Daniel Harlow
Massachusetts Institute of Technology
Lecture notes for Physics 8.323: Relativistic Quantum field theory I, spring 2024

Contents

1 Why quantum field theory?

The goal of this class is to teach you **quantum field theory** (QFT), which is the central foundation (together with general relativity) of most of contemporary theoretical physics. In 2024 you cannot claim to understand the laws of physics without knowing some QFT, and here is an opportunity for you to learn it. The full class is three semesters long, so in this semester we are just getting started.

QFT is not an easy subject to study: there are many subtle arguments and long calculations, and moreover, as we will see throughout the class, QFT rests on somewhat shaky mathematical ground, which can make it difficult to know which results are really solid. It is often said that nobody truly understands QFT, and many current research seminars around the world are devoted to trying to understand how to formulate it better. A consequence of this state of affairs is that, unlike for older subjects such as classical electromagnetism, there is no settled way to teach QFT and the various textbooks are all written from quite different perspectives. I'll say a bit more about the perspective of this class at the end of the section.

Given the difficulty of the subject, it is important to understand at the outset where we are going and why. The goal of this section is to present the basic conceptual motivations for thinking about QFT, aiming to get some intuition for why it is a good idea to think about the quantum mechanics of fields. There are three main motivations that we will consider:

- Quantum field theory is (likely) the only way to create a quantum theory of interacting relativistic particles. This is why quantum field theory is of great importance in particle physics: the standard model of particle physics, which governs the interactions of elementary particles through the electromagnetic, strong, and weak forces, is a quantum field theory. For example one of the great triumphs of the standard model of particle physics is its successful description of something called the **anomalous magnetic moment of the electron**:

$$\begin{aligned} \text{Theory}: &\quad a_e = 0.001159652181643(764) \\ \text{Experiment}: &\quad a_e = 0.00115965218073(28). \end{aligned} \tag{1.1}$$

  This theory calculation is a tour-de-force of quantum field theory, and we will compute the first few digits of it next semester in QFT 2.

- Quantum field theory is the natural language for describing the low-energy physics of many-body quantum systems with local interactions. This is why quantum field theory is of great importance in condensed matter physics: many important solid-state phenomena such as superconductivity, phase transitions in magnets, and the fractional quantum hall effect are quantitatively understood using the machinery of quantum field theory. For example in an Ising magnet the spontaneous magnetization M scales as

$$M \propto (T_c - T)^\beta \tag{1.2}$$

  for temperatures just below the critical temperature T_c, with the "critical exponent" β being given by

$$\beta = \begin{cases} \frac{1}{8} & d_{spatial} = 2 \\ .326419(3) & d_{spatial} = 3 \\ \frac{1}{2} & d_{spatial} = 4 \end{cases} \tag{1.3}$$

  These exponents can be computed in quantum field theory: by the end of this semester we will be able to understand the $d_{spatial} = 2$ and $d_{spatial} = 4$ cases, while the $d_{spatial} = 3$ case (the hardest) is an area of active ongoing research!

- Quantum field theory arises ubiquitously in our most promising approach to combining quantum mechanics and gravity, which consists of a set of related ideas under the general umbrella of "string theory". There it arises both as the low-energy description of brane systems and also as the "holographic dual" of non-perturbative quantum gravity in spacetimes with negative cosmological constant.

One of the big successes of the latter is its confirmation in many cases of the **Bekenstein-Hawking formula** for the entropy of a black hole:

$$S = \frac{A_{horizon} c^3}{4G\hbar}. \tag{1.4}$$

We will now discuss each of these motivations in turn, focusing on the first two since the third is mostly beyond the scope of this class. If you are new to QFT (as I hope many of you are), the arguments may go by a bit fast for you. If that is the case do not worry: we will do most of these manipulations again in much more detail in later sections. Our goal here is to paint in broad strokes, getting a flavor of what is to come in the weeks and months ahead!

1.1 Combining quantum mechanics and special relativity

By the end of the 1920s non-relativistic quantum mechanics was on a fairly firm mathematical foundation. For example for a system of N particles of mass m interacting via a potential $V(\vec{x}_1, \dots \vec{x}_M)$ we can (at least in principle) determine everything we want to know about the system by solving the many-body Schrodinger equation

$$\left[-\frac{\hbar^2}{2m} \sum_{i=1}^{N} \nabla_i^2 - i\hbar\partial_t + V(\vec{x}_1, \dots \vec{x}_M) \right] \psi(\vec{x}_1, \dots, \vec{x}_N; t) = 0. \tag{1.5}$$

This equation however has two problems from the point of view of special relativity:

(i) The kinetic terms are non-relativistic, being compatible with $E = \frac{|p|^2}{2m}$ instead of $E = \sqrt{|p|^2 c^2 + m^2 c^4}$.

(ii) The potential interactions are instantaneous, which is not compatible with the relativistic principle that nothing can move faster than light.

Problem (i) is not too difficult to solve: it isn't pretty, but we can just make the replacement

$$-\frac{\hbar^2}{2m} \sum_{i=1}^{N} \nabla_i^2 \to \sum_{i=1}^{N} \sqrt{-\hbar^2 c^2 \nabla_i^2 + m^2 c^4} \tag{1.6}$$

in equation (1.5). In this way it is fairly straightforward to make a relativistic theory of non-interacting quantum particles, for example in the one-particle case the solutions of the Schrodinger equation can be expanded in a basis of energy eigenstates

$$\psi(\vec{x}, t) = e^{i\vec{p}\cdot\vec{x} - i\sqrt{|p|^2 c^2 + m^2 c^4} t}. \tag{1.7}$$

Problem (ii) however is more serious: in order for quantum dynamics to be compatible with special relativity, we need to make sure that all interactions are local in spacetime. Based on our experience with electromagnetism, which is after all a relativistic theory, we can guess that the natural way to incorporate spacetime locality is to introduce **fields**. When we move an electric charge here in Cambridge, it is not really true that there is a physically-detectable Coulomb potential that immediately adjusts what is going on in the Andromeda galaxy. What happens instead is that we create a ripple in the electromagnetic field which then propagates outwards at the speed of light, updating the Coulomb field as it goes along. What is perhaps more surprising is that it turns out that we need to introduce fields for the charges as well: an electron field, a proton field, and so on.

1.1.1 A notational aside

Before proceeding further, it is time for the deep realization that the factors of $\hbar$ and c in the previous paragraph are unnecessary and distracting. We can get rid of the former by measuring time in units of

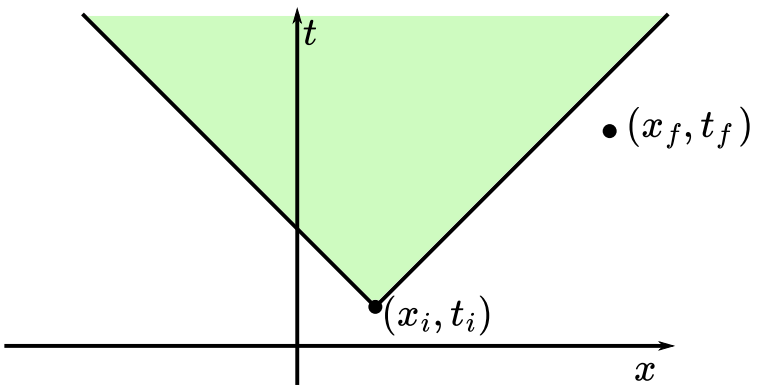


Figure 1: Propagation inside the lightcone in $1+1$ dimensions: in a theory where nothing is faster than light a disturbance at (x_i, t_i) should not be able to reach a point (x_f, t_f) which is spacelike separated.

inverse energy, with $\hbar$ as the conversion factor, and we can get rid of the latter by measuring distance in units of seconds, with c as the conversion factor. From now on we will therefore work in units where

$$\hbar = c = 1, \tag{1.8}$$

with all dimensionful quantities having units which are some power of energy. In particular length and time are both measured in units of inverse energy, while mass is measured in units of energy. For example the radius of the earth is

$$\frac{R_\oplus}{\hbar c} = \frac{1}{4.9 \times 10^{-33}\,\mathrm{J}} \tag{1.9}$$

and the acceleration due to gravity at the Earth's surface is

$$\frac{g\hbar}{c} = 3.43 \times 10^{-42}\,\mathrm{J}. \tag{1.10}$$

These units are clearly not so practical for daily life, but in situations where both relativity and quantum mechanics are important they are indispensable.

1.1.2 Relativistic propagator and causality

Let's now try to understand in more detail why particles need to be replaced by fields. Although the non-interacting theory based on the wave functions (1.7) is both relativistic and quantum, there is a sense in which it allows information to propagate faster than light. To be concrete, let's work in $1+1$ dimensions and consider the propagator

$$G(x_f, t_f; x_i, t_i) := \langle x_f | e^{-iH_0(t_f - t_i)} | x_i \rangle. \tag{1.11}$$

Here $|x_i\rangle$ and $|x_f\rangle$ are eigenstates of the particle position: in the p basis the wave functions of such eigenstates are given by

$$\langle p | x \rangle = e^{-ipx}. \tag{1.12}$$

The physical meaning of the propagator is that its absolute value squared is the probability to find the particle at position x_f at time t_f given that it was located at position x_i at time t_i (its phase has information about the same question for initial and final momenta). We'll focus on the situation where $x_f > x_i$, $t_f > t_i$, and $t_f - t_i < x_f - x_i$, in which case (x_f, t_f) and (x_i, t_i) are spacelike separated so no signal which is not faster than light can propagate between them. More geometrically, the point (x_f, t_f) is outside of the future lightcone of (x_i, t_i) (see figure 1). We'll now show that the propagator in this situation is nonzero, which

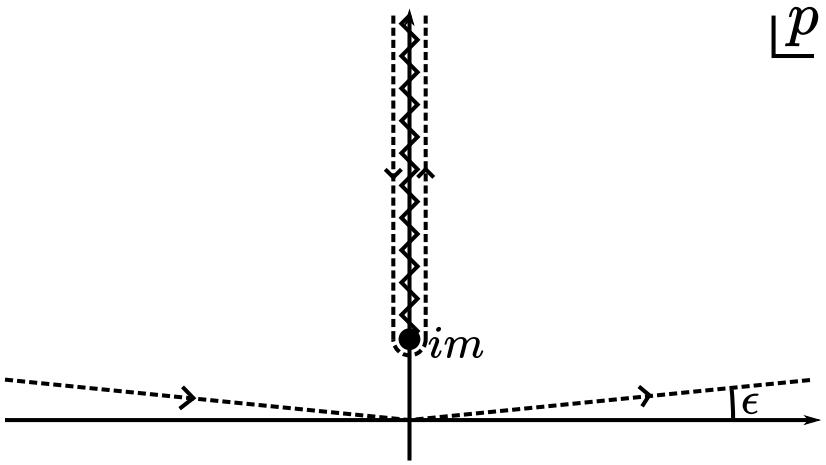


Figure 2: Deforming the contour in the complex p-plane. The defining contour is the lower dashed line, which can be smoothly deformed via a two large circle segments at infinity to the upper contour which wraps the branch cut along the positive imaginary axis.

shows that the information that there is a particle located at position x_i at time t_i propagates faster than light!

We can evaluate the propagator (1.11) by inserting a complete set of momentum eigenstates:

$$G(x_f, t_f; x_i, t_i) = \int_{-\infty}^{\infty} \frac{dp}{2\pi} e^{ip(x_f - x_i) - i\sqrt{p^2+m^2}(t_f - t_i)}. \tag{1.13}$$

Going forward we might as well use translation invariance to set $t_i = x_i = 0$ and relabel $t_f = t$, and $x_f = x$, in which case we can consider the simpler function

$$G(x, t) := \int_{-\infty}^{\infty} \frac{dp}{2\pi} e^{ipx - i\sqrt{p^2+m^2}t}. \tag{1.14}$$

This integral is not so easy to evaluate analytically, and one can worry if it even converges due to the oscillatory behavior at infinity. Since we are assuming that $0 < t < x$, the convergence of this integral is controlled by the ipx term in the exponent. To make sure it is convergent, we can slightly rotate the phase of the p integral so that it goes off to infinity at a small angle ϵ above the real p-axis in both directions (see figure 2). This is convergent since at large positive p we have

$$e^{ie^{i\epsilon}px} \approx e^{-\epsilon px + ipx} \tag{1.15}$$

while at large negative p we have

$$e^{ie^{-i\epsilon}px} \approx e^{\epsilon px + ipx}. \tag{1.16}$$

Moreover by Cauchy's theorem the answer is independent of ϵ since we can rotate the contour freely from one ϵ to the next (the circle segments at infinity do not contribute since the integrand is exponentially suppressed), and so we can take $\epsilon \to 0$ to recover the propagator. On the other hand to estimate the value of the propagator, it is more convenient to instead rotate the contour up to wrap around the branch cut that runs along the imaginary axis from $p = im$ off to $p = i\infty$ (see figure 2). Evaluating the integral on this contour, we see that we have

$$\begin{aligned} G(x, t) &= i \int_m^{\infty} \frac{d\lambda}{2\pi} e^{-\lambda x} \left(e^{\sqrt{\lambda^2 - m^2}t} - e^{-\sqrt{\lambda^2 - m^2}t} \right) \\ &= \frac{i}{\pi} \int_m^{\infty} d\lambda e^{-\lambda x} \sinh\left(\sqrt{\lambda^2 - m^2} t \right). \end{aligned} \tag{1.17}$$

The integrand here (ignoring the factor of i) is strictly positive for $\lambda > m$, and so we see that the propagator is indeed nonzero outside of the lightcone! On the other hand it isn't *very* nonzero: by using the monotonicity of $\sinh y$ for $y > 0$ and then ignoring the negative exponential we have

$$\int_m^\infty d\lambda e^{-\lambda x} \sinh\left(\sqrt{\lambda^2 - m^2}t\right) < \frac{1}{2}\int_m^\infty d\lambda e^{-\lambda(x-t)} = \frac{e^{-m(x-t)}}{2(x-t)}, \tag{1.18}$$

and thus

$$0 < |G(x,t)| < \frac{e^{-m(x-t)}}{2\pi(x-t)}. \tag{1.19}$$

Therefore the propagator of a massive relativistic particle is suppressed exponentially outside the lightcone, but it isn't zero.

In a relativistic theory, there is something deeply wrong with being able to send information faster than light. Indeed by doing a boost we can change the time ordering of any pair of spacelike-separated events, so if we can communicate faster than light then we can also communicate backwards in time. In the presence of interactions it is even worse: by sending a message to a point outside of your future lightcone and then receiving a message back you can communicate directly with points in your own own past lightcone. Such things are called violations of **causality**, which is the principle that you shouldn't be able to send signals to your own past. Physics seems unlikely to make much sense in situations where causality is violated, so we had better find a way to fix this.

1.1.3 Creation and annihilation operators on multi-particle Hilbert space

There is a useful way of re-organizing the above discussion of relativistic particles, historically called **second quantization**, which helps points the way towards how to restore causality. The idea is to introduce a larger Hilbert space, called **Fock space**, where any number of relativistic particles, including zero, is allowed. The nicest way to do this is by introducing an **annihilation operator** $a(x)$, which removes a particle from the system at point x if one exists and otherwise annihilates the state. Its adjoint $a^\dagger(x)$ creates a particle at x. The algebra of creation and annihilation operators is given by

$$\begin{aligned}
[a(x), a(x')] &= 0 \\
[a^\dagger(x), a^\dagger(x')] &= 0 \\
[a(x), a^\dagger(x')] &= \delta(x - x'),
\end{aligned} \tag{1.20}$$

and the zero-particle state $|\Omega\rangle$ is defined to be the one which is annihilated by all $a(x)$. Other states are created from $|\Omega\rangle$ by acting with creation operators, for example a one-particle state with wave function $\psi(x)$ is represented in this language by

$$|\psi\rangle = \int dx \psi(x) a^\dagger(x)|\Omega\rangle. \tag{1.21}$$

For practice we can check that the norm works out:

$$\begin{aligned}
\langle\psi|\psi\rangle &= \int dx' \psi^*(x') \int dx \psi(x) \langle\Omega|a(x')a(x)^\dagger|\Omega\rangle \\
&= \int dx' \psi^*(x') \int dx \psi(x) \langle\Omega|[a(x'), a(x)^\dagger]|\Omega\rangle \\
&= \int dx' \psi^*(x') \int dx \psi(x) \delta(x - x') \\
&= \int dx |\psi(x)|^2 \\
&= 1.
\end{aligned} \tag{1.22}$$

More generally a multi-particle state $\psi(x_1, x_2, \dots x_N)$ is represented by

$$|\psi\rangle = \frac{1}{\sqrt{N!}} \int dx_1 \dots dx_n \psi(x_1, \dots, x_N) a^\dagger(x_1) \dots a^\dagger(x_N) |\Omega\rangle. \tag{1.23}$$

We note in passing that since the $a^\dagger(x_i)$ all commute with each other, the particles we are describing are bosons:[1]

$$a^\dagger(x) a^\dagger(x') |\Omega\rangle = a^\dagger(x') a^\dagger(x) |\Omega\rangle. \tag{1.24}$$

The Hamiltonian is given by

$$\begin{aligned} H_0 &= \int dx\, a^\dagger(x) \sqrt{-\partial_x^2 + m^2}\, a(x) \\ &= \int \frac{dp}{2\pi} \sqrt{p^2 + m^2}\, a^\dagger(p) a(p), \end{aligned} \tag{1.25}$$

where in the second line we have introduced the Fourier-transformed annihilation operator

$$a(p) = \int dx e^{-ipx} a(x). \tag{1.26}$$

It may feel like we have suddenly introduced an entire new form of quantum mechanics, but except for introducing a rule that the particles are bosons (which we couldn't see before since we only considered one-particle states), this is really just a different bookkeeping for the same old multi-particle quantum mechanics. In particular the second expression for the Hamiltonian shows that the energy eigenstates are states of the form

$$a^\dagger(p_1) \dots a^\dagger(p_N) |\Omega\rangle, \tag{1.27}$$

with the total energy just being

$$E = \sum_{i=1}^{N} \sqrt{p_i^2 + m^2}. \tag{1.28}$$

In this language we can rewrite the propagator (1.14) as

$$\begin{aligned} G(x,t) &= \langle \Omega | a(x) e^{-iH_0 t} a^\dagger(0) | \Omega \rangle \\ &= \langle \Omega | a(x,t) a^\dagger(0) | \Omega \rangle \\ &= \langle \Omega | [a(x,t), a^\dagger(0)] | \Omega \rangle. \end{aligned} \tag{1.29}$$

where we have introduced the Heisenberg picture annihilation operator

$$a(x,t) = e^{iH_0 t} a(x) e^{-iH_0 t}. \tag{1.30}$$

In fact this is true as an operator equation:

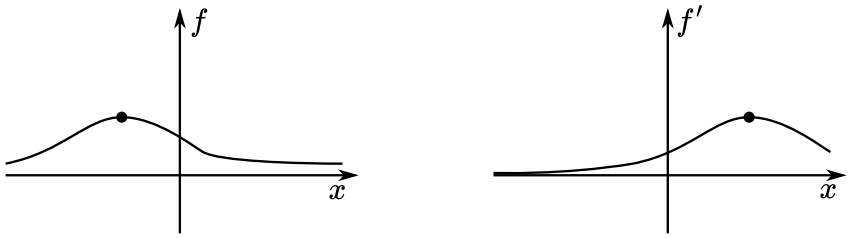
$$[a(x,t), a^\dagger(0)] = G(x,t). \tag{1.31}$$

What we have learned from our discussion of the propagator is therefore that creation and annihilation operators in the Heisenberg picture do not commute at spacelike separation. It is a bit tedious to work out, but this also implies that the number operator

$$N(x,t) = a^\dagger(x,t) a(x,t), \tag{1.32}$$

which counts how many particles there are at position x and time t, does not commute with itself at spacelike separation. Unlike the creation/annihilation operators, the number operator is hermitian and thus should be observable. If we are going to save causality, we thus need to argue that the number of particles at position x and time t cannot actually be measured by someone in the vicinity of x and t!

[1] If we want to get fermions, we should instead impose **anticommutation relations** $\{a(x), a(x')\} = \{a^\dagger(x), a^\dagger(x') = 0$ and $\{a(x), a^\dagger(x')\} = \delta(x - x')$, where $\{A, B\} = AB + BA$ is called the anticommutator of A and B. We will discuss fermions in more detail later in the semester.



$$f'(x) = f(x-a)$$

Figure 3: Translation of a function $f(x)$ by a. Note that here f' is the transformed function, not the derivative, and that it is the inverse of the translation which appears in the argument of the function.

