Contents

0.1 Introductory remarks ................................................. 3
0.2 Conventions .......................................................... 5

1 From particles to fields to particles again 6
  1.1 Quantum sound: Phonons ........................................ 6
  1.2 Scalar field theory ................................................ 15
  1.3 Quantum light: Photons ......................................... 20
  1.4 Lagrangian field theory .......................................... 24
  1.5 Casimir effect: vacuum energy is real ....................... 29

2 The path integral makes some things easier 33
  2.1 Fields mediate forces ........................................... 33
  2.2 Euclidean path integral and Wick rotation ................. 37
  2.3 Feynman diagrams from the path integral ............. 41

3 Lorentz invariance and causality 45
  3.1 Causality and antiparticles .................................. 47
  3.2 Propagators, Green’s functions, contour integrals .... 51
  3.3 Interlude: where is 1-particle quantum mechanics? .... 57

4 Interactions, not too strong 60
  4.1 Where does the time dependence go? ....................... 60
  4.2 S-matrix ............................................................ 64
  4.3 Time-ordered equals normal-ordered plus contractions .... 66
  4.4 Time-ordered correlation functions ...................... 69
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 QFT.

A consequence of relativity is that the number of particles isn’t fixed. That is: there are processes where the number of particles changes in time. This is a crucial point of departure for QFT, 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 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), which are rays in \( \mathcal{H}_1 \).

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, for example, 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), in the sense that \( i\hbar \partial_t |\psi\rangle = \mathbf{H} |\psi\rangle \). 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 in the eigenbasis of \( \mathbf{A} \): \( \mathbf{A} |a\rangle = a |a\rangle \), \( |\psi\rangle = \sum_a |a\rangle \langle a|\psi\rangle \); 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. But it’s not something that can happen within $\mathcal{H}_1$, since the number of particles changes during the process. 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: (1) space can go on forever (the infrared (IR)), and (2) in the continuum, in between any two points of space are more points (the ultraviolet (UV)). The former may be familiar from statistical mechanics, where it is associated with the thermodynamic limit, which is where interesting things happen. For our own safety, we’ll begin our discussion in a padded room, protected on both sides from the terrors of the infinite.

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

Peskin and Schroeder, An introduction to quantum field theory (Wiley)
Zee, Quantum Field Theory in a Nutshell (Princeton, 2d Edition)
Le Bellac, Quantum Physics (Cambridge)
Schwartz, Quantum field theory and the standard model (Cambridge)
David Tong’s lecture notes

Many other bits of wisdom come from the Berkeley QFT 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 $dx^\mu \equiv (dt, \vec{d}x)^\mu$, the Lorentz-invariant distance is:

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

with

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

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

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

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

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

$$dk \equiv \frac{dk}{2\pi}.$$

I will also write $\delta^d(q) \equiv (2\pi)^d \delta^{(d)}(q)$. I will try to be consistent about writing Fourier transforms as

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

IFF $\equiv$ if and only if.

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

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

$+O(x^n) \equiv$ 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. When I say ‘Peskin’ I mean ‘Peskin & Schroeder’.

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 arrange themselves in 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:

$$H = \sum_{n=1}^{N} \left( \frac{p^2_n}{2m} + \frac{1}{2}\kappa (q_n - q_{n-1})^2 \right) + \lambda q^4.$$  \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, 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 $k^1$:

$$q_k = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{ikx_n} q_n, \quad p_k = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{ikx_n} p_n,$$

(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 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 $T$ defined by $Tf(x) = f(x + a)$:

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

**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$^2$:

$$0 < k \leq \frac{2\pi}{a} \quad \text{or} \quad -\frac{\pi}{a} < k \leq \frac{\pi}{a}. \quad (1.3)$$

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.

---

$^1$The inverse transformation is:

$$q_n = \frac{1}{\sqrt{N}} \sum_{k>0}^{2\pi/a} e^{-ikx_n} q_k, \quad p_n = \frac{1}{\sqrt{N}} \sum_{k>0}^{2\pi/a} e^{-ikx_n} p_k. \quad (1.2)$$

$^2$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.
The whole Hamiltonian is a bunch of decoupled oscillators, labelled by these funny wave numbers:

$$H = \sum_k \left( \frac{p_k p_{-k}}{2m} + \frac{1}{2} m \omega_k^2 q_k q_{-k} \right) \quad (1.4)$$

where the frequency of the mode labelled $k$ is

$$\omega_k \equiv 2 \sqrt{\frac{\kappa}{m}} \sin \frac{|k|a}{2}. \quad (1.5)$$

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:

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

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 $\omega^2 = \omega_k^2 = \left( \frac{2\kappa}{m} \right) \sin^2 \frac{|k|a}{2}$ for the allowed modes. We see that (the classical version of) this system describes waves:

$$0 = (\omega^2 - \omega_k^2) q_{k,\omega} \quad \forall k \ll 1/a \quad \overset{\approx}{\sim} \quad (\omega^2 - v_s^2 k^2) q_{k,\omega} .$$

The result for small $k$ is the Fourier transform of the wave equation:

$$\left( \partial_t^2 - v_s^2 \partial_x^2 \right) q(x, t) = 0 . \quad (1.6)$$

$v_s$ 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

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

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.

QM. So far the fact that quantumly $[q_n, p_n'] = i\hbar \delta_{nn'} \mathbb{1}$ hasn’t really mattered in our analysis (go back and check – we could have derived the wave equation classically).

In case you are rusty, or forget the numerical factors like I do, here is a concise summary of the operator solution of the quantum harmonic oscillator:

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

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:

$$Q \equiv \left( \frac{m \omega}{\hbar} \right)^{1/2} q, \quad P \equiv \left( \frac{1}{m \hbar \omega} \right)^{1/2} p.$$

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

$$\langle n | a a^\dagger | n \rangle = \langle n | N | n \rangle = n \langle n | n \rangle$$

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 $H = \hbar \omega (N + \frac{1}{2})$) are

$$|0\rangle, \quad |1\rangle \equiv a^\dagger |0\rangle, \quad \ldots |n\rangle = c_n (a^\dagger)^n |0\rangle \ldots$$

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

\[9\]
For the Fourier modes, this implies the commutator
\[ [q_k, p_{k'}] = \sum_{n, n'} U_{k n} U_{k' n'} [q_n, p_{n'}] = i \hbar \| \sum_{n} U_{k n} U_{k' n} = i \hbar \delta_{k, -k'} \| . \]
(In the previous expression I called \( U_{k n} = \frac{1}{\sqrt{N}} e^{i k x_n} \) 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

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

They satisfy
\[ [a_k, a_{k'}^\dagger] = \delta_{k, k'}. \]

In terms of these, the Hamiltonian is
\[ H = \sum_k \hbar \omega_k \left( a_{k}^\dagger a_k + \frac{1}{2} \right) + \frac{p_0^2}{2 m} \]
– it is a sum of decoupled oscillators, and a free particle describing the center-of-mass.

The discovery of Fock space. 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
\[ a_k^\dagger |0\rangle \propto |\text{one phonon with momentum } \hbar k\rangle \quad (1.7) \]
has energy \( \hbar \omega_k \). (This is an eigenstate because \( [N_k, a_k^\dagger] = a_k^\dagger \], where \( N_k = a_k^\dagger a_k \).) It is called a phonon with momentum \( \hbar k \). 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:
\[ |\text{one phonon at position } x\rangle = \sum_k e^{i k x} |\text{one phonon with momentum } \hbar k\rangle \sim \sum_k e^{i k x} a_k^\dagger |0\rangle . \]

This is the state which in single-particle QM we would have called \( |x\rangle \).
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 – for this reason we could also call it the ‘vacuum’. 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 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

$$\left(a_{k_1}^\dagger\right)^{n_{k_1}}\left(a_{k_2}^\dagger\right)^{n_{k_2}} \cdots |0\rangle = |\{n_{k_1}, n_{k_2}, \ldots\}\rangle$$

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 wavenumber. 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. (Further evidence for this is that $a_k^\dagger a_{k'}^\dagger = a_{k'}^\dagger a_k^\dagger$.)

[End of Lecture 1]

**Lessons from the hard work of the first lecture:**

- Starting from a collection of particles, we chained them together, and made a field; treating this system quantumu, 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. The state of such indistinguishable particles is determined merely by specifying a collection of positions (or momenta) and of spin states – we don’t need to say which is which (and in fact, we cannot).

- 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 (roughly $a \to 0$ with the correct quantities held fixed), to which we devolve soon, this maximum (which occurs at $k = \frac{\pi}{a}$) goes off to the sky.
• **Lorentz invariance can emerge.** The dispersion relation for the long-wavelength \((ka \ll 1)\) sound mode was \(\omega^2 = v^2k^2\). This is the fourier transform of

\[
(\partial_t^2 - v^2\vec{\nabla}^2)\phi(x) = \partial_{\mu}\partial^{\mu}\phi(x) = 0,
\]

s a wave equation which has Lorentz symmetry (if \(v\) is the speed appearing in the Minkowski metric). We had to ignore the \(O(a^4k^4)\) terms in the long-wavelength expansion of the dispersion relation, \(2(1 - \cos(ka))\). 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.

**Phonons are real: heat capacity of (insulating) solids.** 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 \(T \to 0\).

**Phonons are real: the Mössbauer effect.** 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 \( E_{\text{recoil}} = \frac{\Delta \vec{p}^2}{2M} \). This can be as big as \( E_{\text{recoil}}^{\max} = \frac{(2\vec{p})^2}{2M} = \frac{(2E_\gamma/c)^2}{2M} \) (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 (its mass is \( NM \) where \( N \) is the number of nuclei!), and therefore does not recoil very much at all (it loses only energy \( \frac{p^2}{2NM} \)).

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

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

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. These expressions are formally identical to the formulae we’ll find below 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 \( \phi \equiv \sqrt{m}q \) would get rid of them. The other difference is because we still have IR and UV regulators in place, so we have sums instead of integrals.

Path integral reminder in a box.

If we use the path integral description, some of these things (in particular the continuum, sound-wave limit) are more obvious-seeming.

