
0 to $\pi/2 - \epsilon$ in $p^0 \equiv e^{i\theta} p^D$. The Feynman contour is the analytic continuation of the euclidean contour, and the $\epsilon$ is the memory of this.

In position space, the Feynman propagator is

$$\Delta_F(x) \equiv \int_{C_F} d^{d+1}p \frac{e^{-ip_x x}}{p^2 - m^2} = \theta(x^0) \Delta^+(x) + \theta(-x^0) \Delta^+(x) = \theta(x^0) \Delta^+(x) - \theta(-x^0) \Delta^-(x).$$

If $x^0 > 0$ ($< 0$), we must close the contour in the LHP (UHP) and get $\Delta^+$ ($\Delta^-$).

Recalling that $\Delta^+(x-y) = \langle 0 | \phi(x) \phi(y) | 0 \rangle$, 

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

The $T$ is the time-ordering symbol: the operators at the earliest times go on the right, so we can regard time as increasing to the left.

**The propagator is a Green’s function.** So we’ve learned that

$$\frac{i}{p^2 - m^2} \equiv \tilde{\Delta}(p)$$

is the Fourier transform of $\Delta(x)$, the momentum-space propagator (either retarded or Feynman). From this we can see that $\Delta(x)$ is a Green’s function for the differential operator $\partial_{\mu} \partial^{\mu} + m^2$ in the sense that

$$(\partial_{\mu} \partial^{\mu} + m^2) \Delta(x) = -i\delta(x)$$

(by plugging in the Fourier expansion of $\Delta$ and of the delta function, $\delta^{(d+1)}(x) = \int d^{d+1}p \ e^{-ipx}$, and differentiating under the integral). Notice that this did not depend on the choice of contour, so this equation in fact has several solutions, differing by the routes around the poles (hence by $\Delta^\pm$, which are solutions to the homogeneous equation, without the delta function). On the homework, you will show this directly from the position space definition.

[End of Lecture 6]

**Physics preview.** Here is a preview of the physics of the Feynman propagator. Imagine we introduce some interactions, such as a cubic term in the Lagrangian, e.g.

$$\mathcal{L} \ni \phi_p(x) \phi_n(x) \phi_{\pi}(x) + \text{h.c.}$$

(2.3)

where the fields appearing here destroy or create particles with the names in the subscripts. Here are two stories we might tell about a collection of such particles.
In both pictures, time goes to the left. In the first picture, a $\Delta^-$ emits a $\pi^-$, becoming a $\Delta^0$ at spacetime point $P$. This $\pi^-$ propagates to $Q$ where it is absorbed by a $p$, which turns into an $n$. In the second picture, a $p$ emits a $\pi^+$ at $Q$, and becomes $n$; that $\pi^+$ is absorbed by a $\Delta^-$ which becomes a $\Delta^0$.

But these two stories are related by a Lorentz boost which exchanges the relative times of the interaction events – they must be the same story. The Feynman propagator includes both automatically.

**Antiparticles going backwards in time.** This story is clearer if we discuss the complex scalar, where particles (created by $a^\dagger$) and antiparticles (created by $b^\dagger$) are distinct:

$$\Phi(x) = \int \frac{d^dp}{\sqrt{2\omega_p}} a_p e^{-ipx} + \int \frac{d^dp}{\sqrt{2\omega_p}} b^\dagger_p e^{ipx}.$$  

The commutator is

$$[\Phi(x), \Phi^*(y)] = [\Phi^+(x), \Phi^{*-}(y)] + [\Phi^-(x), \Phi^{*+}(y)]$$

from $[a, a^\dagger]$, particles from $[b, b^\dagger]$, antiparticles

$$= \Delta^+_a(x-y) + \Delta^*_b(x-y)$$

$$= \int_{C_+} d^{d+1} e^{-ip(x-y)} \frac{i}{p^2 - m^2} + \int_{C_-} d^{d+1} e^{-ip(x-y)} \frac{i}{p^2 - m^2}.$$  

The propagator that we'll really need to compute $S$-matrix elements is

$$\Delta_F(x-y) \equiv \langle 0 | T(\Phi(x)\Phi^*(y)) | 0 \rangle$$

$$= \theta(x^0 - y^0) \langle 0 | \Phi(x)\Phi^*(y) | 0 \rangle$$

$$= \langle 0 | \Phi^+(x)\Phi^{*-}(y) | 0 \rangle = \langle 0 | [\Phi^+(x), \Phi^{*-} (y)] | 0 \rangle = \Delta^+_a(x-y), \text{ particles}$$
\[
\theta(y^0 - x^0) \langle 0 | \Phi^*(y) \Phi(x) | 0 \rangle
\]
\[
= -\langle 0 | \Phi^+(y) \Phi^-(x) | 0 \rangle = -\Delta^+(x-y), \quad \text{antiparticles}
\]
2.3 Interlude: where is 1-particle quantum mechanics?

[Tong, §2.8] Consider a relativistic complex free massive scalar field $\Phi$, with mass $m$. The minimum energy of a single-particle state is $\omega_{p=0} = m$ (above the vacuum); in its rest frame, the wavefunction is $e^{-imt}$. Consider the change of variables:

$$\Phi(\vec{x}, t) = e^{-imt} \Psi(\vec{x}, t) \sqrt{2m}.$$

The Klein-Gordon equation is

$$0 = \partial_t^2 \Phi - \nabla^2 \Phi + m^2 \Phi = \sqrt{2m} e^{-imt} \left( \ddot{\Psi} - 2im \dot{\Psi} - \nabla^2 \Psi \right),$$

where the terms with $m^2$ cancel; so far we've just changed variables. The non-relativistic limit is $|\vec{p}| \ll m^2$ which on-shell, $\omega_p = \sqrt{\vec{p}^2 + m^2} \simeq m + \frac{\vec{p}^2}{2m} + \cdots$, implies that $|\dot{\Psi}| \ll m|\ddot{\Psi}|$ so we can ignore that term in the KG equation, leaving

$$i\partial_t \Psi = -\frac{1}{2m} \nabla^2 \Psi. \quad (2.5)$$

This looks like the Schrödinger equation for a particle in no potential, in position space, but that is an illusion: $\Psi^* \Psi \propto \Phi^* \Phi$ need not be positive, hence is not a probability density. This equation (2.5) is the eom for the lagrange density

$$\mathcal{L} = i \Psi^* \dot{\Psi} - \frac{1}{2m} \nabla \Psi^* \cdot \nabla \Psi$$

from which $\pi_{\Psi} = i \Psi^*$, $\pi_{\Psi^*} = 0$. The ETCRs are then

$$[\Psi(\vec{x}), \Psi^*(\vec{y})] = \delta^d(\vec{x} - \vec{y}), \quad [\Psi, \Psi] = 0 = [\Psi^*, \Psi^*]$$

and the Hamiltonian is

$$H = \int d^d x \frac{1}{2m} \nabla \Psi^* \cdot \nabla \Psi.$$

The solution in terms of creation operators is then

$$\Psi(x) = \int d^d p e^{i\vec{p} \cdot \vec{x}} a_p, \quad \Psi^*(x) = \int d^d p e^{-i\vec{p} \cdot \vec{x}} a_p^\dagger$$

with $[a_p, a_q^\dagger] = (2\pi)^d \delta^d(p - q)$ as before. The hamiltonian is then

$$H = \int d^d p \frac{\vec{p}^2}{2m} a_p^\dagger a_p \quad (2.6)$$

(with no normal-ordering constant – the vacuum of this non-relativistic theory is extremely boring and has no dance of birth and death of the fluctuating particle-antiparticle pairs).
The crucial point here is that the antiparticles are gone, despite the fact that the field is complex. In the limit we’ve taken, we don’t have enough energy to make any. The states are
\[ a_{p_1}^\dagger \cdots a_{p_n}^\dagger |0\rangle \equiv |\{p\}\rangle, \quad a_p |0\rangle = 0 \]
and these are energy eigenstates with energy \( H |\{p\}\rangle = \sum_i \frac{p^2}{2m} |\{p\}\rangle \), a nice non-relativistic (NR) dispersion for each particle. Notice that the NR limit required a complex field (otherwise we can’t multiply by a phase). The particle-number symmetry is still present \( \Psi \to e^{-i\alpha}\Psi \), but now the current is
\[ j^\mu = \left( \Psi^* \Psi, \frac{i}{2m} \Psi^* \nabla \Psi + \text{h.c.} \right)^\mu = (\rho, j) \]

Now we can find the QM of a single particle which cannot go away (since we got rid of the antiparticles), with some position and momentum operators. In fact the momentum operator is just the charge associated with translation invariance, and takes the form (just like on the homework)
\[ \vec{P} = \int d^d p \vec{p} a_p^\dagger a_p \]
and \( \vec{P} |\{p\}\rangle = \sum_a \vec{p}_a |\{p\}\rangle \). What’s the position operator? A state with a particle at position \( \vec{x} \) is
\[ |\vec{x}\rangle = \Psi^*(x) |0\rangle = \int d^d p e^{-i\vec{p} \cdot \vec{x}} a_p^\dagger |0\rangle. \]
If we let
\[ \vec{X} \equiv \int d^d x \Psi^*(x) \vec{x} \Psi(x) \]
then indeed \( \vec{X} |\vec{x}\rangle = \vec{x} |\vec{x}\rangle \). To see that the Heisenberg algebra \([\vec{X}, \vec{P}] = \vec{i}\) works out, consider the general 1-particle state
\[ |\psi\rangle = \int d^d x \psi(x) |x\rangle. \]
The function \( \psi(x) \) here is the usual position-basis Schrödinger wavefunction. You can check on the homework that
\[ \vec{X}^i |\psi\rangle = \int d^d x x^i \psi(x) |x\rangle, \quad \vec{P}^i |\psi\rangle = \int d^d x \left( -i \frac{\partial}{\partial x^i} \psi(x) \right) |x\rangle \]
which implies the Heisenberg commutator. Finally, the hamiltonian (2.6) gives the time evolution equation
\[ i\partial_t \psi = -\frac{\nabla^2}{2m} \psi \]
which really is the Schrödinger equation.

Many particles which one studies in NR QM are actually fermions (e, p, n...) and therefore not described by a scalar field. But in the 1-particle sector, who can tell? No one. Later we’ll see the NR limit of the Dirac equation, which is basically the same, but with some extra juicy information about spin.

Next we will speak about ‘interactions’. This term is used in two ways. In NR QM, it is sometimes used to describe an external potential $V(x)$ appearing as an extra term in the Schrödinger equation

$$i\partial_t \psi = -\frac{\nabla^2}{2m}\psi + V(x)\psi(x).$$

Such a term explicitly violates translation symmetry. It can be accomplished by adding to the action the quadratic term

$$\Delta S_V = -\int d^d x \Psi^*(x) \Psi(x) V(x) = -\int d^d x \rho(x) V(x).$$

This says that nonzero density of particles at $x$ costs energy $V(x)$. A second sense of ‘interaction’ which is how it will be used forever below is interaction between particles. With only one particle this cannot happen. NR QM theory does accommodate more than one particle, and we can consider an interaction between them like

$$\Delta S = -\int d^d x \int d^d y \Psi^*(x) \Psi(x) V(x-y) \Psi^*(y) \Psi(y).$$

If $V(x-y) = \delta^d(x-y)$, this interaction is local.
3 Interactions, not too strong

3.1 Where does the time dependence go?

[Peskin chapter 4.2] Now we must get straight where to put the time dependence. Different ways of doing the book-keeping are called different ‘pictures’. At some reference time, say \( t = 0 \), they all agree: label states by \( |\psi, 0\rangle \) and operators by \( \mathcal{O}(0) \). At a later time, in picture \( P \), these evolve to \( |\psi, t\rangle_P, \mathcal{O}_P(t) \). Physics, such as any amplitude like

\[
P \langle \psi, t | \mathcal{O}_P(t) | \psi, t \rangle_P
\]

is independent of \( P \). Let’s assume the hamiltonian \( H \) has no explicit time dependence.