1.1.4 Quantum fields

A good way to proceed is to consider what kind of interactions we could add to the Hamiltonian (1.25) that wouldn't violate causality. We'd like to somehow build the interaction Hamiltonian V out of the creation and annihilation operators, but in a way where (now working in d spacetime dimensions) it is an integral

$$V(t) = \int d^{d-1}x \mathcal{H}_{int}(t, \vec{x}) \tag{1.33}$$

of an interaction density $\mathcal{H}_{int}(t, \vec{x})$ that commutes with itself at spacelike separation (otherwise we could violate causality by measuring the energy density at spacelike separation). The easiest way to achieve this is to construct an interaction density which commutes with itself at *spatial* separation, and then also demand that it transform as a Lorentz scalar in the sense that[2]

$$U(\Lambda, a)^\dagger \mathcal{H}_{int}(x) U(\Lambda, a) = \mathcal{H}_{int}(\Lambda^{-1}(x-a)), \tag{1.34}$$

where $U(\Lambda, a)$ is the unitary operator on Hilbert space which implements the Poincaré transformation

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu \tag{1.35}$$

on the Hilbert space of the theory. This ensures commutativity at spacelike separation since if x and y are spacelike separated there is always a Poincaré transformation that sends them to the same time slice and we have assumed that $\mathcal{H}_{int}(x)$ commutes with itself at spatial separation. It may be puzzling that we used the inverse Poincaré transformation in the argument of $\mathcal{H}_{int}$, the idea behind this is shown in figure 3: we want to define the symmetry transformation to "move the scalar along with the symmetry", meaning that the "new" scalar at x should be equal to the "old" scalar at the point $\Lambda^{-1}(x-a)$ where x "came from". You will also show in the homework that defining things this way is necessary for us to have two successive Poincaré transformations combine in the natural way.

I'll note in passing that by using time-dependent perturbation theory we can write a formula for the particle scattering matrix in a theory with an interaction of this form (see chapter three of Weinberg volume I) as

$$S = 1 + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int d^d x_1 \dots d^d x_n T\{\mathcal{H}_{int}(x_1) \dots \mathcal{H}_{int}(x_n)\}. \tag{1.36}$$

If $\mathcal{H}_{int}$ is a Lorentz scalar then this is manifestly Lorentz-invariant except for the time-ordering symbol T. As long as $\mathcal{H}_{int}$ commutes with itself at spacelike separation however, then the time ordering is also independent of Lorentz frame and so S will indeed be Lorentz-invariant

$$U(\Lambda, a)^\dagger S U(\Lambda, a) = S. \tag{1.37}$$

[2]Here I'll introduce a standard notation for the rest of the class: when I write $\vec{x}$ I mean a point in space, while when I write x I mean a point $(t, \vec{x})$ in spacetime.

We will discuss scattering theory in more detail later in the class.

How then can we build an $\mathcal{H}_{int}$ which is a Lorentz scalar that commutes with itself at spacelike separation? The only idea which seems to work is that it should be built out of **fields**: linear combinations of the creation and annihilation operators of the form[3]

$$\phi_i(x) = \sum_{\sigma,n} \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \Big(u_i(x;p,\sigma,n) a(p,\sigma,n) + v_i(x;p,\sigma,n) a^\dagger(p,\sigma,n) \Big), \tag{1.38}$$

with the coefficient functions u_i and v_i carefully constructed to ensure that[4]

$$[\phi_i(x), \phi_j(y)]_\pm = [\phi_i(x), \phi_j^\dagger(y)]_\pm = 0 \qquad (x-y)^2 > 0, \tag{1.39}$$

and also that under Poincaré transformations we have a simple transformation law

$$U(\Lambda,a)^\dagger \phi_i(x) U(\Lambda,a) = \sum_j D_{ij}(\Lambda) \phi_j(\Lambda^{-1}(x-a)). \tag{1.40}$$

Here we have allowed for multiple species of particle labeled by n, and also for the particles have spin σ, in which case the creation and annihilation operators need to be labeled by n and σ in addition p. You will show in the homework that the consistently composing Poincaré transformations requires the matrices $D_{ij}(\Lambda)$ to furnish a **representation** of the Lorentz group in the sense that

$$\sum_j D_{ij}(\Lambda_1) D_{jk}(\Lambda_2) = D_{ik}(\Lambda_1 \Lambda_2). \tag{1.41}$$

The requirement (1.39) of (anti)commutativity at spacelike separation is sometimes called **microcasaulity**. Given fields obeying (1.39) and (1.40), it is then a straightforward matter to construct interaction Hamiltonians which are Lorentz scalars. For example given a vector field $V^\mu(x)$ transforming as

$$U(\Lambda,a)^\dagger V^\mu(x) U(\Lambda,a) = \Lambda^\mu{}_\nu V^\nu(\Lambda^{-1}(x-a)), \tag{1.42}$$

some local interactions we could write down which are Lorentz scalars are

$$(V^\mu V_\mu)^2 \qquad V^\mu V^\nu \partial_\mu V_\nu \qquad V^\mu V_\mu \partial_\nu V^\nu, \tag{1.43}$$

which all commute with each other at spacelike separation since $V^\mu(x)$ does.

So far you might be tempted to view this construction as just more bookkeeping: we are still working in our old multi-particle Hilbert space and constructing things using creation and annihilation operators (albeit in nice linear combinations). How can bookkeeping fix a problem with causality? The key point is that we now make a fundamental shift in how we physically interpret all of the above equations:

- ⋆ In quantum field theory, we postulate that the observables that can be measured in the vicinity of a spacetime point x are those constructed from the fields at x, **not** those constructed from the position-space creation and annihilation operators $a^\dagger(x)$ and $a(x)$.

We will see in section 3 below that the $a(x)$ and $a^\dagger(x)$ are non-local when expressed in terms of the fields, so the apparent failure of causality we saw above is really just a consequence of failing to identify the right physical degrees of freedom. If we build a detector here in this room right now, the claim is that what it really couples to are the fields and not the particles.

We are already in a position to see two of the most remarkable consequences of relativistic quantum field theory:

[3]Here we work in the "interaction picture", where operators evolve under the free Hamiltonian H_0. Heisenberg picture fields in interacting theories cannot be decomposed in this way, and when interactions are strong the interaction picture is not useful so this motivation needs some revisiting (see the next section).

[4]Here $\pm$ indicates that for fields which create fermions we actually want anticommutativity instead of commutativity at spacelike separation.

- **In interacting quantum field theories, the number of particles is not conserved:** we have not yet specified the functions u_i and v_i, but we will see soon that both most be nonzero in order to preserve commutativity at spacelike separation. Interactions which are polynomials of the fields will thus always heuristically have the form $(a + a^\dagger)^n$, which necessarily includes terms that do not have the same number of creation and annihilation operators and thus do not conserve particle number. We should therefore expect that particle scattering in field theory can create any set of particles for which there is sufficient energy, at least as long as the final particles have the same symmetry charges as the initial particles. The idea that energy can be freely converted into particles is quite natural from the point of view of Einstein's equation $E = mc^2$.

- **Every particle must have an antiparticle of equal mass and opposite charge:** we will see soon that the time-dependence of u_i and v_i is given by $e^{\mp i\sqrt{p^2+m^2}t}$, so to preserve commutativity at spacelike separation for all times, which requires a cancellation between terms involving both u_i and v_i, these must have the same time-dependence and thus multiply creation/annihilation operators for particles of the same mass. On the other hand they must have opposite charge under any continuous internal symmetry. This is because in order to have an internal symmetry of the Lagrangian we need the field to have a simple transformation law

$$e^{-iQ\theta}\phi_i(x)e^{iQ\theta} = e^{i\theta q}\phi_i(x), \tag{1.44}$$

  where Q is the charge operator for the symmetry and q is the charge of the field, which means that the annihilation and creation operators appearing in ϕ_i must both transform by a factor $e^{i\theta q}$. This means that the particles created by the creation operator have opposite charge of those annihilated by the annihilation operator (you will show this in the homework). A particle can be its own antiparticle, but only if it has $q = 0$ for all continuous symmetries.

There are other important general consequences we will understand later, including:

- **Spin-statistics theorem:** in constructing u_i and v_i, it turns out that commutativity at spacelike separation is only possible when the particles that are created/annihilated have integer spin. For particles of half-integer spin, we instead need to impose **anticommutativity** at spacelike separation. As mentioned above, commutativity leads to bosons and anticommutativity leads to fermions. Hence we see that bosons must have integer spin and fermions must have half-integer spin.

- $\mathcal{CRT}$ **theorem**: In any relativistic quantum field theory, it turns out that there is always a symmetry that exchanges particles and antiparticles ($\mathcal{C}$), reflects a spatial direction ($\mathcal{R}$), and reverses time ($\mathcal{T}$).

All of these predictions have been confirmed to remarkable precision by experiment, for example colliding two photons at high energy can produce an electron-positron pair, a neutron decays to a proton, an electron, and a neutrino, the antiparticle of the electron is the positron, electrons are fermions of spin 1/2 while photons are bosons of spin one, and it was recently confirmed that hydrogen and antihydrogen have the same rate for the $2s \to 1s$ transition, as required by $\mathcal{CRT}$ symmetry.

1.2 Many-body quantum systems with local interactions

There is another way to motivate quantum field theory. Let's imagine a physical system with a large number of independent degrees of freedom that are arrayed in a lattice pattern in space, as in figure 4. What it means to say that the degrees of freedom are independent is that the Hilbert space of the theory has a tensor product form

$$\mathcal{H} = \bigotimes_i \mathcal{H}_i, \tag{1.45}$$

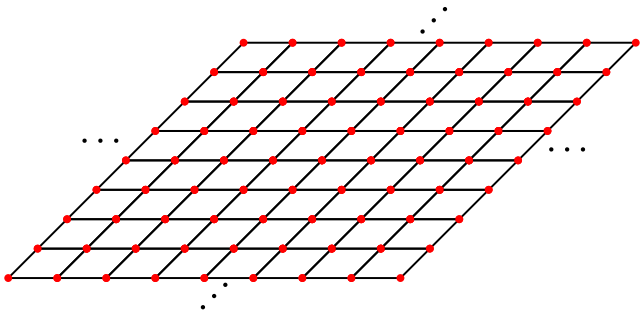


Figure 4: A lattice system with two spatial dimensions. There are independent degrees of freedom at each of the red sites, and each term in the Hamiltonian only couples degrees of freedom on nearby sites.

where i labels the sites on the lattice.[5] We say that an operator is a **local operator** at site i if it is the tensor product of an operator on $\mathcal{H}_i$ with the identity operator on all of the other sites. We are then interested in Hamiltonians of the form

$$H = \sum_i O_i, \tag{1.46}$$

where each O_i is built from local operators at sites in the vicinity of i, meaning sites that are an $O(1)$ number of links away (as opposed to something that grows with the total size of the system). Such Hamiltonians are called **local Hamiltonians**. For example our lattice could be ions in a crystal, and the degrees of freedom at the sites could describe local displacements of the ions. Another example we will come back to repeatedly is the quantum Ising model in a transverse field, where each $\mathcal{H}_i$ is a two-level system and the Hamiltonian is given by

$$H = -\sum_i \sigma_x(i) - \lambda \sum_{\langle ij \rangle} \sigma_z(i)\sigma_z(j), \tag{1.47}$$

where $\langle ij \rangle$ indicates nearest neighbors on the lattice.

In studying many-body local quantum systems we are usually not interested in the details of what is happening on the lattice scale. For example if you look at the atomic scale a superconductor is a giant mess: it is only when you zoom out and look at the long-distance behavior that you can see that something remarkable is going on. As we do this zooming out process, it becomes harder and harder to see that there is really a lattice and the system starts to look continuous in space. In other words it starts to look like a system whose Hamiltonian has the form

$$H = \int d^{d-1}x \mathcal{H}(\vec{x}), \tag{1.48}$$

with the energy density $\mathcal{H}(\vec{x})$ being built out of operators localized at $\vec{x}$. Moreover local operators at $\vec{x}$ will commute with local operators at $\vec{y}$, since they live on different tensor factors of the microscopic Hilbert space (1.46). Thus it starts to look like a quantum field theory! This zooming out process is called the **renormalization group**, and it is an idea of fundamental importance for any dynamical system (including classical systems) with local interactions at short distances.

It is often the case that the interesting long-distance excitations of a many-body quantum system look rather different than the fundamental lattice degrees of freedom. For example:

- In a crystal, the fundamental degrees of freedom are protons and electrons interacting through Coulomb forces but at long distances the excitations are **phonons**, which are ripples made out of vibrations in the lattice structure.

[5] This is not the most general possibility, as we could also add degrees of freedom on the links of the lattice, faces, etc, and also perhaps constrain the physical states by imposing some kind of local constraint. We will see these generalizations arise later when we consider gauge theories.

- In quantum chromodynamics (QCD), which is the fundamental theory of strong interactions, the fundamental degrees of freedom are quarks and gluons but the long-distance excitations are **hadrons** such as protons, neutrons, and pions.

- In $1+1$ dimensions the fundamental degrees of freedom of the quantum Ising model (1.47) are Pauli spins, but at the "critical point" $\lambda = 1$ the long-distance excitations are pairs of non-interacting massless fermions. This is also the essence of why the two-dimensional *classical* Ising model is solvable.

These examples illustrate an important weak point of our above argument that we need quantum field theory to combine special relativity and quantum mechanics: when the interactions of the fundamental fields appearing in the Lagrangian are strong, such as in QCD at low energies, there need not be any simple relationship between these fundamental fields and the low-energy particle excitations. What the argument leading to (1.38) really constructs is a "low-energy effective field theory", whose fields create and annihilate the low-energy excitations. In quantum field theory the basic question we are often really trying to answer is the following: given some short-distance formulation of the theory using local fields, what are the long-distance excitations and how do they interact?

It is also worth emphasizing that, although we began this discussion by talking about particles, not all quantum field theories lead to particles. Field theories without particles include "conformal field theories", which are more naturally understood in terms of correlation functions of local operators with simple scaling transformations, and "topological field theories", which are more naturally understood in terms of the algebra of extended "surface" operators which can be freely deformed in spacetime. Moreover these are not weird esoteric theories: the long-distance description of any second-order phase transition is a conformal field theory, and the fractional quantum hall effect is described by a topological quantum field theory. Even in the standard model of particle physics, there are "infrared divergences" arising from the presence of massless particles such as the photon and dealing with these correctly requires us to consider asymptotic states which are clouds of infinite numbers of particles rather than individual particles. In quantum field theory it is the fields that are essential, not the particles.

There is an important caveat to mention here: we are quite confident that the laws of nature are relativistic, so in high-energy physics we are for the most part only interested in relativistic quantum field theories. In condensed matter physics however Lorentz invariance can be broken by the existence of the material we are studying, so the field theories that show up in condensed matter physics do not need to be relativistic. Sometimes they are however, for example in the case of the quantum Ising model or the fractional quantum hall system, and the methods you learn in this class generalize to the non-relativistic case without much difficulty.

1.3 Quantum field theory in quantum gravity

We have seen that quantum field theory gives a way to successfully combine quantum mechanics and special relativity. The next frontier in fundamental physics is learning how to combine quantum field theory and *general* relativity, which is Einstein's theory of gravity. General relativity has been quite successful in explaining gravitational phenomena in astrophysics and cosmology, but it is a classical theory and so far attempts to "quantize" it in the conventional way (which we will review next time) have not been successful. A number of lines of reasoning have led to the idea that a theory which combines gravity and quantum mechanics will need to be "holographic", in the sense that its fundamental formulation lives in fewer spacetime dimensions than are perceived by long-distance observers such as ourselves. The most concrete example of this phenomenon is the "AdS/CFT" correspondence, which says that quantum gravity in a universe with negative cosmological constant is equivalent to a standard quantum field theory (in fact a conformal field theory, hence "CFT") living in one fewer spacetime dimension. This idea has been realized (and in fact was discovered) within the broader framework of "string theory", which is a speculative proposal for a theory of quantum gravity based on dynamical objects called "branes" (short for membrane), which have spatial volumes of various dimension. Often these branes have the feature that at long distances the gravity in the ambient space the live in can be ignored, in which case their long-distance excitations are again controlled

by quantum field theory. In fact AdS/CFT correspondence arises in string theory in precisely this way. This also gives a novel way of constructing interesting quantum field theories that so far are not accessible by the more conventional techniques based on Lagrangians that we will use in this class.

1.4 Mathematical difficulties

One of the difficulties of learning quantum field theory is that many of the standard manipulations are difficult to justify in a mathematically rigorous manner. These mathematical problems arise because in quantum field theory there is formally an infinite number of degrees of freedom: at least one for each point in space. This leads to two different kinds of divergences in mathematical expressions, short-distance or "UV" divergences and long-distance or "IR" divergences. The former arise because in a finite volume of space there are infinitely many degrees of freedom due to the continuous nature of space. These divergences can be regularized by working on a spatial lattice as in figure 4. Conventionally the lattice spacing is called a, and its inverse

$$\Lambda := \frac{1}{a} \tag{1.49}$$

is called the **UV cutoff**. IR divergences instead arise because the volume of flat space is infinite: these divergences are present even in the presence of a spatial lattice, so to regulate them we need to make the spatial volume V finite. In most cases we are only really able to make sense of quantum field theory in a mathematically rigorous way when both Λ and V are finite. We then need to master the art of constructing appropriate observables that stay finite in the limit that Λ and V both go to infinity. For example at finite temperature the total energy goes to infinity as $V \to \infty$ but the energy density stays finite. Similarly the fluctuations of a field $\phi(x)$ go to infinity as $\Lambda \to \infty$ but the fluctuations of a "smeared" field

$$\phi_f = \int d^d x f(x)\phi(x) \tag{1.50}$$

stays finite. In the latter f is taken to be a smooth function of compact support.

1.5 What this course is and is not

Finally I'll make a few comments about the philosophy of this class. The traditional approach to teaching quantum field theory is based on getting to perturbative calculations of scattering processes in the standard model of particle physics as quickly as possible. For this approach see e.g. the books of Peskin and Schroeder or Schwartz. This will not be the approach we take here. Although particle physics was the original arena of interest for quantum field theory, today it has grown far beyond these beginnings. Quantum field theorists today study many different quantum field theories in a variety of spacetime dimensions, and it would be a mistake to focus so narrowly on one particular quantum field theory in one particular spacetime dimension.[6] Indeed the traditional approach to teaching field theory is a bit like designing an electromagnetism class to get to capacitors as quickly as possible, and then staying there for months. Our approach will instead be to study quantum field theory as a general framework for analyzing many-body quantum systems. In illustrating quantum field theory phenomena we will typically go for the simplest model that exhibits them, and although we will sometimes use perturbation theory we will work non-perturbatively whenever it is possible. Many of the most exciting quantum field theory phenomena such as confinement and duality are fundamentally non-perturbative, and an overly perturbative class along the traditional lines would miss them. We will eventually talk about the standard model of particle physics, since after all it is good to know the fundamental laws of nature, but we will view it as one application among many rather than the main goal of our labors.

[6] In fact even for particle physics applications it turns out to be useful to work in d spacetime dimensions, as this enables us to use 't Hooft and Veltman's loony (but brilliant) dimensional regularization method for computing Feynman diagrams.

1.6 Homework

1. Let's get some practice using natural units:

   (a) What is the mass of the sun measured in Joules? What about in electron volts? Recall that $1\,\mathrm{eV} = 1.6 \times 10^{-19}\mathrm{J}$.

   (b) What is one meter in inverse electron volts?

   (c) The mass of a proton is $1.67 \times 10^{-27}\,\mathrm{kg}$. What is this in electron volts? What do we get if we convert it to an inverse length? How does this length compare to the size of a nucleus?

   (d) The mass of an electron is $9.1 \times 10^{-31}\,\mathrm{kg}$. What is this in electron volts? What do we get if we convert it to an inverse length? How does this length compare to the size of an atom? Any thoughts about how this comparison went versus the one for the nucleus?

   (e) Say that a force is quoted to you in units of eV^2. What factors of c and $\hbar$ should you supply to convert it back to Newtons?

   (f) Say that an energy flux is quoted to you in units of eV^4. What factors of c and $\hbar$ should you supply to convert it back to Joules per meter squared per second?

   If you are having trouble with these, a good way to proceed is to remember that $\hbar$ has units of energy times time and c has units of length over time. So you can use c to convert all lengths to times, and then use $\hbar$ to convert all times to energies. Masses can be converted to energy by multiplying by c^2,

2. Show that the multiparticle states (1.23) are normalized correctly. You will need to use the creation/annihilation algebra. If you are having trouble I recommend showing it recursively.

3. If an operator a annihilates particles of charge q, what is the commutator of the symmetry charge Q with a and $a^\dagger$? What are $e^{-iQ\theta}ae^{iQ\theta}$ and $e^{-iQ\theta}a^\dagger e^{iQ\theta}$?

4. Show that the second line of (1.25) follows from the first

5. Show that if we apply the Poincaré transformations (Λ_1, a_1) and then (Λ_2, a_2) in succession, the resulting Poincaré transformation is $(\Lambda_2\Lambda_1, \Lambda_2 a_1 + a_1)$. Then show that the field transformation (1.40) is consistent with the composition rule $U(\Lambda_2, a_2)U(\Lambda_1, a_1) = U(\Lambda_2\Lambda_1, \Lambda_2 a_1 + a_2)$ provided that the matrix D obeys the Lorentz representation condition (1.41).