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

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

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

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

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

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

\[ \mathbb{1} = \mathbb{1}^2 = \left( \int dp \langle p \rangle \right) \left( \int dq \langle q \rangle \right) \]
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.** [Zee §1.3] The path integral for our collection of oscillators is

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

with \( S[q] = \int dt \left( \sum_n \frac{1}{2} m_n q_n^2 - V(\{q\}) \right) \equiv \int dt L(q, \dot{q}). \) The potential is \( V(\{q\}) = \sum_n \frac{1}{2} \kappa (q_{n+1} - q_n)^2. \) Now suppose we have poor eyesight and can’t resolve the individual atoms in the chain; rather we’re only interested in the long-wavelength (small-wavenumber) physics. So let’s try to take the continuum limit \( a \to 0, N \to \infty. \)

Basically the only thing we need is to think of \( q_n = q(x = na) \) as defining a smooth function:

![Diagram of continuum limit](Diagram.png)

[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’ll use

\[ (q_n - q_{n-1})^2 \sim a^2 (\partial_x q)^2 \big|_{x = na}, \quad a \sum_n f(q_n) \sim \int dx f(q(x)). \]

The path integral becomes:

\[ Z = \int [dq] e^{iS[q]} \]

with \([dq]\) now representing an integral over all configurations \( q(t, x) \) (defined by this limit) and

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

where I’ve introduced some parameters \( \mu, v_s, r, u \) determined from \( m, \kappa, a \ldots \) in some ways that we needn’t worry about, except to say that they are finite in the continuum limit. The \( \cdots \) includes terms like \( a^4 (\partial_x q)^4 \) which are small when \( k \ll \frac{1}{a} \), so we ignore them. \( L \) 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 \( \frac{\partial}{\partial q_n} \).  

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 \( q_n \rightarrow q_n + \epsilon \). 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 \( \pi(x) = \frac{\partial L}{\partial \dot{q}} = \mu \dot{q} \) and doing the Legendre transformation:

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

(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 we’ll discuss next. (Note that vibrations of a crystal in three dimensions actually do involve vector indices! We omit this complication.)

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

---

7 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

\[
\phi_n = \phi(x_n)
\]

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.
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

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

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

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

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

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

(1.11)

This is almost what we would have found for the long-wavelength \((ka \ll 1)\) description of a \( d \)-dimensional lattice of masses on springs, like a mattress (except that there would have been one \( \phi \) for each direction in which the atoms can wiggle). The equation of motion is

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

(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,

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

(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 \( \mathcal{L} = \frac{1}{2} (\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2) \). This describes free continuum real massive relativistic scalar quantum field theory. (Match the adjectives to the associated features of the lagrangian; collect them all!)

The canonical momentum is \( \pi = \frac{\partial L}{\partial \dot{\phi}} = \dot{\phi} \) and the Hamiltonian (which we can instantly promote to a quantum operator by using boldface symbols) is then

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

Note that all these terms are positive.
A translation invariant linear problem is solved by Fourier transforms: \( \phi(x) = \int d^d k \, e^{-i\vec{k} \cdot \vec{x}} \phi_k, \) and \( \pi(x) = \int d^d k \, e^{-i\vec{k} \cdot \vec{x}} \pi_k, \) this is

\[
H = \int d^d k \left( \frac{1}{2} \pi_k \pi_{-k} + \frac{1}{2} \left( v_s^2 k^2 + m^2 \right) \phi_k \phi_{-k} \right)
\]

where \( k^2 = ( -i \vec{k} ) \cdot ( i \vec{k} ) = \vec{k} \cdot \vec{k}. \) 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

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

Their commutator follows from \([\phi(x), \pi(y)] = i\hbar \delta^{(d)}(x - y)\). In terms of the ladder operators,

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

The field operators

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

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

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 \( a^\dagger_k \), since only the combination \( a_k + a^\dagger_{-k} \) multiplies \( e^{i\vec{k} \cdot \vec{x}} \). The

---

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.
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

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

(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:

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

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 might be appropriate here, but it will be clearer later when we talk about examples of particles that are not their own antiparticles.)

[End of Lecture 2]

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 $e^{-i\vec{P} \cdot \vec{x}}$. Here $\vec{P}$ is the last in a long list of objects in this section 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 it 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

$$\phi(x^\mu) = e^{i\vec{P}_\mu x^\mu} \phi(0) e^{-i\vec{P}_\mu x^\mu}$$

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

19
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 (with $c = 1$) are:

\[ \varepsilon_{\mu\nu\rho\sigma} \partial_\nu F_{\rho\sigma} = 0 \]
\[ \partial^\mu F_{\mu\nu} = 4\pi j_\nu \]
\[ \vec{\nabla} \cdot \vec{B} = 0, \quad \vec{\nabla} \times \vec{E} = -\partial_t \vec{B}, \quad (1.16) \]
\[ \vec{\nabla} \cdot \vec{E} = 4\pi \rho, \quad \nabla \times \vec{B} = \partial_t \vec{E} + 4\pi \vec{j} \quad (1.17) \]

(where the familiar electric and magnetic fields are $E_i = -F_0^i$ and $\varepsilon^{ijk} B^k = -F^{ij}$).

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 $A_\mu = (\Phi, \vec{A})_\mu$ which determine the fields by taking derivatives and which automatically solve the constraints (1.16):

\[ F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \]
\[ \text{aka} \quad \vec{E} = -\vec{\nabla} \Phi - \partial_t \vec{A}, \quad \vec{B} = \vec{\nabla} \times \vec{A}. \]

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

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

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 $\nabla \cdot \vec{A} = 0$. In fourier space $\vec{A}(x) = \sum_k e^{i\vec{k} \cdot \vec{x}} \vec{A}(k)$ this condition is

$$0 = \vec{k} \cdot \vec{A}(k)$$

– the vector field is *transverse*.

An action which gives rise to Maxwell’s equations is

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

Note that we must regard $A$ as the dynamical variable to obtain (1.17) by $0 = \delta S / \delta A^\mu(x)$. The canonical momentum of $A$ is then $\Pi_A^i = \partial \mathcal{L}_{\text{Maxwell}} / \partial \dot{A}^i = E^i$. So the Hamiltonian is\(^9\):

$$H = \frac{1}{2} \int d^3 x \left( \vec{E}^2 + c^2 \vec{B}^2 \right). \quad (1.18)$$

Here $\vec{E} = -\partial_t \vec{A}$ plays the role of field momentum $\pi(x)$ in (1.10), and $\vec{B} = \nabla \times \vec{A}$ 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

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

where $i, j = 1..3$ are spatial indices\(^10\). So we can immediately write down an expression for the quantum Maxwell field in terms of independent creation and annihilation operators:

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

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

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

\(^9\)You may also recall that the energy density of a configuration of Maxwell fields is $u = \frac{1}{2} \left( \vec{E}^2 + \vec{B}^2 \right)$. This result can be obtained either by Legendre transformation of $\mathcal{L}_{\text{Maxwell}}$, or from $T^0_0$, the energy momentum tensor as you’ll see on the HW.

\(^10\)As a check, note that using this Hamiltonian and the canonical commutator, we can reproduce Maxwell’s equations using Ehrenfest’s theorem:

$$\left\langle \partial_t^2 \vec{A} \right\rangle = \partial_t \left\langle \vec{E} \right\rangle = -\frac{i}{\hbar} \left\langle [H, \vec{E}] \right\rangle = \left\langle c^2 \nabla^2 \vec{A} \right\rangle.$$
Also, the magnetic field operator is
\[
\vec{B} = \nabla \times \vec{A} = \sum_k \sum_s \sqrt{\frac{\hbar}{2\epsilon_0\omega_k L^3}} i\vec{k} \times \left( a_{k,s}(\hat{k}) e^{i\vec{k} \cdot \vec{r}} - a_{k,s}^\dagger(\hat{k}) e^{-i\vec{k} \cdot \vec{r}} \right);
\]
the magnetic field is analogous to \( \nabla \phi \) in the scalar field theory\(^{11}\). Plugging these expressions into the Hamiltonian (1.18), we can write it in terms of these oscillator modes (which create and annihilate \textit{photons}). As for the scalar field, the definitions of these modes were designed to make this simple: It is:
\[
H = \sum_{k,s} \hbar \omega_k \left( a_{k,s}^\dagger a_{k,s} + \frac{1}{2} \right).
\]

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

The vacuum energy is
\[
E_0 = \frac{1}{2} \sum_{k,s} \hbar \omega_k = \frac{L^3}{(2\pi)^3} \int d^3khck.
\]
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.

\(^{11}\)I should say a little more about the \textit{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 polarization vectors of a given \( \vec{k} \) can be chosen to satisfy the following completeness relation:
\[
\sum_s e_{s_i}(\hat{k}) e_{s_j}^\dagger(\hat{k}) = \delta_{ij} - \hat{k}_i \hat{k}_j. \tag{1.19}
\]
This says that they span the plane perpendicular to \( \hat{k} \).
**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 \( \text{Re} \equiv \frac{1}{2} (A + A^\dagger) \) as usual. The normalization constant is \( \mathcal{N} = \frac{1}{2} \sqrt{\frac{\hbar}{2m\omega}} \).

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

\[
[q, p] = i\hbar \quad \Rightarrow \quad [a, a^\dagger] = \mathbb{1}.
\]

\[
q = \text{Re} \mathcal{N} a, \quad p = m \text{Im} \omega \mathcal{N} a.
\]

\[
\mathbf{H}_{\text{scalar}} = \int dx \left( \frac{1}{2\mu} \pi^2 + \frac{1}{2} \mu c^2 (\partial_x \phi)^2 \right) = \sum_k \hbar \omega_k \left( a_k^\dagger a_k + \frac{1}{2} \right)
\]

\[
[\phi(x), \pi(x')] = i\hbar \delta(x - x') \quad \Rightarrow \quad [a_k, a_{k'}^\dagger] = i\hbar \delta_{kk'}.
\]

\[
\phi(x) = \text{Re} \left( \sum_k \mathcal{N}_k e^{ikx} a_k \right), \quad \pi(x) = \mu \text{Im} \left( \sum_k \omega_k \mathcal{N}_k e^{ikx} a_k \right).
\]

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

\[
[A_i(x), E_j(x')] = i\hbar \delta^3(x - x') \delta_{ij} \quad \Rightarrow \quad [a_{ks}, a_{k's}^\dagger] = \hbar \delta_{kk'} \delta_{ss'}.
\]

\[
\mathbf{A}(x) = \text{Re} \left( \sum_k \mathcal{N}_k e^{ikx} a_{ks} \bar{e}_s(k) \right), \quad \mathbf{E}(x) = \mu \text{Im} \left( \sum_k \omega_k \mathcal{N}_k e^{ikx} a_{ks} \bar{e}_s(k) \right).
\]

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 got 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 \( p_\mu = \hbar k_\mu \) have mass dimension +1. The space and time coordinates \( x^\mu \) have mass dimension -1. The action goes in the exponential of the path integral measure \( \int [D\phi] e^{\frac{iS}{\hbar}} \) and so must be dimensionless.