In Heisenberg picture \( (P = H) \), \( |\psi, t\rangle_H \equiv |\psi, 0\rangle \) for all time, and the burden of the time dependence is all on the operators

\[
\mathcal{O}_H(t) = U_H^\dagger(t) \mathcal{O}(0) U_H(t).
\]

The Heisenberg equations of motion are

\[
\text{i} \dot{\mathcal{O}}_H = [\mathcal{O}_H, H_H]
\]
so in particular \( \dot{H}_H = 0 \) so \( H_H(t) = H_H(0) = H \). Then (3.2) is solved by \( U_H(t) = e^{-iHt} \).

For example, another example of an operator is the free field:

\[
\Phi_H(\vec{x}, t) = \int \frac{d^d p}{\sqrt{2\omega_p}} \left( a_p e^{-ipx} + b_p^\dagger e^{ipx} \right).
\]

In fact, this equation is basically the whole story of free field theory. The field makes particles which don’t care about each other.

In Schrödinger picture \( (P = S) \), \( \frac{d}{dt} \mathcal{O}_S = \partial_t \mathcal{O}_S \) time dependence of operators comes only from explicit, external dependence in the definition of the operator (which will not happen here), so \( \mathcal{O}_S(t) = \mathcal{O}(0) \), and (3.1) then requires

\[
|\psi, t\rangle_S = U_H(t) |\psi, 0\rangle.
\]

And the unitary evolution operator is

\[
U_H(t) = e^{-iH(0)t} = e^{-iH_{st}t} = U_S = U
\]
so does not require a picture label.

**Interactions.** So, in an interacting field theory, all we need to do is to find \( U \) to figure out what it does. For better or worse, this isn’t a realistic goal in general. In
this class we are going to focus on the special case where the interactions are weak, so that the hamiltonian takes the form

\[ H = H_0 + V \]

where \( H_0 \) is quadratic in fields (linear terms are allowed but annoying) and we assume that the interaction term \( V \) is proportional to a small parameter. This by no means exhausts all interesting questions in field theory; on the other hand, a surprisingly enormous amount of physics can be done using this assumption.

**Interaction picture.** \((P = I)\) In this case, it is very convenient to use a hybrid picture where the time-dependence of the operators is as in the Heisenberg picture for the hamiltonian with \( V \to 0 \). This free field evolution is solvable:

\[ O_I(t) \equiv U_0^\dagger O(0) U_0, \quad U_0(t) \equiv e^{-iH_0 t}. \]

Note that in this picture, \( H_0(t) = H_0(0) = H_0 \). Equivalently, \( i\dot{O}_I = [O_I, H_0] \), where in this expression, crucially, \( H_0 = H_0(\Phi_I) \) is made from interaction picture fields, whose evolution we know from above; for example, for a complex scalar,

\[ \Phi_I(\vec{x}, t) = \int \frac{d^d p}{\sqrt{2\omega_p}} \left(a_p e^{-ipx} + b_p^\dagger e^{ipx}\right). \]

The catch is that the interaction-picture states are still time-dependent:

\[ H\langle \varphi, t | O_H(t) | \psi, t \rangle_H \left(\overset{(3.1)\dagger}{=} i \langle \varphi, t | O_I(t) \right| \overset{=}{} = U_0^\dagger O(0) U_0(t) \mid \psi, 0 \rangle = U_0^\dagger U_I(t) O(0) U_0(t) \mid \psi, 0 \rangle. \]

\( \forall \varphi, \psi \) which says that

\[ \langle \psi, t | I = U_0^\dagger(t) U_H(t) | \psi, 0 \rangle \equiv U_I(t) | \psi, 0 \rangle. \]

In the interaction picture, the interaction hamiltonian itself evolves according to

\[ i \frac{d}{dt} V_I = [V_I, H_0] \implies V_I(t) = U_0^\dagger V(0) U_0. \]

So for example, if \( V(0) = \int d^d x g \phi^3(x, 0) \), then using \( 1 = U_0^\dagger U_0 \) repeatedly,

\[ U_0^\dagger V(0) U_0 = g \int d^d x U_0^\dagger \phi^3(x, 0) U_0 = g \int d^d x U_0^\dagger U_0^\dagger \phi(x, 0) U_0 U_0^\dagger \phi(x, 0) U_0 = g \int d^d x (\phi_I(x, t))^3. \]

This trick wasn’t special to \( \phi^3 \) and works for any local interaction:

\[ (V(t))_I = V|_{t=0} (\phi_I(t)) \]
– just stick the interaction-picture-evolved fields into the form of the interaction at
\( t = 0 \), easy.

How do the states evolve? Notice that \([U_0^\dagger, U_H] \neq 0\), if the interactions are interesting. So

\[
\partial_t |\psi, t\rangle_I = \partial_t (U_I(t) |\psi, 0\rangle) = \partial_t \begin{pmatrix} U_0^\dagger(t) & U_H(t) \\ e^{-iH(0)t} & e^{-i(H_0+V)t} \end{pmatrix} |\psi, 0\rangle
\]

\[
= U_0^\dagger \begin{pmatrix} iH_0 - iH \\ -iV(0) \end{pmatrix} U_H |\psi, 0\rangle
\]

\[
= -iU_0^\dagger V(0) U_H U_0 |\psi, t\rangle_I = -iU_0^\dagger V(0) |\psi, t\rangle_I.
\] (3.4)

That is

\[
i\partial_t |\psi, t\rangle_I = V(t) |\psi, t\rangle_I.
\]

Alternatively, the interaction-picture evolution operator satisfies

\[
i\partial_t U_I(t) = V(t) U_I(t).
\]

Notice how this differs from the Heisenberg evolution equation (3.2): although the full
\( H \) is time-independent, \( V_I(t) \) actually does depend on \( t \), so \([V(t), V(t')] \neq 0\), and so the
solution is not just a simple exponential. We’ll find a nice packaging for the solution
next in the form of Dyson’s expansion.

Peskin’s notation for this object is \( U_I(t) = U(t, t_0)|_{t_0=0} \). We can figure out how to
change the reference time from zero as follows:

\[
|\psi, t\rangle_I = U_I(t) |\psi, 0\rangle, \quad |\psi, t'\rangle_I = U_I(t') |\psi, 0\rangle \implies |\psi, 0\rangle = U_I^\dagger(t) |\psi, t'\rangle
\]

\[
|\psi, t\rangle_I = \underbrace{U_I(t) U_I^\dagger(t')}_{=U(t,t')} |\psi, t'\rangle_I
\]

From which we infer that

\[
U(t, t') = U_0^\dagger(t) U_H(t) U_H^\dagger(t') U_0(t') = e^{iH_0 t} e^{-iH(t-t')} e^{-H_0 t'}.
\]

From now on we drop the \( P = I \) subscripts: everything is \( I \).

**Definition of S-matrix.** What are we going to do with the evolution operator? Here is a basic (only slightly naive) three-step framework for doing particle physics (not yet for making predictions).
• At time \( t_i \), specify (e.g., measure) all the particle types, spins, momenta in the form of an initial state \(|i\rangle\) in the QFT Hilbert space.

• Wait. At time \( t \), the state is

\[
U(t, t_i) |i\rangle = |\psi, t\rangle .
\]

• At time \( t_f \), measure all the particle data and get some state \(|f\rangle\).

Quantum mechanics says that the probability for this outcome is

\[
|\langle f | U(t_f, t_i) |i\rangle|^2 .
\]

(One way in which this is significantly naive is that the space of outcomes is continuous, so we must instead make probability distributions. More soon.) Because particle interactions are a fast-moving business, a useful idealization is to take \( t_i \to -\infty \) and \( t_f \to \infty \), and let

\[
S_{fi} \equiv \langle f | \hat{S} |i\rangle , \quad \hat{S} \equiv U(\infty, -\infty)
\]

the \( S \)-matrix (‘\( S \)’ is for ‘scattering’).

This has only three ingredients: initial state, final state, and time evolution operator. Let’s focus on the last one:

**Dyson expansion.** We need to solve the equation

\[
\partial_t |\psi, t\rangle = -iV(t) |\psi, t\rangle , \quad \text{with initial condition} \quad |\psi, t_i\rangle = |i\rangle .
\]

Here’s a “solution”:

\[
|\psi, t\rangle = |i\rangle + (-i) \int_{t_i}^t dt_1 V(t_1) |\psi, t_1\rangle .
\]

The only small problem is that we don’t know \(|\psi, t_1\rangle\). But we can use this expression for that too:

\[
|\psi, t\rangle = |i\rangle + (-i) \int_{t_i}^t dt_1 V(t_1) \left( |i\rangle + (-i) \int_{t_i}^{t_1} dt_2 V(t_2) \right)
\]

\[
= |i\rangle + (-i) \int_{t_i}^t dt_1 V(t_1) |i\rangle + (-i)^2 \int_{t_i}^t dt_1 \int_{t_i}^{t_1} dt_2 V(t_1)V(t_2) |\psi, t_2\rangle
\]

Why stop there? Two comments: (1) This is a good idea when \( V \propto \lambda \ll 1 \). (2) Notice the time-ordering: the range of integration restricts \( t_1 \geq t_2 \), and the earlier operator \( V(t_2) \) is to the right. The ad absurdum limit is

\[
|\psi, t\rangle = \sum_{n=0}^\infty (-i)^n \int_{t_i}^t dt_1 \int_{t_i}^{t_1} dt_2 \cdots \int_{t_i}^{t_{n-1}} dt_n V(t_1)V(t_2) \cdots V(t_n) |i\rangle = U(t, t_i) |i\rangle
\]

which since this is true for any \(|i\rangle\) tells us a formula for \( U \).
To review, the equation we are trying to solve is:

\[
\frac{i\partial}{\partial t} |\psi, t\rangle = V(t) |\psi, t\rangle = i\partial_t U U^\dagger |\psi, t\rangle.
\]

This is true for all $|\psi, t\rangle$, so it means $i\partial_t U U^\dagger = V$. Multiplying the BHS on the right by $U$ gives

\[\Rightarrow \partial_t U = -i V U.\]

We might expect that an equation like this has a solution which is something like $U \sim e^{-i V t}$.