2 Lagrangian field theory

Having motivated the idea of quantum fields from various directions, we now commence studying them in detail. We will begin with the *classical* theory of fields, starting from the Lagrangian point of view.[7]

2.1 Particle Lagrangians

We'll first briefly consider the Lagrangian mechanics of interacting particles, whose trajectories are parametrized by functions $x^a(t)$ with $a = 1, 2, \ldots, N$. Depending on how we interpret a we can think of this as N particles moving in one spatial dimension or as $N/(d-1)$ particles moving in $(d-1)$ spatial dimensions. We can think of $x^a(t)$ as an N-component vector evolving in time, which we will notate as $x(t)$. The dynamics are determined by the Lagrangian function $L\big(x, \dot{x}; t\big)$, with the rule being that physical trajectories are those around which the action functional

$$S[x] := \int_{t_i}^{t_f} dt L\big(x(t), \dot{x}(t); t\big) \tag{2.1}$$

is stationary up to terms at the future/past boundaries.[8] Note that the Lagrangian is **local** in time: at time t it only depends on the positions and velocities of the particles at time t. We have included t as a separate argument in the Largangian to allow it to have some explicit time-dependence, for example through a time-dependent background field that the particle is moving in. To study stationarity, we insert an infinitesimal variation

$$x'(t) = x(t) + \delta x(t) \tag{2.2}$$

into the action:

$$\begin{aligned} S[x'] =& S[x] + \sum_a \int_{t_i}^{t_f} dt \left[\frac{\partial L}{\partial x^a} \delta x^a(t) + \frac{\partial L}{\partial \dot{x}^a} \delta \dot{x}^a(t) \right] \\ =& S[x] + \sum_a \int_{t_i}^{t_f} dt \left[\frac{\partial L}{\partial x^a} - \frac{\dot{\partial} L}{\partial \dot{x}^a} \right] \delta x^a(t) + \sum_a \left(\frac{\partial L}{\partial \dot{x}^a} \delta x^a(t) \right) \Bigg|_{t_i}^{t_f}. \end{aligned} \tag{2.3}$$

The third term in the second line consists of a future boundary term and a past boundary term, so stationarity means that the second term should vanish for all variations $\delta x^a(t)$. In other words the Euler-Lagrange equations

$$\frac{\partial L}{\partial x^a} = \frac{\dot{\partial} L}{\partial \dot{x}^a} \tag{2.4}$$

must hold. For example if we have

$$L = \frac{m}{2} \dot{x}^2 - V(x), \tag{2.5}$$

then we must have

$$m \ddot{x}^a = -\frac{\partial V}{\partial x^a}. \tag{2.6}$$

We can pass to the Hamiltonian formalism by introducing the canonical momenta

$$p_a = \frac{\partial L}{\partial \dot{x}^a}, \tag{2.7}$$

[7] I'll present the traditional approach that assumes the Lagrangian depends only on the fields and their first derivatives. Later in the semester we will also be interested in theories with more derivatives in the Lagrangian: the traditional method for dealing with this is to introduce auxiliary fields to rewrite the Lagrangian in a way that only involves first derivatives. For a more modern approach that works directly with the original fields see my paper 1906.08616 with Jie-qiang Wu.

[8] You may have been taught that the action should be stationary without qualification. This is true if we fix boundary conditions at t_f/t_i, but doing that amounts to singling out some particular set of initial/final conditions. We are trying to characterize the theory as a whole, so we shouldn't bias the discussion by picking out some particular state of the system.

and the Hamiltonian is given by

$$H = \sum_a p_a \dot{x}^a - L. \tag{2.8}$$

To quantize we replace the Poisson bracket

$$\{x^a, p_b\} = \delta^a_b \tag{2.9}$$

by a commutator

$$[x^a, p_b] = i\delta^a_b, \tag{2.10}$$

which we represent on a Hilbert space spanned by eigenstates $|x\rangle$ of the X^a operator:

$$X^a|x\rangle = x^a|x\rangle. \tag{2.11}$$

2.2 Field Lagrangians

Let's now generalize from particles to fields. We will consider fields living in d-dimensional Minkowski space, with spatial points labeled by a $(d-1)$-vector

$$\vec{x} = (x^1, x^2, \ldots, x^{d-1}). \tag{2.12}$$

The field trajectories are given by functions $\phi^a(t, \vec{x})$, where a is a label that runs over some finite number of fields. This can be viewed as a generalization of the previous subsection in two different ways. The first way is that we are now allowing the trajectories to depend on space as well as time, in which case we go back to the particle case by taking $d = 1$. The second way is that we can think of each field at each point in space as a distinct particle, in which case we have generalized the previous subsection to an infinite number of particles. Our notation is more closely aligned to the former interpretation, but the latter is valuable conceptually because it makes clear that fundamentally we shouldn't have to do anything for fields that we didn't already do for particles.

To specify the field dynamics we need a Lagrangian. As we are interested in constructing field theories which respect microcausality (i.e. commutativity at spacelike separation), we should take this Lagrangian to be an integral over space of a local Lagrangian density:

$$L[\phi; t] = \int d^{d-1}x\, \mathcal{L}\Big(\phi(t, \vec{x}), \dot{\phi}(t, \vec{x}), \vec{\nabla}\phi(t, \vec{x}); t, \vec{x}\Big). \tag{2.13}$$

Here $\mathcal{L}(\phi, \dot{\phi}, \vec{\nabla}\phi; t, \vec{x})$ is constructed out the fields and their derivatives at $(t, \vec{x})$, and we have allowed for explicit dependence on space and time. I've written $L[\phi; t]$ with square brackets to emphasize that it is a **functional**: it is a function of the functions ϕ^a and $\dot{\phi}^a$ throughout timeslice at time t. A simple example of a Lagrangian density is the free scalar field Lagrangian, where we have a single field $\phi(t, \vec{x})$ with

$$\mathcal{L}(\phi, \dot{\phi}, \vec{\nabla}\phi) = \frac{1}{2}\left(\dot{\phi}^2 - \vec{\nabla}\phi \cdot \vec{\nabla}\phi - m^2\phi^2\right), \tag{2.14}$$

where m is a parameter that we will see next time gives the mass of the particles created by this field. We could introduce explicit space and time dependence by letting m depend on t and $\vec{x}$.

To find the equations of motion we adopt the same principle as before: the action

$$S := \int_{t_i}^{t_f} dt d^{d-1}x\, \mathcal{L}\Big(\phi(t, \vec{x}), \dot{\phi}(t, \vec{x}), \vec{\nabla}\phi(t, \vec{x}); t, \vec{x}\Big) \tag{2.15}$$

should be stationary about physical field configurations up to future and past boundary terms. Computing the variation, we find

$$\begin{aligned}\delta S =& \sum_a \int_{t_i}^{t_f} dt \int d^{d-1}x \Big[\frac{\partial \mathcal{L}}{\partial \phi^a} \delta\phi^a(t,\vec{x}) + \frac{\partial \mathcal{L}}{\partial \dot{\phi}^a} \delta\dot{\phi}^a(t,\vec{x}) + \frac{\partial \mathcal{L}}{\partial \vec{\nabla}\phi^a} \cdot \vec{\nabla}\delta\phi^a(t,\vec{x}) \Big] \\ =& \sum_a \int_{t_i}^{t_f} dt \int d^{d-1}x \Big[\frac{\partial \mathcal{L}}{\partial \phi^a} - \frac{\dot{\partial}\mathcal{L}}{\partial \dot{\phi}^a} - \vec{\nabla} \cdot \left(\frac{\partial \mathcal{L}}{\partial \vec{\nabla}\phi^a} \right) \Big] \delta\phi^a(t,\vec{x}) \\ &+ \sum_a \left(\int d^{d-1}x \frac{\partial \mathcal{L}}{\partial \dot{\phi}^a} \delta\phi^a \right) \Big|_{t_i}^{t_f} + \sum_a \int_{t_i}^{t_f} dt \int_{\mathbb{S}^{d-2}_\infty} d\vec{A} \cdot \frac{\partial \mathcal{L}}{\partial \vec{\nabla}\phi^a} \delta\phi^a(t,\vec{x}). \end{aligned} \tag{2.16}$$

The third line consists of future/past boundary terms and a spatial boundary term at the $(d-2)$-sphere $\mathbb{S}^{d-2}_\infty$ at spatial infinity. The former are acceptable, but the latter need to vanish in order for the theory to make sense. The usual way to deal with this is to impose spatial boundary conditions requiring the fields to vanish at infinity, in which case the variations $\delta\phi^a$ must also vanish at infinity and so this term vanishes.[9] We therefore see that the action will be stationary (up to future/past terms) if the Euler-Lagrange equations

$$\frac{\dot{\partial}\mathcal{L}}{\partial \dot{\phi}^a} + \vec{\nabla} \cdot \left(\frac{\partial \mathcal{L}}{\partial \vec{\nabla}\phi^a} \right) = \frac{\partial \mathcal{L}}{\partial \phi^a} \tag{2.17}$$

are satisfied throughout spacetime. For example for our free scalar field Lagrangian we have

$$\ddot{\phi} - \nabla^2\phi = -m^2\phi, \tag{2.18}$$

which is a massive version of the wave equation known as the **Klein-Gordon equation**.

As in the particle case we can also introduce a canonical momentum

$$\pi_a \equiv \frac{\partial \mathcal{L}}{\partial \dot{\phi}^a}, \tag{2.19}$$

in terms of which the Hamiltonian is given by

$$H[\phi;t] = \int d^{d-1}x\, \mathcal{H}\Big(\phi(t,\vec{x}), \dot{\phi}(t,\vec{x}), \vec{\nabla}\phi(t,\vec{x}); t, \vec{x} \Big) \tag{2.20}$$

with Hamiltonian density

$$\mathcal{H}(\phi, \dot{\phi}, \vec{\nabla}\phi; t, \vec{x}) = \sum_a p_a \dot{\phi}^a - \mathcal{L}(\phi, \dot{\phi}, \vec{\nabla}\phi; t, \vec{x}). \tag{2.21}$$

To complete the construction of the Hamiltonian formalism we need to solve (2.19) to determine $\dot{\phi}$ in terms of ϕ and p. Sometimes this is not possible due to constraints, in which case more sophisticated methods are needed that we will return to later. For the free scalar field there is no problem, we simply have

$$\pi = \dot{\phi} \tag{2.22}$$

and

$$\mathcal{H} = \frac{1}{2} \left(\pi^2 + \vec{\nabla}\phi \cdot \vec{\nabla}\phi + m^2\phi^2 \right). \tag{2.23}$$

[9] It is also interesting to consider field theories in finite volume, in which case there is more to say about this term. For example we could impose Neumann boundary conditions $\hat{n} \cdot \frac{\partial \mathcal{L}}{\partial \vec{\nabla}\phi^a} = 0$ instead of Dirichlet boundary conditions $\delta\phi^a = 0$ and it would still vanish. More generally we can take the action to include additional boundary terms at spatial infinity, whose variation is designed to cancel the spatial boundary term in (2.16) when we impose the boundary conditions of interest. Ultimately the choice of spatial boundary conditions is part of the definition of the theory: a box with Dirichlet boundary conditions is a different physical system from one with Neumann boundary conditions.

Once the Hamiltonian formalism is constructed, we can then quantize the theory by converting the equal-time Poisson brackets

$$\begin{aligned}\{\phi^a(t,\vec{x}),\phi^b(t,\vec{y})\} &= 0\\ \{\pi_a(t,\vec{x}),\pi_b(t,\vec{y})\} &= 0\\ \{\phi^a(t,\vec{x}),\pi_b(t,\vec{y})\} &= \delta^{d-1}(\vec{x}-\vec{y})\end{aligned} \tag{2.24}$$

to commutators in the usual way and then representing them on a Hilbert space spanned by field eigenstates $|\phi\rangle$ obeying

$$\Phi^a(0,\vec{x})|\phi\rangle = \phi^a(\vec{x})|\phi\rangle. \tag{2.25}$$

We will carry out this procedure in detail for the free scalar field next time, but we note that in a Lorentz-invariant theory the commutativity at spatial separation we see here extends to commutativity at spacelike separation.

2.3 Relativistic notation

It is now convenient to introduce relativistic notation by combining space and time into a d-vector

$$x = (t,\vec{x}). \tag{2.26}$$

We will write the components of $\vec{x}$ as x^i, with $i = 1,2,\ldots,d-1$, and the components of x as x^μ, with $\mu = 0,1,\ldots,d-1$. By definition we have

$$x^0 = t. \tag{2.27}$$

If we were mathematicians we would be zealous in adhering to using x for the vector and x^μ for its components, but in physics there is a longstanding tradition of conflating the two since writing x^μ instead of x has the convenient feature of reminding us what kind of object we are talking about (and after all if you know its components then you know the vector).[10] The inner product of two d-vectors v^μ and u^μ is given by

$$u\cdot v = u^\mu v^\nu \eta_{\mu\nu}, \tag{2.28}$$

where[11]

$$\eta_{\mu\nu} := \begin{pmatrix} -1 & 0 & 0 & \ldots & 0\\ 0 & 1 & 0 & \ldots & 0\\ 0 & 0 & 1 & \ldots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \ldots & 1\end{pmatrix} \tag{2.29}$$

is the d-dimensional Minkowski metric and we are using the **Einstein summation convention** that sums over pairs of repeated indices automatically. This inner product is preserved under Lorentz transformations

$$\begin{aligned}u'^\mu &= \Lambda^\mu{}_\nu u^\nu\\ v'^\mu &= \Lambda^\mu{}_\nu v^\nu,\end{aligned}$$

where $\Lambda^\mu{}_\nu$ is any $d\times d$ matrix that obeys

$$\Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \eta_{\mu\nu} = \eta_{\alpha\beta}. \tag{2.30}$$

[10] This abuse of notation is sometimes formalized by using "abstract index notation", which writes the abstract vector as x^a and its components as x^μ. This is done for example in Wald's book. We already have enough kinds of index to be getting on with however, so we will stick to being somewhat cavalier about the difference between x and x^μ (and analogously $\vec{x}$ and x^i). A similar remark applies about the difference between a function f and the evaluation $f(x)$ of that function on an element x of its domain, which we have already conflated several times.

[11] Some benighted particle theorists use a horrid "mostly-minus" convention for $\eta_{\mu\nu}$ that reverses its overall sign, and in this context our convention is called "mostly-plus". In general life is not improved by increasing the number of minus signs, and that is absolutely the case here.

Indeed we have

$$\begin{aligned} u' \cdot v' &= \Lambda^{\mu}{}_{\alpha} u^{\alpha} \Lambda^{\nu}{}_{\beta} v^{\beta} \eta_{\mu\nu} \\ &= \eta_{\alpha\beta} u^{\alpha} v^{\beta} \\ &= u \cdot v. \end{aligned} \tag{2.31}$$

The set of $d \times d$ matrices obeying (2.30) is called the **Lorentz group**, and we will have lots to say about it later.

It is also convenient to introduce d-component objects with a lowered Lorentz index, called **one-forms**, which transform as

$$\omega'_{\mu} = \Lambda_{\mu}{}^{\nu} \omega_{\nu}. \tag{2.32}$$

Here $\Lambda_{\nu}{}^{\mu}$ indicates the transpose of the inverse of $\Lambda^{\nu}{}_{\mu}$, meaning that it obeys

$$\Lambda^{\nu}{}_{\mu} \Lambda_{\nu}{}^{\alpha} = \delta^{\alpha}_{\mu} \tag{2.33}$$

with δ^{α}_{μ} being the **Kronecker delta** that is equal to one if $\alpha = \mu$ and zero otherwise. A simple example of a one-form is the scalar gradient

$$\partial_{\mu} \phi = (\dot{\phi}, \vec{\nabla} \phi), \tag{2.34}$$

which transforms with the inverse-transpose of Λ because the partial derivative transforms opposite to the spacetime coordinates. We can compute the inner product of two one-forms by using the inverse metric $\eta^{\mu\nu}$, which again is diagonal with diagonal elements $(-1, 1, \ldots, 1)$:

$$\omega \cdot \sigma = \omega_{\mu} \sigma_{\nu} \eta^{\mu\nu}. \tag{2.35}$$

You will show on the homework that we can use the metric to turn a vector into a one-form and the inverse metric to turn a one-form into a vector by "lowering" and "raising" the indices

$$\begin{aligned} u_{\mu} &:= \eta_{\mu\nu} u^{\nu} \\ \omega^{\mu} &:= \eta^{\mu\nu} \omega_{\nu}, \end{aligned} \tag{2.36}$$

and also that the inverse-transpose Lorentz transformation $\Lambda_{\nu}{}^{\mu}$ is indeed obtained by raising/lowering the indices of the original Lorentz transformation $\Lambda^{\nu}{}_{\mu}$ in this way.

Using this notation we can write the free scalar field Lagrangian density more elegantly in a few different ways,

$$\begin{aligned} \mathcal{L}(\phi, \partial\phi) &= -\frac{1}{2} \left(\partial_{\mu} \phi \partial_{\nu} \phi \eta^{\mu\nu} + m^2 \phi^2 \right) \\ &= -\frac{1}{2} \left(\partial_{\mu} \phi \partial^{\mu} \phi + m^2 \phi^2 \right) \\ &= -\frac{1}{2} \left(\partial\phi \cdot \partial\phi + m^2 \phi^2 \right), \end{aligned} \tag{2.37}$$

and the Klein-Gordon equation becomes

$$\left(\partial^2 - m^2 \right) \phi = 0. \tag{2.38}$$

More generally we can write the action as

$$S[\phi; x] = \int d^d x \, \mathcal{L}\Big(\phi(x), \partial\phi(x); x \Big) \tag{2.39}$$

and the Euler-Lagrange equations as

$$\partial_{\mu} \left(\frac{\partial \mathcal{L}}{\partial \partial_{\mu} \phi^a} \right) = \frac{\partial \mathcal{L}}{\partial \phi^a}. \tag{2.40}$$

2.4 Symmetries and currents

One of the most important advantages of the Lagrangian formalism is the close relationship between symmetries and conserved quantities. By definition a symmetry in the Lagrangian formalism is an invertible change of variables

$$\phi'^a(x) = F[\phi] \tag{2.41}$$

which leaves the action invariant up to future/past boundary terms. A particularly interesting kind of symmetry is a **continuous symmetry**, which is a family of symmetries F_θ labeled by a continuous parameter θ such that when $\theta = 0$ the transformation F_θ is the identity. In particular we can take θ to be infinitesimal, in which case we'll call it ϵ, and we then have a field theory version of Noether's theorem: *any infinitesimal transformation of the fields that leaves the action invariant up to future/past boundary terms leads to a conserved current.* Indeed consider an infinitesimal transformation[12]

$$\phi'^a(x) = \phi^a(x) + \epsilon\delta_S\phi^a(x) \tag{2.42}$$

that to first order in ϵ leaves the action (2.39) invariant up to future/past boundary terms.[13] Here $\delta_S\phi^a$ is some function of ϕ^a and its derivatives at x, and possibly also x itself explicitly. One way for this to happen is for the Lagrangian density itself to be invariant, but more generally its transformation could be the divergence of a vector since that would still integrate to a future/past boundary term (assuming that the spatial boundary terms vanish). More explicitly, in order for the transformation (2.42) to be a symmetry we need

$$\delta_S\mathcal{L} = \sum_a \left(\frac{\partial\mathcal{L}}{\partial\phi^a}\delta_S\phi^a + \frac{\partial\mathcal{L}}{\partial\partial_\mu\phi^a}\partial_\mu\delta_S\phi^a \right) = \partial_\mu\alpha^\mu, \tag{2.43}$$

where α^μ is some local function of ϕ and $\partial\phi$. We can rewrite this expression as

$$\partial_\mu \left(\sum_a \frac{\partial\mathcal{L}}{\partial\partial_\mu\phi^a}\delta_S\phi^a - \alpha^\mu \right) = \sum_a \left(\partial_\mu \frac{\partial\mathcal{L}}{\partial\partial_\mu\phi^a} - \frac{\partial\mathcal{L}}{\partial\phi^a} \right) \delta_S\phi^a, \tag{2.44}$$

and then observe that the right-hand side vanishes for field configurations $\phi^a(x)$ that obey the Euler-Lagrange equations (2.40). In other words we see that the **Noether current**

$$J^\mu(x) := -\sum_a \frac{\partial\mathcal{L}}{\partial\partial_\mu\phi^a}\delta_S\phi^a(x) + \alpha^\mu(x) \tag{2.45}$$

obeys the conservation law

$$\partial_\mu J^\mu = 0. \tag{2.46}$$

Writing this equation in non-relativistic notation we have

$$\vec{\nabla}\cdot\vec{J} = -j^0, \tag{2.47}$$

which is precisely the continuity equation familiar from electromagnetism. In the usual way it implies that the charge

$$Q(t) = \int d^{d-1}x\, J^0(t,\vec{x}) \tag{2.48}$$

is independent of time. We have chosen the overall sign and normalization of J^μ so that Q is the generator of the symmetry in the sense that for any observable O we have the Poisson bracket[14]

$$\{Q, O\} = \delta_S O. \tag{2.49}$$

[12] It is not obvious, but every infinitesimal symmetry can be "exponentiated" to produce a continuous symmetry so the two ideas are equivalent.

[13] It is important here that the action needs to be invariant for any $\phi^a(x)$, not just solutions of the equations of motion. In the latter case the action is always invariant to first order under any continuous transformation!

[14] This Poisson bracket is easy to derive when $\delta_S\phi$ depends only on ϕ and not its derivatives and $\alpha = 0$. The general case is tricky and I haven't found a textbook discussion, the only derivation I know is given in section 4.2 of 1906.08616.

After quantization this becomes

$$[Q, O] = i\delta_S O. \tag{2.50}$$

As a simple example of this construction let's return to our free scalar theory and now set the mass m to zero. The Lagrangian density is then invariant under the **shift symmetry**

$$\phi'(x) = \phi(x) + \epsilon, \tag{2.51}$$

so we have a continuous symmetry with

$$\begin{aligned} \delta_S \phi &= 1 \\ \alpha^\mu &= 0. \end{aligned} \tag{2.52}$$

The Noether current is given by

$$J^\mu = \partial^\mu \phi, \tag{2.53}$$

and the conservation law follows immediately from the (massless) wave equation $\partial^2 \phi = 0$.