So the Lagrangian density has mass dimension \( d+1 \). This means that the KG field has mass dimension \( \frac{d-1}{2} \) (and the mass \( m \) has mass dimension 1 (yay!)). In \( d+1 = 3+1 \) dimensions, \( E \sim \dot{A}, B \sim \nabla A \) have mass dimension 2 and \( A \) has mass dimension one (and \( e \) is dimensionless). This is nice because then the covariant derivative \( \partial_\mu + A_\mu \) has mass dimension one. Notice that \( E^2 + B^2 \) has dimension 4 which is good for an energy per unit volume.

The equation of motion is

\[
0 = \frac{\delta S}{\delta \phi_r(x)}.
\]
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. Note that since we are interested here in the bulk equations of motion, we ignore boundary terms unless we are interested in field theory on a space with boundary. That is a worthy subject but an unnecessary complication for now.

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

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

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 next 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

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial (\partial_t \phi(x))}.$$  

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 $\frac{\partial}{\partial \phi(x)} \mathcal{L} = d^d x \pi(x)$ (as opposed to $\mathcal{L}$). I will often forget to say ‘density’ here.

The hamiltonian is then

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

**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 trans-
formation 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:

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

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):

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

where in the last step we used Stokes’ theorem.

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

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

but on the other hand, by Taylor expansion,

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

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

$$j^\mu = \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_r)} \Delta \phi_r - \mathcal{J}^\mu.$$  \hfill (1.20)

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

$$\delta \phi \equiv i \epsilon [Q, \phi].$$  \hfill (1.21)

We’ll say that $Q$ generates the symmetry, for the following reason. (1.21) is the infinitesimal version of the finite transformation

$$\phi \rightarrow \phi' \equiv e^{iQ} \phi e^{-iQ}.$$  

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

Some examples will be useful:

- For example, suppose $S[\phi]$ only depends on $\phi$ through its derivatives, for example, $S[\phi] = \int \frac{1}{2} \partial_\mu \phi \partial^\mu \phi$. Then there is a shift symmetry $\phi \rightarrow \phi' \equiv \phi + \epsilon$. Letting $\epsilon$ depend on spacetime, the variation of the action is $S[\phi + \epsilon(x)] = \int \epsilon \partial_\mu \partial^\mu \phi$, 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

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

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

- On the homework you’ve seen a complex scalar $\Phi$, with $S[\Phi, \Phi^\star]$ is invariant under $\Phi \rightarrow e^{i\kappa} \Phi = \Phi + i \epsilon \Phi + \mathcal{O}(\epsilon^2)$. This $U(1)$ phase transformation can be rewritten
in terms of the real and imaginary parts as an $\text{SO}(2)$ rotation. The charge can be written as
\[ Q = \int d^d x j^0 = \int d^d p \left( a^\dagger_p a_p - b^\dagger_p b_p \right) \]
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 solutions.) So the particles created by $a$ and $b$ have opposite charge (this is essential given the mode expansion $\Phi_k \sim a_k + b^\dagger_{-k}$) and can be interpreted as each others' antiparticles: there can be symmetry-respecting processes where an $a$ particle and $b$ particle take each other out.

- Consider spacetime 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 $0 \neq \frac{d}{dx^\mu} L \overset{\text{chain rule}}{=} \partial_\nu \phi \frac{\partial L}{\partial \partial_\nu \phi} + \partial_\mu \partial_\nu \phi \frac{\partial L}{\partial \partial_\mu \phi}$). Let’s use the prosaic method for this one: the shift in the Lagrangian density also can be found by Taylor expansion
\[ L \mapsto L + a^\mu \frac{d}{dx^\mu} L = L + a^\nu \partial_\nu (\delta^\mu_\nu L). \]
So the formula (1.20) gives
\[ T^\mu_\nu = \frac{\partial L}{\partial (\partial_\mu \phi)} \partial_\nu \phi \overset{\Delta \phi}{\Delta \phi} - L \delta^\mu_\nu. \]
For the time translation, the conserved charge $T^0_0$ gives back the hamiltonian density $\mathfrak{h} = \pi \dot{\phi} - L$ 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 $|\vec{k}_1, \ldots, \vec{k}_n\rangle$ 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

$$P_i = -\int d^d x \, \pi \partial_i \phi = \int d^d k \, k_i a^\dagger_{\vec{k}} a_{\vec{k}}$$

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:

\[
\begin{array}{c}
\vdash d \rightarrow \vdash L - d \rightarrow \\
\end{array}
\]

(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

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

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 \( \frac{1}{2}\hbar \omega s \) 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

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

We have done something wrong. What?

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 \( \omega_j \gg \frac{\pi}{a} \) high frequency, where \( a \) is some short time\(^{12}\). Replace

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

\]

\(^{12}\)You can think of \( a \) as the time it takes the waves to move by one lattice spacing. If we work in units where the velocity is \( c = 1 \), this is just the lattice spacing. I will do so for the rest of this discussion.
\[ F = -\partial_d E_0 = -(f'(d) - f'(L - d)) \]

\[ \approx -\frac{\pi \hbar}{2d} \left( \frac{1}{d^2} - \frac{1}{(L - d)^2} \right) \]

\[ \approx \frac{\pi \hbar c}{24d^2} (1 + O(d/L)) . \]
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

$$ \sum_{n,\vec{k}}' \equiv \frac{1}{2} \sum_{n=0,\vec{k}} + \sum_{n\geq1,\vec{k}} $$

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

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

2 The path integral makes some things easier

2.1 Fields mediate forces

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

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

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

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

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

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

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

To do this integral, we retreat to the case where the action is quadratic in $\phi$, so that

$$\mathcal{L}(\phi) = \frac{1}{2} \left( \partial_{\mu} \phi \partial^{\mu} \phi - m^2 \phi^2 \right) \overset{\text{IBP}}{=} -\frac{1}{2} \phi \left( \partial^2 + m^2 \right) \phi + \text{total derivative.} \quad (2.1)$$

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

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

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

$$- (\partial^2 + m^2) D(x - y) = \delta(x - y). \quad (2.2)$$
This equation says that $D$ is a *Green’s function* for the operator $-(\partial^2 + m^2)$. The fact that there is no special point in spacetime says $A_{xy}^{-1} = D(x - y)$ only depends on the difference of its arguments.

Does this integral actually converge? The integral you saw on the homework was of the form $\int dqe^{-\frac{1}{2}qAq}$, which surely converges if $A$ is a positive matrix. Actually, this is overkill — it is enough to replace $m^2 \to m^2 - i\epsilon$ to make all the integrals converge. Here $\epsilon$ is an *infinitesimal*, which means $\epsilon^2 = 0$ and $c\epsilon = \epsilon$ for any positive $c$. Then each $\int dq_{\mu\nu}$ will have a factor of $e^{-\epsilon\int q_{\mu\nu}}$ which suppresses the integrand in the dangerous large-field region.

[End of Lecture 4]

The equation (2.2) is translation-invariant and linear so you should not be surprised that it is solved by going to Fourier space (in space and time):

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

in terms of which (2.2) becomes the algebraic equation $1 = (k^2 - m^2 + i\epsilon) D_k$. Hence

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

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

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

is zero when

$$\omega = \pm \sqrt{\bar{k}^2 + m^2 - i\epsilon} = \pm (\omega_k - i\epsilon), \quad \omega_k \equiv \sqrt{\bar{k}^2 + m^2}.$$ 

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

We can then do the $\omega$ integral by contours: if $t > 0$ ($t < 0$), we can close the contour in the UHP (LHP) since the integrand goes like $e^{-\operatorname{Im}\omega t}$, and the integral equals

---

13 You might notice that I am swindling you here: there could be other ways to make the integral well-defined (there are). Another thing you might be bothered by is the boundary conditions on the fields and their relation to the initial and final states. In the next subsection, we’ll say more.

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

---

34
the residue of the pole at $\omega = \omega_k \mp i \epsilon$ (times $2\pi i$):

$$D(x) = -i \int d^d k \left( \theta(t) \frac{e^{-i(\omega_k t - \vec{k} \cdot \vec{x})}}{2\omega_k} + \theta(-t) \frac{e^{i(\omega_k t - \vec{k} \cdot \vec{x})}}{2\omega_k} \right). \quad (2.3)$$

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

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

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

Putting back all the labels, the same manipulations show that

$$D(x - y) \equiv \langle 0| \phi(x)\phi(y)|0 \rangle \quad (2.4)$$

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

(Notice that if the system is Lorentz invariant (which starting from (2.1) it is) then since $D(x)$ is a scalar quantity, it can only depend on $x$ through Lorentz invariants made from $x^\mu$, namely the proper distance $x^2 = t^2 - \vec{x}^2$, and the sign of $t$.)

Why the ‘?’ in (2.4)? For one thing, $\phi(x)$ and $\phi(y)$ are operators – the order matters. How do I know which order in which to write them? To reproduce (2.3) the thing to do is to time-order them:

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

To see this, plug in the mode expansion (1.14) to see e.g.

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

(where $k^0 = \omega_k$, $q^0 = \omega_q$ to satisfy the KG equation), which reproduces the first term in (2.3)\textsuperscript{15}.

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

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

\textsuperscript{15}The other ways of making the path integral well-defined correspond to other ways of ordering the $\phi$s, and other initial and final states.
Here \( J(x) = \int d^{d+1} k e^{i k x} J_k \), \( J_k = J_{-k} \). We get to pick \( J(x) \). Choose \( J = J_1 + J_1 \) to describe (in Zee’s words) two lumps sitting still on the mattress: \( J_a(x) = \delta^3(x-x_a), a = 1,2 \). Then \( J_k = \int d x^0 e^{-i k^0 x^0} \left( e^{i k x_1} + e^{i k x_2} \right) \). The interaction between the two lumps mediated by the mattress field \( \phi \) will then be described by the \( J_1 J_2 \) cross-terms in \( W[J] \):

\[
W[J] = -\frac{1}{2} \int d x^0 \int d y^0 \int d k^0 e^{i k^0 (x^0-y^0)} \int d^3 k \frac{e^{i k (\vec{x}_1 - \vec{x}_2)}}{k^2 - m^2 + i \epsilon} + ..
\]

Notice that we can drop the \( i \epsilon \) now, because this integrand is nonsingular for real \( \vec{k} \).
In \( d = 1 \), there are poles at \( k = \pm i m \), and we can close the contour in the UHP for free to get\(^{16}\)

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

Notice that we can drop the \( i \epsilon \) now, because this integrand is nonsingular for real \( \vec{k} \). In \( d = 1 \), there are poles at \( k = \pm i m \), and we can close the contour in the UHP for free to get\(^{16}\)

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

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

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

which falls off exponentially with the separation between the lumps. The range of the potential goes inversely with the mass of the ‘force carrier’ \( \phi \). The 3d version of this potential \( e^{-m x} \) (see footnote \(^{16}\)) is called the Yukawa potential.