Now we must deal with what Lawrence Hall calls "the wretched $n$". Starting from our series solution (3.5),

\[
U(t, t_i) = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{t_i}^{t} dt_1 \int_{t_i}^{t_1} dt_2 \cdots \int_{t_i}^{t_{n-1}} dt_n V(t_1) V(t_2) \cdots V(t_n)
\]

\[= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{t_i}^{t} dt_1 \int_{t_i}^{t_1} dt_2 \cdots \int_{t_i}^{t_{n-1}} dt_n T(V(t_1)V(t_2) \cdots V(t_n))
\]

\[= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{t_i}^{t} dt_1 \int_{t_i}^{t_1} dt_2 \cdots \int_{t_i}^{t_{n-1}} dt_n T(V(t_1)V(t_2) \cdots V(t_n))
\]

(3.6)

In the first step I used the fact that the operators are already time ordered (this followed from the differential equation we are solving, since the $V$ always acts from the left). In the second step we used the fact that the time-ordered integrand doesn’t change if we permute the labels on the times. So we can just average over the $n!$ possible orderings of $n$ times. If we pull out the time-ordering symbol, this is an exponential series:

\[U(t, t_i) = T \left( e^{-i \int_{t_i}^{t} dt V(t')} \right).\]

The time-ordered exponential is defined by its Taylor expansion.

### 3.2 $S$-matrix

Taking the times to $\pm \infty$ in the previous equation gives an expression for the $S$-matrix:

\[\hat{S} = U(-\infty, \infty) = T \left( e^{-i \int_{-\infty}^{\infty} dt V(t)} \right).\]

(3.7)

The practical value of these expressions is that they give a (compact) recipe for evaluating the time evolution operator as a series in powers of the small parameter in front of $V(0)$: we know $V(t)$ in terms of things like $a, a^\dagger$, can pull them down term-by-term.
I should have called the previous expression the ‘$S$-operator’, since the thing we are after is really the $S$-matrix elements, $\langle f | \hat{S} | i \rangle$, for which we still need $\langle i |$ and $| f \rangle$.

Here we encounter some small trouble. Can we just use the states like $\sqrt{2\omega_p}a_p^\dagger |0\rangle$ (the eigenstates of the free hamiltonian) which we’ve grown to love? In fact, even the vacuum $|0\rangle$ is not an eigenstate of the actual $H_0 + V$ (since $[H_0, V] \neq 0$), so it will not stay where we put it. The vacuum of the interacting theory $|\Omega\rangle$ is itself an object of mystery (a boiling sea of virtual particles and antiparticles), and the stationary excited states are too (a particle carries with it its disturbance of the vacuum). We’ll learn to deal with this in perturbation theory, but here’s an expedient: pick a function $f(t)$ which is zero at one end, one in the middle, and then zero again at the far end. Now replace the interaction hamiltonian $V$ with $f(t)V(t)$. Then, if we take $t_i <$ the time before which the interaction turns on, and $t_f >$ the time after which we turn it off, then we can use the free hamiltonian eigenstates. This is in fact wrong in detail, but it will get us started.

**Example.** Let’s return to the ‘scalar Yukawa theory’ that we briefly encountered earlier in (2.3). Simplifying the notation a bit, the whole Lagrangian density is

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

(3.8)

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

The mode expansions are

$$\phi = \int \frac{d^4p}{\sqrt{2\omega_p}} \left( a_p e^{-ipx} + a_p^\dagger e^{ipx} \right) |p^\rho = \omega_p\rangle$$

$$\Phi = \int \frac{d^4p}{\sqrt{2E_p}} \left( b_p e^{-ipx} + c_p^\dagger e^{ipx} \right) |p^\rho = E_p\rangle$$

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

$$q = N_c - N_b.$$  

But the $\phi$ particles are not conserved.$^{17}$

**Artisanal meson decay.** [Tong §3.2.1] Consider

$$|i\rangle = \sqrt{2\omega_p}a_p^\dagger |0\rangle, \quad |f\rangle = \sqrt{2E_{q_1}2E_{q_2}}b_{q_1}^\dagger c_{q_2}^\dagger |0\rangle.$$  

$^{17}$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 $S$-matrix element between these states is
\[
\langle f|\hat{S}|i\rangle = \langle f| \left( 1 + (-i) \int d^{d+1}x g\phi_x \Phi_x^* \Phi_x + O(g^2) \right) |i\rangle
\]
where all operators are in interaction picture. The first term dies because $\langle f|i\rangle = 0$. The $\phi \sim a + a^\dagger$ takes a one-particle state into a superposition of states with zero and two $\phi$-particles. We need to end up with zero $\phi$-particles. The leading-order nonzero term is
\[
= -ig \int d^{d+1}x e^{-ipx} \langle 0| b_{q_1} c_{q_2} \sqrt{4E_{q_1} E_{q_2}} \int \frac{d^d k_1}{\sqrt{2E_{k_1}}} e^{ik_1 x} b_{k_1}^\dagger \int \frac{d^d k_2}{\sqrt{2E_{k_2}}} e^{ik_2 x} c_{k_1}^\dagger |0\rangle
\]
This is a small victory. This delta function imposes conservation of energy and momentum on the transition amplitude. In the $\phi$ rest frame, $p^\mu = (M, 0)$ which says the amplitude is only nonzero when $q_1 = -q_2$ and when $M = 2\sqrt{|q_1|^2 + m^2}$. Notice that this can only happen if $M \geq 2m$.

How do we get from this amplitude to a probability? We have to square it:
\[
P_{fi} \sim |S_{fi}|^2 = g^2 \left( \delta^{d+1}(p_f - p_i) \right)^2.
\]
The square of a delta function is infinity. What did we do wrong?

[End of Lecture 8]

Not so much, we just asked a dumb question. Here is where it helps to be a physicist. Consider:
\[
\left( \delta^{d+1}(p) \right)^2 = \delta^{d+1}(p)\delta^{d+1}(0) = \delta^{d+1}(p) \int d^{d+1}x e^{i0x} = \delta^{d+1}(p)VT
\]
where $VT$ is the volume of spacetime – the size of the box times how long we’re willing to wait. There is a nonzero probability per unit time per unit volume that a $\phi$ particle in a plane wave state will decay. We’ll get its lifetime out momentarily.

For more complicated examples, it will help to streamline this process, which is the job of §3.3.

### 3.3 Time-ordered equals normal-ordered plus contractions

We have an expression for $\hat{S}$ in (3.7) involving only time-ordered products of operators. If we stick this between states with just a few particles, the annihilation operators in
there very much want to move to the right so they can get at the vacuum and annihilate it, as is their wont. Wick’s theorem tells us how to do this, along the following lines:

\[
\text{Wick: } \quad \mathcal{T}(\phi \ldots \phi) = :\phi \ldots \phi: + ?
\]

In the previous schematic non-equation, I introduced a notation for a normal-ordered product \( :\phi \cdots \phi:\) which means each term has all the annihilation operators to the right of all the creation operators, for example

\[
: \phi(x)\phi(y) : = \phi^{-}(x)\phi^{+}(y) + \phi^{-}(y)\phi^{+}(x) + \phi^{+}(x)\phi^{+}(y) + \phi^{-}(x)\phi^{-}(y).
\]  

This operation is by definition linear.

**Normal-ordering difficulty.** Actually, there is a trap here which was brought to my attention by Chuncheong Lam: If we want the normal-ordering operation to be linear, then we must have

\[
: (a^\dagger a) : = : aa^\dagger : - : 1 :.
\]

But \( \langle 0 | : a^\dagger a : |0 \rangle = \langle 0 | : aa^\dagger : |0 \rangle = 0 \) so we must have \( \langle 0 | : 1 : |0 \rangle = 0 \), which means we must have the shocking-looking equation:

\[
: 1 : = 0
\]

that is: the normal-ordered product of a c-number must be zero. This definition (which, beware, differs from Peskin’s) has the advantage that my statement below that the vacuum expectation value (VEV) of any normal ordered product is zero (with no exceptions for c-numbers). The price is that we cannot put the normal-ordering symbol around the c-number bits, as Peskin does.

More generally, let \( A, B, C \) be the positive- and negative-frequency bits of some fields. Then

\[
: ABC \cdots : = : A'B'C' \cdots : (\text{only } a^\dagger \text{s} \quad \text{only } a \text{s})
\]

Peskin writes \( N(\cdots) : = \cdots : \). Notice that \( \langle 0 | : \text{anything} : |0 \rangle = 0 \).
A comment about fermions. Later we will use anticommuting operators, which have
\[
\{c_k, c_p^\dagger\} = \delta(k - p), \quad \{c_k^\dagger, c_p^\dagger\} = 0.
\]
In particular, the equation \((c_p^\dagger)^2 = 0\) is an algebraic realization of the Pauli principle. The cost is that even the \(\phi^-\) bits generate signs when they move through each other. In that case, we define the normal ordered product as
\[
: ABC \cdots : \equiv \left( \begin{array}{c}
A'B'C' \cdots \\
\text{only } a^\dagger \text{ as}
\end{array} \right) (-1)^P
\]
where \(P\) is the number of fermion interchanges required to get from \(ABC \cdots\) to \(A'B'C' \cdots\). Keeping track of these signs, and replacing commutators with anticommutators, everything below goes through for fermion fields.

Let’s go back to (3.9). Because \([\phi^\pm, \phi^\pm] = 0\), the order in the last two terms doesn’t matter. This can differ from the time-ordered product only in the first or second term. If \(y^0 > x^0\), it differs by \([\phi^-(x), \phi^+(y)] = -\Delta^+(x - y)\), and if \(x^0 > y^0\), it differs by \([\phi^-(y), \phi^+(x)] = +\Delta^+(x - y)\). Altogether:
\[
:\phi(x)\phi(y) := T(\phi(x)\phi(y)) - \Delta_F(x - y) \equiv T(\phi(x)\phi(y)) - \sqrt{1}(\phi(x)\phi(y)). \tag{3.10}
\]

More generally, writing \(\phi_a \equiv \phi_a(x_a)\), Wick’s theorem says
\[
:\phi_1 \cdots \phi_n := T(\phi_1 \cdots \phi_n) + (\text{all contractions})
\]
where a contraction is defined as the price for moving a pair of operators through each other to repair the time ordering, as in (3.10), and denoted by the symbol in (3.10).

For example, for four fields, the theorem says
\[
T(\phi_1 \cdots \phi_4) =: \left(\phi_1 \cdots \phi_4 + (\sqrt{1} \phi_1\phi_2\phi_3\phi_4 + 5 \text{ more})\right) + (\sqrt{1} \phi_1\phi_2\phi_3\phi_4 + 2 \text{ more})
\]
Notes: The fully-contracted bits are numbers, so they are outside the normal-ordering symbol. For a product of \(n\) fields, there are \(\binom{n}{2} + \binom{n}{4} + \binom{n}{6} + \binom{4}{2} \cdots + \binom{n}{\lfloor n/2 \rfloor} (= \text{many})\) contractions. But if we take the vacuum expectation value (VEV) of the BHS, most terms go away.
Here’s the idea of the proof (Peskin page 90), which is by induction on the number of fields \(m\) in the product. We showed \(m = 2\) above. Assume WLOG that \(x_1^0 \geq \cdots \geq x_m^0\), or else relabel so that this is the case. Wick for \(\phi_2 \cdots \phi_m\) says

\[
\mathcal{T} (\phi_1 \cdots \phi_m) = \underbrace{\phi_1}_{=\phi_1^+ + \phi_1^-} : \phi_2 \cdots \phi_m : + (\text{all contractions w/o } \phi_1 )
\]

The \(\phi_1^-\) term is already in the right place and can slip for free inside the normal-ordering sign. The \(\phi_1^+\) needs to move past all the uncontracted \(\phi_{j \geq 2}\); this process will add a term for every possible contraction involving \(\phi_1\).