2.5 Noether currents for Poincaré symmetry

A more sophisticated example of a continuous symmetry in field theory is **Poincaré symmetry**, which is the full set of symmetries obtained by combining Lorentz transformations and spacetime translations. A general Poincaré transformation can be put in the form

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\nu, \tag{2.54}$$

where Λ is a Lorentz transformation obeying (2.30) and a is an arbitrary vector, but to interpret it as a dynamical symmetry for Noether's theorem we need to recast it as a transformation of the fields rather than the coordinates.[15] On a scalar field the transformation is simple to write down: we have

$$\phi'(x) = \phi(\Lambda^{-1}(x - a)). \tag{2.55}$$

Here the inverse transformation appears inside the field so that the composition of Poincaré transformations works out correctly, as you showed on the previous homework.

To apply Noether's theorem we need to understand infinitesimal Lorentz transformations, meaning we should write

$$\Lambda^\mu{}_\nu = \delta^\mu_\nu + \epsilon \omega^\mu{}_\nu \tag{2.56}$$

and substitute into (2.30) to see what the constraints are on $\omega^\mu{}_\nu$. Indeed we have

$$\left(\delta^\mu_\alpha + \epsilon \omega^\mu{}_\alpha\right)\left(\delta^\nu_\beta + \epsilon \omega^\nu{}_\beta\right)\eta_{\mu\nu} = \eta_{\alpha\beta} + \epsilon\left(\omega_{\alpha\beta} + \omega_{\beta\alpha}\right) + O(\epsilon^2), \tag{2.57}$$

so for (2.30) to hold we need ω with both indices down to be antisymmetric:

$$\omega_{\beta\alpha} = -\omega_{\alpha\beta}. \tag{2.58}$$

Including also an infinitesimal translation $a^\mu = \epsilon b^\mu$ we therefore have

$$\begin{aligned} \phi(\Lambda^{-1}(x - a)) &= \phi(x - \epsilon(b + \omega x) + O(\epsilon^2)) \\ &= \phi(x) - \epsilon(b^\nu + \omega^\nu{}_\alpha x^\alpha)\partial_\nu \phi(x) + O(\epsilon^2), \end{aligned} \tag{2.59}$$

[15] The viewpoint where the fields transform and the coordinates stay the same is sometimes called the **active** viewpoint, to be distinguished from a **passive** viewpoint where the coordinates transform and the fields stay the same. As in many situations, here it is better to be active.

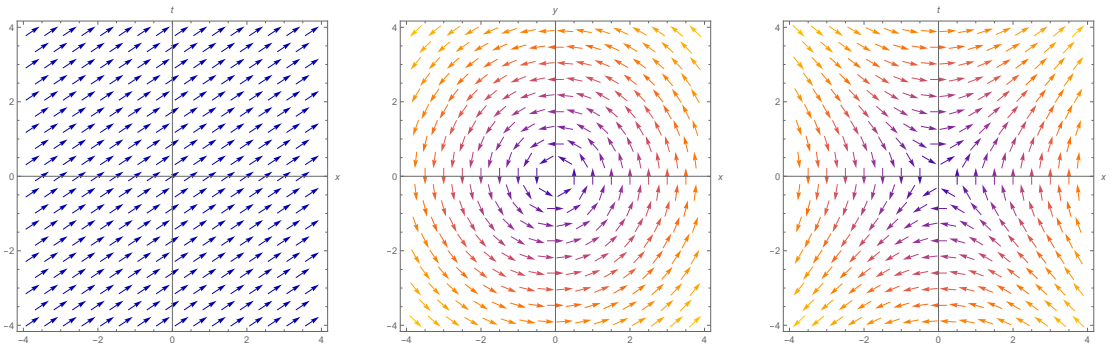


Figure 5: Killing vector fields for a spacetime translation, a spatial rotation, and a boost.

and thus

$$\begin{aligned}\delta_S\phi(x) &= -(b^\mu + \omega^\mu{}_\alpha x^\alpha)\partial_\mu\phi(x) \\ &= -\xi^\mu(x)\partial_\mu\phi(x),\end{aligned} \tag{2.60}$$

where in the second line we have introduced a **Killing vector field**

$$\xi^\mu(x) := b^\mu + \omega^\mu{}_\alpha x^\alpha \tag{2.61}$$

that we can view as pointing in the direction of the infinitesimal Poincaré transformation in question. For example an infinitesimal boost in the x^1 direction has $\omega^0{}_1 = \omega^1{}_0 = 1$ and thus

$$\xi^\mu = (x^1, x^0, 0, \ldots, 0), \tag{2.62}$$

while a pure rotation in the 12 plane has $\omega^2{}_1 = -\omega^1{}_2 = 1$, and thus

$$\xi^\mu = (0, -x^2, x^1, 0, \ldots, 0). \tag{2.63}$$

These Killing vector fields are illustrated in figure 5. More generally a Killing vector field is by definition a vector field for which

$$\partial_\mu\xi_\nu + \partial_\nu\xi_\mu = 0, \tag{2.64}$$

as you can easily check is the case here.[16] Contracting this equation with the inverse metric, we also see that

$$\partial_\mu\xi^\mu = 0. \tag{2.65}$$

In Minkowski space (2.61) gives the full set of Killing vectors. It is spanned by $d-1$ infinitesimal boosts, $(d-1)(d-2)/2$ infinitesimal rotations, and d infinitesimal spacetime translations. For $d = 4$ this gives three boosts (in the x, y, and z directions), three rotations (in the xy, yz, and zx planes), three space translations (in the x, y, and z directions), and one time translation.

By definition a theory which is Poincaré-invariant is one whose Lagrangian density is a scalar under Poincaré transformations, meaning that

$$\delta_S\mathcal{L} = -\xi^\mu\partial_\mu\mathcal{L} \tag{2.66}$$

for any Killing vector ξ^μ. You will check this equation (somewhat laboriously) for our free scalar theory in the homework. Since ξ^μ is a Killing vector, by equation (2.65) we have

$$\xi^\mu\partial_\mu\mathcal{L} = \partial_\mu(\xi^\mu\mathcal{L}) \tag{2.67}$$

[16]The motivation for this definition is that an infinitesimal coordinate transformation $x'^\mu = x^\mu + \xi^\mu(x)$ leaves the spacetime metric $\eta_{\mu\nu}$ invariant if and only if ξ^μ obeys (2.64).

and thus

$$\delta_S \mathcal{L} = \partial_\mu \alpha^\mu \tag{2.68}$$

with

$$\alpha^\mu = -\xi^\mu \mathcal{L}. \tag{2.69}$$

For any Killing vector ξ^μ we therefore have a conserved Noether current

$$J^\mu_\xi(x) = -\sum_a \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi^a} \delta_S \phi^a(x) - \xi^\mu \mathcal{L}(x). \tag{2.70}$$

In the free scalar theory we can evaluate this, giving

$$\begin{aligned} J^\mu_\xi &= -\xi^\alpha \partial^\mu \phi \partial_\alpha \phi + \xi^\mu \frac{1}{2} \left(\partial_\alpha \phi \partial^\alpha \phi + m^2 \phi^2 \right) \\ &= -\xi_\nu T^{\mu\nu}, \end{aligned} \tag{2.71}$$

where

$$T^{\mu\nu} := \partial^\mu \phi \partial^\nu \phi - \frac{1}{2} \eta^{\mu\nu} \left(\partial_\alpha \phi \partial^\alpha \phi + m^2 \phi^2 \right) \tag{2.72}$$

is called the **energy-momentum tensor**. It has two nice properties:

(1) Symmetry:

$$T^{\mu\nu} = T^{\nu\mu} \tag{2.73}$$

(2) Conservation:

$$\partial_\mu T^{\mu\nu} = 0. \tag{2.74}$$

Indeed any tensor obeying these two properties has the feature that contracting it with a Killing vector gives a conserved current:

$$\partial_\mu (\xi_\nu T^{\mu\nu}) = \partial_\mu \xi_\nu \, T^{\mu\nu} + \xi_\nu \partial_\mu T^{\mu\nu} = 0. \tag{2.75}$$

It is not obvious from (2.70) that we can in general write J^μ_ξ in terms of a symmetric conserved energy-momentum tensor in this way, since when there are fields that are not scalars $\delta_S \phi^a$ can involve derivatives of ξ^μ,but it turns out that when such derivatives appear they can always be removed by shifting J^μ_ξ by local term whose divergence is identically zero.[17] The resulting energy-momentum tensor has a more elegant equivalent definition as the derivative of the action with respect to the spacetime metric:

$$S[\phi, \eta_{\mu\nu} + \epsilon h_{\mu\nu}] = S[\phi, \eta_{\mu\nu}] + \frac{\epsilon}{2} \int d^d x T^{\mu\nu}(x) h_{\mu\nu}(x) + O(\epsilon^2). \tag{2.76}$$

The metric is a symmetric tensor so this $T^{\mu\nu}$ obeys condition (1) automatically, and with a little more differential geometry than we are requiring for this class you can also show that it obeys condition (2) provided that there are no Lorentz-violating background fields.

We can understand the physical meaning of the energy momentum tensor by looking at the Noether currents for pure spacetime translations with $\omega_{\mu\nu} = 0$. By definition the total momentum vector P^λ, which is the generator of spacetime translations, is given by

$$\int d^{d-1} x \, J^0_\xi = -\xi_\lambda P^\lambda. \tag{2.77}$$

Therefore we have

$$P^\mu = \int d^{d-1} x \, T^{0\mu}, \tag{2.78}$$

[17] See section 7.4 of Weinberg for the case where the Lagrangian has only first derivatives, as we've been considering here, or appendix A of 2108.04841 for the general case.

so we can think of T^{00} as the energy density and T^{0i} as the momentum density (hence the name of the tensor). And indeed for our free scalar theory, from (2.72) we have

$$T^{00} = \frac{1}{2}\left(\dot{\phi}^2 + \vec{\nabla}\phi \cdot \vec{\nabla}\phi + m^2\phi^2\right), \tag{2.79}$$

consistent with the Hamiltonian density (2.23). We can also define generators of pure Lorentz transformations (with $b^\mu = 0$) via

$$\int d^{d-1}x J^0_\xi = \frac{1}{2}\omega_{\mu\nu}J^{\mu\nu}, \tag{2.80}$$

which gives

$$J^{\mu\nu} = \int d^{d-1}x \left(x^\mu T^{0\nu} - x^\nu T^{0\mu}\right). \tag{2.81}$$

Here J^{ij} is the angular momentum for a rotation in the ij plane, while $J^{i0} \equiv K^i$ is the generator of a boost in the i direction.

There are some annoying sign conventions in the previous paragraph, which perhaps are worth drawing attention to. From (2.77) we see that P^0 is the generator of time translation in the sense of equation (2.49), but the generator of spatial translations is actually $-P^i$. Similarly K^i is the generator of boosts in the i direction, but the generator of rotations in the ij plane is actually $-J^{ij}$. These sign differences arise because we really should have defined the generator of time translations to be P_0 instead of P^0, in which case we could have switched the sign of (2.49) in order to have momentum and angular momentum indeed be the generators of spatial translations and rotations. Unfortunately it is too late to change the sign of energy so we are stuck with these sign differences whatever choice we make for the sign in (2.49).[18]

[18] You might object that if we switch the sign then the energy becomes bounded from above instead of below, and the ground state becomes the state of highest energy. In fact I think this is preferable, as the ground state is really quite special and fine-tuned. We could perhaps instead call it the "pinnacle" or "paragon" state. Flipping the sign of the energy also makes more sense in statistical mechanics, for example what is usually called negative temperature is actually hotter than what is usually called positive temperature. Alas this is a battle which cannot possibly be won, so I will rest my case here.

2.6 Homework

1. Show that ${\Lambda_\mu}^\nu = \eta_{\mu\alpha}\eta^{\nu\beta}{\Lambda^\alpha}_\beta$ is indeed the inverse of ${\Lambda^\mu}_\nu$, in the sense that ${\Lambda^\mu}_\lambda{\Lambda_\nu}^\lambda = \delta^\mu_\nu$.

2. Show that the gradient $\partial_\mu\phi$ transforms as a one-form under the Poincaré transformation $\phi'(x) = \phi(\Lambda^{-1}(x-a))$.

3. Show that if V^μ is a vector then $V_\mu = \eta_{\mu\nu}V^\nu$ transforms as a one-form, and also that if ω_μ is a one-form then $\omega^\mu = \eta^{\mu\nu}\omega_\nu$ transforms as a vector.

4. The Lagrangian density for Maxwell theory is

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

where

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

is the field strength tensor and A_μ is a one-form usually called the gauge potential or gauge field. The relationship between A_μ and the usual scalar potential ϕ and vector potential $\vec{A}$ is that $A = (-\phi, \vec{A})$.

   (a) Write out the Euler-Lagrange field equations which follow from the Maxwell Lagrangian. Use the relativistic variables A_μ and $F_{\mu\nu}$.

   (b) For $d = 4$, give expressions for the components of $F_{\mu\nu}$ in terms of the usual electric and magnetic fields $\vec{E}$ and $\vec{B}$, and use these to rewrite the equations of motion in terms of $\vec{E}$ and $\vec{B}$. How do these relate to Maxwell's equations? Did you get all four equations, and if not where do the others come from?

   (c) Now add a term $A_\mu J^\mu$ to the Lagrangian density, where $J^\mu = (\rho, \vec{J})$ is the spacetime electric current. Show how this modifies the equations of motion, and check that for $d = 4$ it gives the correct charge and current terms in Maxwell's equations. In this part you can view J^μ as a "background" current, meaning that when you compute the variation of the action you can take its variation to be zero. Eventually we will build J^μ out of other fields which create charged particles, but this does not effect the equation of motion obtained by varying A_μ.

5. The Langrangian density for a *complex* free scalar field is given by

$$\mathcal{L} = -\partial^\mu\phi^*\partial_\mu\phi - m^2\phi^*\phi.$$

   (a) Find the Euler-Lagrange equations for this action. In principle in computing variations you should treat the independent fields as the real and imaginary parts of ϕ, but your life will be easier if you can convince yourself that you can instead treat ϕ and ϕ^* as the independent variables. Convince yourself that you indeed can do this for a general Lagrangian density $\mathcal{L}(\phi, \phi^*, \partial\phi, \partial\phi^*)$.

   (b) Show that the transformation $\phi'(x) = e^{i\theta}\phi(x)$ is a symmetry for any θ, write out its infinitesimal version (i.e. to linear order in θ), and construct the associated Noether current. Confirm explicitly that this current is conserved as a consequence of the equations of motion. You again will do better to view ϕ and ϕ^* as the independent fields.

   (c) Write an expression for the conserved symmetry charge Q, and check that it indeed generates the symmetry transformation as in equation (2.49).

6. Show explicitly that the free real scalar Lagrangian density obeys the invariance condition (2.66) under the infinitesimal transformation $\delta_S\phi = -\xi^\mu\partial_\mu\phi$ for any Killing vector ξ^μ.

7. (extra credit) The action of a free scalar field in a general metric $g_{\mu\nu}$ is given by

$$S = -\frac{1}{2}\int d^dx\sqrt{-g}\left(\partial_\mu\phi\partial_\nu\phi g^{\mu\nu} + m^2\phi^2\right),$$

where g indicates the determinant of the matrix $g_{\mu\nu}$ and $g^{\mu\nu}$ is its inverse. Show that if we take $g_{\mu\nu} = \eta_{\mu\nu} + \epsilon h_{\mu\nu}$, the energy momentum tensor we construct as in equation (2.76) is the same one we found from the Noether current. To do this you need to look up or derive how the determinant and inverse of a matrix respond to a small change in the matrix.

8. (extra credit) The Maxwell action in a general metric is

$$S = -\frac{1}{4}\int d^d x \sqrt{-g} F_{\mu\nu} F_{\alpha\beta} g^{\mu\alpha} g^{\nu\beta}.$$

What is energy-momentum tensor which follows from varying this action with respect to $g_{\mu\nu}$? For $d = 4$ write T_{00} in terms of the electric and magnetic fields; does the answer look familiar?

3 Quantization of a free scalar field

The previous section was rather formal. Formalism is good for organizing one's thinking, but to really understand things you need to get your hands dirty. In this section we will carry out in detail the canonical quantization of a free scalar field in d spacetime dimensions, with Lagrangian density

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{m^2}{2}\phi^2. \tag{3.1}$$

The word "free" here means that the Lagrangian is quadratic in the fields - we'll see that this implies that the particles in this theory are non-interacting. For now we will take ϕ to be real-valued, we will discuss soon how to generalize to the case of complex ϕ. The free scalar field is both simple and profound: it is exactly solvable, and yet it illustrates many of the deep aspects of quantum field theory that we will return to again and again. Before beginning it is worth emphasizing that this model is not only of interest as an example: it has many physical realizations. Some examples in various dimensions:

- The Higgs boson in the Standard Model of particle physics, discovered in 2012 at the LHC, is to first approximation described by a free scalar field with $d = 4$ and $m = 125\,\text{GeV}$.
- Helium 4 (He_4) at low temperature and standard pressure is a special kind of liquid, called a **superfluid**, which flows with zero viscosity. The low-energy excitations of this liquid are density waves called phonons, and they are described by a free scalar field theory with $d = 4$ and $m = 0$. If we confine Helium-4 to a two-dimensional surface, then it is described by a free scalar field with $d = 3$ and $m = 0$.
- The protons and neutrons in nuclei are held together by exchanging particles called pions, and these pions are governed at low-energy by free scalar fields with $d = 4$ and $m = 134\,\text{MeV}$ (for the π^0) and $m = 139\,\text{MeV}$ (for the $\pi^\pm$). The π^0 is a real scalar field, while the $\pi^\pm$ are complex (as we will introduce below).
- In string theory the embedding of the string worldsheet into spacetime is described using free scalar fields with $d = 2$ and $m = 0$.

In fact the 2016 Nobel prize in physics was awarded in substantial part for understanding the $d = 2$ version of this theory!

3.1 Canonical commutation relations and wave functionals

Let's first recall the Hamiltonian formulation of the free scalar: defining a canonical conjugate momentum

$$\pi = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = \dot{\phi}, \tag{3.2}$$

we have the Hamiltonian density

$$\mathcal{H} = \frac{1}{2}\pi^2 + \frac{1}{2}\nabla\phi\cdot\nabla\phi + \frac{m^2}{2}\phi^2. \tag{3.3}$$

Lifting the classical fields $\phi(x)$ and $\pi(x)$ to quantum operators $\Phi(x)$ and $\Pi(x)$ and their Poisson brackets to commutators, we have the algebra[19]

$$\begin{aligned}
&[\Phi(t,\vec{x}),\Phi(t,\vec{y})] = 0\\
&[\Pi(t,\vec{x}),\Pi(t,\vec{y})] = 0\\
&[\Phi(t,\vec{x}),\Pi(t,\vec{y})] = i\delta^{d-1}(\vec{x}-\vec{y}).
\end{aligned} \tag{3.4}$$

[19] You may wonder why the third commutator has a δ-function on the right-hand side instead of some kind of Kronecker-δ with continuous indices. This is because we defined π as a partial derivative of the Lagrangian density, as opposed to a partial derivative of the Lagrangian. The latter actually vanishes since it is multiplied by the infinitesimal $d^{d-1}x$, so in field theory it is better to use the former. With a lattice regulator they are related by a power of the lattice spacing a, as we will see in a moment.

The first step of canonical quantization is to represent this algebra on a Hilbert space, which in the particle case we take to be the vector space of square-normalizeable wave functions. We can do the same thing here, but we need to introduce a space of normalizeable **wave functionals**

$$\Psi[\phi] = \langle \phi | \Psi \rangle, \tag{3.5}$$

where $|\phi\rangle$ labels a complete eigenbasis for $\Phi(\vec{x}) := \Phi(0, \vec{x})$:

$$\Phi(\vec{x})|\phi\rangle = \phi(\vec{x})|\phi\rangle. \tag{3.6}$$

Note that the states $|\phi\rangle$ are labeled by functions $\phi : \mathbb{R}^{d-1} \to \mathbb{R}$, so Ψ is indeed a functional (a function of a function). In order to compute the inner product between two wave functionals Ψ_1 and Ψ_2, we need to compute a **functional integral**

$$\langle \Psi_2 | \Psi_1 \rangle := \int \mathcal{D}\phi \Psi_2[\phi]^* \Psi_1[\phi]. \tag{3.7}$$

Functional integrals are rather delicate mathematical objects, as we will discuss in more detail when we get to path integrals. Roughly speaking the idea is to define the measure as

$$\mathcal{D}\phi := \prod_{\vec{x}} d\phi(\vec{x}), \tag{3.8}$$

so in other words we integrate independently over the value of ϕ at each point in space.