\(^{16}\)For convenience, here’s the integral in 3d:

\[
\int d^3 k \frac{e^{i \vec{k} \cdot \vec{x}}}{k^2 + M^2} \equiv \cos \theta = \frac{1}{(2 \pi)^2} \int_0^\infty \frac{k^2 dk}{k^2 + M^2} \int_{-1}^1 dy e^{i k y r} = \frac{1}{(2 \pi)^2 r} \int_{-\infty}^{\infty} dk k \sin k r = \frac{1}{2 \pi} \int_{-\infty}^{\infty} dk \frac{k e^{i k r}}{k^2 + M^2} \qquad \text{close contour in UHP for free}
\]

\[
= \frac{1}{(2 \pi)^2 r} \frac{1}{2i} \int_{-\infty}^{\infty} dk \frac{k e^{i k r}}{k^2 + M^2 + h.c.} = \frac{1}{(2 \pi)^2 r} \frac{1}{2i} \int_{-\infty}^{\infty} \left( \frac{1}{2 i M} \right) e^{i k r} = \frac{1}{4 \pi r}.
\]
2.2 Euclidean path integral and Wick rotation

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

For simplicity, let us focus on a single mode of the field – a single harmonic oscillator

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

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

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

Then the path integral becomes

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

This integrand suppresses configurations with large $q$, and large $\partial_\tau q$, and the integral is therefore totally well-defined. The euclidean action is\(^{17}\)

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

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

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

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

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

and we have

$$G(\sigma) = \int d\omega \frac{e^{i\omega \sigma}}{\omega^2 + \Omega^2} = e^{-\Omega|\sigma|} \frac{1}{2\Omega}. \quad (2.8)$$

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

\(^{17}\)It is called euclidean because the $(\partial_\tau q)^2$ has the same sign as the spatial derivatives $(\partial_x q)^2$, so this is the action we get in euclidean spacetime with metric $\delta_{\mu\nu}$, rather than $\eta_{\mu\nu}$. Exercise: put back the spatial derivative terms and check that this is the case.
I claim that the real-time calculation which keeps the oscillator in its groundstate is the analytic continuation of the one we did above, where we replace

\[ \omega_{\text{Mink}} = e^{-i(\pi/2-\epsilon) \omega_{\text{above}}} \]

where \( \epsilon \) is a familiar infinitesimal. In the picture of the euclidean frequency plane in Fig. 1, this is a rotation by nearly 90 degrees. We don’t want to go all the way to 90 degrees, because then we would hit the poles at \( \pm i\Omega \). The replacement (2.9) just means that if we integrate over real \( \omega_{\text{Mink}} \), we rotate the contour in the integral over \( \omega \) as follows:

\[ \omega_{\text{eucl}} = -i \omega_{\text{Mink}}, \quad \omega_{\text{eucl}} t_{\text{eucl}} = \omega_{\text{Mink}} t_{\text{Mink}}, \quad t_{\text{eucl}} = +i t_{\text{Mink}}. \] 

So this is giving us a contour prescription for the real-frequency integral. The result is the Feynman contour, and it is the same as what we got from \( m^2 \rightarrow m^2 - i\epsilon \):
Figure 2: The Feynman contour in the $\omega_{\text{Mink}}$ complex plane.

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

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

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

because of (2.10). This means the path integrand is $e^{-S_{\text{eucl}}} = e^{iS_{\text{Mink}}}$.

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

$$
\int_{q(\beta)=q_0} [dq] e^{-S[q]} = \langle q_0 | e^{-\beta H} | ... \rangle = \psi_{\text{gs}}(q_0)
$$

(2.11)

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

More slowly: the path integral representation for the real-time propagation amplitude is

$$
\langle f | e^{-iHt} | i \rangle = \int_{f\leftarrow i} [dq] e^{\int_{0}^{t} dt L}.
$$

On the RHS here, we sum over all paths between $i$ and $f$ in time $t$ (i.e. $q(0) = q_i$, $q(t) = f$).
$q_f$, weighted by a phase $e^{i \int dt L}$. But that means you also know a representation for

$$\sum_f \langle f | e^{-\beta H} | f \rangle \equiv \text{tr} e^{-\beta H}$$

– namely, you sum over all periodic paths $q_i = q_f$ in imaginary time $t = -i \beta$. So:

$$Z(\beta) = \text{tr} e^{-\beta H} = \oint [dq] e^{-\int_0^\beta dr L} = \oint [dq] e^{-S_{\text{eucl}}[q]}$$

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

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

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

But we found a special one which arises by analytic continuation from the euclidean Green’s function, which is unique\(^1\). It is

$$G(s, t) = \langle \mathcal{T} (q(s) q(t)) \rangle ,$$

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

Let’s pursue this one more step. The same argument tells us that the generating functional for real-time correlation functions of $Q$ is

$$Z[J] = \langle \mathcal{T} e^{i \int J q} \rangle = \langle 0 | \mathcal{T} e^{i \int J q} | 0 \rangle .$$

\(^1\) Another important perspective on the uniqueness of the euclidean Green’s function and the non-uniqueness in real time: in euclidean time, we are inverting an operator of the form $-\partial_x^2 + \Omega^2$ which is positive (≡ all its eigenvalues are positive) – recall that $-\partial_x^2 = \hat{p}^2$ is the square of a hermitian operator. If all the eigenvalues are positive, the operator has no kernel, so it is completely and unambiguously invertible. This is why there are no poles on the axis of the (euclidean) $\omega$ integral in (2.8). In real time, in contrast, we are inverting something like $+\partial_t^2 + \Omega^2$ which annihilates modes with $\partial_t = i\Omega$ (if we were doing QFT in $d > 0 + 1$ this equation would be the familiar $p^2 - m^2 = 0$). These are called ‘on-shell states’, they are actual states in the spectrum of the Hamiltonian of the system. So the operator we are trying to invert has a kernel and this is the source of the ambiguity. In frequency space, this is reflected in the presence of poles of the integrand on the contour of integration; the choice of how to negotiate them encodes the choice of Green’s function.
In the second step I just emphasized that the real time expectation value here is really a vacuum expectation value. This quantity has the picturesque interpretation as the vacuum persistence amplitude, in the presence of the source \( J \).

So we see that in general, the correlation functions that are computed by this “\( \mathbf{i} \epsilon \) prescription” of Wick rotating from Euclidean spacetime are time-ordered:

\[
\frac{1}{Z} \int [D\phi] e^{iS_{m^2 \rightarrow m^2 - \mathbf{i} \epsilon} f(\phi)} = \langle 0 | \mathcal{T} f(\phi) | 0 \rangle .
\]

Causality. In other treatments of this subject, you will see the Feynman contour motivated by ideas about causality. We'll discuss this in §3. This was not the logic of our discussion but it is reassuring that we end up in the same place. Note that even in 0+1 dimensions there is a useful notion of causality: effects should come after their causes.

[End of Lecture 5]

2.3 Feynman diagrams from the path integral

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

Suppose we want to do the integral

\[
Z(J) = \int_{-\infty}^{\infty} dq \, e^{-\frac{1}{2}m^2q^2 + Jq} = \int dq \, e^{-S(q)} . \quad (2.12)
\]

It is the path integral for \( \phi^4 \) theory with fewer labels. For \( g = 0 \), this is a gaussian integral which we know how to do. For \( g \neq 0 \) it’s not an elementary function of its arguments. We can develop a (non-convergent!) series expansion in \( g \) by writing it as

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

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

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

So:

\[
Z(J) = \sqrt{\frac{2\pi}{m^2}} e^{-\frac{g}{4!} \left( \frac{\partial}{\partial J} \right)^4 e^{\frac{1}{2} J \frac{1}{m^2} \mathcal{J}}} .
\]
This is a double expansion in powers of $J$ and powers of $g$. The process of computing the coefficient of $J^n g^m$ can be described usefully in terms of diagrams. There is a factor of $1/m^2$ for each line (the propagator), and a factor of $(-g)$ for each 4-point vertex (the coupling), and a factor of $J$ for each external line (the source). For example, the coefficient of $g J^4$ comes from:

$$\sim \left( \frac{1}{m^2} \right)^4 g J^4.$$

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

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

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

In perturbation theory this is:

$$G \simeq 1 + \frac{1}{2} g m^{-4} + \frac{2}{3} g^2 m^{-8} + O(g^3)$$

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

42
Brief comments about large orders of perturbation theory.

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

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

- The integral actually does have a name – it’s a Bessel function:

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

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

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

\[
Z(0) \simeq \sqrt{\frac{2\pi}{m^2}} \sum_n \frac{(-1)^n 2^{2n+\frac{1}{2}}}{n! (4!)^n} \Gamma \left( 2n + \frac{1}{2} \right) \left( \frac{g}{m^4} \right)^n. \tag{2.14}
\]

- The expansion for \( G \) is of the form

\[
G \simeq m^{-2} \sum_{n=0}^{\infty} c_n \left( \frac{g}{m^4} \right)^n.
\]

When \( n \) is large, the coefficients satisfy \( c_{n+1} \gtrsim 1 - \frac{2}{3} n c_n \) (you can see this by looking at the coefficients in (2.14)) so that \(|c_n| \sim n!\). This factorial growth of the number of diagrams is general in QFT and is another way to see that the series does not converge.

- The fact that the coefficients \( c_n \) grow means that there is a \textit{best} number of orders to keep. The errors start getting bigger when \( c_{n+1} \left( \frac{g}{m^4} \right) \sim c_n \), that is, at order \( n \sim \frac{3m^4}{2g} \). So if you want to evaluate \( G \) at this value of the coupling, you should stop at that order of \( n \).
I said above that the fact that the perturbation series doesn’t converge means that it doesn’t define the field theory amplitudes. What does it miss? To answer this, consider trying to do the integral (2.12) by saddle point (at \( J = 0 \) for simplicity):

\[
0 = S'(q_\star) = m^2 q_\star + \frac{g}{3!} q_\star^3.
\]

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

\[
q_\star = 0, \quad q_\star = \pm i \sqrt{\frac{3! m^2}{g}}.
\]

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

\[
S\left(q_\star = \pm i \sqrt{\frac{3! m^2}{g}}\right) = -\frac{3 m^4}{2 g},
\]

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

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

\[
B(z) \equiv \sum_{m=0}^{\infty} \frac{c_m}{n!} z^m
\]

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

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

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

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

This procedure requires both that the series in \( B(z) \) converges and that the Laplace transform can be done. In fact this procedure works in this case.
The fact that the number of diagrams at large order grows like \( n! \) is correlated with the existence of saddle-point contributions to \( Z(g) \) which go like \( e^{-a/g} \). This is because they are associated with singularities of \( B(z) \) at \( z = a \); such a singularity means the sum of \( c_n z^n \) must diverge at \( z = a \). (More generally, non-perturbative effects which go like \( e^{-a/g^{1/p}} \) (larger if \( p > 1 \)) are associated with (faster) growth like \( (pn)! \). See this classic work.)