### 3.4 Time-ordered correlation functions

Time-ordered correlation (or Green’s) functions of local operators will be useful:

\[
G^{(n)}(x_1 \cdots x_n) \equiv \langle \Omega | \mathcal{T} (\phi_1^H(x_1) \cdots \phi_n^H(x_n)) | \Omega \rangle .
\]

Here, the operators are in Heisenberg picture for the full hamiltonian, and \(\Omega\) is its actual lowest-energy eigenstate, \(H | \Omega \rangle = E_0 | \Omega \rangle\). The fourier transform is also useful:

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

In the free theory of a real scalar, we know something about these:

\[
G^{(2)}_{\text{free}}(x_1, x_2) = \Delta_F(x_1 - x_2) = \frac{i}{p^2 - m^2 + i\epsilon} = \frac{1}{(p_1 + p_2) \cdot \epsilon} .
\]  

The higher correlations are Gaussian, in the sense that they are sums of products of the two point functions:

\[
G^{(4)}_{\text{free}}(x_1 \cdots x_4) = \Delta_F(12)\Delta_F(34) + \Delta_F(13)\Delta_F(24) + \Delta_F(14)\Delta_F(23) = \frac{1}{(p_1 + p_2) \cdot \epsilon} .
\]  

**Expectations.** Our next goal is to construct a perturbative expansion in the case of \(V = \int d^d z \frac{\lambda}{4!} \phi^4(z)\). We expect a correction of order \(\lambda\) of the form: \(\chi^* \chi^* \chi \chi\). In momentum space, we have blobs (unspecified sums of diagrams) with external lines
labelled by $p_i^\mu$: Notice that there is no need to restrict their values to the mass shell $p_i^2 = m^2$, that is, $\tilde{G}(p)$ is nonzero even when $p_i$ are off-shell: these Green’s functions contain “off-shell” information, more information than is available in just the scattering matrix. However, something special will happen when the external legs are on-shell. As you can see from the free two-point function, (3.11), they blow up on the mass-shell. The existence of a singularity of $\tilde{G}$ on the mass-shell is a general fact, and their residues give the $S$-matrix elements:

$$\tilde{G}(p_1 \cdots p_n) \xrightarrow{p_i^2 \to m_i^2} \prod_i \frac{i}{p_i^2 - m_i^2 + i\epsilon} S(p_1 \cdots p_n).$$

(This is the LSZ theorem, about which more later.)

**Perturbative expansion of time-ordered correlators.** We’ll do this in three steps: (1) Relate $|\Omega\rangle$ to $|0\rangle$. (2) Relate $\phi_H$ to $\phi_I$. (3) Wick expand and organize the diagrams.

**Step (1):** [Peskin page 86-87] Some preparations:

- Fix the additive normalization of the hamiltonian by $H_0 |0\rangle = 0$.
- Label the spectrum of $H$ by $|n\rangle$, so $1 = \sum_n |n\rangle \langle n|$. This is a very scary sum over the whole QFT Hilbert space, really an integral.
- Assume that $\langle \Omega |0\rangle \neq 0$. A necessary condition for this is that the actual Hamiltonian is in the same phase as the $H_0$. Also, let’s keep the volume finite for awhile.

Now consider

$$\langle \Omega | e^{-iHT} = \sum_n \langle 0|n\rangle \langle n| e^{-iHT} = \sum_{n \neq \Omega} \langle 0|n\rangle \langle n| e^{-iE_nT} + \langle 0|\Omega\rangle \langle \Omega| e^{-iE_0T}.$$ Since $E_0 < E_n$ for all other $n$, by given $T$ a large negative imaginary part, $T \to \infty(1-\text{i}\epsilon)$ we can make the contribution of $\Omega$ arbitrarily larger than the others. Multiplying by $e^{iE_0T}/\langle 0|\Omega\rangle$ gives

$$\langle \Omega \rangle = \lim_{T \to \infty(1-\text{i}\epsilon)} \frac{\langle 0| e^{-iHT} e^{iE_0T} \rangle}{\langle 0|\Omega\rangle} = \lim_{T \to \infty(1-\text{i}\epsilon)} \frac{\langle 0| e^{iE_0T} e^{-iHT} e^{iE_0T} \rangle}{\langle 0|\Omega\rangle}.$$ Since $T$ is infinite anyway, we can shift it to $T \to T - t_0$ without change:

$$\langle \Omega \rangle = \lim_{T \to \infty(1-\text{i}\epsilon)} \frac{\langle 0| e^{iH_0(T-t_0)} e^{-iH(T-t_0)} e^{iE_0(T-t_0)} \rangle}{\langle 0|\Omega\rangle} = \lim_{T \to \infty(1-\text{i}\epsilon)} \frac{\langle 0| U_I(t, t_0) e^{iE_0(T-t_0)} \rangle}{\langle 0|\Omega\rangle}. \quad (3.13)$$
Do the same for $|\Omega\rangle$.

Now step (2): Consider

$$G^{(2)}(x, y) = \langle \Omega | T \left( \phi_H(x)\phi_H(y) \right) | \Omega \rangle$$

$$= \lim_{T \to \infty(1-i\epsilon)} \left[ (e^{-iE_0(T-t_0)} \langle 0|\Omega \rangle)^{-1} \langle 0|U_I(T, t_0) \right.$$ 
$$\cdot \frac{U_I^\dagger(x^0, t_0)\phi_I(x)U_I(x^0, t_0) \cdot U_I^\dagger(y^0, t_0)\phi_I(y)U_I(y^0, t_0)}{\phi_H\phi_H} \cdot U_I(t_0, -T) | 0 \rangle \left( e^{-iE_0(T+t_0)} \langle \Omega|0 \rangle \right)^{-1} \right]$$

$$= \lim_{T \to \infty(1-i\epsilon)} \left( e^{-i2E_0T} \langle 0|\Omega \rangle^2 \right)^{-1} \langle 0|U(T, x^0)\phi(x)U(x^0, y^0)\phi(y)U(y^0, -T) | 0 \rangle$$

$$= \mathcal{T}(\phi(x)\phi(y)U(T, -T))$$

In the last expression I’ve gone back to implying the $I$ subscripts on the interaction picture fields. In observing that the big underbraced product is time ordered we are appealing to the Dyson formula for the interaction-picture evolution operators, e.g. $U_I(t, t’) = \mathcal{T} \left( e^{-i\int_{t'}^t dt’V(t’)} \right)$ – so it is a sum of time-ordered products, evaluated between the times in the argument. Notice that I did something slippery in the first step by combining the factors into one big limit; this is OK if each factor converges separately.

What’s the denominator? The norm of the vacuum is one, but we can assemble it from these ingredients (3.13):

$$1 = \langle \Omega|\Omega \rangle = \lim_{T \to \infty(1-i\epsilon)} \left( e^{-i2E_0T} \langle 0|\Omega \rangle^2 \right)^{-1} \langle 0|U(T, t_0)U(t_0, -T) | 0 \rangle$$

Therefore

$$G^{(2)}(x, y) = \lim_{T \to \infty(1-i\epsilon)} \frac{\langle 0| \mathcal{T} \left( \phi(x)\phi(y)e^{-i\int_{-T}^{T} dt’V(t’)} \right) | 0 \rangle}{\langle 0| \mathcal{T} \left( e^{-i\int_{-T}^{T} dt’V(t’)} \right) | 0 \rangle}$$

[End of Lecture 9]

The same methods give the analogous formula for $G^{(n)}(x_1 \cdots x_n)$ for any number of any local operators. Now we can immediately perturbate to our hearts’ content by expanding the exponentials. Let’s do some examples, then I will comment on the familiarity of the prescription for $T$, and we will see that the denominator is our friend because it cancels annoying (disconnected) contributions in the numerator.

**Examples.** For $V = \frac{1}{4!} \phi^4$, let’s study the numerator of $G^{(2)}(x, y)$ in the first few orders of $\lambda$:

$$G^{(2)}_{\text{num}}(x, y) = \langle 0| \mathcal{T} \left( \phi(x)\phi(y)e^{-i\int_{-T}^{T} dt’d^{4+1}z_\lambda \phi^4(z)} \right) | 0 \rangle$$

58
\[ \langle 0 | T \phi(x)\phi(y) | 0 \rangle + \frac{-i\lambda}{4!} \int d^{d+1}z \langle 0 | T (\phi(x)\phi(y)\phi(z)\phi(z)\phi(z)\phi(z)) | 0 \rangle + \mathcal{O}(\lambda^2) \]

\[ = \Delta_F(x-y) + \frac{-i\lambda}{4!} \int d^{d+1}z \left( 3\phi(x)\phi(y)\phi(z)\phi(z)\phi(z)\phi(z) + 4 \cdot 3\phi(x)\phi(z)\phi(y)\phi(z)\phi(z)\phi(z) + \mathcal{O}(\lambda^2) \right) \]

The \( \mathcal{O}(\lambda^2) \) contribution is

\[ \frac{1}{2!}(-i\lambda)^2 \int d^{d+1}z_1 d^{d+1}z_2 \langle 0 | T (\phi(x)\phi(y)\phi(z_1)\phi(z_2)^4) | 0 \rangle \]

With ten fields, there will be five propagators in each diagram. The ingredients which we must connect together are:

The answer is

\[ (\cdots \phi \cdots)^4 \cdot (\phi) \cdot (\phi) + \phi + \phi + \phi \]

For example,

\[ \phi(x)\phi(z_1)\phi(y)\phi(z_1) \left( \phi(z_1)\phi(z_2) \right)^2 \phi(z_2)\phi(z_2) \propto \]

up to the symmetry factor.

**Feynman rules for \( \phi^4 \) theory in position space.** The set of diagrams is made by drawing external vertices for each \( x_i \), and \( n \) internal vertices, and connecting them in all possible ways with propagators,

\[ \{\text{diagrams}\} \equiv \{A\} = \{A_0\} \cup \{A_1\} \cup \cdots \]

where \( A_m \) gives contributions proportional to \( \lambda^m \). Then the associated amplitude is \( \mathcal{M}_A \) and the Green’s function is \( G^{(n)}(x_1 \cdots x_n) = \sum_A \mathcal{M}_A \). To get \( \mathcal{M}_A \),

- Put a \(-i\lambda \int d^{d+1}z_a\) for each vertex (notice no \( \frac{1}{4!} \)).
• Put a $\Delta_F(y_i - y_j)$ for each propagator $\mathcal{A} \rightarrow \mathcal{B}$, where $y$ may be an internal point $x_i$ or an internal point $z_a$.

• Multiply by the symmetry factor $s(A)$. The symmetry factor is defined to be $s(A) = |\text{Aut}(A)|^{-1}$, the inverse of the order of the automorphism group of the diagram. Symmetries of the diagram mean that the sum over contractions fails to completely cancel Dyson’s wretched $\frac{1}{n!}$. For example:

$$s \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right) = \frac{1}{4!} \cdot 3 = \frac{1}{8}, \quad s \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right) = \frac{1}{4!} \cdot 4 \cdot 3 = \frac{1}{2}.$$ 

Do not get hung up on this right now.