To represent the algebra (3.4) on this Hilbert space, imitating nonrelativistic quantum mechanics we can take

$$\Pi(\vec{x}) := \Pi(0, \vec{x}) = -i \frac{\delta}{\delta \phi(\vec{x})}, \tag{3.9}$$

where the quantity appearing on the right-hand side is the **functional derivative** defined by

$$\frac{\delta}{\delta \phi(\vec{x})} \phi(\vec{y}) = \delta^{d-1}(\vec{x} - \vec{y}). \tag{3.10}$$

We can easily check the canonical commutation relation:

$$\phi(\vec{x}) \cdot \left(-i \frac{\delta}{\delta \phi(\vec{y})} \right) - \left(-i \frac{\delta}{\delta \phi(\vec{y})} \right) \cdot \phi(\vec{x}) = i \delta^{d-1}(\vec{x} - \vec{y}). \tag{3.11}$$

Proceeding as in non-relativistic quantum mechanics, the next step is then to construct energy eigenstates by solving the **functional Schrödinger equation**

$$\frac{1}{2} \int d^{d-1}x \left[-\frac{\delta^2}{\delta \phi(\vec{x})^2} + \nabla \phi \cdot \nabla \phi(\vec{x}) + m^2 \phi(\vec{x})^2 \right] \Psi[\phi] = E \Psi[\phi]. \tag{3.12}$$

In principle solving this equation (including interactions and other types of fields) is "all there is" to quantum field theory.[20]

We can make the functional Schrödinger formalism more rigorous by **regularizing** the theory using a spatial lattice, so that the field variable is only defined on a discrete set of spatial points $\vec{x}$ which are part of a lattice L. Taking L to be a cubic lattice, this more explicitly looks like

$$H = \frac{1}{2} \sum_{\vec{x} \in L} a^{d-1} \left[\Pi(\vec{x})^2 + \sum_{\vec{\delta}} \left(\frac{\Phi(\vec{x} + \vec{\delta}) - \Phi(\vec{x})}{a} \right)^2 + m^2 \Phi(\vec{x})^2 \right], \tag{3.13}$$

[20] More carefully this is all there is to field theories which are constructed from Lagrangians. There are some exotic field theories that do not seem constructable in this way, and studying them requires techniques that are mostly beyond the scope of this class.

with

$$[\Phi(\vec{x}), \Pi(\vec{y})] = ia^{-(d-1)}\delta_{\vec{x},\vec{y}} \tag{3.14}$$

and thus

$$\Pi(\vec{x}) = -ia^{-(d-1)}\frac{\partial}{\partial\phi(\vec{x})}. \tag{3.15}$$

Here a is the lattice spacing, and $\vec{\delta}$ ranges over the orthogonal lattice displacements $a\hat{x}_1, a\hat{x}_2, \ldots, a\hat{x}_{d-1}$. The functional Schrödinger equation then becomes

$$\frac{1}{2}\sum_{\vec{x}\in L} a^{d-1}\left[-a^{-2(d-1)}\frac{\partial^2}{\partial\phi(\vec{x})^2} + \sum_{\vec{\delta}}\left(\frac{\phi(\vec{x}+\vec{\delta}) - \phi(\vec{x})}{a}\right)^2 + m^2\phi(\vec{x})^2\right]\Psi[\phi] = E\Psi[\phi], \tag{3.16}$$

which is now just a second-order partial differential equation in many variables. If we also work in finite volume, so that the total number of points is finite, then we can (at least in principle) try solving this equation on a computer. In free theories this is not necessary since the theory can be solved exactly (see below), but if we include interactions (such as say a ϕ^4 term in the Hamiltonian) then this approach can be viable.[21]

3.2 Heisenberg fields and particle states

Wave functionals are conceptually important in quantum field theory because they make it clear that ultimately we are still doing the same quantum mechanics we learned in the non-relativistic case. Unfortunately however they are somewhat unwieldy objects, as we have already seen, and indeed in quantum field theory the wave functional approach is not so useful in practice. It turns out to be a better idea to study the field operators directly, rather than the states, especially in the Heisenberg picture.

Let's first recall that by definition in the Heisenberg picture the time-dependence of an operator is given by

$$O(t) = e^{iHt}O(0)e^{-iHt}. \tag{3.17}$$

Taking the time derivative (and being a bit cavalier about operator ordering in the second step) we see that

$$\dot{O}(t) = i[H, O(t)] = -\{H, O(t)\}, \tag{3.18}$$

which is precisely the classical equation of motion (in Hamiltonian form) for $O(t)$. Thus in the free scalar theory (where there is no issue of operator ordering since there are no terms in the Hamiltonian involving both Φ and Π) the Heisenberg field

$$\Phi(t, \vec{x}) = e^{iHt}\Phi(\vec{x})e^{-iHt} \tag{3.19}$$

should obey its classical equation of motion, namely the Klein-Gordon equation

$$(\partial^2 - m^2)\Phi = 0. \tag{3.20}$$

As discussed in the last section we will impose boundary conditions requiring the fields to vanish at spatial infinity, and any solution of the Klein-Gordon equation which vanishes at spatial infinity can be expanded in terms of a plane-wave basis set of solutions given by

$$f_{\vec{k}}(t, \vec{x}) = \frac{1}{\sqrt{2\omega_{\vec{k}}}}e^{i\vec{k}\cdot\vec{x} - i\omega_{\vec{k}}t} \tag{3.21}$$

and its complex conjugate. Here we have defined

$$\omega_{\vec{k}} = \sqrt{|k|^2 + m^2}, \tag{3.22}$$

[21] In practice there are often much better numerical techniques available however, with "monte carlo" evaluation of the path integral being the long-standing champion for many theories (including this one). Newer approaches which are gaining ground are the "numerical bootstrap" and quantum simulation.

where $|k| = \sqrt{\vec{k}\cdot\vec{k}}$, and we have included the factor of $\frac{1}{\sqrt{2\omega_{\vec{k}}}}$ for future convenience (it ensures that we end up with properly-normalized annihilation/creation operators below). Defining a spacetime momentum vector

$$k^\mu = (\omega_{\vec{k}}, \vec{k}), \tag{3.23}$$

in relativistic notation we have

$$f_{\vec{k}}(x) = \frac{1}{\sqrt{2k^0}} e^{ik\cdot x}. \tag{3.24}$$

Expanding the Heisenberg field in terms of these solutions we have

$$\begin{aligned}\Phi(x) &= \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \left[f_{\vec{k}}(x) a_{\vec{k}} + f^*_{\vec{k}}(x) a^\dagger_{\vec{k}} \right] \\ &= \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left[e^{ik\cdot x} a_{\vec{k}} + e^{-ik\cdot x} a^\dagger_{\vec{k}} \right],\end{aligned} \tag{3.25}$$

where $a_{\vec{k}}$ and $a^\dagger_{\vec{k}}$ are operator coefficients in the mode expansion of the operator $\Phi(x)$. The operator coefficients of $f_{\vec{k}}$ and $f^*_{\vec{k}}$ are hermitian conjugates because ϕ is a real field and so Φ needs to be a hermitian operator. The factor of $\frac{1}{(2\pi)^{d-1}}$ is included as a matter of convenience: it has to appear somewhere due to the way that Fourier transforms work, and this turns out to be the best place to put it. There is a mantra for remembering where it goes which we'll call **Coleman's rule**:

- ⋆ Whenever you integrate over momentum there is a factor of $1/(2\pi)$ for each component, and whenever you have a momentum-conserving δ-function then it comes with a factor of 2π for each component.

So far we haven't actually done much, but let's now see what the canonical commutation relations (3.4) have to say about the algebra of $a_{\vec{k}}$ and $a^\dagger_{\vec{k}}$. The easiest way to do this is to use the Fourier transform to extract $a_{\vec{k}}$ and $a^\dagger_{\vec{k}}$ from the $t = 0$ fields $\Phi(\vec{x})$ and $\Pi(\vec{x})$. In doing such calculations there are two crucial identities:

$$\begin{aligned}\int d^{d-1}x e^{-i\vec{k}\cdot\vec{x}} &= (2\pi)^{d-1}\delta^{d-1}(\vec{k}) \\ \int \frac{d^{d-1}k}{(2\pi)^{d-1}} e^{i\vec{k}\cdot\vec{x}} &= \delta^{d-1}(\vec{x}),\end{aligned} \tag{3.26}$$

where we have placed the factors of 2π in accordance with Coleman's rule. Using the first of these we have

$$\begin{aligned}\int d^{d-1}x e^{-i\vec{p}\cdot\vec{x}} \Phi(\vec{x}) &= \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \int d^{d-1}x e^{-i\vec{p}\cdot\vec{x}} \left[e^{i\vec{k}\cdot\vec{x}} a_{\vec{k}} + e^{-i\vec{k}\cdot\vec{x}} a^\dagger_{\vec{k}} \right] \\ &= \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left[(2\pi)^{d-1}\delta^{d-1}(\vec{k}-\vec{p}) a_{\vec{k}} + (2\pi)^{d-1}\delta^{d-1}(\vec{k}+\vec{p}) a^\dagger_{\vec{k}} \right] \\ &= \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left(a_{\vec{p}} + a^\dagger_{-\vec{p}} \right)\end{aligned} \tag{3.27}$$

and

$$\begin{aligned}\int d^{d-1}x e^{-i\vec{p}\cdot\vec{x}} \Pi(\vec{x}) &= \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \frac{-i\omega_{\vec{k}}}{\sqrt{2\omega_{\vec{k}}}} \int d^{d-1}x e^{-i\vec{p}\cdot\vec{x}} \left[e^{i\vec{k}\cdot\vec{x}} a_{\vec{k}} - e^{-i\vec{k}\cdot\vec{x}} a^\dagger_{\vec{k}} \right] \qquad (3.28) \\ &= -i \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \sqrt{\frac{\omega_{\vec{k}}}{2}} \left[(2\pi)^{d-1}\delta^{d-1}(\vec{k}-\vec{p}) a_{\vec{k}} - (2\pi)^{d-1}\delta^{d-1}(\vec{k}+\vec{p}) a^\dagger_{\vec{k}} \right] \\ &= -i\sqrt{\frac{\omega_{\vec{p}}}{2}} \left(a_{\vec{p}} - a^\dagger_{-\vec{p}} \right), \qquad (3.29)\end{aligned}$$

and thus

$$a_{\vec{p}} = \frac{1}{\sqrt{2}} \int d^{d-1}x e^{-i\vec{p}\cdot\vec{x}} \left(\sqrt{\omega_{\vec{p}}}\, \Phi(\vec{x}) + \frac{i}{\sqrt{\omega_{\vec{p}}}} \Pi(\vec{x}) \right)$$
$$a^\dagger_{\vec{p}} = \frac{1}{\sqrt{2}} \int d^{d-1}x e^{i\vec{p}\cdot\vec{x}} \left(\sqrt{\omega_{\vec{p}}}\, \Phi(\vec{x}) - \frac{i}{\sqrt{\omega_{\vec{p}}}} \Pi(\vec{x}) \right). \tag{3.30}$$

We can then use these expressions together with the canonical commutation relations (3.4) to show that:

$$\begin{aligned}
[a_{\vec{p}}, a_{\vec{p}'}] &= \frac{i}{2} \int d^{d-1}x \int d^{d-1}y e^{-i\vec{p}\cdot\vec{x} - i\vec{p}'\cdot\vec{y}} \Big([\Phi(\vec{x}), \Pi(\vec{y}] - [\Phi(\vec{y}), \Pi(\vec{x})] \Big) = 0 \\
[a^\dagger_{\vec{p}}, a^\dagger_{\vec{p}'}] &= -[a_{\vec{p}}, a_{\vec{p}'}]^\dagger = 0 \\
[a_{\vec{p}}, a^\dagger_{\vec{p}'}] &= -\frac{i}{2} \int d^{d-1}x \int d^{d-1}y e^{-i\vec{p}\cdot\vec{x} + i\vec{p}'\cdot\vec{y}} \Big([\Phi(\vec{x}), \Pi(\vec{y}] + [\Phi(\vec{y}), \Pi(\vec{x})] \Big) \\
&= \int d^{d-1}x e^{i(\vec{p}' - \vec{p})\cdot\vec{x}} \\
&= (2\pi)^{d-1} \delta^{d-1}(\vec{p} - \vec{p}').
\end{aligned} \tag{3.31}$$

These results should look familiar: they are the algebra of creation and annihilation operators for an infinite number of harmonic oscillators, with the oscillators labeled by the spatial momentum $\vec{p}$. They are also the momentum space version of the creation/annihilation operators on multi-particle Fock space that we introduced back in the section 1. Defining a vacuum state $|\Omega\rangle$ by the property that

$$a_{\vec{p}} |\Omega\rangle = 0 \tag{3.32}$$

for all $\vec{p}$ (we will show in a moment that this is indeed the ground state of the Hamiltonian), we have one-particle states of the form

$$a^\dagger_{\vec{p}} |\Omega\rangle, \tag{3.33}$$

two-particle states of the form

$$a^\dagger_{\vec{p}} a^\dagger_{\vec{p}'} |\Omega\rangle, \tag{3.34}$$

and so on.

To justify the words "vacuum" and "particle" here however, we need to study the Hamiltonian. This is given by

$$H = \frac{1}{2} \int d^{d-1}x \left[\Pi(\vec{x})^2 + |\vec{\nabla}\Phi(\vec{x})|^2 + m^2 \Phi(\vec{x})^2 \right], \tag{3.35}$$

into which we should substitute our expression (3.25) for the Heisenberg field. This calculation is a bit tedious, I'll compute the first term here and you'll do the other two in the homework:

$$\begin{aligned}
\frac{1}{2} \int d^{d-1}x \Pi(\vec{x})^2 &= \frac{1}{2} \int d^{d-1}x \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{1}{2\sqrt{\omega_{\vec{k}}\omega_{\vec{p}}}} \left(-i\omega_{\vec{k}} e^{i\vec{k}\cdot\vec{x}} a_{\vec{k}} + i\omega_{\vec{k}} e^{-i\vec{k}\cdot\vec{x}} a^\dagger_{\vec{k}} \right) \left(-i\omega_{\vec{p}} e^{i\vec{p}\cdot\vec{x}} a_{\vec{k}} + i\omega_{\vec{p}} e^{-i\vec{p}\cdot\vec{x}} a^\dagger_{\vec{p}} \right) \\
&= \frac{1}{4} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \omega_{\vec{k}} \left(a^\dagger_{\vec{k}} a_{\vec{k}} + a_{\vec{k}} a^\dagger_{\vec{k}} - a_{\vec{k}} a_{-\vec{k}} - a^\dagger_{\vec{k}} a^\dagger_{-\vec{k}} \right).
\end{aligned} \tag{3.36}$$

Combining all three terms, we find

$$H = \frac{1}{2} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \omega_{\vec{k}} \left(a^\dagger_{\vec{k}} a_{\vec{k}} + a_{\vec{k}} a^\dagger_{\vec{k}} \right). \tag{3.37}$$

This looks quite a bit like the harmonic oscillator Hamiltonian, and we can make it look more so by using the algebra (3.31):

$$\begin{aligned} H &= \frac{1}{2}\int \frac{d^{d-1}k}{(2\pi)^{d-1}}\omega_{\vec{k}}\left(a^\dagger_{\vec{k}}a_{\vec{k}} + a_{\vec{k}}a^\dagger_{\vec{k}}\right) \\ &= \frac{1}{2}\int \frac{d^{d-1}k}{(2\pi)^{d-1}}\omega_{\vec{k}}\left(a^\dagger_{\vec{k}}a_{\vec{k}} + [a_{\vec{k}}, a^\dagger_{\vec{k}}] + a^\dagger_{\vec{k}}a_{\vec{k}}\right) \\ &= \int \frac{d^{d-1}k}{(2\pi)^{d-1}}\omega_{\vec{k}}a^\dagger_{\vec{k}}a_{\vec{k}} + \frac{1}{2}\int d^{d-1}k\omega_{\vec{k}}\delta^{d-1}(0). \end{aligned} \tag{3.38}$$

The first term here is just what we would like: the operator $a^\dagger_{\vec{k}}a_{\vec{k}}$ is the number operator that counts how many particles there are of momentum $\vec{k}$, so this term says that each particle of momentum $\vec{k}$ contributes $\omega_{\vec{k}}$ to the energy. For example if we act on a one-particle state we have

$$\left(\int \frac{d^{d-1}k}{(2\pi)^{d-1}}\omega_{\vec{k}}a^\dagger_{\vec{k}}a_{\vec{k}}\right)a^\dagger_{\vec{p}}|\Omega\rangle = \int \frac{d^{d-1}k}{(2\pi)^{d-1}}\omega_{\vec{k}}a^\dagger_{\vec{k}}(2\pi)^{d-1}\delta^{d-1}(\vec{k}-\vec{p})|\Omega\rangle = \omega_{\vec{p}}\,a^\dagger_{\vec{p}}|\Omega\rangle, \tag{3.39}$$

so one-particle states $a^\dagger_{\vec{p}}|\Omega\rangle$ are eigenstates of this term with eigenvalue $\omega_{\vec{p}}$. Ignoring the second term, we thus have succeeding in finding the eigenstates of the Hamiltonian!

What however are we to say about the second term in (3.38)? On the one hand it does not involve the creation/annihilation operators and thus is proportional to the identity, which means that the eigenstates we just found are also eigenstates of the full Hamiltonian. On the other hand it is embarassingly infinite, for two different reasons. The first reason is the δ-function evaluated at zero, which is an "infrared (IR) divergence" arising because the momentum $\vec{k}$ is a continuous parameter. If we were to work in finite volume V, then the momentum would be discrete and we would find $\delta^{d-1}(0) \sim V$. The second reason is the integral over $\vec{k}$, which diverges at large $\vec{k}$ since in continuum field theory we can have particles of arbitrarily high momentum. This is called an "ultraviolet (UV) divergence", and it is regulated if we introduce a lattice with lattice spacing a since then it does not make sense to consider momenta larger than of order the "UV cutoff"

$$\Lambda := \frac{1}{a}. \tag{3.40}$$

With both cutoffs in place we therefore have

$$\frac{1}{2}\int d^{d-1}k\omega_{\vec{k}}\delta^{d-1}(0) \sim V\Lambda^d, \tag{3.41}$$

which you can check indeed has units of energy. What are we to make of this term? The essential point is that since it is proportional to V, we can write it as a local integral of a constant over space:

$$\frac{1}{2}\int d^{d-1}k\omega_{\vec{k}}\delta^{d-1}(0) \sim \Lambda^d\int d^{d-1}x. \tag{3.42}$$

We would thus precisely get a term of this form if from the beginning we had taken the Lagrangian to include a "cosmological constant" term

$$\Delta\mathcal{L} = -\rho_0, \tag{3.43}$$

and so the term (3.41) is usually called a **renormalization of the cosmological constant**. Somehow the dynamics of our free scalar field have generated a gigantic energy density filling the universe! This is a quite remarkable prediction, but unfortunately it is also quite inconsistent with our understanding of the world. In the absence of gravity such an energy density would have no measurable effect, but gravity responds to the total energy density and such a gigantic positive energy density would lead to a universe that tore itself apart via exponential expansion on a timescale of order $\frac{1}{\Lambda}$. We don't quite know what the scale of Λ should

be, but from the Large Hadron Collider it should at least be bigger than ~ 10TeV and this already tells us that $\frac{1}{\Lambda} \lesssim \frac{\hbar}{10TeV} \sim 6 \times 10^{-29}$ s. Not good. No es bueno. 很不好.

What should we do? There is only one way out: we need to introduce an additional "bare" cosmological constant term in the original Lagrangian,

$$\mathcal{L}_{ct} \sim \Lambda^d, \tag{3.44}$$

called a **counterterm**, whose coefficient is precisely tuned to *cancel* the cosmological constant generated by our free scalar field. The full Hamiltonian is then just be given by

$$H_{ren} = \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \omega_{\vec{k}} a^\dagger_{\vec{k}} a_{\vec{k}}, \tag{3.45}$$

so the vacuum has zero energy as hoped. This is our first example of a procedure called **renormalization**, by which we carefully tune the coefficients in the Lagrangian in a Λ-dependent way to cancel UV divergences. This may seem like a rather ugly fix. Why should the Lagrangian be fine-tuned in this way? How do we know that there won't be other UV divergences that can't be canceled in this way? These are excellent questions, and we will discuss them in considerable detail in later sections.

3.3 Non-locality of the annihilation operator in position space

In the first section we tried (and failed) to build a quantum theory of relativistic particles using annihilation and creation operators labeled by position. We can now straightforwardly see why this did not work: taking the Fourier transform of (3.30), we see that

$$a_{\vec{x}} = \int \frac{d^{d-1}p}{(2\pi)^{d-1}} e^{i\vec{p}\cdot\vec{x}} a_{\vec{p}} = \frac{1}{\sqrt{2}} \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \int d^{d-1}y e^{i\vec{p}\cdot(\vec{x}-\vec{y})} \left(\sqrt{\omega_{\vec{p}}}\, \Phi(\vec{y}) + \frac{i}{\sqrt{\omega_{\vec{p}}}} \Pi(\vec{y}) \right). \tag{3.46}$$

This is not a local function of Φ and Π since the inverse Fourier transforms of $\sqrt{\omega_{\vec{p}}}$ and $\frac{1}{\sqrt{\omega_{\vec{p}}}}$ do not vanish away from zero, and this non-locality is the origin of the apparent acausality we found in the first section. In the non-relativistic limit however these functions become constants, in which case their inverse Fourier transforms are δ-functions so $a_{\vec{x}}$ indeed becomes local:

$$a_{\vec{x}} \to \sqrt{\frac{m}{2}} \left(\Phi(\vec{x}) + \frac{i}{m} \Pi(\vec{x}) \right). \tag{3.47}$$

For this reason non-relativistic systems are often formulated using $a_{\vec{x}}$ and $a^\dagger_{\vec{x}}$ instead of $\Phi(\vec{x})$ and $\Pi(\vec{x})$.