- The function \( G(g) \) can be analytically continued in \( g \) away from the real axis, and can in fact be defined on the whole complex \( g \) plane. It has a branch cut on the negative real axis, across which its discontinuity is related to its imaginary part. The imaginary part goes like \( e^{-\frac{a}{|g|}} \) near the origin and can be computed by a tunneling calculation like (2.15).

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

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

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

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

### 3 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 \( |\vec{p}_1, \ldots, \vec{p}_n\rangle \propto a^\dagger_{\vec{p}_1} \cdots a^\dagger_{\vec{p}_n} |0\rangle \) 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 \( |p \rangle \equiv c_p a_p^\dagger |0 \rangle \). 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 \( c_p = 1 \). Then

\[
\langle \vec{k}'|\vec{p} \rangle \equiv \langle 0|a_{\vec{k}}a_{\vec{p}}^\dagger|0 \rangle = \langle 0|[a_{\vec{k}}, a_{\vec{p}}^\dagger]|0 \rangle = (2\pi)^{d} \delta^{(d)}(\vec{k} - \vec{p}) \equiv \delta(\vec{k} - \vec{p}).
\]

Suppose the previous equation is true in my rest frame \( F \). Since \( 1 = \int d^d p \delta(p - k) \), we see that \( d^d p \delta(p - k) \) 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, \( p'_\mu = \Lambda^\mu_\nu p_\nu \),

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

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

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

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

There is an easy fix:

\[ |\vec{p} \rangle \equiv \sqrt{2\omega_p a_{\vec{p}}^\dagger} |0 \rangle. \]

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

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

46
while
\[
\langle \vec{k}' | \vec{p}' \rangle = 2 \omega_p \frac{\omega_p'}{\omega_p} \delta^{(d)}(k' - p').
\]

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, \( \mathbb{1}_1 \equiv \sum_i |i\rangle \langle i| \). I claim that the following expression does this and makes Lorentz symmetry manifest:

\[
\mathbb{1}_1 \equiv \frac{1}{\sqrt{2}} \int d^d k \theta(k^0) \frac{1}{2\omega_k} \delta^{(d)}(k^2 - m^2) \left| k \right\rangle \left\langle k \right|
\]

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

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

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

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

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

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

A second reason to study commutators is ...

### 3.1 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\(^{19}\), call it \( B \), localized near \( B \), to the shared state of their quantum many body system \( |\psi\rangle_{ABE} \). (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 \( [B, A(t^\eta)] = 0 \) 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\(^{20}\). [End of Lecture 6]

So nonvanishing commutators are essential for sending information. Notice that entangled states of well-separated local degrees of freedom (such as a Bell pair of distant spins \(|\uparrow_0\rangle - |\downarrow_0\rangle \in \mathcal{H}_0 \otimes \mathcal{H}_r \)) are insufficient for sending signals: whatever we do to the spin at 0 has no effect on the outcomes of measurements of the spin at \( r \). Here’s a proof: the spin operators at 0, \( S_0^i \) and the spin operators at \( r \), \( S_r^j \) commute: \([S_0^i, S_r^j] = 0\), since \( S_0^i \) acts the identity on \( \mathcal{H}_r \) and vice versa. The groundstate of a quantum field is somewhat similar: there is quite a bit of entanglement between nearby points in space, but the equal-time commutators \([\phi(\vec{x}), \phi(\vec{y})] = 0\) vanish, so it can’t be used to send instantaneous signals.

---

**Causality in relativistic QFT.** [Peskin §2.4] In a Lorentz invariant system, ‘precede’ is sometimes a frame-dependent notion. If \( A \) is in the future lightcone of \( B \),\(^{19}\) Applying an operator’ is more complicated than it seems. Actually it means ‘changing the Hamiltonian so that the time evolution operator is \( B' \).

\(^{20}\) 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 \( p(x) = \sqrt{2a^3/\pi} \) 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

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

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 Chuncheong Lam for help with this point.

---

\(^{19}\) Applying an operator’ is more complicated than it seems. Actually it means ‘changing the Hamiltonian so that the time evolution operator is \( B' \).

\(^{20}\) 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 \( p(x) = \sqrt{2a^3/\pi} \) 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

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

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 Chuncheong Lam for help with this point.
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 in which they occur at the same time, and frames where they occur in either order. This is the dangerous case.

So: causality will follow if, for all local operators \(A, B\), \([A(x_A), B(x_B)] = 0\) 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 statement will follow from considering commutators of

\[
\phi(x) = \int \frac{d^d \hat{p}}{\sqrt{2\omega_{\hat{p}}}} \left( a_\mu e^{-i p_\mu x^\mu} + a_\mu^+ e^{i p_\mu x^\mu} \right) | p^\mu = \omega_{\hat{p}} \equiv \phi^{(+)}. 
\]

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 \([a, a^\dagger] = 1\),

\[
[\phi(x), \phi(y)] = \int \frac{d^d \hat{p}}{2\omega_{\hat{p}}} \left( e^{-i p_\mu (x-y)^\mu} - e^{i p_\mu (x-y)^\mu} \right) = \int d^{d+1} \hat{p} 2\pi \delta(p^2 - m^2) \theta(p^0) \left( e^{-i p_\mu (x-y)^\mu} - e^{i p_\mu (x-y)^\mu} \right) \tag{3.1}
\]

Here comes the slippery stuff: Suppose \(x - y\) is spacelike. Let’s choose a frame where they (the points labelled by \(x, y\)) 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 \(\hat{p}^\mu = \Lambda^\mu_\nu \tilde{p}^\nu\), so that \(p_\mu (x-y)^\mu = \tilde{p}_\mu \tilde{x}^\mu\). Then

\[
[\phi(x), \phi(y)] \overset{(x-y)^2 < 0}{\approx} \int d^{d+1} \hat{p} 2\pi \delta(p^2 - m^2) \theta(p^0) \left( e^{-i \hat{p}^\mu \Delta x} - e^{+i \hat{p}^\mu \Delta x} \right) = 0. \tag{3.2}
\]

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 \((x - y)^2 < 0\) is important.

So: vanishing equal-time commutators (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. It is possible to send signals inside the light cone.
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 \([a, a^\dagger] \propto \mathbb{I}\), \(\hat{\Delta}^\pm\) is a c-number, independent of what state it acts on. So, for any normalized state \(|\psi\rangle\),

\[
\Delta^+(x - y) = \langle \psi | \hat{\Delta}^+(x - y) | \psi \rangle = \langle 0 | \hat{\Delta}^+(x - y) | 0 \rangle
= \langle 0 | [\phi^{(+)}(x), \phi^{(-)}(y)] | 0 \rangle
= \langle 0 | \phi^{(+)}(x) \phi^{(-)}(y) | 0 \rangle - \langle 0 | \phi^{(-)}(y) \phi^{(+)}(x) | 0 \rangle
= \langle 0 | \phi(x) \phi(y) | 0 \rangle \equiv D(x - y)
\]

where in the last step we again used \(\phi^+ |0\rangle = 0\). This is the vacuum-to-vacuum amplitude, or propagator\(^{22}\), in the sense that

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

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

That is, \(\Delta^+\) is the amplitude to propagate the excitation created by the field from \(x\) to \(y\).

\[
\Delta^+(x - y) = \sum_p (\langle x | - - - \to - - - \cdot y \rangle) = \int \frac{d^d p}{2\omega_p} e^{-ip \cdot (x - y)} |p^0 = \omega_p\rangle.
\]

This integral can be done in terms of functions with names\(^{23}\), 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.

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

You can read more about how to arrive at these expressions in Peskin (page 27); the spacelike case related by a Lorentz boost (to the rest frame) to the calculation of the Yukawa potential that we did in the last section.

Notice that this quantity \(\langle \phi(x) \phi(y) \rangle\) is not zero outside the lightcone. There are nonzero spatial correlations (and indeed entanglement) in the scalar field vacuum.

---

\(^{21}\)Note from the definition that \(\Delta^\pm(x - y) = [\phi^{(+)}(x), \phi^{(\mp)}(y)] = -\Delta^\mp(y - x)\).

\(^{22}\)We’ll understand its connection to the time-ordered propagator just below.

\(^{23}\)Specifically, in four spacetime dimensions and spacelike separation, \((x - y)^2 \equiv -r^2\), \(\Delta^+(x - y) = \frac{m}{2\pi^2} K_1(r)\).
What gives? Nothing. Causality only requires the vanishing of \textit{commutators} outside the lightcone, which we already showed in (3.2).

The cancellation in (3.2) can be interpreted as destructive interference between particles and antiparticles. It’s clearer for the complex scalar field, where

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

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

So consider the analogous\textsuperscript{24}

\[ D(x - y) \equiv \langle 0 | \Phi(x) \Phi^*(y) | 0 \rangle = \langle 0 | [\Phi^+(x), \Phi^-(y)] | 0 \rangle = \underbrace{\Delta_a^+(x - y)}_{\text{from } [a, a^\dagger]} \]

\[ D^*(y - x) \equiv \langle 0 | \Phi^*(y) \Phi(x) | 0 \rangle = \langle 0 | [\Phi^{+*}(y), \Phi^-(x)] | 0 \rangle = -\underbrace{\Delta_b^-(y - x)}_{\text{from } [b, b^\dagger]} \]

So if we consider the commutator,

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

\[ = \sum_p \left( (x \cdot - - \leftarrow p \rightarrow - - \cdot y) \text{ particle} - (x \cdot - - p \rightarrow - - \cdot y) \text{ antiparticle!} \right) \]

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 (3.2). Antimatter makes QFT causal.

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

[This subsection is mostly in a different color because I am mostly skipping it in lecture: we already learned all these things from the point of view of the path integral in the last section. I’m leaving it in because it has nice pictures, and because it has a possibly-useful summary of the discussion of choosing the contour for the frequency integrals in the propagator.]

\textsuperscript{24}Notice the extremely annoying fact that the positive frequency component of \(\Phi^*\), \(\Phi^{*(+)}\), is actually the dagger of the \textit{negative} frequency component of \(\Phi\):

\[ (\Phi^*)^{(+)} = (\Phi^{(-)})^*. \]
I claim that the propagator for a real free scalar field can be represented as:

\[
\Delta(x) = \int_C \frac{d^{d+1}p}{p^2 - m^2} e^{-ip_{\mu}x^\mu} \frac{i}{p^2 - m^2},
\]

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

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

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 \(C_+\) and \(C_-\), which are small circles going clockwise around the poles at \(\pm \omega_{\vec{p}}\).