Let’s do the numerator of $G^{(4)}$ through order $\lambda^2$:

$$\mathcal{O}(\lambda^0) : \quad \frac{1}{4} + \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right)$$

$$\mathcal{O}(\lambda^1) : \quad \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right) + \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right) + \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right) + \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right) + \cdots + \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right)$$

Notice that only the last term here is “fully connected” in the sense that you can’t divide the diagram up into disjoint pieces without cutting propagators. The other diagrams follow a simple pattern: the first three are obtained from the $\mathcal{O}(\lambda^0)$ diagrams by multiplying by a figure-eight bubble. The second set is obtained by multiplying $G^{(2)}_0 \cdot G^{(2)}_1$, where $G^{(m)}$ denotes the an order-$\lambda^m$ contribution to $G^{(n)}$.

$$\mathcal{O}(\lambda^2) : \quad \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right) + \cdots + \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right) + \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right) + \cdots + \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right)$$

$$\quad \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right) + \cdots + \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right) + \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right) + \cdots + \left( \begin{array}{c}
\mathcal{A} \\
\mathcal{B}
\end{array} \right)$$

(3.14)
Here’s an easy one: \( G^{(n)} = 0 \) when \( n \) is odd. Technically, we can see this from the fact that there is always a \( \phi \) left over after all contractions, and \( \langle 0 | \phi | 0 \rangle = 0 \). Slightly more deeply, this is because of the \( \phi \to -\phi \) symmetry.

**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:

![Diagram](image)

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) and multiplicities \( n_i \) (e.g. \( V_1 \) could be a figure eight, and there could be \( n_1 = 2 \) of them, as in the second term indicated in (3.14)). 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

\[
\mathcal{M}_A = \mathcal{M}_{A_c} \cdot \frac{V_1^{n_1}}{n_1!} \cdot \frac{V_2^{n_2}}{n_2!} \cdots \frac{V_{\alpha}^{n_{\alpha}}}{n_{\alpha}!}
\]

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

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

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

\[
G^{(0)}_{\text{numerator}} = \langle 0 | \mathcal{T} e^{-i \int V} | 0 \rangle = 1 \cdot e^{\sum_i V_i}
\]
But we care about this because it is the denominator of the actual Green’s function:

\[ G^{(n)} = \frac{\langle 0 | T(\phi_1 \cdots \phi_n e^{-iH}) | 0 \rangle}{\langle 0 | T e^{-iH} | 0 \rangle} = \frac{G^{(n)}_{\text{numerator}}}{G^{(0)}_{\text{numerator}}} = \sum_{A_c} M_{A_c}. \]  

(3.16)

And with that we can forget all about the bubbles. So for example,

\[ G^{(2)} = = \ldots + \bigcirc + \bigcirc + \bigcirc + \bigcirc + \bigcirc + \bigcirc + \bigcirc + \bigcirc \]

Notice that in this manipulation (3.16) we are adding terms of many orders in perturbation theory in the coupling \( \lambda \). If we want an answer to a fixed order in \( \lambda \), we can regard anything of higher order as zero, so for example, it makes perfect sense to write

\[ G^{(2)} = \frac{\zeta_i \cdot (1 + 8 + 88 + \cdots)}{(1 + 8 + 88 + \cdots)} + O(\lambda) = \frac{\zeta_i \cdot e^V}{e^V} + O(\lambda) = \zeta_i \zeta_l + O(\lambda). \]

(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. Let’s think about

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

Again, this an off-shell Green’s function, a function of general \( p \), not necessarily \( p^2 = m^2 \). It will, however, vanish unless \( \sum_i p_i^2 = 0 \) by translation invariance. Consider a fully-connected contribution to it, at order \( \lambda^N \). (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 \( \Delta_F(y_A - y_B) = \int d^{d+1}q_r e^{-i(y_A - y_B)q_r} \frac{1}{q_r^2 - m^2 + i\epsilon} \). Since each propagator has two ends, the number of internal lines (by the fully-connected assumption) is

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

The associated amplitude is then

\[ A_{FC}^N = \int d^{d+1}x_1 \cdots d^{d+1}x_n e^{-i \sum_i p_i x_i} (\frac{-i\lambda}{s(FC)})^{N} \int d^{d+1}z_1 \cdots \int d^{d+1}z_N \prod_{r=1}^{N_I} \int d^{d+1}q_r e^{-i(y_A - y_B)q_r} \frac{i}{q_r^2 - m^2 + i\epsilon}. \]
For example, consider the particular contribution with \( n = 4 \) external legs and \( N = 2 \) interaction vertices:

\[
\begin{align*}
\text{has } N_I &= \frac{4 + 2 \cdot 4}{2} = 6
\end{align*}
\]

Notice that we are doing a silly thing here of labelling the momenta of the external lines (the first \( n \) momenta \( q_{i=1\ldots n} \)). Here’s why it’s silly: Look at the integral over \( x_1 \). Where is the dependence on \( x_1 \)? There is the external factor of \( e^{-i p_1 x_1} \) that we put to do the Fourier transform, and there is the propagator taking \( x_1 \) to \( z_1 \), 
\[
\Delta_F(x_1 - z_1) = \int \! d^{d+1} q_1 e^{-i (x_1 - z_1) q_1} \frac{1}{q_1^2 - m^2 + i \epsilon}.
\]
So the integral over \( x_1 \)
\[
\int \! d^{d+1} x_1 e^{-i x_1 (p_1 - q_1)} = \delta^{d+1} (p_1 - q_1)
\]
just sets \( p_1 = q_1 \), and eats the \( \int \! d^{d+1} q_1 \). The same thing happens for each external line, and reduces the number of momentum integrals to \( N_I - n \).

Where is the dependence on \( z_2 \)?
\[
\int \! d^{d+1} z_2 e^{-i z_2 (q_3 + q_4 + q_5 + q_6)} = \delta^{d+1} (q_3 + q_4 + q_5 + q_6)).
\]
Similarly, the \( z_1 \) dependence is all in the exponentials:
\[
\int \! d^{d+1} z_1 e^{-i z_1 (-q_3 - q_4 + q_1 + q_2)} = \delta^{d+1} (q_3 + q_4 - q_1 - q_2)).
\]
These two factors combine to set \( q_1 + q_2 = q_3 + q_4 = -q_5 - q_6 \); momentum is conserved at the vertices. Notice that in the example \( q_5 - q_6 \) is \textit{not} determined.

Each internal vertex reduces the number of undetermined momenta by one, except one combination is already fixed by overall momentum conservation so we have left
\[
N_I - n - (N - 1) = N - \frac{n}{2} + 1
\]
momentum integrals. This number is \( \geq 0 \) for fully connected diagrams, and it is the number of \textit{loops} in the diagram. (This counting is the same as in a Kirchoff’s law resistor network problem.) In the example, \( N_L = 2 - 2 + 1 = 1 \) which agrees with one undetermined momentum integral.

[End of Lecture 10]
In practice now, we need not introduce all those extra $q$s. Label the external lines by $p_1 \cdots p_n$, and the loop momenta by $k_\alpha, \alpha = 1..N_L$. In the example, we might do it like this: for which the amplitude is

$$A_{FC}(p_1 \cdots p_n) = \frac{(-i\lambda)^N}{s(FC)} \delta^{(d+1)}(\sum p_i) \int \prod_{\text{loops},\alpha=1}^{N_L} d^{d+1}k_\alpha \prod_{\text{lines},r}^{i} \frac{i}{q_r^2 - m^2 + i\epsilon}$$

$$= \frac{(-i\lambda)^2}{2!} \delta^{d+1} \left( \sum_{i=1}^{4} p_i \right)^{n=4} \int \prod_{i=1}^{d+1} d_k i \frac{i}{k^2 - m^2 + i\epsilon} \int \prod_{i=1}^{d+1} d_k i \frac{i}{(p_1 + p_2 + k)^2 - m^2 + i\epsilon}$$

(You might notice that the integral over $k$ is in fact formally infinite, since at large $k$ it goes like $\int d^d k \sim \log(\Lambda)$. Try to postpone that worry.) For now, let’s celebrate my successful prediction, for this particular graph, that there would be poles when the external particles are on-shell, $p_i^2 = m^2$. (It would be more correct to call it Lehmann, Symanzik and Zimmerman’s successful prediction.)

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

$$\tilde{G}^{(2)} = \sum_{n_k=0}^{\infty} + \sum_{n_k=1}^{\infty} + \sum_{n_k=2}^{\infty} + \sum_{n_k=3}^{\infty} + \mathcal{O}(\lambda^3)$$

(3.17)

I draw the blue dots to emphasize the external propagators. So here are the momentum space Feynman rules for Green’s function in $\phi^4$ theory:

- An internal line is $\mathcal{I} = \frac{i}{p^2 - m^2 + i\epsilon} = \tilde{\Delta}_F(p)$. Notice that since $\Delta_F(x-y) = \Delta_F(y-x)$, the choice of how we orient the lines is not so fateful.

- An external line at fixed position, $\mathcal{E} = e^{-ipx}$ More generally, external vertices are associated with the wavefunctions of the states we are inserting; here they are plane waves.

- An internal vertex gives $\mathcal{V} \sim (-i\lambda) \int d^{d+1}z e^{-i\sum n_z} = (-i\lambda)\delta^{d+1}(\sum_i p_i)$, 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 conservation, the remaining consequence of the vertex is $\mathcal{V} = -i\lambda$. 64
• 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.

Comment on \( T \to \infty(1 - i\epsilon) \). What happened to the limit on \( T \)? It’s hidden in the integrals over the vertices:

\[
\int d^{d+1}z \ e^{-i\epsilon z_0(z_q)} \ldots = \lim_{T \to \infty(1 - i\epsilon)} \int dz_0 d^d z \ e^{-i \epsilon (z_0(z_q - \bar{z}_q - \sum_i \bar{z}_q))} \ldots
\]

One end of the integral \( z_0 = \pm \infty \) is going to be infinite unless \( \sum_i p_0 z^0_i \in i\mathbb{R} \), in which case it just oscillates. This seems scary. We can make ourselves feel better about it if we just replace every \( p_0 \) with \( p_0(1 + i\epsilon) \) for some infinitesimal \( \epsilon \). This means that the integrals will look like:

\[
\int \bar{d}^{d+1}p \ e^{-i\epsilon p_0} = \int \bar{d}^{d+1}p \ e^{-i\epsilon p_0(1 + i\epsilon)}
\]

That is:

if we use the Feynman contour for every propagator

\[
\Delta_F(x) = \int_{c_F} d^{d+1}p e^{-ipx} \frac{i}{p^2 - m^2 + i\epsilon_F}
\]

with \( \epsilon_F = \epsilon \) then this problem goes away.

The factors of \( T \) give another perspective on the exponentiation of the vacuum bubbles. Consider the diagram:

\[
\Delta_F(x) = (-i\lambda)^2 \prod_{i=1}^4 \int d^{d+1} p_i \phi^{d+1}(p_1 + p_2) \phi^{d+1}(p_1 + p_2) \ldots
\]

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

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

where \( V \) is the volume of space. 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. But the free energy \( \propto \log Z = \log G(0) \) should be extensive, \( \propto VT \). Therefore, the vacuum bubbles must exponentiate.
3.5 Interlude: old-fashioned perturbation theory

[Schwartz, chapter 4] I want to take a brief break from the inexorable building of theoretical machines to demonstrate some virtues of those machines. It will explain what I was really mumbling about when I said that the Feynman propagator involves antiparticles going backwards in time.