3.4 Lorentz transformations and microcausality revisited

We'll now make an aside to see more explicitly how the scalar field theory we have constructed avoids the problems we saw in the first section with a particle-based relativistic quantum mechanics. There we motivated fields by looking for linear combinations of creation and annihilation operators that

(1) have simple Lorentz transformation properties

(2) commute at spacelike seperation.

In the free field theory we have been studying these conditions follow automatically from the canonical commutation relations together with the invariance of the action under Lorentz transformations which act on Φ as a scalar, but it is instructive to see how they arise from the point of view of the creation and annihilation operators.

Let's first consider Lorentz transformations. In the next section we will show that any Lorentz transformation Λ that does not reverse the direction of time must be implemented on the Hilbert space by a unitary operator $U(\Lambda)$, which we will take to leave the ground state invariant:

$$U(\Lambda)|\Omega\rangle = |\Omega\rangle. \tag{3.48}$$

To understand how $U(\Lambda)$ acts on the rest of the Hilbert space, we need to understand its action on the creation and annihilation operators. Before deciding this it is convenient to first understand the Lorentz transformation properties of the measure $\frac{d^{d-1}p}{(2\pi)^{d-1}}$. The easiest way to do this is to note that the full measure $\frac{d^dp}{(2\pi)^d}$ which integrates over p^0 as well as $\vec{p}$ is Lorentz-invariant, since Lorentz transformations preserve the Minkowski metric $\eta_{\mu\nu}$. We however only want to integrate over Lorentz vectors p^μ which obey the on-shell condition $p^0 = \omega_{\vec{p}}$. We can implement this using a Lorentz-invariant δ-function, leading to a manifestly Lorentz-invariant measure

$$\frac{d^dp}{(2\pi)^d}\, 2\pi\delta(-p^2+m^2)\,\Theta(p^0). \tag{3.49}$$

The Heaviside Θ function here is one for $p^0 > 0$ and zero for $p^0 < 0$, and is there to make sure that the δ function only picks out $p^0 = \omega_{\vec{p}}$ (as opposed to $p^0 = -\omega_{\vec{p}}$). $\Theta(p^0)$ is Lorentz invariant because we are only considering Lorentz transformations that do not reverse time. We can then relate this measure to $\frac{d^{d-1}p}{(2\pi)^{d-1}}$ via

$$\begin{aligned}\frac{d^dp}{(2\pi)^d}\, 2\pi\delta(-p^2+m^2)\,\Theta(p^0) &= \frac{d^{d-1}p}{(2\pi)^{d-1}}\frac{dp^0}{2\pi}\,\frac{2\pi}{2p^0}\delta(p^0-\omega_{\vec{p}}) \\ &= \frac{d^{d-1}p}{(2\pi)^{d-1}}\frac{1}{2\omega_{\vec{p}}},\end{aligned} \tag{3.50}$$

so if we define

$$\Lambda^\mu{}_\nu p^\nu = (p^0_\Lambda, \vec{p}_\Lambda) \tag{3.51}$$

then

$$\frac{d^{d-1}p_\Lambda}{(2\pi)^{d-1}}\frac{1}{2\omega_{\vec{p}_\Lambda}} = \frac{d^{d-1}p}{(2\pi)^{d-1}}\frac{1}{2\omega_{\vec{p}}}. \tag{3.52}$$

This also shows that we have

$$\omega_{\vec{p}_\Lambda}(2\pi)^{d-1}\delta^{d-1}(\vec{p}\,'_\Lambda - \vec{p}_\Lambda) = \omega_{\vec{p}}(2\pi)^{d-1}\delta^{d-1}(\vec{p}\,' - \vec{p}). \tag{3.53}$$

Proceeding to consider the action of $U(\Lambda)$, we can guess that

$$U(\Lambda)a^\dagger_{\vec{p}}|\Omega\rangle = N_{\vec{p},\Lambda}a^\dagger_{\vec{p}_\Lambda}|\Omega\rangle \tag{3.54}$$

for some constant $N_{\vec{p},\Lambda}$ that we can determine by requiring $U(\Lambda)$ to be unitary. Indeed we want that

$$\begin{aligned}|N_{\vec{p},\Lambda}|^2(2\pi)^{d-1}\delta^{d-1}(\vec{p}\,'_\Lambda - \vec{p}_\Lambda) &= \langle\Omega|a_{\vec{p}\,'}U(\Lambda)^\dagger U(\Lambda)a^\dagger_{\vec{p}}|\Omega\rangle \\ &= \langle\Omega|a_{\vec{p}\,'}a^\dagger_{\vec{p}}|\Omega\rangle \\ &= (2\pi)^{d-1}\delta^{d-1}(\vec{p}\,' - \vec{p}),\end{aligned} \tag{3.55}$$

so from (3.53) we see that

$$N_{\vec{p},\Lambda} = \sqrt{\frac{\omega_{\vec{p}_\Lambda}}{\omega_{\vec{p}}}} \tag{3.56}$$

is consistent with unitarity. We therefore have the Lorentz transformations

$$\begin{aligned}U(\Lambda)a_{\vec{p}}U(\Lambda)^\dagger &= \sqrt{\frac{\omega_{\vec{p}_\Lambda}}{\omega_{\vec{p}}}}a_{\vec{p}_\Lambda} \\ U(\Lambda)a^\dagger_{\vec{p}}U(\Lambda)^\dagger &= \sqrt{\frac{\omega_{\vec{p}_\Lambda}}{\omega_{\vec{p}}}}a^\dagger_{\vec{p}_\Lambda}.\end{aligned} \tag{3.57}$$

We can use this to work out the Lorentz transformations of the field:

$$\begin{aligned}
U(\Lambda)\Phi(x)U(\Lambda)^\dagger &= \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left[e^{ip\cdot x} U(\Lambda) a_{\vec{p}} U(\Lambda)^\dagger + e^{-ip\cdot x} U(\Lambda) a_{\vec{p}}^\dagger U(\Lambda)^\dagger \right] \\
&= \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{\sqrt{\omega_{\vec{p}_\Lambda}}}{\sqrt{2}\omega_{\vec{p}}} \left[e^{ip\cdot x} a_{\vec{p}_\Lambda} + e^{-ip\cdot x} a_{\vec{p}_\Lambda}^\dagger \right] \\
&= \int \frac{d^{d-1}p_\Lambda}{(2\pi)^{d-1}} \frac{\sqrt{\omega_{\vec{p}_\Lambda}}}{\sqrt{2}\omega_{\vec{p}_\Lambda}} \left[e^{ip\cdot x} a_{\vec{p}_\Lambda} + e^{-ip\cdot x} a_{\vec{p}_\Lambda}^\dagger \right] \\
&= \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left[e^{i(\Lambda^{-1}p)\cdot x} a_{\vec{p}} + e^{-i(\Lambda^{-1}p)\cdot x} a_{\vec{p}}^\dagger \right] \\
&= \Phi(\Lambda x).
\end{aligned} \tag{3.58}$$

Going from the first to the second line we used (3.57), going from the second to the third we used (3.52), going from the third to the fourth we relabeled the integration variable $\vec{p}_\Lambda \to \vec{p}$, and in going from the fourth to the fifth we used that

$$(\Lambda^{-1}p)\cdot x = \Lambda_\alpha{}^\beta p^\alpha x_\beta = p_\alpha \Lambda^\alpha{}_\beta x^\beta = p\cdot(\Lambda x). \tag{3.59}$$

Thus we see that indeed we have succeeded in constructing a Lorentz scalar out of creation/annihilation operators which themselves have more complicated transformations, at least for Lorentz transformations that do not reverse time. We will discuss time-reversal symmetry in the next section, where we will see that it needs to be represented on Hilbert space by an antiunitary operator instead of a unitary operator.

Turning now to microcausality, let's compute the commutator of Φ at spatial separation:

$$\begin{aligned}
[\Phi(\vec{x}), \Phi(\vec{y})] &= \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \frac{1}{2\sqrt{\omega_{\vec{p}}\omega_{\vec{k}}}} \left(e^{i\vec{p}\cdot\vec{x} - i\vec{k}\cdot\vec{y}} [a_{\vec{p}}, a_{\vec{k}}^\dagger] - e^{-i\vec{p}\cdot\vec{x} + i\vec{k}\cdot\vec{y}} [a_{\vec{k}}, a_{\vec{p}}^\dagger] \right) \\
&= \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{1}{2\omega_{\vec{p}}} \left(e^{i\vec{p}\cdot(\vec{x}-\vec{y})} - e^{-i\vec{p}\cdot(\vec{x}-\vec{y})} \right) \\
&= 0,
\end{aligned} \tag{3.60}$$

where in going from the first to the second line we used (3.31) and in going from the second to the third we flipped the sign of the integration variable in the second term. The point to notice however is that the vanishing of this commutator required a nontrivial cancellation between two terms. For example if we had tried to make the field Φ using only annihilation operators, then its commutator with its hermitian conjugate would not vanish at spatial separation:

$$\left[\int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{1}{\sqrt{2\omega_{\vec{p}}}} e^{i\vec{p}\cdot\vec{x}} a_{\vec{p}}, \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \frac{1}{\sqrt{2\omega_{\vec{k}}}} e^{-i\vec{k}\cdot\vec{y}} a_{\vec{k}}^\dagger \right] = \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{1}{2\omega_{\vec{p}}} e^{i\vec{p}\cdot(\vec{x}-\vec{y})} \neq 0 \tag{3.61}$$

It is the requirement of microcausality that requires us to use fields that involve both creation and annihilation operators, leading to the distinctive predictions of particle number non-conservation and the existence of antiparticles as discussed in the first section.

3.5 Quantization of a complex scalar field, antiparticles

There is a simple but instructive generalization of the free scalar field we have been discussing so far, where ϕ is taken to be complex and the Lagrangian density to be

$$\mathcal{L} = -\partial^\mu \phi^* \partial_\mu \phi - m^2 \phi^* \phi. \tag{3.62}$$

You will show on the homework that the equation of motion for this theory is again just

$$\partial^2 \phi = m^2 \phi, \tag{3.63}$$

but when we expand the field in terms of solutions there is no longer a reason for the creation and annihilation operators to be related. We thus should write

$$\Phi(x) = \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left[e^{ip\cdot x} a_{\vec{p}} + e^{-ip\cdot x} b^\dagger_{\vec{p}} \right], \tag{3.64}$$

where $a_{\vec{p}}$ and $b_{\vec{p}}$ are not related. The canonical commutation relations follow from the observation that

$$\Pi(x) = \dot{\Phi}^\dagger(x), \tag{3.65}$$

so we have

$$[\Phi(\vec{x}), \dot{\Phi}^\dagger(\vec{y})] = [\Phi^\dagger(\vec{x}), \dot{\Phi}(\vec{y})] = i\delta^{d-1}(\vec{x} - \vec{y}) \tag{3.66}$$

with all other commutators vanishing. In the homework you will show that these imply that $a_{\vec{p}}, a^\dagger_{\vec{p}}$ and $b_{\vec{p}}, b^\dagger_{\vec{p}}$ give two independent sets of annihilation/creation operators. This theory thus has two species of particles, both with mass m. You will also show that these particles have opposite charge under the symmetry $\phi'(x) = e^{i\theta}\phi(x)$, and indeed one is the antiparticle of the other.

3.6 Correlation functions I: Definition and physical meaning

We have now solved the theory of a free scalar field. What we haven't done however is compute anything interesting with it. We've acknowledged that the functional Schrödinger formalism is not so useful in practice, so what kinds of questions *are* interesting in quantum field theory? Long experience has shown that the physics of quantum fields is most elegantly packaged into *vacuum expectation values of products of Heisenberg fields*, otherwise known as **correlation functions**.

The simplest correlation function for any field $O(x)$ is its **one-point function** in the ground state:

$$\langle O(x) \rangle := \langle \Omega | O(x) | \Omega \rangle. \tag{3.67}$$

If $O(x)$ is hermitian then the physical interpretation of this quantity is clear: it is the expectation value for what we get if we measure $O(x)$ in the ground state. In quantum field theory it is often (but not always) the case that the one-point functions of the fields vanish. Usually more interesting is the **two-point function**: for any two fields $O_1(x_1)$ and $O_2(x_2)$ we have

$$\langle O_2(x_2) O_1(x_1) \rangle := \langle \Omega | O_2(x_2) O_1(x_1) | \Omega \rangle. \tag{3.68}$$

The two point function is important for many physical questions. Perhaps the most direct physical interpretation is that when x_1 and x_2 are spacelike separated and O_1 and O_2 have vanishing one-point functions, the two-point function is a measure of how correlated the fluctuations are in measurements of the independent observables O_1 and O_2 (we need to assume spacelike separation to ensure the operators are independent, i.e. commuting). More generally if their one-point functions don't vanish we can still quantify the amount of correlation using the **connected two-point function**

$$\langle O_2(x_2) O_1(x_1) \rangle_c := \langle O_2(x_2) O_1(x_1) \rangle - \langle O_2(x_2) \rangle \langle O_1(x_1) \rangle. \tag{3.69}$$

The two-point function also has a physical interpretation when x_1 and x_2 are not spacelike separated: it tells us about the **linear response** of the theory to an external source. Indeed let's say we have a field theory with Hamiltonian H, and then we turn on a position-dependent source $J(x)$ for a field $O_1(x)$ such that the Schrödinger picture Hamiltonian becomes:

$$H(t) = H_0 + V(t) \tag{3.70}$$

with

$$V(t) := \lambda \int d^{d-1}x J(t, \vec{x}) O_1(\vec{x}). \tag{3.71}$$

Here λ is a parameter controlling the strength of the source that we will take to be small. The question we will ask is the following: assuming that J goes to zero at early times, if we start in the ground state of H_0 at early times, what is the expectation value of a field O_2 as a function of space and time? We can answer this question using time-dependent perturbation theory. Indeed including this interaction we have

$$\begin{aligned}\langle\Omega|O_2(t_2,\vec{x}_2)|\Omega\rangle &= \langle\Omega|(Te^{-i\int_{-\infty}^{t_2}dt'H(t')})^\dagger O_2(\vec{x}_2)Te^{-i\int_{-\infty}^{t_2}dt'H(t')}|\Omega\rangle \\ &= \langle\Omega|U_I(t_2)^\dagger e^{iH_0t_2}O_2(\vec{x}_2)e^{-iH_0t_2}U_I(t_2)|\Omega\rangle, \end{aligned} \tag{3.72}$$

where

$$U_I(t) = e^{iH_0t}Te^{-i\int_{-\infty}^{t}dt'H(t')} = Te^{-i\int_{-\infty}^{t}dt'e^{iH_0t}V(t)e^{-iH_0t}} \tag{3.73}$$

is the **interaction picture time-evolution operator** (you can check that these two expressions are equivalent by showing they have the same time derivative and obey the same initial condition at $t=-\infty$). The letter T here is the **time-ordering symbol**, it means that earlier operators go to the right. Expanding in λ we have

$$U_I(t) = 1 - i\lambda\int_{-\infty}^{t}dt'\int d^{d-1}x' J(t,\vec{x}')e^{iH_0t}O_1(\vec{x}')e^{-iH_0t} + O(\lambda^2), \tag{3.74}$$

and thus to linear order in λ we have (assuming that O_2 has vanishing one-point function in the unperturbed theory)

$$\langle\Omega|O_2(t_2,\vec{x}_2)|\Omega\rangle = -i\lambda\int_{-\infty}^{t_2}dt'\int d^{d-1}x' J(t',\vec{x}')\langle[O_2(t_2,\vec{x}_2),O_1(t',\vec{x}')]\rangle_0. \tag{3.75}$$

Here $\langle\rangle_0$ indicates the vacuum expectation value of Heisenberg operators in the unperturbed theory. In particular if we take J to be a delta function localized at $(t_1,\vec{x}_1)$, then we have

$$\langle\Omega|O_2(t_2,\vec{x}_2)|\Omega\rangle = -i\lambda\Theta(t_2-t_1)\langle[O_2(t_2,\vec{x}_2),O_1(t_1,\vec{x}_1)]\rangle_0. \tag{3.76}$$

Thus we see that the response of a quantum field theory to a local perturbation is determined by a difference of two-point functions at arbitrary separation. The Θ function arises because if $t_1 > t_2$ then the source is outside of the region of t' intregration so the δ-function never contributes. This response vanishes unless x_2 is in the future lightcone of x_1, as it had better, which by the way is another illustration of the fact that by introducing fields we have solved the causality problems of relativistic particle quantum mechanics.

A simple example of an application of this calculation is the following: we can create a source for the scalar field theory describing liquid helium-4 by firing a high-energy neutron at a bubble of liquid helium, and then (3.76) describes how the local density of helium atoms in the bubble responds. In the homework you will play with this and see how the response depends on whether the sample has two or three spatial dimensions.

Higher-point correlation functions are also interesting. At spacelike separation they quantify *conditional* fluctuations such as knowing how likely we are to see correlation between two operators given that we measured a third to have some value, while at timelike separation they give more information about how the theory responds to perturbations. We will also see later that for quantum field theories with particles at low energies, higher-point correlation functions can be used to extract the scattering matrix.

3.7 Correlation functions II: Calculation

Having introduced the idea of correlation functions, let's compute some in our free scalar field theory.

The one-point function of the scalar field Φ is easy:

$$\langle\Phi(x)\rangle = 0 \tag{3.77}$$

since we can view the $a_{\vec{p}}$ in Φ as annihilating $|\Omega\rangle$ and the $a^\dagger_{\vec{p}}$ as annihilating $\langle\Omega|$.

The two-point function

$$G(x_2,x_1) := \langle\Phi(x_2)\Phi(x_1)\rangle \tag{3.78}$$

is more interesting. Note that the two-point function we have defined has $\Phi(x_2)$ to the left of $\Phi(x_1)$ regardless of the time-ordering of x_1 and x_2: it is to be distinguished from the **Feynman propagator**, which is defined to include a time-ordering symbol

$$G_F(x_2, x_1) := \langle T\Phi(x_2)\Phi(x_1)\rangle. \tag{3.79}$$

In quantum field theory correlation functions without time ordering such as (3.78) are sometimes called **Wightman functions** to distinguish them from correlation functions that are time-ordered. We can easily write the Feynman propagator in terms of the Wightman two-point function:

$$G_F(x_2, x_1) = \Theta(t_2 - t_1)G(x_2, x_1) + \Theta(t_1 - t_2)G(x_1, x_2). \tag{3.80}$$

It is harder to go the other way (you need to do some nontrivial analytic continuation), so in quantum field theory it is usually a good idea to view the Wightman functions as the fundamental objects of the theory. In particular we emphasize that the linear response (3.76) involves two-point functions with both time orderings and thus requires the Wightman two-point function. The Feynman propagator is important in perturbative calculations, as we will see in later sections.

In the free scalar field theory we can compute the (Wightman) two-point function:

$$\begin{aligned} G(x_2, x_1) &= \int \frac{d^{d-1}p_1}{(2\pi)^{d-1}} \frac{d^{d-1}p_2}{(2\pi)^{d-1}} \frac{1}{2\sqrt{\omega_{\vec{p}_1}\omega_{\vec{p}_2}}} e^{ip_2\cdot x_2 - ip_1\cdot x_1} \langle\Omega|a_{\vec{p}_2} a^\dagger_{\vec{p}_1}|\Omega\rangle \\ &= \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{1}{2\omega_{\vec{p}}} e^{ip\cdot(x_2 - x_1)}, \end{aligned} \tag{3.81}$$

where in the first line we observed that the only non-vanishing term involves an annihilation operator to the left of a creation operator and in the second line we used the algebra (3.31). We won't spend valuable class time doing this integral since we will later have a better way to compute the same quantity using the path integral,[22] but the result is

$$G(x_2, x_1) = \frac{1}{(2\pi)^{d/2}} \left(\frac{m}{\sqrt{(x_2 - x_1)^2 + is_{21}\epsilon}} \right)^{\frac{d-2}{2}} K_{\frac{d-2}{2}}(m\sqrt{(x_2 - x_1)^2 + is_{21}\epsilon}) \tag{3.82}$$

where s_{21} is equal to one if $t_2 - t_1$ is positive and minus one if it is negative and ϵ is a small positive quantity whose purpose is to define the branch of the square root when $(x_2 - x_1)^2 < 0$ but should otherwise be taken to zero. This is an example of what is called an "$i\epsilon$ prescription", which we will see again and again. $K_\alpha(x)$ is a modified Bessel function of the second kind: the only things worth knowing about it at the moment are its asymptotics:[23]

$$K_\alpha(x) \approx \begin{cases} \frac{\Gamma(\alpha)2^{\alpha-1}}{x^\alpha} & 0 < |x| \ll 1 \\ \sqrt{\frac{\pi}{2x}} e^{-x} & x \gg 1 \end{cases}. \tag{3.84}$$

In particular at general separations which are small compared to the inverse mass we have

$$G(x_2, x_1) \approx \frac{\Gamma(d/2 - 1)}{2\pi^{d/2}} \frac{1}{\left((x_2 - x_1)^2 + is_{21}\epsilon\right)^{\frac{d-2}{2}}}, \tag{3.85}$$

[22] If you want to try it, I recommend first considering the $d = 2$ case. You can deform the p contour to wrap around one of the cuts on the imaginary p axis as in figure 2, which leads to one of the standard integral representations of $K_0(m|x_2 - x_1|)$. In the general case you need to first do an angular integral, after which you can do the same manipulation.