These basic integrals are (for a reminder about Cauchy’s theorem see footnote 14):

\[
\int_{C_+} d^{d+1}p \frac{i}{p^2 - m^2} e^{-ip_{\mu}x^\mu} = \int d^dp \frac{1}{2\omega_{\vec{p}}} e^{-ipx} |_{p_0 = \omega_{\vec{p}}} = \Delta^+(x).
\]

\[
\int_{C_-} d^{d+1}p \frac{i}{p^2 - m^2} e^{-ip_{\mu}x^\mu} = \int d^dp \frac{1}{2\omega_{\vec{p}}} e^{-ipx} |_{p_0 = -\omega_{\vec{p}}} = -\Delta^+(-x) \text{ let } \vec{q} \equiv -\vec{p} \Delta^-(x).
\]

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

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

This is one particular choice of contour, and others are also interesting. Consider the *retarded* propagator,

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

This is the Green’s function which governs linear response:

\[
\frac{\delta \langle \phi(x) \rangle}{\delta J(y)} = D_R(x - y).
\]

(I’ll explain how to see this later; for now, notice that the \(\theta(x^0 - y^0)\) guarantees that the response happens after the perturbation that causes it.) We can reproduce \(D_R\) 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 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

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

This is called the Feynman contour, $C_F$, and we saw in the last section that it arises by continuation from Euclidean spacetime. Consider again the euclidean propagator, where we get rid of the relative sign in the metric:

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

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.
To be more precise about this ‘smooth rotation’ of the contour, we can use Cauchy’s theorem again, with the contour at right (the figure is from Matthew Schwarz’s book). The semicircular arcs are at \(|k^0| = \Lambda\), and give a contribution that decays like \(\frac{1}{\Lambda}\), and we must take \(\Lambda \to \infty\). The integrand is analytic inside the contour, no poles, so the integral \(\int \frac{e^{-i k x}}{k^2 - m^2} = 0\). Therefore the real-frequency integral (the real axis) and euclidean integral (minus the imaginary axis contribution, with \(x^D = i t\)) are equal.

In position space, the Feynman propagator is

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

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

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

\[\Delta_F(x-y) = \langle 0| (\theta(x^0-y^0)\phi(x)\phi(y) + \theta(y^0-x^0)\phi(y)\phi(x)) |0\rangle \equiv \langle 0| T(\phi(x)\phi(y)) |0\rangle.\]

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, \(\delta^{(d+1)}(x) = \int d^{d+1} p \ e^{-ip \cdot x}\), 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.

**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, \(e.g.\)

\[
\mathcal{L} \ni \phi_p(x)\phi_n(x)\phi_\pi(x) + h.c.
\]
where the fields appearing here destroy or create particles with the names in the sub-
scripts. 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:

$$\Phi(x) = \int \frac{d^d p}{\sqrt{2\omega_p}} a_p e^{-i p x} = \Phi^+(x)$$

$$\Phi(x) = \int \frac{d^d p}{\sqrt{2\omega_p}} b^\dagger_p e^{i p x} = \Phi^-(x)$$

The commutator is

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

$$\quad = \Delta^+_{a^\dagger} (x-y) + \Delta^-_{b^\dagger} (x-y)$$

$$\quad = \int_{C_+} d^{d+1} p e^{-i p (x-y)} \frac{i}{p^2 - m^2} + \int_{C_-} d^{d+1} p e^{-i p (y-x)} \frac{i}{p^2 - m^2}.$$  \hfill (3.5)

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

$$\Delta_F(x-y) \equiv \langle 0 | T (\Phi(x)\Phi^*(y)) | 0 \rangle$$
\[
\begin{align*}
&= \theta(x^0 - y^0) \langle 0 | \Phi(x)\Phi^*(y) | 0 \rangle \\
&= (0|\Phi^+(x)\Phi^*- (y)|0) = (0|[\Phi^+(x),\Phi^*- (y)]|0) = \Delta^+_{\mathfrak{a}}(x-y), \text{ particles} \\
&+ \theta(y^0 - x^0) \langle 0 | \Phi^*(y)\Phi(x) | 0 \rangle \\
&= -(0|\Phi^*+(y),\Phi^-(x)|0) = -\Delta^+_{\mathfrak{a}}(x-y), \text{ antiparticles}
\end{align*}
\]
3.3 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:

$$\Phi(\vec{x}, t) = e^{-imt} \frac{1}{\sqrt{2m}} \Psi(\vec{x}, t).$$

(3.6)

The Klein-Gordon equation is

$$0 = \partial_t^2 \Phi - \nabla^2 \Phi + m^2 \Phi = \sqrt{2me^{-imt}} \left( \frac{\Psi}{\sqrt{2m}} - 2im\frac{\frac{\partial}{\partial t} \Psi}{\sqrt{2m}} - \nabla^2 \Psi \right),$$

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 says $\omega_p = \sqrt{|\vec{p}|^2 + m^2} \simeq m + \frac{p^2}{2m} + \cdots$. The point of factoring out the phase $e^{-imt}$ is that the time evolution of $\Psi$ looks like $e^{-i(E - m)t} \sim e^{-i\frac{\vec{p}}{2m}t}$; so in the NR limit, $|\Psi| \ll m|\Psi|$ meaning that we can ignore that term in the KG equation. The remaining equation of motion is

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

(3.7)

This looks like the Schrödinger equation for a particle in no potential, in position space, but that is a coincidence: $\Psi$ is not a wavefunction. This equation (3.7) is the eom associated with the lagrange density

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

from which $\pi_\Psi = i\Psi^*, \pi_{\Psi^*} = 0$ (which you can also get by plugging (3.6) into $\partial_\mu \Phi^* \partial^\mu \Phi$). 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

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

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

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

(3.8)
(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. They disappeared when we dropped the second-order time derivative (recall that a first-order equation has half as many integration constants as a second-order equation). In the limit we’ve taken, we don’t have enough energy to make any. (More precisely, we’re promising to only study states where there is not enough energy to make any antiparticles.) 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 \(H |\{p\}\rangle = \sum_i \frac{\vec{p}_i^2}{2m} |\{p\}\rangle\), a nice non-relativistic (NR) dispersion for each particle. The particle-number symmetry is still present \(\Psi \to e^{-i\alpha} \Psi\), but now the Noether current is

\[
j^\mu = \left(\Psi^* \Psi, \frac{i}{2m} \Psi^* \vec{\nabla} \Psi + h.c.\right)^\mu = (\rho, \vec{j})^\mu.
\]

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

\[
|\vec{x}\rangle = \Psi^*(x) |0\rangle = \int d^d x e^{-i\vec{p}\cdot \vec{x}} a_p^\dagger |0\rangle.
\]

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 commuator. Finally, the hamiltonian (3.8) 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 \textit{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 \textit{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. \hfill \text{[End of Lecture 7]}
4 Interactions, not too strong

4.1 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

\[
P \langle \phi,t| \mathcal{O}_P(t)|\psi,t\rangle_P \tag{4.1}
\]

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

\[
\text{i}\dot{\mathcal{O}}_H = [\mathcal{O}_H, H_H] \tag{4.2}
\]

so in particular \( \dot{H}_H = 0 \) so \( H_H(t) = H_H(0) = H \). Then (4.2) is solved by \( U_H(t) = e^{-\text{i}Ht} \).

Another example of an operator is the free field:

\[
\Phi_H(x,t) = \int \frac{d^d p}{\sqrt{2\omega_p}}(a_p e^{-\text{i}px} + b_p^\dagger e^{\text{i}px})
\]

which you time-evolved this way on the homework. 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 (4.1) then requires

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

And the unitary evolution operator is

\[
U_H(t) = e^{-\text{i}H(0)t} = e^{-\text{i}Hst} = U_S = U
\]

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:

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

\[ (4.3) \]

Note that in this picture, \( H_0(t) = H_0(0) = H_0 \). Equivalently, \( i\dot{\Phi}_I = [\mathcal{O}_I, H_0] \), 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,

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

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

\[ \langle \varphi, t | [H, \mathcal{O}_I(t)] | \psi, t \rangle_H = \langle \varphi, t | \mathcal{O}_I(t) | \psi, t \rangle_I = \langle \varphi, t | \mathcal{O}_I(t) \rangle = \langle \psi, 0 | U_0^\dagger(t) \mathcal{O}(0) U_H(t) \psi, 0 \rangle = \langle \psi, 0 | U_0^\dagger(t) \mathcal{O}(0) U_0(t) \psi, 0 \rangle \]

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

\[ |\psi, t\rangle_I = U_0^\dagger(t) U_H(t) |\psi, 0\rangle \equiv U_I(t) |\psi, 0\rangle, \quad \text{that is,} \quad U_I(t) = U_0^\dagger(t) U_H(t). \]  

\[ (4.4) \]

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 \( 1 = U_0 U_0^\dagger \) repeatedly,

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

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 \([ U^\dagger_0, U_H ] \neq 0 \), if the interactions are interesting. So

\[
\partial_t |\psi,t \rangle_I = \partial_t ( U_I(t) |\psi,0 \rangle ) = \partial_t \left( \begin{pmatrix} U^\dagger_0(t) & U_H(t) \end{pmatrix} \right) |\psi,0 \rangle
\]

\[
= U^\dagger_0 \left( \begin{pmatrix} \mathbf{i}H_0 - \mathbf{i}H \end{pmatrix} \right) U_H |\psi,0 \rangle = -\mathbf{i}U^\dagger_0 V(0) U_H |\psi,t \rangle_I
\]

\[
= -\mathbf{i}U^\dagger_0 V(0) U_H U^\dagger_0 |\psi,t \rangle_I = -\mathbf{i}U^\dagger_0 V(0) U_0 |\psi,t \rangle_I .
\]

That is

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

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 (4.2): 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:

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

\[
|\psi,t \rangle_I = U_I(t U_I^\dagger(t') |\psi,t' \rangle_I
\]

From which we infer that

\[
U(t,t') = U^\dagger_0(t) U_H(t) U_H^\dagger(t') U_0(t') = e^{iH_0 t} e^{-iH(t-t')} e^{-iH_0 t'} .
\]

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 \(|i⟩\) in the QFT Hilbert space.