Consider a system which is a small perturbation of a solvable system $H = H_0 + V$. Suppose that the initial system $H_0$ has a continuous spectrum, so that there are eigenstates at every nearby energy. Then given an eigenstate of $H_0$, $H_0 |\varphi\rangle = E |\varphi\rangle$, we expect an eigenstate of $H$ with the same energy $H |\psi\rangle = E |\psi\rangle$. Palpating the previous equation appropriately gives

$$|\psi\rangle = |\varphi\rangle + \frac{1}{E - H_0 + i\epsilon} V |\psi\rangle$$

(the Lippmann-Schwinger equation). This represents the perturbed eigenstate as the free one plus a scattering term, in terms of the ‘propagator’ $\Pi$. The $i\epsilon$ is a safety factor which helps us negotiate the fact that $E - H_0$ is not actually invertible. To write this entirely in terms of the known free state $|\varphi\rangle$, iterate. Let $V |\psi\rangle \equiv T |\varphi\rangle$ where $T$ is the transfer matrix:

$$|\psi\rangle = |\varphi\rangle + \Pi T |\varphi\rangle.$$

Now act on both sides with $V$ to get $V |\psi\rangle = T |\varphi\rangle = V |\varphi\rangle + V\Pi T |\varphi\rangle$, which is true for any $\varphi$ so

$$T = V + V\Pi T = V + V\Pi(V + V\Pi T) = V + V\Pi V + V\Pi V\Pi V + \cdots = \left(1 - V\Pi\right)^{-1} V$$

Given a complete set of eigenstates of $H_0$, with $\sum_i |\varphi_i\rangle \langle \varphi_i| = 1$,

$$T_{fi} \equiv \langle \varphi_f | T |\varphi_i\rangle = V_{fi} + V_{fj}\Pi(j)V_{ji} + V_{fj}\Pi(j)V_{jk}\Pi(k)V_{ki} + \cdots$$

where $V_{fi} \equiv \langle \varphi_f | V |\varphi_i\rangle$ gives the first Born approximation, and $\Pi(j) \equiv \frac{1}{E - E_j}$, and $E = E_i = E_f$, energy is conserved.

For a vivid example, consider the mediation of a force by a boson field. Let

$$V = \frac{1}{2}e \int d^4x \Psi_e(x)\phi(x)\Phi_e(x)$$

where ‘$\phi$’ is for ‘photon’ and ‘$e$’ is for ‘electron’ but we’ve omitted spin and polarization information, and got the statistics of the electron wrong, for simplicity. Consider the free eigenstates $|i\rangle = |\vec{p}_1, \vec{p}_2\rangle$, $|f\rangle = |\vec{p}_3, \vec{p}_4\rangle$. Then

$$T_{fi} = V_{fi} + \sum_n V_{fn} \frac{1}{E_i - E_n} V_{ni} + \cdots.$$
What are the possible intermediate states $|n\rangle$? It has to be two $e$ and one $\phi$, as in the following visualization (not a Feynman diagram in the sense we’ve been discussing):

Time goes to the left, as always. You see that there are two classes of possibilities: $|n_R\rangle = |p_3, p_\gamma; p_2\rangle$, $|n_A\rangle = |p_1; p_\gamma, p_4\rangle$. Consider them from the point of view of particle 2. In the first (R) case, $e_2$ feels a photon emitted by particle 1, after the emission happens:

$$
V_{ni}^R = \langle p_3^3, p_\gamma, p_2^2 \mid V \mid p_1, p_2 \rangle = \langle p_3^3, p_\gamma \mid V \mid p_1 \rangle \langle p_2^2 \rangle
$$

$$
= \frac{e}{2} \int d^4x \langle p_3^3, p_\gamma \mid \Psi_e(x) \phi(x) \Psi_e(x) \mid p_1^1 \rangle
$$

$$
= \frac{e}{2} \int d^4x \langle p_\gamma \mid \phi(x) \mid 0 \rangle \langle p_3^3 \mid \Psi(x) \rangle \langle p_1^1 \rangle
$$

$$
= e \delta^d (p_1 - p_3 - p_\gamma)
$$

– momentum is conserved. Note that energy is not, $E_i \neq E_m$ (or else the denominator is zero).

The other possibility is $|n_A\rangle = |p_1; p_\gamma, p_4\rangle$, which means $e_2$ feels the effects of a photon it emitted, which is later absorbed by $e_1$ (!!).

$$
V_{ni}^A = \langle p_4^4 p_\gamma \mid V \mid p_2^2 \rangle = e \delta^d (p_2 - p_4 - p_\gamma).
$$

 Altogether, to leading nonzero order,

$$
T_{fi} = \sum_n V_{fn} \frac{1}{E_i - E_n} V_{ni} = \sum_{n=R,A} \int d^4p_\gamma \delta^d (p_1 - p_3 - p_\gamma) \delta^d (p_2 - p_4 + p_\gamma) \frac{e^2}{E_i - E_n}.
$$

A bit of kinematics: let the $\phi$ have mass $m_\gamma$, so for a given $p_\gamma$, $E_\gamma = \sqrt{|p_\gamma|^2 + m_\gamma^2}$. Notice that these are real particles, they satisfy the equations of motion. For the $R$ case, the intermediate energy is

$$
E_n^R = E_3 + E_\gamma^R + E_2 = E_3 + \sqrt{|p_1 - p_3|^2 + m_\gamma^2}
$$
\[ E_i - E_R^{\gamma} = E_1 + E_2 - (E_3 + E_R^{\gamma} + E_2) = E_1 - E_3 - E_R^{\gamma} = -\Delta E - E_\gamma \]

where \(\Delta E \equiv E_3 - E_1 = E_4 - E_2\) (by overall energy conservation). Momentum conservation means \(p_2 - p_4 = p_1 - p_3\) so

\[ E_\gamma^A = \sqrt{(|\vec{p}_1 - \vec{p}_3|^2 + m_\gamma^2)} = \sqrt{(|\vec{p}_2 - \vec{p}_4|^2 + m_\gamma^2)} = E_R^{\gamma} \equiv E_\gamma. \]

Therefore

\[ E_i - E_n^A = \Delta E - E_\gamma. \]

The sum of these factors is

\[
\sum_{n=R,A} \frac{e^2}{E_i - E_n} = \frac{e^2}{-\Delta E - E_\gamma} + \frac{e^2}{\Delta E - E_\gamma} = \frac{2E_\gamma e^2}{\Delta E^2 - E_\gamma^2} = 2E_\gamma \frac{e^2}{k^2 - m_\gamma^2}.
\]

Here we defined \(k^\mu = (\Delta E, \vec{p}_\gamma)^\mu\), and \(k^2 = k_\mu k^\mu\) is a Lorentz-invariant inner product. Ignoring the normalization factor \(2E_\gamma\), this is the Lorentz-invariant momentum-space propagator for the \(\phi\) particle with four-momentum \(k^\mu\). Notice that there is no actual particular with that four-momentum! It is a superposition of a real particle going forward in time and its (also real) antiparticle going backward in time. If we followed the \(i\epsilon\) that would work out, too, to give the Feynman propagator.

### 3.6 From correlation functions to the \(S\) matrix

Now we resume our inexorable progress towards observable physics (such as cross sections and lifetimes). 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.

Recall our expression for the momentum-space two-point function (3.17) in terms of a sum of connected diagrams, ordered by the number of powers of \(\lambda\). Let’s factor out the overall delta function by writing:

\[ \tilde{G}^{(2)}(p_1, p_2) \equiv \rho^{d+1}(p_1 + p_2)\tilde{G}^{(2)}(p_1). \]
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, \( \Omit[300]\) is 1PI, but \( \Omit[100]\) is not; rather, the latter contributes to the bit with two 1PI insertions. Then

\[
\tilde{G}^{(2)}(p) = \quad + 
\]

So that we may write equations without pictures, let

\[-i\Sigma(p) \equiv \quad \]

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

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

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

\[
= \frac{i}{p^2 - m_0^2} \frac{1 - \Sigma}{1 - \frac{\Sigma}{p^2 - m_0^2}} = \frac{i}{p^2 - m_0^2 - \Sigma(p)}
\]

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, \( m_0^2 + \Sigma(p) \big|_{\text{pole}} \) 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

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

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 absent in the free theory. Later (in 215C?) we will see that unitarity requires \( Z < 1 \). Notice that if we know \( \Sigma \) only to some order in perturbation theory, then (3.19) 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

\[
\tilde{G}_{\text{1PI}}^{(n)}(p_1 \cdots p_n) \equiv \nonumber
\]

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 \( G_{\text{1PI}} \) does not include diagrams like:

\[
\text{or } \nonumber
\]

Notice that 1PI diagrams are amputated – their external limbs have been cut off.

This is almost what we need to make S-matrix elements. If we multiply the \( n \)-point function by \( \prod_{i=1}^{n} \frac{p_i^2 - m^2}{\sqrt{Z}} \) we cancel out the propagators from the external legs. 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:
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):

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

(3.21)

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; I also like Schwartz’ chapter 5; 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:

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

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

(3.22)

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

2. Now, recall our brontosaurus expedient (introduced previously after (3.7)): turn the interactions off at $t = \pm \infty$.\(^{18}\) This allows us to write the field in terms of some pretend free fields

$$\phi(x) \begin{cases} 
\overset{\text{at } -\infty}{\sim} & Z_{\frac{1}{2}} \phi_{\text{in}}(x) \\
\overset{\text{at } +\infty}{\sim} & Z_{\frac{1}{2}} \phi_{\text{out}}(x)
\end{cases}$$

\(^{18}\)Here’s why this is really bad: everything we might scatter is a boundstate. For example: atoms, nuclei, nucleons etc... But if there are no interactions there are no boundstates.
The factors of $\sqrt{Z}$ are required to get the correct two point functions (3.20) near the mass shell. The mode operators for $\phi_{\text{in}}$ are called $a^{(\text{in})}$ etc. $\phi_{\text{in, out}}$ are free fields: their full hamiltonian is $H_0$. They are in Heisenberg picture, and the reference time for $\phi_{\text{in, out}}$ is $\pm \infty$ respectively. Since they are free fields, we can write use (3.22) to write

$$\sqrt{2\omega_k}a^{(\text{in})\dagger} = -i \int d^d x \ e^{-ikx} (i\omega_k + \partial_0) \phi_{\text{in}}(x) = -iZ^{-1/2} \int_{t=-\infty}^{t=\infty} d^d x \ e^{-ikx} (i\omega_k + \partial_0) \phi(x)$$

where in the second step we used the independence on time in (3.22), even though $\phi(x)$ is not a free field.