[23] Another thing that is perhaps worth knowing is that it simplifies when α is a half-integer, which here means that d is odd. For example for $d = 3$ we simply have $K_{1/2}(x) = \sqrt{\frac{\pi}{2x}} e^{-x}$ and thus

$$G(x_2, x_1) = \frac{e^{-m\sqrt{(x_2 - x_1)^2 + is_{21}\epsilon}}}{4\pi\sqrt{(x_2 - x_1)^2 + is_{21}\epsilon}}. \tag{3.83}$$

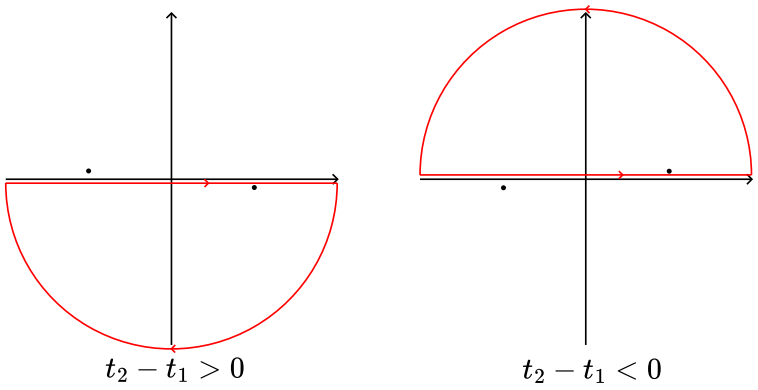


Figure 6: Closing the contour in the complex p^0 plane for the integral (3.90).

while at spacelike separations which are large compared to the inverse mass we have

$$G(x_2, x_1) \approx \frac{m^{d-2}}{2^{\frac{d+1}{2}} \pi^{\frac{d-1}{2}} (m|x_2 - x_1|)^{\frac{d-1}{2}}} e^{-m|x_2 - x_1|}. \tag{3.86}$$

There is quite a bit of physics in these expressions, here are some key points:

(1) The two-point function is nonzero at spacelike separation, so independent fields are correlated with each other in the ground state. Correlation between independent (i.e. commuting) degrees of freedom in a pure quantum state is called **entanglement**, so what we are seeing is that in quantum field theory the vacuum is a highly-entangled state. Indeed since the two-point function diverges in the limit $x_2 \to x_1$, the amount of entanglement is infinite!

(2) In the massless limit (3.85) becomes exact so the correlation function (for $d > 2$) decays as an inverse power of the distance between the points. You will study the $d = 2$ case in the homework.

(3) In the massive case the correlation decays exponentially with distance at spacelike separations which are large compared to m^{-1}.

This discussion illustrates something of a general maxim about correlation functions in quantum field theory: the physics is more clear in position space, but the formulas are simpler in momentum space. More pithily, *in quantum field theory you should think in position space but compute in momentum space.*

The short-distance divergence of the two-point function also has an important mathematical consequence: it shows that the field $\Phi(x)$ is not actually a good quantum operator, since acting on the vacuum (or indeed any other state of finite energy) we get a state of infinite norm. In order to get something which is a good operator, we need to **smear** $\Phi(x)$ against a smooth function of compact support:

$$\Phi_f = \int d^d x f(x) \Phi(x). \tag{3.87}$$

This statement is sometimes formalized by saying that in quantum field theory the fields themselves are **operator-valued distributions**. We will show in the next section that this smearing indeed produces a well-defined operator.

You may have found it annoying that our expression (3.81) for the two-point function involves integrals over only the spatial components of momentum; wouldn't it be nice to have a more manifestly covariant expression? Of course we did already show that the measure $\frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{1}{2\omega_{\vec{p}}}$ is Lorentz-invariant, but there our demonstration involved the non-analytic objects $\Theta(p^0)$ and $\delta(p^2 + m^2)$. It turns out to be a very good idea

to come up with an expression for the two-point function that is manifestly both covariant and analytic in momentum. We can do this by showing that

$$\frac{1}{2\omega_{\vec{p}}}e^{-i\omega_{\vec{p}}(t_2-t_1)} = \lim_{\epsilon\to 0}\int_{-\infty}^{\infty}\frac{dp^0}{2\pi}\frac{-is_{21}}{p^2+m^2-i\epsilon s_{21}}e^{-ip^0(t_2-t_1)}, \tag{3.88}$$

where s_{21} again is one if $t_2 - t_1$ is positive and minus one if it is negative, since from (3.81) we then have

$$G(x_2,x_1) = \lim_{\epsilon\to 0^+}\int\frac{d^dp}{(2\pi)^d}\frac{-is_{21}}{p^2+m^2-i\epsilon s_{21}}e^{ip\cdot(x_2-x_1)}. \tag{3.89}$$

The appearance of the vanishingly small quantity $\epsilon > 0$ here is another example of an $i\epsilon$ prescription. To demonstrate (3.88), it is convenient to rewrite the integral on the right-hand side as

$$-s_{21}\int_{-\infty}^{\infty}\frac{dp^0}{2\pi i}\frac{e^{-ip^0(t_2-t_1)}}{(p^0-(\omega_{\vec{p}}-i\epsilon s_{21}))(p^0+(\omega_{\vec{p}}-i\epsilon s_{21}))}, \tag{3.90}$$

where we have used that

$$(p^0-(\omega_{\vec{p}}-i\epsilon s_{21}))(p^0+(\omega_{\vec{p}}-i\epsilon s_{21})) = -(p^2+m^2-2i\omega_{\vec{p}}\epsilon s_{21})+O(\epsilon^2) \tag{3.91}$$

and then redefined $2\omega_{\vec{p}}\epsilon \to \epsilon$ since the only thing we care about ϵ is that it is small and positive. The integral (3.90) can be computed using the residue theorem. Indeed recall that if $f(z)$ is an analytic function in a region R containing a point z_0, then we have[24]

$$\frac{1}{2\pi i}\int_{\partial R}\frac{f(z)}{z-z_0} = f(z_0) \tag{3.92}$$

where the integral is taken in the counter-clockwise direction about z_0. Said differently, the function $\frac{f(z)}{z-z_0}$ has a simple pole at $z = z_0$ and the integral around this pole extracts the **residue** $f(z_0)$. The integrand (3.90) has two simple poles, at

$$p^0 = \pm\left(\omega_{\vec{p}} - i\epsilon s_{21}\right). \tag{3.93}$$

We can evaluate the integral using the residue theorem by closing the integration contour along the real axis at infinity in the lower or upper half plane depending on whether s_{21} is positive or negative respectively (see figure 6). Either way the integral picks up the residue of the pole at $p^0 = \omega_{\vec{p}} - i\epsilon s_{21}$, but there is a sign difference since in the former case the integral is clockwise while in the latter case it is counter clockwise. We therefore have

$$\int\frac{dp^0}{2\pi i}\frac{e^{-ip^0(t_2-t_1)}}{(p^0-(\omega_{\vec{p}}-i\epsilon s_{21}))(p^0+(\omega_{\vec{p}}-i\epsilon s_{21}))} = -s_{21}\frac{1}{2\omega_{\vec{p}}}e^{-i\omega_{\vec{p}}(t_2-t_1)}, \tag{3.94}$$

so multiplying by $-s_{21}$ we recover (3.88).

Finally it will be convenient later to have a formula similar to (3.89) for the Feynman propagator. Proceeding as in the derivation of (3.81), we have

$$\begin{aligned} G_F(x_2,x_1) &= \int\frac{d^{d-1}p}{(2\pi)^{d-1}}\frac{1}{2\omega_{\vec{p}}}e^{is_{21}p\cdot(x_2-x_1)} \\ &= \int\frac{d^{d-1}p}{(2\pi)^{d-1}}\frac{1}{2\omega_{\vec{p}}}e^{i\vec{p}\cdot(\vec{x}_2-\vec{x}_1)-is_{21}\omega_{\vec{p}}(t_2-t_1)} \end{aligned} \tag{3.95}$$

[24]The intuition for this is essentially the divergence theorem in two dimensions, although to make it rigorous the logic goes the other way since the divergence theorem requires continuous partial derivatives and showing that an analytic function has continuous partial derivatives is usually done using the residue theorem.

where in the first line there is an s_{21} in the exponent because depending on the time-ordering which field contributes an $a_{\vec{p}}$ and which contributes an $a^\dagger_{\vec{p}}$ switches and in going from the first line to the second line we flipped the direction of the integral over $\vec{p}$. The quantity $s_{21}(t_2-t_1)$ is always positive, so in this integral we can use the identity (3.88) replacing $(t_2-t_1)\to s_{21}(t_2-t_1)$ and setting s_{21} to one on the right hand side. Flipping the direction of the p^0 integral we get

$$G_F(x_2,x_1)=\lim_{\epsilon\to 0}\int\frac{d^dp}{(2\pi)^d}\frac{-i}{p^2+m^2-i\epsilon}e^{ip\cdot(x_2-x_1)}, \tag{3.96}$$

which is a bit simpler than the expression (3.89) for the two-point function. In particular the Feynman propagator has the nice property that it is a Green's function for the Klein-Gordon operator:

$$\begin{aligned}(\partial_2^2-m^2)G_F(x_2,x_1)&=\lim_{\epsilon\to 0}\int\frac{d^dp}{(2\pi)^d}\frac{i(p^2+m^2)}{p^2+m^2-i\epsilon}e^{ip\cdot(x_2-x_1)}\\&=i\int\frac{d^dp}{(2\pi)^d}e^{ip\cdot(x_2-x_1)}\\&=i\delta^d(x_2-x_1).\end{aligned} \tag{3.97}$$

This would not have worked for the Wightman function since the derivative acting on s_{21} would have generated additional terms.

You may be wondering why we stopped with two-point functions: what about three-point functions, four-point functions, and so on? In free field theory the answer is simple: these end up either vanishing or just being combinations of two-point functions. Indeed the n-point function

$$\langle\Phi(x_1)\Phi(x_2)\dots\Phi(x_n)\rangle \tag{3.98}$$

vanishes when n is odd since there are no terms with an equal number of creation and annihilation operators. When n is even we simply pair them up to get a sum of products of two-point functions. For example to compute the four-point function we introduce annihilation and creation parts of $\Phi(x)$ as

$$\begin{aligned}\Phi_-(x)&=\int\frac{d^{d-1}p}{(2\pi)^{d-1}}\frac{1}{2\omega_{\vec{p}}}e^{ip\cdot x}a_{\vec{p}}\\\Phi_+(x)&=\int\frac{d^{d-1}p}{(2\pi)^{d-1}}\frac{1}{2\omega_{\vec{p}}}e^{-ip\cdot x}a^\dagger_{\vec{p}},\end{aligned} \tag{3.99}$$

observe that

$$[\Phi_-(x),\Phi_+(y)]=\int\frac{d^{d-1}p}{(2\pi)^{d-1}}\frac{1}{2\omega_{\vec{p}}}e^{ip(x-y)}=G(x,y), \tag{3.100}$$

and then compute

$$\begin{aligned}\Big\langle\Phi(x_4)\Phi(x_3)\Phi(x_2)\Phi(x_1)\Big\rangle&=\Big\langle\Phi_-(x_4)\big(\Phi_+(x_3)+\Phi_-(x_3)\big)\big(\Phi_+(x_2)+\Phi_-(x_2)\big)\Phi_+(x_1)\Big\rangle\\&=\Big\langle\big([\Phi_-(x_4),\Phi_+(x_3)]+\Phi_-(x_4)\Phi_-(x_3)\big)\big([\Phi_-(x_2),\Phi_+(x_1)]+\Phi_+(x_2)\Phi_+(x_1)\big)\Big\rangle\\&=G(x_4,x_3)G(x_2,x_1)+\Big\langle\Phi_-(x_4)\Phi_-(x_3)\Phi_+(x_2)\Phi_+(x_1)\Big\rangle\\&=G(x_4,x_3)G(x_2,x_1)+\Big\langle\Phi_-(x_4)\big([\Phi_-(x_3),\Phi_+(x_2)]+\Phi_+(x_2)\Phi_-(x_3)\big)\Phi_+(x_1)\Big\rangle\\&=G(x_4,x_3)G(x_2,x_1)+G(x_3,x_2)G(x_4,x_1)+G(x_4,x_2)G(x_3,x_1).\end{aligned} \tag{3.101}$$

This pattern continues to higher orders: the n-point function with even n is given by the sum over all pairings of n of the products of two-point functions of the pairing, with the order of the operators in each pair given by their order in the full n-point function. The same is true for the time-ordered n-point function, but with the two-point function replaced by the Feynman propagator.

3.8 Homework

1. Evaluate the other two terms in our expression (3.35) for the free scalar Hamiltonian, confirming that this leads to (3.37).

2. Find the vacuum wave functional for a free scalar field. Hint: the answer has the form

$$\Psi[\phi] \propto \exp\left[-\frac{1}{2}\int d^{d-1}x d^{d-1}y K(\vec{x}-\vec{y})\phi(\vec{x})\phi(\vec{y})\right],$$

   so you just need to find the function $K(\vec{x}-\vec{y})$. The condition you need to satisfy is that this wave functional is annihilated by $a_{\vec{p}}$ for all momenta $\vec{p}$, and you can use the expression (3.30) for $a_{\vec{p}}$ and also the definition (3.9) of the canonical momenta acting on wave functionals. Your life will be easiest if you transform K and ϕ to momentum space, but extra credit if you can give a position-space expression for K in $d = 4$ (Bessel functions are involved).

3. The response of superfluid liquid helium to a localized perturbation with source $O_1 = \phi(t_1, \vec{x}_1)$ is given by equation (3.76), with the two Wightman functions appearing in the commutator given by (3.85). Taking the perturbation at $t_1 = 0$ and $\vec{x}_1 = 0$ and taking the measured operator O_2 to be $\phi(t, \vec{x})$, plot the response $\langle\Omega|\phi(t, \vec{x})|\Omega\rangle$ as a function of t and the spatial radius $r = |x|$ for $d = 3$ and $d = 4$. Is there a qualitative difference between two cases?

4. Starting from the expression (3.64) for a complex scalar field and the canonical commutators (3.66), calculate the commutators of the operators $a_{\vec{p}}$, $b_{\vec{p}}$, $a^\dagger_{\vec{p}}$, $b^\dagger_{\vec{p}}$. Derive an expression for the Hamiltonian H in terms of these creation/annihilation operators, and also give an expression for the symmetry charge Q for the symmetry $\phi' = e^{i\theta}\phi$ that you derived in the last homework. What are the charges of the particles in this theory?

5. Expand the massless two-point function (3.85) in the limit $d \to 2$. You will find a series in $(d-2)$ that begins with a divergence that goes like $1/(d-2)$ followed by a term that is finite and nonzero as $d \to 2$. What is this correction term? Do you see anything strange about it?

6. Extra credit: evaluate the momentum integral (3.81) for the two-point function assuming that x_1 and x_2 are spacelike separated in the cases $d = 2$, $d = 3$, and $d = 4$. You will likely need to consult some reference on Bessel integrals, e.g. Gradshteyn and Ryzhik or Abramowitz and Stegun, both of which are available as pdfs online. If the experience leaves you enthusiastic you can try the case of timelike separation as well; this is actually a bit easier since you can go to a frame where $\vec{x}_2 - \vec{x}_1 = 0$.

4 Algebras and symmetries in quantum field theory

In this section we return to formalism, introducing a general algebraic language that we can use to precisely define the idea of symmetry in quantum field theory. We will learn about the difference between internal symmetries and spacetime symmetries, learn more about global structure of the Lorentz group, and study how correlation functions in quantum field theory are constrained by global symmetries.

4.1 The algebraic approach to field theory

In the Lagrangian approach to field theory we have been pursuing thus far, there is a set of "fundamental" fields $\phi^a(x)$ appearing as dynamical variables in the Lagrangian. Other local operators such as ϕ^2 and $\partial_\mu\phi\partial_\nu\phi$ are constructed out of these fundamental fields and their derivatives. In strongly-interacting theories however it is often the case that the fundamental fields are not so closely related to the interesting physics at long distances. Indeed sometimes the same quantum field theory has multiple presentations in terms of different choices of fundamental fields, which is a phenomenon called **duality**. It is therefore sometimes useful to adopt a language for quantum field theory that de-emphasizes the fundamental fields and treats all local operators on equal footing. This is the **algebraic approach** to quantum field theory.

The basic idea of algebraic field theory is that for each open spatial region R there is an *algebra* of operators $\mathcal{A}[R]$ associated to that region. Roughly speaking $\mathcal{A}[R]$ consist of all the operators made out of sums and products of the fields in R and their derivatives. There are various opinions about how general the spatial regions R should be, in this class we will require that each R lies within a constant time slice in some Lorentz frame.[25] The algebras obey three natural axioms:

- **Nesting:** If $R_1 \subset R_2$, then $\mathcal{A}[R_1] \subset \mathcal{A}[R_2]$.
- **Causality:** If R_1 and R_2 are spacelike separated, then $\mathcal{A}[R_1] \subset \mathcal{A}'[R_2]$. Here the symbol $\mathcal{A}'[R]$ indicates the **commutant** of $\mathcal{A}[R]$, meaning the set of (bounded) operators that commute (or anticommute in the case of fermions) with everything in $\mathcal{A}[R]$.
- **Haag Duality:** For any region R we have $\mathcal{A}'[R] = \mathcal{A}[\overline{R}]$, where $\overline{R}$ is the interior of the spatial complement of R in the time slice it lives in.

Nesting, also sometimes called "isotony", formalizes the idea that you cannot make more operators by restricting which fields you can use, causality is a consequence of the (anti)commutativity of fields at spacelike separation, and Haag duality expresses the idea that the algebra is "complete" in the sense that $\mathcal{A}[R]$ contains everything you can build out of the fields.[26]

Conceptually these axioms are all we will need from the algebraic approach to field theory, but there are some mathematical subtleties in making the definition of $\mathcal{A}[R]$ precise which are worth discussing. Don't worry if the rest of this section goes by too fast, the goal is to make you aware of these things rather than to turn you into a master practitioner. The first problem is that we saw in the last section that the fields themselves are not actually genuine operators. For example if we act with a free scalar field on the vacuum we get a state of infinite norm:

$$\langle\Omega|\Phi(x)\Phi(x)|\Omega\rangle = G(x,x) = \infty. \tag{4.1}$$

To get a good operator we need to smear against a smooth (meaning infinitely-differentiable) function $f : \mathbb{R}^d \to \mathbb{R}$ of compact support:

$$\Phi_f = \int d^dx f(x)\Phi(x). \tag{4.2}$$

[25] We do this to avoid needing to discuss quantization on curved slices. More generally R can be any open subset of a Cauchy slice.

[26] Haag duality should not be confused with the "duality" mentioned in the previous paragraph, whereby the same quantum field theory can have two seemingly different presentations. Unfortunately both usages are completely standard.

To see that this makes the norm finite, we can first note that we have

$$\begin{aligned}\Phi_f &= \int dx f(x) \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left[e^{ip\cdot x} a_{\vec{p}} + e^{-ip\cdot x} a_{\vec{p}}^\dagger \right] \\ &= \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left[\widetilde{f}(\omega_{\vec{p}}, \vec{p})^* a_{\vec{p}} + \widetilde{f}(\omega_{\vec{p}}, \vec{p}) a_{\vec{p}}^\dagger \right], \end{aligned} \tag{4.3}$$

where

$$\widetilde{f}(k) = \int d^d x e^{-ik\cdot x} f(x) \tag{4.4}$$

is the (d-dimensional) Fourier transform of f. It is useful to recall two facts about Fourier transforms:

- If $f : \mathbb{R}^d \to \mathbb{R}$ is a smooth function that is bounded in absolute value by $\frac{C}{1+|x|^{d+1}}$ for some $C > 0$ (here $|x|$ is the Euclidean length on $\mathbb{R}^d$), and moreover which has the property that when acted on by any finite number of partial derivatives it continues obey this bound (possibly with different C for different sets of derivatives), then the Fourier transform $\widetilde{f}(k)$ exists and decays faster than any power at large $|k|$. The proof of this is fairly simple: by differentiating under the integral sign and integration by parts we have

$$\begin{aligned} k_{\mu_1} \dots k_{\mu_m} \widetilde{f}(k) &= \int d^d x k_{\mu_1} \dots k_{\mu_m} e^{-ik\cdot x} f(x) \\ &= i^m \int d^d x \partial_{\mu_1} \dots \partial_{\mu_m} (e^{-ik\cdot x}) f(x) \\ &= (-i)^m \int d^d x e^{-ik\cdot x} \partial_{\mu_1} \dots \partial_{\mu_m} f(x), \end{aligned} \tag{4.5}$$

  and the third line vanishes at large $|k|$ by the Riemann-Lebesgue lemma (see Wikipedia) since by assumption $\partial_{\mu_1} \dots \partial_{\mu_m} f(x)$ is integrable since it is smooth and bounded in absolute value by $\frac{C}{1+|x|^{d+1}}$.

- If $f : \mathbb{R}^d \to \mathbb{R}$ is a continuous function of compact support then its Fourier transform $\widetilde{f}(k)$ is an entire function, meaning that it is analytic for arbitrary complex k. This is because we can simply define the derivative of the Fourier transform by

$$\frac{\partial \widetilde{f}}{\partial k_\mu} = \int_S d^d x (ix^\mu) e^{-ik\cdot x} f(x), \tag{4.6}$$

  which is convergent since f is continuous and S (the support of f) is compact.