• Wait. At time \( t \), the state is

\[
U(t, t_i) |i⟩ = |ψ, t⟩ .
\]

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

Quantum mechanics says that the probability for this outcome is

\[
| ⟨f| U(t_f, t_i) |i⟩|^2 .
\]

(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 \( t_i → -∞ \) and \( t_f → ∞ \), and let

\[
S_{fi} ≡ ⟨f| \hat{S} |i⟩, \quad \hat{S} ≡ U(∞, -∞)
\]

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

This has only three ingredients: initial state, time evolution operator, and final state. Let’s focus on the middle of the sandwich some more:

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

\[
∂_t |ψ, t⟩ = -iV(t)|ψ, t⟩ , \quad \text{with initial condition} \quad |ψ, t_i⟩ = |i⟩ .
\]

Here’s a “solution”:

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

as you can check by the fundamental theorem of calculus. The only small problem is that we don’t know \(|ψ, t_1⟩\). But we can use this expression for that too:

\[
|ψ, t⟩ = |i⟩ + (-i) \int_{t_i}^{t} dt_1 V(t_1) \left( |i⟩ + (-i) \int_{t_i}^{t_1} dt_2 V(t_2) \right) \\
= |i⟩ + (-i) \int_{t_i}^{t} dt_1 V(t_1) |i⟩ + (-i)^2 \int_{t_i}^{t} dt_1 \int_{t_i}^{t_1} dt_2 V(t_1) V(t_2) |ψ, t_2⟩
\]

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

\[
|ψ, t⟩ = ∑_{n=0}^{∞} (-i)^n \int_{t_i}^{t} dt_1 \int_{t_i}^{t_1} dt_2 \cdots \int_{t_i}^{t_{n-1}} dt_n V(t_1) V(t_2) \cdots V(t_n) |i⟩ = U(t, t_i) |i⟩
\]

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

\[ i \partial_t \langle \psi, t \rangle = i \partial_t U \langle \psi, t_i \rangle = i \partial_t U U^\dagger \langle \psi, t \rangle. \]

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

\[ \implies \partial_t U = -iVU. \]

We might expect that an equation like this has a solution which is something like \( U \sim e^{-iVt} \).

Now we must deal with what Lawrence Hall calls “the wretched \( n! \)”. Starting from our series solution (4.6),

\[
U(t, t_i) = \sum_{n=0}^{\infty} (-i)^n \int_{t_i}^{t} dt_1 \int_{t_i}^{t_1} dt_2 \cdots \int_{t_i}^{t_{n-1}} dt_n V(t_1)V(t_2) \cdots V(t_n) \\
= \sum_{n=0}^{\infty} (-i)^n \frac{1}{n!} \int_{t_i}^{t} dt_1 \int_{t_i}^{t} dt_2 \cdots \int_{t_i}^{t} dt_n \mathcal{T} (V(t_1)V(t_2) \cdots V(t_n)) \\
= \sum_{n=0}^{\infty} (-i)^n \frac{1}{n!} \int_{t_i}^{t} dt_1 \int_{t_i}^{t} dt_2 \cdots \int_{t_i}^{t} dt_n \mathcal{T} (V(t_1)V(t_2) \cdots V(t_n)) \tag{4.7}
\]

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:

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

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

### 4.2 S-matrix

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

\[ \hat{S} = U(-\infty, \infty) = \mathcal{T} \left( e^{-i \int_{-\infty}^{\infty} dt V(t)} \right). \tag{4.8} \]

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 (3.4). Simplifying the notation a bit, the whole Lagrangian density is

$$L = \frac{1}{2} \partial_\mu \Phi^\ast \partial^\mu \Phi - \frac{1}{2} m^2 \Phi^\ast \Phi + \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} M^2 \phi^2 + L_I$$

with $L_I = -g \Phi^\ast \Phi \phi$.

The mode expansions are

$$\phi = \int \frac{d^d p}{\sqrt{2 \omega_p}} \left( a_p e^{-ipx} + a_p^\dagger e^{ipx} \right) |_{p^0 = \omega_p}$$

$$\Phi = \int \frac{d^d p}{\sqrt{2 E_p}} \left( b_p e^{-ipx} + c_p^\dagger e^{ipx} \right) |_{p^0 = E_p}$$

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

$$q = N_c - N_b.$$  

But the $\phi$ particles are not conserved.\(^{25}\)

**Artisanal meson decay.** [Tong §3.2.1] The interaction $L_I$ can change the number of $\phi$ particles. What is the amplitude for a $\phi$ particle to decay to a $\Phi$-anti-$\Phi$ pair? So consider

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

\(^{25}\)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

$$\langle f | \hat{S} | i \rangle = \langle f | T \left( 1 + (-i) \int d^{d+1}x g \phi_x \Phi_x + \mathcal{O}(g^2) \right) | i \rangle$$

where all operators are in interaction picture. The first term dies because $\langle f | i \rangle = 0$. In the $\mathcal{O}(g)$ term, the time-ordering doesn’t matter because all the operators are at the same time. 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

$$= -ig \int d^{d+1}x \langle f | \Phi^\dagger \Phi \rangle \int \frac{d^d k}{\sqrt{2\omega_k}} e^{-ikx} \frac{a_k a_p^\dagger}{\delta^{(d-1)}(k-p)\delta(0)}$$

$$= -ig \int d^{d+1}x e^{-ipx} \langle 0 | b_{q_1} c_{q_2} \sqrt{4E_{q_1} E_{q_2}} \int \frac{d^d k_1}{\sqrt{2E_{k_1}}} e^{ik_{1x}} b_{k_1}^\dagger \int \frac{d^d k_2}{\sqrt{2E_{k_2}}} e^{ik_{2x}} c_{k_2}^\dagger | 0 \rangle$$

$$= -ig \int d^{d+1}x e^{i(q_1 + q_2 - p)x} = -ig(2\pi)^{d+1}\delta^{d+1}(q_1 + q_2 - p)$$

This is a small victory. The 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:

$$P_{fi} \sim |S_{fi}|^2 = g^2 (\delta^{d+1}(p_f - p_i))^2$$

The square of a delta function is infinity. What did we do wrong?

Not so much, we just asked a dumb question. Here is where it helps to be a physicist. Consider:

$$(\delta^{d+1}(p))^2 = \delta^{d+1}(p)\delta^{d+1}(0) = \delta^{d+1}(p) \int d^{d+1}x e^{i0x} = \delta^{d+1}(p)VT$$

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 §4.3.

[End of Lecture 8]

4.3 Time-ordered equals normal-ordered plus contractions

We have an expression for $\hat{S}$ in (4.8) 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:

\[ T(\phi...\phi) = :\phi...\phi: + ? \quad (4.10) \]

In the previous schematic non-equation, I introduced a notation for a \textit{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

\[ :\phi(x)\phi(y) : = \phi^-(x)\phi^+(y) + \phi^-(y)\phi^+(x) + \phi^+(x)\phi^+(y) + \phi^+(x)\phi^-(y). \quad (4.11) \]

**Normal-ordering difficulty.** There is a sticky point in the definition of normal ordering\(^{26}\). Notice that we are defining normal ordering as a lexicographic operation of moving symbols around, not as an operator. If we tried to define the normal-ordering operation to be linear, then we would need to have

\[ : \left( \begin{array}{c} a^\dagger a \\ = aa^\dagger -1 \end{array} \right) : \equiv : aa^\dagger : - : 1 : . \]

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

\[ : 1 : \equiv 0 \]

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

Instead we’ll define \( : c := c \) for \( c \) a c-number. More generally, let \( A, B, C \) be the positive- and negative-frequency bits of some fields. Then

\[ : ABC \cdots : = \left( \begin{array}{c} A^\dagger B^\dagger C^\dagger \cdots \\ \text{only } a^\dagger \text{s} \quad \text{only } a \text{s} \end{array} \right) \]

Peskin writes \( N(\cdots) \equiv : \cdots : \). Notice that \( \langle 0| : \text{anything except a c-number} : |0\rangle = 0 \).
A comment about fermions. Later we will use anticommuting operators, which have
\[ c_k c_p^\dagger + c_p^\dagger c_k = \delta(k-p), \quad \{c_k^\dagger, c_p^\dagger\} = 0. \]
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
\[
: A B C \cdots : \equiv \left( \begin{array}{c}
A' B' C' \cdots \phantom{\cdots} \\
\text{only a's} \end{array} \right) (-1)^P
\]
where \(P\) is the number of fermion interchanges required to get from \(A B C \cdots\) 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 (4.11) and compare with (4.10). Because \([\phi^-, \phi^+] = 0\), 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:
\[
: \phi(x) \phi(y) := T(\phi(x)\phi(y)) - \Delta_F(x-y) \equiv T(\phi(x)\phi(y)) - \phi(x)\phi(y). \quad (4.12)
\]
More generally, writing \(\phi_a \equiv \phi_a(x_a)\), Wick’s theorem says
\[
: \phi_1 \cdots \phi_n := T(\phi_1 \cdots \phi_n) - \text{(all contractions)} \quad (4.13)
\]
where a contraction is defined as the price for moving a pair of operators through each other to repair the time ordering, as in (4.12), and denoted by the symbol in (4.12).

For example, for four fields, the theorem says
\[
T(\phi_1 \cdots \phi_4) =: \left( \phi_1 \cdots \phi_4 + \phi_1^\dagger \phi_2 \phi_3 \phi_4^\dagger + 5 \text{ more} \right) : + (\phi_1^\dagger \phi_2 \phi_3 \phi_4^\dagger + 2 \text{ more})
\]

Notes: The fully-contracted bits are numbers, so (with Peskin’s convention) it doesn’t matter if they are inside the normal-ordering symbol. For a product of \(n\) fields, there are \(\binom{n}{2} + \binom{n}{4} \left( \frac{4}{2} \right) + \binom{n}{6} \left( \frac{6}{4} \right) \left( \frac{4}{2} \right) \cdots + \left( \binom{n}{\lfloor n/2 \rfloor} \right) \) (= 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 of (4.13) [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
\[
T(\phi_1 \cdots \phi_m) =: \phi_1^{\dagger} : \phi_2 \cdots \phi_m : + \text{(all contractions w/o } \phi_1 \text{ )}
\]
\[
= \phi_1^{\dagger} + \phi_1^-
\]
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 $\phi^-_{j \geq 2}$s; this process will add a term for every possible contraction involving $\phi_1$.

4.4 Time-ordered correlation functions

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

$$G^{(n)}(x_1 \cdots x_n) \equiv \langle \Omega | T (\phi^H_1(x_1) \cdots \phi^H_n(x_n)) | \Omega \rangle.$$  

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:

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

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

$$G^{(2)}_{\text{free}}(x_1, x_2) = \Delta_F(x_1 - x_2) = \frac{i}{p^2 - m^2 + i\epsilon}.$$  

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

$$G^{(4)}_{\text{free}}(x_1 \cdots x_4) = \Delta_F(12)\Delta_F(34) + \Delta_F(13)\Delta_F(24) + \Delta_F(14)\Delta_F(23).$$

Expectations. Our next goal is to construct a perturbative expansion in the case of $V = \int d^d z \lambda \phi^4(z)$. We expect a correction of order $\lambda$ of the form: In momentum space, we have blobs (unspecified sums of diagrams) with external lines labelled by $p^\mu_i$:

Notice that there is no need to restrict their values to the mass shell $p^2_i = m^2$, 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, (4.14), they blow up on the mass-shell. The existence of a singularity of $\hat{G}$ on the mass-shell is a general fact, and their residues give the $S$-matrix elements:

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

This is the content of 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 of space finite for awhile.