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

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

$$= iZ^{-1/2} \int d^{d+1} x \left( e^{-ikx} \phi_0 - \phi \cdot \partial_0^2 e^{-ikx} \right) \left( \frac{1}{\sqrt{2 - m^2}} e^{-ikx} \right)$$

$$= \text{IBP} \ iZ^{-1/2} \int d^{d+1} x e^{-ikx} \left( \Box + m^2 \right) \phi(x) \quad (3.23)$$

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 $p_j$.

$$\langle p_1 \cdots p_n | S | k_1 \cdots k_m \rangle$$

$$= \prod_{p, k} \sqrt{2\omega} \langle 0 | \prod_{p, k} a^{\text{out}}_p S \prod_{k} a^{\text{in}\dagger}_k | 0 \rangle$$

$$= \prod_{p, k} \sqrt{2\omega} \langle 0 | \mathcal{T} \left( \prod_{p, k} a^{\text{out}}_p S \prod_{k} a^{\text{in}\dagger}_k \right) | 0 \rangle$$

$$= \prod_{p, k} \sqrt{2\omega} \langle 0 | \mathcal{T} \left( \prod_{p} a^{\text{out}}_p S \left( a^{\text{in}\dagger}_{k_1} - a^{\text{out}\dagger}_{k_1} \right) \prod_{k} a^{\text{in}\dagger}_k \right) | 0 \rangle$$

$$(3.23)$$

$$= iZ^{-1/2} \int d^{d+1} x e^{-ik_1 x_1} \langle 0 | \mathcal{T} \left( \prod_{p} a^{\text{out}}_p S \left( \Box + m^2 \right) \phi(x_1) \prod_{k} a^{\text{in}\dagger}_k \right) | 0 \rangle$$

$$= iZ^{-1/2} \int d^{d+1} x e^{-ik_1 x_1} \left( \Box + m^2 \right) \langle 0 | \mathcal{T} \left( \prod_{p} a^{\text{out}}_p S \phi(x_1) \prod_{k} a^{\text{in}\dagger}_k \right) | 0 \rangle + X$$

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 $k^2 \rightarrow m^2$, I promise.
5. Now do this for every particle to get

$$\langle p_1 \cdots p_n | S | k_1 \cdots k_m \rangle = \prod_{j=1}^m \int d^{d+1} y_j \ e^{+ip_j y_j} Z^{-1/2} (m^2)$$

$$\prod_{i=1}^n \int d^{d+1} x_i \ e^{-ik_i x_i} Z^{-1/2} (m^2) \ 0 \ | T \phi(x_1) \cdots \phi(y_j) S | 0 \rangle + X$$

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 (3.21).

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 $\mathcal{O}$:

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

There is a more general statement of LSZ:

$$\prod_{a \in i} \left( Z_a^{-1/2} i \int d^{d+1} x_a e^{-ip_a x_a} (m_a^2) \right) \prod_{b \in f} \left( Z_b^{-1/2} i \int d^{d+1} x_b e^{+ip_b x_b} (m_b^2) \right) G^{(n)}_{\mathcal{O}}(1 \cdots n)$$

$$= \langle \{p_f\} | S | \{p_a\} \rangle$$

This more general statement follows as above if we can write $\mathcal{O}_a \underset{t \rightarrow -\infty}{\sim} \sqrt{Z_a} \phi_{in}$.

Here is a summary of the long logical route connecting Feynman diagrams to measurable quantities:
One step is left.

\[ G^{(n)} = \langle 0 | \mathcal{T} \left( \phi_{\mu} \phi_{\nu} \phi_{\mu} \phi_{\nu} \right) | \Sigma \rangle \]

\[ \tau \to \infty \left( 1 - i\epsilon \right), \]

Feynman contour

\[ \langle 0 | \mathcal{T} \left( \phi \phi \phi \phi e^{-i\int V} \right) | 0 \rangle \]

Wick, exponentiation of vacuum bubbles


[End of Lecture 12]
\textbf{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 $\mathcal{M}$ defined by

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

(3.25)

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

The rules for the Feynman diagram calculation of $\mathcal{M}$ (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 (3.19). In particular, a diagram like $\times$ is already included by using the correct $Z$ and the correct $m$.

\textbf{Example: Nucleon scattering.} [Here we follow Tong §3.5 very closely] Let’s return to the example with a complex scalar field $\Phi$ and a real scalar field $\phi$ with Lagrangian (3.8). Relative to $\phi^4$ theory, the differences are: we have two kinds of propagators, one of which is oriented, and instead of a 4-point vertex which costs $-i\lambda$, we have a 3-point vertex for $\phi\Phi^\ast\Phi$ which costs $-ig$.

Let’s consider $2 \to 2$ scattering of $\Phi$ particles [recall HW 5 or see Tong §3.3.3 for the artisanal version of this calculation], so

$$
|i\rangle = |\vec{p}_1, \vec{p}_2\rangle, |f\rangle = |\vec{p}_3, \vec{p}_4\rangle \quad \text{with} \quad |\vec{p}_i, \vec{p}_j\rangle \equiv \sqrt{2E_{\vec{p}_i}}\sqrt{2E_{\vec{p}_j}} b_{\vec{p}_i}^\dag b_{\vec{p}_j}^\dag |0\rangle .
$$

75
The Feynman rules above give, to leading nonzero order,

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

(3.26)

The diagrams depict two ‘snucleons’ \( \Phi \) (solid lines with arrows indicating snucleons versus antimolecules) exchanging a meson \( \phi \) (double gray line, with no arrow) with

\[
momentum \ k \equiv p_1 - p_3 = p_2 - p_4 \quad \text{(first term)} \quad \text{or} \quad k \equiv p_1 - p_4 = p_2 - p_3 \quad \text{(second term)}.
\]

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.

The meson in these diagrams is virtual, or off-shell, in the sense that it does not satisfy its equation of motion \( k^2 \neq M^2 \). As we saw in 3.5, each of these diagrams is actually the sum of retarded and advanced exchange of real on-shell particles. The two diagrams included in (3.26) 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 \( |i\rangle = \sqrt{2E_{\vec{p}_1}} \sqrt{2E_{\vec{p}_2}} b_{\vec{p}_1}^\dagger c_{\vec{p}_2}^\dagger |0\rangle \), then the leading diagrams are

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

(3.27)

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 nucleon:

\[ iM = \begin{align*} & \begin{array}{c} i \frac{g^2}{(p+k)^2-m^2+i\epsilon} + i \frac{g^2}{(p-k')^2-m^2+i\epsilon} \end{array} \end{align*} \tag{3.28} \]

Now the intermediate state is a nucleon.

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.

### 3.7 Observable physics from the \( S \)-matrix

Now, finally, we extract some physics from all the machinery we’ve built.

**Mediation of forces.** Consider the non-relativistic (NR) limit of the nucleon-nucleon scattering amplitude (3.26). In the center-of-mass frame \( \vec{p} \equiv \vec{p}_1 = -\vec{p}_2 \) and \( \vec{p}' \equiv \vec{p}_3 = -\vec{p}_4 \). In the NR limit, \( |\vec{p}| \ll m \), and so \( p^0_1 = m(1 + \frac{1}{2} \left( \frac{|\vec{p}|^2}{m^2} \right) + \cdots ) \).

Energy-momentum conservation says \( p_1 + p_2 = p_3 + p_4 \), so \( |\vec{p}'| = |\vec{p}| \ll m \) as well. In this limit, the meson propagator (in the first diagram) depends on \( (p_1 - p_3)^2 = (p_1^0 - p_3^0)^2 - (\vec{p} - \vec{p}')^2 = -(\vec{p} - \vec{p}')^2 \), so the amplitude reduces to

\[ iM = +i g^2 \left( \frac{1}{(\vec{p} - \vec{p}')^2 - M^2} + \frac{1}{(\vec{p} + \vec{p}')^2 - M^2} \right). \]

Now compare to NR QM. The scattering amplitude in the COM frame for two particles with relative position \( \vec{r} \) and potential \( U(\vec{r}) \) is, in the first Born approximation, \(^{19}\)

\[ iA_{\text{Born}}(\vec{p} \to \vec{p}') = \text{NR} \langle \vec{p}' | U(\vec{r}) | \vec{p} \rangle_{\text{NR}} = -\int d^dr U(\vec{r}) e^{-i(\vec{p}' - \vec{p}) \cdot \vec{r}} \]

where the two-particle state with NR normalization is

\[ |\vec{p}\rangle_{\text{NR}} = \frac{1}{\sqrt{2E_1 \sqrt{2E_2}}} |p_1, p_2\rangle = \frac{1}{2m} |p_1, p_2\rangle. \]

The two diagrams in the relativistic answer come from Bose statistics, which means we can’t distinguish \( \vec{p} \to \pm \vec{p}' \) from each other; to infer the potential we can just compare

\(^{19}\)Please beware my signs here; thanks to Ken Intriligator for pointing out a problem here.
the first diagram, \((2m)^2iA_{\text{Born}}(\vec{p} \to \vec{p}') = +ig^2\frac{1}{(\vec{p}-\vec{p}')^2-M^2}\) to find:

\[
\int d^d r \ U(\vec{r})e^{-i(\vec{p}-\vec{p}')\cdot \vec{r}} = -\int d^d q\mathcal{M}(q)e^{i\vec{q}\cdot \vec{r}} = -\frac{(\frac{g}{2m})^2}{(\vec{p} - \vec{p}')^2 + M^2}
\]

which means, in \(d = 3\),

\[
U(\vec{r}) = -\frac{(\frac{g}{2m})^2}{4\pi r} e^{-Mr}
\]

This is the Yukawa potential. (You encountered this potential on the homework, for the same reason, by a different approach.) It has a range, \(M^{-1}\), determined by the mass of the exchanged particle. If we take \(M \to 0\), it becomes the Coulomb potential. The sign means that it is attractive, even though this is the potential between particle and particle; this is a general consequence of scalar exchange. Notice that in \(d = 3\), the Yukawa coupling between scalars has \(1 = \int d^4xg\Phi^2 = -4 + 3 + [g]\) so \(g/m\) is dimensionless.

A brief warning: while it is satisfying to make contact with something familiar here, the way we actually measure any such potential is by scattering the particles and measuring cross-sections.

[End of Lecture 13]

\[\text{For convenience, here’s the integral:}\]

\[
\int d^4k \ e^{i\vec{k}\cdot \vec{r}} \frac{1}{k^2 + M^2} \ e^{-y}\cos \theta = \frac{1}{(2\pi)^2} \int_0^\infty \frac{k^2 \ dk}{k^2 + M^2} \int_{-1}^1 \ dy e^{ikyr} = \frac{1}{(2\pi)^2} \int_{-\infty}^\infty \frac{dk}{k^2 + M^2} \sin kyr
\]

\[
= \frac{1}{(2\pi)^2} \int_{-\infty}^\infty \frac{dk}{k^2 + M^2} \frac{k e^{ikyr}}{k^2 + M^2 + h.c.} = \frac{1}{2\pi} \int_{-\infty}^\infty \frac{dk}{k^2 + M^2} \frac{k e^{ikyr}}{k^2 + M^2},\ 	ext{close contour in UHP for free}
\]

\[
= \frac{1}{2\pi i} \int_{-\infty}^\infty \frac{dk}{k^2 + M^2} \cdot 2 = \frac{e^{-Mr}}{4\pi r}.
\]
Lifetimes. [Schwartz, chapter 4] How do we compute the lifetime of an unstable particle in QFT? Consider such a particle in its rest frame, \( p^\mu = (M, \vec{0})^\mu \). Let \( dP \) be the probability that the particle decays (into some set of final states \( f \)) during a time \( T \). The decay rate is then \( d\Gamma \equiv \frac{1}{T} dP \), the probability per unit time. I put a \( d\Gamma \) to indicate a differential decay rate into some particular set of final states. If we sum over all possible final states, we can make a practical, frequentist definition of the decay rate, with the idea that we have a big pile of particles and we just count how many go away in some time window:

\[
\Gamma \equiv \frac{\text{# of decays per unit time}}{\text{# of particles}} \equiv \frac{1}{\tau} \tag{3.29}
\]

where \( \tau \) is the lifetime.