Now consider

$$\langle 0| e^{-iHT} = \sum_n \langle 0|n\rangle \langle n| e^{-iHT} = \sum_{n \neq \Omega} \langle 0|n\rangle \langle n| e^{-iE_nT} + \langle 0|\Omega\rangle \langle \Omega| e^{-iE_0T}. $$

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

$$\langle \Omega \rangle = \lim_{T \to \infty(1-i\epsilon)} \left( \frac{\langle 0| e^{-iHT} e^{iE_0T}}{\langle 0|\Omega\rangle} \right) = \lim_{T \to \infty(1-i\epsilon)} \left( \frac{\langle 0| e^{iH_0T} e^{-iHT} e^{iE_0T}}{\langle 0|\Omega\rangle} \right) \overset{(4.16)}{=} \lim_{T \to \infty(1-i\epsilon)} \left( \frac{\langle 0| U_I(t, t_0) e^{iE_0(T-t_0)}}{\langle 0|\Omega\rangle} \right).$$

$^{27}$Big picture comment: This is a long chapter. We are working our way towards a useful and correct perturbative expansion of the $S$-matrix, from which we can extract observable physics.
Do the same for $|\Omega\rangle$ (rather than just taking the dagger of (4.16)):

$$|\Omega\rangle = \lim_{T \to \infty (1-\epsilon)} \frac{U_I(t_0, -T) |0\rangle}{e^{-iE_0 T} \langle \Omega | 0 \rangle}.$$ 

Now step (2), relate $\phi_H$ and $\phi_I$: in terms of the field $\phi$ at some reference time $t_0$ (could be zero), we have $\phi_I = U_0^\dagger \phi U_0$, $\phi_H = U_H^\dagger \phi U_H$ and therefore $\phi_H = U_H^\dagger U_0^\dagger \phi_I U_0 U_H = U_I^\dagger \phi_I U_I$, using (4.4). Now put together the Green’s function. First assume $x^0 > y^0$:

$$G^{(2)}(x, y) = \langle \Omega | T \left( \phi_H(x) \phi_H(y) \right) |\Omega\rangle_{\text{(4.16)}}$$

$$= \lim_{T \to \infty (1-\epsilon)} \left[ (e^{-iE_0(T-t_0)} \langle 0 | \Omega \rangle)^{-1} \langle 0 | U_I(T, t_0) \right.$$ 

$$\cdot U_I^\dagger(x^0, t_0) \phi_I(x) U_I(x^0, t_0) \cdot U_I^\dagger(y^0, t_0) \phi_I(y) U_I(y^0, t_0)$$

$$\cdot U_I(t_0, -T) |0\rangle \left( e^{-iE_0(T+t_0)} \langle \Omega | 0 \rangle \right)^{-1} \right]$$

$$= \lim_{T \to \infty (1-\epsilon)} \left( e^{-iE_0 T} \langle 0 | \Omega \rangle \right)^{-1} \langle 0 | U(T, x^0) \phi(x) U(x^0, y^0) \phi(y) U(y^0, -T) |0\rangle$$

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. $U_I(t, t') = T \left( e^{-i \int_t^{t'} dt' V(t') \right)}$ – 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. If $y^0 > x^0$, the relation between the first and last expressions is unchanged.

What’s the denominator? The norm of the vacuum is one, but we can assemble it from these ingredients (4.16):

$$1 = \langle \Omega | \Omega \rangle = \lim_{T \to \infty (1-\epsilon)} \left( e^{-iE_0 T} \langle 0 | \Omega \rangle \right)^{-1} \langle 0 | U(T, t_0) U(t_0, -T) |0\rangle_{=U(T,-T)}$$

Therefore

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

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 $\lambda$:

$$G_{\text{num}}^{(2)}(x, y) = \langle 0 | \mathcal{T} \left( \phi(x) \phi(y) e^{-i \int_T d^{d+1}z \frac{1}{2} \phi^4(z)} \right) | 0 \rangle$$

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

$$= \Delta_F(x - y) + \frac{-i \lambda}{4!} \int d^{d+1}z \left( 3 \phi(x) \phi(y) \phi(z) \phi(z) \phi(z) + 4 \cdot 3 \phi(x) \phi(z) \phi(y) \phi(z) \phi(z) \phi(z) \right) + \cdots$$

The $O(\lambda^2)$ contribution is

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

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

The answer is

$$(\xrightarrow{\phi \phi \phi \phi \phi}) \cdot (\xrightarrow{\phi \phi \phi \phi \phi}) \cdot (\xrightarrow{\phi \phi}) \cdot (\xrightarrow{\phi \phi \phi}) + \xrightarrow{\phi \phi} + \xrightarrow{\phi \phi \phi} + \xrightarrow{\phi \phi \phi \phi \phi} +$$

For example,

$$\phi(x) \phi(z_1) \phi(y) \phi(z_1) \left( \phi(z_1) \phi(z_2) \right)^2 \phi(z_2) \phi(z_2) \propto \xrightarrow{\phi \phi \phi \phi \phi}$$

up to the symmetry factor.

**Feynman rules for $\phi^4$ theory in position space.** The set of diagrams is made by drawing one external vertex for each $x_i$, and $m$ 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 $\lambda^m$. Let’s call $\mathcal{M}_A$ the amplitude associated to diagram $A$, 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 \( \bigstar \) (notice no \( \frac{1}{\pi} \)).

• Put a $\Delta_F(y_i - y_j)$ for each propagator \( \rightarrow \rightarrow \). 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, that is the number of ways of permuting the ingredients of the diagram which map the diagram to itself. Symmetries of the diagram mean that the sum over contractions fails to completely cancel Dyson’s wretched \( \frac{1}{n!} \). For example:

\[
s \left( \begin{array}{c} \bigcirc \\ \bigcirc \\ \bigcirc \\ \bigcirc \end{array} \right) = \frac{1}{4!} \cdot 3 = \frac{1}{8}, \quad s \left( \begin{array}{c} \bigcirc \\ \bigcirc \end{array} \right) = \frac{1}{4!} \cdot 4 \cdot 3 = \frac{1}{2}.
\]

Do not get hung up on this right now.

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

\[
\mathcal{O}(\lambda^0): \quad \begin{array}{c} \bigstar \\ \bigstar \\ \bigstar \end{array} 
\]

\[
\mathcal{O}(\lambda^1): \quad \begin{array}{c} \bigstar \\ \bigstar \end{array} + \begin{array}{c} \bigstar \\ \bigstar \end{array} + \begin{array}{c} \bigstar \\ \bigstar \end{array} + \begin{array}{c} \bigstar \\ \bigstar \end{array} + \cdots + \bigstar
\]

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 $G^{(n)}_m$ denotes the an order-$\lambda^m$ contribution to $G^{(n)}$.

\[
\mathcal{O}(\lambda^2): \quad \begin{array}{c} \bigstar \\ \bigstar \end{array} + \begin{array}{c} \bigstar \\ \bigstar \end{array} + \begin{array}{c} \bigstar \\ \bigstar \end{array} + \begin{array}{c} \bigstar \\ \bigstar \end{array} + \begin{array}{c} \bigstar \\ \bigstar \end{array} + \begin{array}{c} \bigstar \\ \bigstar \end{array}
\]

\[
(4.17)
\]

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 $\phi \rightarrow -\phi$ 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:

Only some of the components are attached to the external legs; for a given diagram $A$, call the factor associated with these components $A_c$ (note that $A_c$ need not be fully connected). The rest of the diagram is made of a pile of ‘bubbles’ of various types $V_i$ (each one internally connected, but disconnected from the external lines) and multiplicities $n_i$ (e.g. $V_1$ could be a figure eight, and there could be $n_1 = 2$ of them, as in the second term indicated in (4.17)). These bubbles (or ‘vacuum bubbles’) would be there even if we didn’t have any external lines, and they would have the same value; they are describing the fluctuations intrinsic to the vacuum. The amplitude associated with the general diagram is then

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

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

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

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

$$= \sum_{A_c} \mathcal{M}_{A_c} \cdot e^{\sum_i V_i}$$

$$= \sum_{A_c} \mathcal{M}_{A_c} e^{\sum_i V_i}$$

(4.18)

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

$$G_0^{(n)}_{\text{numerator}} = \langle 0 | \mathcal{T} e^{-\int V} | 0 \rangle = 1 \cdot e^{\sum_i V_i}$$

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

$$G^{(n)} = \frac{\langle 0 | \mathcal{T} \left( \phi_1 \cdots \phi_n e^{-\int V} \right) | 0 \rangle}{\langle 0 | \mathcal{T} e^{-\int H} | 0 \rangle} = \frac{C^{(n)}_{\text{numerator}}}{C^{(0)}_{\text{numerator}}} = \sum_{A_c} \mathcal{M}_{A_c}.$$  

(4.19)
And with that we can forget all about the bubbles. So for example,

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

\[ G^{(4)} = \_ + \_ + \_ + (\_ \_ \_ \_\_\_) + \_ + \_ + \quad + \quad + \quad + \quad + \quad + \]

Notice that in this manipulation (4.19) 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

\[ G^{(2)} = \frac{x_1 \cdot (1 + 8 + 88 + \ldots)}{(1 + 8 + 88 + \ldots)} + \mathcal{O}(\lambda) = \frac{x_1 \cdot e^V}{e^V} + \mathcal{O}(\lambda) = \frac{x_1}{x_1} + \mathcal{O}(\lambda). \]

(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

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

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

\[ A_{FC}^N = \int d^{d+1}x_1 \cdots d^{d+1}x_n e^{-\sum_i p_i x_i} \frac{(-i\lambda)^N}{s(FC)} \int d^{d+1}z_1 \cdots \int d^{d+1}z_N \prod_{r=1}^{N_I} \int d^{d+1}q_r e^{-i(y_A - y_B)q_r} \frac{i}{q_r^2 - m^2 + i\epsilon}. \]

For example, consider a particular contribution with \( n = 4 \) external legs and \( N = 2 \) interaction vertices:

![Diagram](image)

has \( N_I = \frac{4 + 2 \cdot 4}{2} = 6 \)
Notice that we are doing a silly thing here of labelling the momenta of the external lines (the first $n$ momenta $q_{i=1...n}$). 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 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 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.