Fortunately for us, particles which are stable in the free theory can decay because of weak interactions; in such a case, we can relate \( dP \) to an \( S \) matrix element for a process which takes one particle to \( n \) particles, \( S_{n \rightarrow 1} \left( \{ p_j \}_{j=1}^n \leftarrow (M, \vec{0}) \right) \). So:

\[
d\Gamma \equiv \frac{1}{T} dP = \frac{1}{T} \frac{|\langle f | S | i \rangle|^2}{\langle f | f \rangle \langle i | i \rangle} d\Pi_f \tag{3.30}
\]

Here are two new ingredients:

1. \( d\Pi_f \) is the volume of the region of final-state phase space, \( d\Pi_f \propto \prod_{j=1}^n d^d p_j \). We are allowing, as we must, for imperfect measurements. We will normalize the density of final states so that \( \int d\Pi = 1 \). Putting back the IR and UV walls of our padded room as in (1.3), we take the continuum limit (\( N \rightarrow \infty \)) of

\[
\Delta x \sum_i = \frac{L}{N} \sum_i^{N \rightarrow \infty , L \text{fixed}} \int dx = L \quad \text{and} \quad \frac{1}{2\pi} \Delta p \sum_i = \frac{1}{2\pi} \frac{2\pi}{L N} \sum_i^{N \rightarrow \infty , L \text{fixed}} \int dp = L^{-1}.
\]

This gives

\[
d\Pi = \prod_{j=1}^n V d^d p_j,
\]

a factor of the volume of space \( V = L^d \) for each final-state particle.

2. The normalization factors \( \langle f | f \rangle \langle i | i \rangle \) are not so innocent as they look, because of our relativistic state normalization. Recall that \( |\vec{p} \rangle = \sqrt{2\omega_p} a^\dagger_\vec{p} |0 \rangle \) the price for
the relativistic invariance of which is
\[ \langle \vec{k} | \vec{p} \rangle = \sqrt{2 \omega_{\vec{k}}^2 \omega_{\vec{p}}^2} \cdot \langle 0 | a_{\vec{k}} a_{\vec{p}}^\dagger | 0 \rangle = 2 \omega_{\vec{p}} \delta^d(\vec{p} - \vec{k}) \]

Therefore,
\[ \langle \vec{p} | \vec{p} \rangle = 2 \omega_{\vec{p}} \delta^d(0) = 2 \omega_{\vec{p}} V. \]

Therefore,
\[ |i\rangle = \sqrt{2M} a_{\vec{0}}^\dagger |0\rangle \quad \Rightarrow \quad \langle i|i\rangle = 2MV \]
\[ |f\rangle = |\{\vec{p}_j\}\rangle \quad \Rightarrow \quad \langle f|f\rangle = \prod_j (2\omega_j V) \quad (3.31) \]

where I’ve abbreviated \( \omega_j \equiv \omega_{\vec{p}_j} \).

Now it is time to square the quantum amplitude
\[ \langle f|S - \mathbb{1}|i\rangle = i \delta^{d+1}(p_T) \langle f|\mathcal{M}|i\rangle \]

\( p_T = \sum p_i - \sum p_f \) is the total momentum change) to get the probability (3.30). Again we encounter a \( \delta^2 \), and again we use \((2\pi)^{d+1} \delta^{d+1}(0) = TV\), so as long as \( f \neq i \), we have
\[ |\langle f|S - \mathbb{1}|i\rangle|^2 = \delta^{d+1}(0) \delta^{d+1}(p_T) |\langle f|\mathcal{M}|i\rangle|^2 = VT \delta^{d+1}(p_T)|\mathcal{M}|^2 \]

so that
\[ dP = TV \delta^{d+1}(p_T) \frac{1}{2M \prod_j (2\omega_j V)} |\mathcal{M}|^2 \prod_j^n Vd^d p_j \]
\[ = \frac{T}{2M} |\mathcal{M}|^2 d\Pi_{LI} \quad (3.32) \]

where all the factors of \( V \) went away (!), and
\[ d\Pi_{LI} \equiv \prod_{\text{final state},j} \frac{d^d p_j}{2\omega_j} \delta^{d+1}(p_T) \]

is a Lorentz-invariant measure on the allowed final-state phase space. You can see that this is the case by the same calculation that led us to stick those \( 2\omega_j \)s in the states.

One more step to physics:
\[ d\Gamma = \frac{1}{T} dP = \frac{1}{T} \frac{T}{2M} |\mathcal{M}|^2 d\Pi_{LI} = \frac{|\mathcal{M}|^2}{2M} \frac{1}{d\Pi_{LI}} \]

80
\[ d\Gamma = |\mathcal{M}|^2 \frac{1}{2M} d\Pi_{LL}. \]

On the RHS is all stuff we know how to calculate (recall the Feynman rules for \( \mathcal{M} \) that we listed after (3.25)), and on the LHS is a particle decay rate.

The boxed formula gives the decay rate in the rest from of the unstable particle. In other frames, the lifetime gets time-dilated. This must be true on general grounds of special relativity, but we can see this directly since in a general frame, the normalization of the initial state is not \( \langle i| i \rangle_{\text{rest frame}} = \sqrt{2m} \) but \( \langle i| i \rangle = \sqrt{2E} \). Therefore

\[ \frac{\Gamma_{\text{rest frame}}}{\Gamma} = \frac{E}{m} = \gamma \leq 1 \]

and \( \tau = \tau_{\text{rest frame}}/\gamma \geq \tau_{\text{rest frame}} \).

**Cross sections.** If we are not in the convenient situation of having in our hands a big pile of particles which are stable in the free theory and decay because of not-too-strong interactions, we need to be more proactive to get physics to come out: we have to smash the particles together. When doing this, we send beams of particles at each other and see what comes out. We will treat these beams as perfectly collimated momentum eigenstates; if something goes wrong, we’ll make a more accurate representation and put them in better-localized wavepackets. A quantity which is good because it is intrinsic to the particles composing the beams is the scattering cross section, \( \sigma \), defined by

\[
\text{Number of events of interest} \equiv \frac{N_A N_B}{A} \sigma
\]

where \( A \) is the common area of overlap of the beams \( A \) and \( B \), and \( N_{A,B} \) are the number of particles in each beam. (Peskin does a bit more worrying at this point, for example, about whether the beams have constant density of particles.) By ‘events of interest’ I mean for example those particles which end up going in a particular direction, for example in a solid angle \( d\Omega(\theta, \varphi) \). Restricting to events of interest in particular direction gives the differential cross section, \( \frac{d\sigma}{d\Omega} \). The notation is motivated by the idea that \( \sigma = \int d\Omega \frac{d\sigma}{d\Omega} \).

The cross-section is the effective cross-sectional area of the beam taken out of the beam and put into the particular state of interest. Here is a picture (adapated from Schwartz’ book) which I think makes the idea behind definition very vivid:

Now we relate \( \sigma \) to the \( S \)-matrix. The scattering rate \( dw_{fi} \equiv \frac{dP_f}{\tau} \) is the scattering probability per unit time, for some fixed initial and final particle states. In a beam,
this is related to the cross section by

\[ dw = j d\sigma \]  

(3.33)

where \( j \) is the particle current density, which for the case of scattering from an initial state with two particles \( A + B \rightarrow ... \) is

\[ j = \frac{\text{relative velocity of } A \text{ and } B}{\text{volume}} = \frac{v_{AB}}{V} \].

The number of particles in each beam does not appear in (3.33) because the BHS is intensive. Putting together these statements, we can relate the cross section to the scattering probability:

\[
d\sigma = \frac{1}{T} \frac{1}{j} \frac{dN}{N_{\text{incoming}}} = \frac{V}{T} \frac{1}{|\vec{v}_A - \vec{v}_B|} dP_{ji}. 
\]  

(3.34)

The first equation in (3.34) is a practical frequentist origin of (3.33), analogous to (3.29) for decay rates. And just as in the discussion of lifetimes above,

\[
dP = \frac{|\langle f | S | i \rangle|^2}{\langle f | f \rangle \langle i | i \rangle} d\Pi_f.
\]

Everything is as before except for the different initial state:

\[
| i \rangle = | \vec{p}_A, \vec{p}_B \rangle \rightarrow \langle i | i \rangle = (2\omega_A V)(2\omega_B V).
\]

Squaring the amplitude gives

\[
dP = \frac{T}{V} \frac{1}{2\omega_A 2\omega_B} |M|^2 d\Pi_{LI};
\]

the only difference is that we replace \( \frac{1}{2M} \) with the factors for the 2-particle initial state. Finally, \( d\sigma = \frac{V}{T} \frac{1}{|\vec{v}_A - \vec{v}_B|} dP \) gives

\[
d\sigma = \frac{1}{2\omega_A 2\omega_B} \frac{1}{|\vec{v}_A - \vec{v}_B|} |M|^2 d\Pi_{LI}.
\]

Again all the IR-divergent factors of \( V \) and \( T \) went away in the intrinsic physical quantity, as they must.
3.7.1 Two-body phase space

[Schwartz] To make the formulae of the previous section more concrete, let’s simplify them for the case of $n = 2$: two particles in the final state, whose momenta we’ll call $p_1, p_2$. Note that overall momentum conservation implies $p_1 + p_2 = p_{CM}$; we can use this to eliminate $p_2$. In that case

$$d\Pi_{LI} = \frac{\delta^{d+1}(p_T)}{2E_1 2E_2} \frac{d^d p_1 d^d p_2}{E_1 E_2}$$

$$E_i \equiv \sqrt{|p_i|^2 + m_i^2}$$

$$d\Pi_{LI} = \frac{1}{4(2\pi)^{d-1}} \frac{d^{d-1} \Omega_{p_1}}{E_1 E_2} \delta(E_1 + E_2 - E_{CM}) \frac{d^d p_1}{=d^{d-1} \Omega_{p_1} d^{d-1} p_1}$$

$$p_1 \equiv |p_1| > 0$$

$$x(p_1) \equiv E_1(p_1) + E_2(p_2 = p_{CM} - p_1) - E_{CM}$$

$$\theta(x) = \frac{E_1 E_2}{E_1 + E_2} d\phi$$

$$dp_1 = \frac{dp_1}{dx} dx = \frac{E_1 E_2}{E_1 + E_2} p_1$$

In the last step, we used the fact that $p_1 \geq 0$ means $E_1(p_1) \geq m_1, E_2(p_2 = p_{CM} - p_1) \geq m_2$.

$2 \rightarrow 2$ scattering in $d = 3$. In the special case where the initial state also consists of two particles, we can also simplify the formula for the cross section. Let the initial momenta be $k_A, k_B$. In particular, the relative velocity factor is

$$|v_A - v_B| \overset{\text{CM}}{=} -k_B \left| \frac{k_A}{E_{k_A}} + \frac{k_B}{E_{k_B}} \right| = \frac{-k_B}{E_{k_A} E_{k_B}}$$

Therefore

$$\frac{d\sigma}{d\Omega} \overset{\text{CM}}{=} \frac{1}{64\pi^2 E_{CM}^2} \frac{|\vec{p}_1|}{|k_A|^2} |\mathcal{M}|^2 \theta(E_{CM} - m_1 - m_2).$$

Warning: for identical particles in the final state, one must be careful about over-counting in the integral over angles, since a rotation by $\pi$ exchanges them. In this case

$$\sigma = \frac{1}{2} \int_{4\pi} d\Omega \frac{d\sigma}{d\Omega}.$$