Results of this type illustrate the general maxim that continuity/differentiability properties of an integrable function f translate into statements about the decay of its Fourier transform at infinity.[27] In particular we have learned that the Fourier transform $\widetilde{f}$ of a smooth function of compact support is a very well-behaved function: it is analytic for all k and decays faster than any power at infinity. These properties ensure that Φ_f is a better-behaved operator than $\Phi(x)$. For example we can compute the norm of the state $\Phi_f|\Omega\rangle$:

$$\langle \Omega | \Phi_f \Phi_f | \Omega \rangle = \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{1}{2\omega_{\vec{p}}} |\widetilde{f}(\omega_{\vec{p}}, \vec{p})|^2. \tag{4.7}$$

This integral is now convergent at large $|p|$ due to the fast decay of $\widetilde{f}$, and for $m > 0$ it is also convergent at $p = 0$ since $\omega_{\vec{p}}$ is finite there and $\widetilde{f}$ is analytic. When $m = 0$ there is an apparent singularity at $p = 0$ due to the $\omega_{\vec{p}}$ in the denominator, but as long as $d > 2$ this is compensated by the volume measure $\frac{d^{d-1}p}{(2\pi)^{d-1}}$

[27] Another useful result which is intermediate between these two is that an integrable function which is analytic in a strip of finite thickness about the real axis has a Fourier transform which decays exponentially at large k.

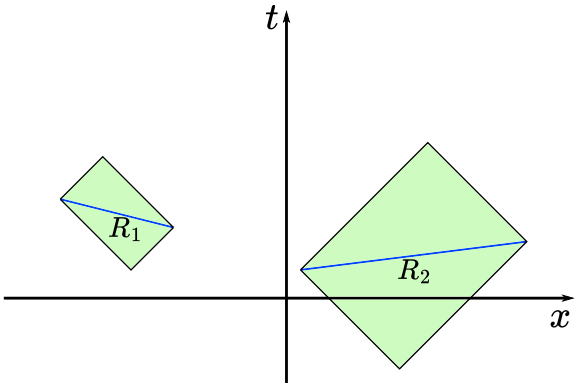


Figure 7: Domains of dependence for spatial regions in Minkowski space. The regions R_1 and R_2 are blue, while their domains of dependence are the green diamond-shaped spacetime regions.

so the integral is still finite. When $d = 2$ there is a logarithmic divergence at $p = 0$ in the massless case, which shows that there is indeed an infrared pathology for a massless scalar in $d = 2$ that cannot be removed by smearing.[28] It is important to emphasize that the introduction of smeared fields Φ_f is not purely a mathematical convenience; no real detector has perfect spatial resolution, so this smearing is really physical - the function f describes the spacetime profile of the detector which couples to $\Phi(x)$.

Which smeared operators ϕ_f can be associated to which spatial regions R? The answer to this question is not completely obvious, since in order to get a good operator we need the support of f to have nontrivial extent in time. On the other hand we should expect that in a relativistic field theory the operators at a location x which lies to the future or past of a timeslice should be expressible solely in terms of the fields on that timeslice which are not spacelike separated from x. Given a spatial region R we therefore introduce the idea of its **domain of dependence** $D[R]$, which is the set of spacetime points x with the property that every timelike curve which intersects x also intersects R. In Minkowski space this is equivalent to the set of points which are spacelike-separated from all points in $\overline{R}$, see figure 7 for an illustration.[29] Moreover this definition has the property that if $R_1 \subset R_2$ then $D[R_1] \subset D[R_2]$. Operators ϕ_f with the support of f contained in $D[R]$ thus will obey nesting and causality, and therefore are thus natural candidates for elements of $\mathcal{A}[R]$.

There is one further issue however that needs to be addressed: although the operator Φ_f is better-behaved than $\Phi(x)$, it still can in general have arbitrary large eigenvalues. An operator whose eigenvalues are unbounded can have rather strong restrictions on its domain, which makes it difficult to include in an algebra since products of unbounded operators are complicated to handle. For example in the simple harmonic oscillator the state

$$|\psi\rangle = \frac{\sqrt{6}}{\pi} \sum_{n=0}^{\infty} \frac{1}{n} |n\rangle \tag{4.8}$$

has unit norm but if we act on this state with the Hamiltonian $H = \sum_n \omega(n + 1/2)|n\rangle\langle n|$ we get a state of infinite norm and the expectation value of H in the state $|\psi\rangle$ is also infinity. This kind of divergence is usually viewed as unphysical however, as given a detector of finite size we can't actually measure an observable with an infinite number of distinct possible outcomes. It is thus standard to restrict $\mathcal{A}[R]$ to only contain operators O which are **bounded** in the sense that there is some constant C such that $\sqrt{\langle\psi|O^\dagger O|\psi\rangle}| \leq C$ for all normalizeable states $|\psi\rangle$. Given smeared fields Φ_f it is not difficult to create bounded operators, for

[28] This has interesting physical consequences, with perhaps the most important being that there cannot be spontaneous breaking of a continuous symmetry in $d = 2$. This statement is called the Mermin-Wagner-Coleman theorem, and we will say more about it when we get to spontaneous symmetry breaking later in the semester.

[29] Another way to motivate the definition of the domain of dependence is that it is the region in which the wave equation (or more generally any well-behaved hyperbolic PDE) should have a unique solution given initial data specified on R. Outside of $D[R]$ the solution will depend also on the initial data on $\overline{R}$.

example if $\Phi(x)$ is hermitian then $e^{i\Phi_f}$ and $\frac{1}{1+\Phi_f^2}$ are both bounded, and so is the spectral projection onto the eigenstates of Φ_f which lie between any two distinct real numbers.

The algebra $\mathcal{A}[R]$ associated to a spatial region R in quantum field theory gives an example of a famous mathematical notion:

Definition 1 *Let $\mathcal{H}$ be a Hilbert space. A set $\mathcal{A}$ of bounded operators on $\mathcal{H}$ is a **von Neumann algebra** if the following things are true:*

(1) $\mathcal{A}$ is closed under addition, multiplication, and hermitian conjugation.

(2) $\mathcal{A}$ contains λI for any $\lambda \in \mathbb{C}$, where I is the identity operator.

(3) $\mathcal{A}$ is closed under "weak limits", meaning that of $O_n \in \mathcal{A}$ are a sequence of operators such that the sequences $\langle\psi|O_n|\phi\rangle$ are convergent for all states $|\psi\rangle, |\phi\rangle \in \mathcal{H}$ then there exists an operator $O \in \mathcal{A}$ such that $\langle\psi|O|\phi\rangle = \lim_{n\to\infty}\langle\psi|O_n|\phi\rangle$ for all $|\psi\rangle, |\phi\rangle \in \mathcal{H}$.

Elements of $\mathcal{A}[R]$ are bounded for the reasons discussed in the previous paragraph, they obey (1) because if we can measure two hermitian operators O_1 and O_2 then we can measure simple functions of them such as $O_1 + O_2$ and $O_1O_2 + O_2O_1$ and $i(O_1O_2 - O_2O_1)$, they obey (2) because we can always measure the identity by doing nothing, and they obey (3) because a limit of measurements should be a measurement. There are many powerful mathematical results about von Neumann algebras with interesting implications for quantum field theory, and in particular there is a classification of von Neumann algebras under which the algebras associated to bounded regions are "type III_1", but this is not a class in mathematical physics we will stop here.

4.2 Symmetry in quantum mechanics

What is a symmetry in quantum field theory? At the classical level we already discussed this in the context of Noether's theorem, where we defined a symmetry as a local transformation of the dynamical fields which leaves the action invariant up to future/past boundary terms. From the path integral point of view (which we have not yet introduced) we could just continue to apply this definition quantum mechanically, but it is useful to also consider how to define symmetries in quantum mechanics directly from the Hilbert space point of view.

A rather minimal requirement for a symmetry in quantum mechanics is that it should at least preserve the probabilistic interpretation of the inner product, meaning that it should be an invertible transformation $f : \mathcal{H} \to \mathcal{H}$ of Hilbert space that preserves instantaneous transition amplitudes

$$|(f(\psi), f(\phi))|^2 = |(\psi, \phi)|^2. \tag{4.9}$$

Here we have temporarily dispensed with Dirac notation and instead used the mathematician notation $(\cdot, \cdot)$ for the inner product on $\mathcal{H}$.[30] We also require that the inverse transformation preserves amplitudes in the same way. It is a fundamental theorem of Wigner (see section 2.A of Weinberg) that the only transformations obeying these requirements arise from unitary or antiunitary operators on $\mathcal{H}$. In other words we must either have a **linear** operator U obeying

$$(U\psi, U\phi) = (\psi, \phi) \tag{4.10}$$

for all ψ and ϕ such that

$$f(\psi) = U\psi, \tag{4.11}$$

or else an **antilinear** operator Θ obeying

$$(\Theta\psi, \Theta\phi) = (\phi, \psi) \tag{4.12}$$

[30] The reasons for this notational change are 1) to write equation (4.9) in Dirac notation we'd need to introduce a dual action of f on bras and 2) Dirac notation is confusing when antilinear operators are involved.

for all ψ and ϕ such that

$$f(\psi) = \Theta\psi. \tag{4.13}$$

A linear operator L is one for which

$$L(a\psi + b\phi) = aL\psi + bL\phi, \tag{4.14}$$

while an antilinear operator A is one for which

$$A(a\psi + b\phi) = a^* A\psi + b^* A\phi. \tag{4.15}$$

Defining the adjoints of linear/antilinear operators by

$$\begin{aligned}(\psi, L^\dagger\phi) &= (L\psi, \phi)\\ (\psi, A^\dagger\phi) &= (\phi, A\psi),\end{aligned} \tag{4.16}$$

we see that a linear operator U is unitary if and only if $U^\dagger U = I$ and an antilinear operator Θ is antiunitary if and only if $\Theta^\dagger\Theta = I$.

Although preserving instantaneous transition amplitudes is a necessary condition to have a symmetry in quantum mechanics, it is clearly not sufficient: otherwise any unitary or antiunitary operator would be a symmetry! There must also be a sense in which the unitaries/antiunitaries which are genuine symmetries preserve more of the structure of the theory. In particular any symmetry of quantum theory should be compatible with its *dynamics*. This requirement is easiest to formalize when the symmetry in question does not affect the direction of time evolution: we then simply require that

$$e^{-iHt}U = Ue^{-iHt}, \tag{4.17}$$

i.e. that transforming and then evolving is the same as evolving and then transforming. Multiplying by $U^\dagger$ on the left, we can also write this as

$$U^\dagger e^{-iHt}U = e^{-iHt}. \tag{4.18}$$

Since either of these equations must be true for all t, they are equivalent to requiring that

$$(iH)U = U(iH). \tag{4.19}$$

So far we have not decided whether U is unitary or antiunitary. Let's first try antiunitary: then (4.19) is equivalent to

$$HU = -UH. \tag{4.20}$$

This however leads to trouble: if ψ_E is an energy eigenstate of energy E, then we have

$$HU\psi_E = -UH\psi_E = -EU\psi_E \tag{4.21}$$

and thus we see that $U\psi_E$ is an energy eigenstate of energy $-E$. Most Hamiltonians of physical interest do not have the property that their spectrum is symmetric about $H = 0$, and in particular in quantum field theory the Hamiltonian is usually bounded from below but not from above. Thus we have learned that *any symmetry which does not affect the direction of time evolution is implemented by a unitary (NOT antiunitary) operator on Hilbert space.* Equation (4.19) then tells us that

$$HU = UH, \tag{4.22}$$

which is the usual maxim that a symmetry in quantum mechanics is a unitary operator that commutes with the Hamiltonian.

The set of all distinct unitaries U that commute with the Hamiltonian form what mathematicians call a **group**, which is a set G whose elements can be multiplied together in such a way that the following conditions are true:

- **Associativity:** For any $g_1, g_2, g_3 \in G$ we have $(g_1 g_2)g_3 = g_1(g_2 g_3)$.
- **Identity:** There exists an element $e \in G$ such that $eg = ge = g$ for all $g \in G$.
- **Inverses:** For each $g \in G$, there exists g^{-1} such that $gg^{-1} = g^{-1}g = e$.

These axioms imply that e and g^{-1} are unique. They are obeyed here because if U_1 and U_2 commute with the Hamiltonian then

$$U_1 U_2 H = U_1 H U_2 = H U_1 U_2, \tag{4.23}$$

and if $UH = HU$ then

$$U^\dagger H = U^\dagger H U U^\dagger = U^\dagger U H U^\dagger = H U^\dagger. \tag{4.24}$$

Hopefully this is not your first time seeing the definition of a group, but if it is then I assure you groups are ubiquitous in physics so best to get started learning about them. Simple examples of groups are the real numbers $\mathbb{R}$ under addition, the group $U(1)$ of complex phases $e^{i\theta}$ under multiplication, the group $U(N)$ of $N \times N$ unitary matrices under matrix multiplication, and the group $SU(N)$ of $N \times N$ unitary matrices of determinant one (again under matrix multiplication).[31] A group G is called **abelian** if it is commutative, meaning that $g_1 g_2 = g_2 g_1$ for all $g_1, g_2 \in G$. $\mathbb{R}$ and $U(1)$ are abelian, while $U(N)$ and $SU(N)$ are non-abelian for $N \geq 2$.

What about symmetries that *do* affect the direction of time evolution? In relativistic theories there are only two such symmetries: we can mix time and space translations using a Lorentz boost, or we can reverse the direction of time using time-reversal symmetry.[32] We have already seen in our free scalar theory that any Lorentz transformation which does not reverse time can be represented by a unitary operator $U(\Lambda)$ which acts on the annihilation operators as

$$U(\Lambda) a_{\vec{p}} U(\Lambda)^\dagger = \sqrt{\frac{\omega_{\vec{p}_\Lambda}}{\omega_{\vec{p}}}} a_{\vec{p}_\Lambda}, \tag{4.25}$$

so in particular this is true for Lorentz boosts. More generally in any quantum field theory we expect that a Lorentz boost in the $\hat{n}$ direction of rapidity η acts on the Hamiltonian as

$$U^\dagger H U = \cosh\eta\, H + \sinh\eta\, \hat{n} \cdot \vec{P}, \tag{4.26}$$

which is a consequence of the fact that the spacetime momentum P^μ transforms as a spacetime vector. Since we are (momentarily) considering the possibility that U could be antiunitary however, we should really require that

$$U^\dagger (iH) U = i(\cosh\eta\, H + \sinh\eta\, \hat{n} \cdot \vec{P}). \tag{4.27}$$

If U is unitary this is equivalent to (4.26), but if it is antiunitary then we should instead require that

$$U^\dagger H U = -(\cosh\eta\, H + \sinh\eta\, \hat{n} \cdot \vec{P}) \tag{4.28}$$

This equation however is not continuous as $\eta \to 0$, so this would be a rather pathological representation of Lorentz symmetry. Moreover it would again have a problem with the spectrum of the Hamiltonian: given a simultaneous eigenstate $\psi_{E,\vec{p}}$ of H and $\vec{P}$, we would have

$$H U \psi_{E,\vec{p}} = -\left(\cosh\eta E + \sinh\eta \hat{n} \cdot \vec{p}\right) U \psi_{E,\vec{p}}. \tag{4.29}$$

In any quantum field theory which can be interpreted as a scattering theory of particles it is quite natural to impose the following requirement:

[31] These examples may misleadingly suggest that all groups are **matrix groups**, meaning groups that can be represented with finite-dimensional matrices. This is true for groups which are topologically compact, but it isn't true in general.

[32] Time-reversal symmetry may not actually be a symmetry by itself, for example in the Standard Model of particle physics it isn't, but we will see in section 6 that there is a combination of time reversal with other transformations, called $\mathcal{CRT}$, which is always a symmetry in any relativistic quantum field theory.

- **Spectrum condition:** In any relativistic quantum field theory we have

$$H \geq \hat{n} \cdot \vec{P}, \tag{4.30}$$

  where $\hat{n}$ is any unit vector and H is defined so that the energy of the ground state is zero. The operator inequality means that $H - \hat{n} \cdot \vec{P}$ is a positive semidefinite operator.

This condition should hold because each particle has energy $\omega = \sqrt{|p|^2 + m^2} \geq |p|$, and when we add up energies there are no cancellations while when we add up momenta there can be.[33] We then have (for $\eta > 0$)

$$\cosh\eta\, E + \sinh\eta\, \hat{n} \cdot \vec{p} \geq \cosh\eta E - \sinh\eta |p| \geq E(\cosh\eta - \sin\eta), \tag{4.31}$$

and so assuming that H is unbounded from above we can again generate energy eigenstates of arbitrarily negative energy by acting with U. From now on we will therefore assume that boosts are implemented by unitary operators.

Finally we can consider time-reversal, which we will take to be represented by an operator $\Theta_{\mathcal{T}}$. This should act on the time evolution operator as

$$\Theta_{\mathcal{T}}^\dagger e^{-iHt} \Theta_{\mathcal{T}} = e^{iHt}, \tag{4.32}$$

and thus obey

$$\Theta_{\mathcal{T}}^\dagger (iH) \Theta_{\mathcal{T}} = -iH. \tag{4.33}$$

If we assume $\Theta_{\mathcal{T}}$ is unitary then we have

$$\Theta_{\mathcal{T}}^\dagger H \Theta_{\mathcal{T}} = -H, \tag{4.34}$$

which we can discard as before since it would require the spectrum of H to be symmetric about zero. We therefore see that we want $\Theta_{\mathcal{T}}$ to be antiunitary, since this gives the more reasonable condition

$$\Theta_{\mathcal{T}}^\dagger H \Theta_{\mathcal{T}} = H. \tag{4.35}$$

For example in the simple harmonic oscillator time reversal is implemented by an antiunitary operator which acts on the X basis

$$\Theta_{\mathcal{T}} |x\rangle = |x\rangle, \tag{4.36}$$

leading to

$$\begin{aligned} \Theta_{\mathcal{T}}^\dagger X \Theta_{\mathcal{T}} &= X \\ \Theta_{\mathcal{T}}^\dagger P \Theta_{\mathcal{T}} &= -P. \end{aligned} \tag{4.37}$$

The energy eigenstates $|n\rangle$ have real wave functions in the X basis, and thus are invariant under time-reversal:

$$\Theta_{\mathcal{T}} |n\rangle = \Theta_T \int dx \langle x|n\rangle |x\rangle = \int dx \langle x|n\rangle \Theta_{\mathcal{T}} |x\rangle = \int dx \langle x|n\rangle |x\rangle = |n\rangle. \tag{4.38}$$

4.3 Internal symmetries in quantum field theory

In quantum field theory there is additional structure which is not present in general quantum systems: the operators are organized into the local algebras $\mathcal{A}[R]$ obeying nesting, causality, and duality. In order for a symmetry in quantum field theory to be useful, it needs to respect this local structure. The simplest kind of symmetry that respects this structure is an **internal** symmetry, which roughly speaking is a symmetry that maps any local (Heisenberg) operator $O(x)$ to another local operator which is located at the same spacetime point. More formally we have a definition:

[33] If we are willing to just assume that boosts are unitary, for example because we reject the discontinuity at $\eta = 0$ in the antiunitary case, then we can give a simpler and more rigorous argument for the spectrum condition: it must be true so that the Hamiltonian in any Lorentz frame is a positive operator.

Definition 2 *An **internal symmetry** of a quantum field theory in d-dimensional Minkowski space is a unitary operator U such that*

(1) For any spatial region R the algebra $\mathcal{A}[R]$ is preserved by conjugation by U and $U^\dagger$, meaning that for any $O \in \mathcal{A}[R]$ we have $U^\dagger O U \in \mathcal{A}[R]$ and $UOU^\dagger \in \mathcal{A}[R]$.

(2) For any spacetime point x the energy-momentum tensor $T_{\mu\nu}(x)$ is invariant under conjugation by U:

$$U^\dagger T_{\mu\nu}(x) U = T_{\mu\nu}(x). \tag{4.39}$$

The first requirement here expresses the idea that the symmetry should preserve the local algebra. The second is a strengthening of the idea that U should commute with the Hamiltonian: it expresses the idea of **local conservation** of the symmetry charge. More concretely, it says that symmetry charge cannot leave a region of space without passing through its edges. This is not obvious, and showing it is a consequence of (4.39) requires more differential geometry than we are using in this class.[34] We can also motivate (4.39) in a more mundane way: a generic quantum field theory shouldn't have more than one energy-momentum tensor, and whatever an internal symmetry sends the energy-momentum tensor to is an equally valid candidate for an energy-momentum tensor and therefore must be the original one. The set of internal symmetries in a quantum field theory forms a group, as you can easily check.

There is an important further classification of internal symmetries based on what kinds of operators they act nontrivially on. The simplest kind of nontrivial internal symmetry is one which acts nontrivially on some local operator $O(x)$. More mathematically, the action of U on $\mathcal{A}[R]$ by conjugation is nontrivial for any region R. Such internal symmetries are called **global** internal symmetries. An example of a global internal symmetry is the phase rotation of a free complex scalar

$$U(\theta)^\dagger \Phi(x) U(\theta) = e^{i\theta}\Phi(x), \tag{4.40}$$

whose symmetry group is clearly isomorphic to the group $U(1)$. Conventionally we say that this theory has a $U(1)$ global symmetry. This semester we will only discuss theories where all operators are built from local operators, so all internal symmetries are global. Next semester we will discuss **gauge theories** such as quantum electrodynamics, where there can be extended operators that are not built from local operators. The reason for this is familiar from Maxwell theory: we cannot create an electrically charged particle without also creating an electric field sourced by it that satisfies Gauss's law, and this electric field must extend out to spatial infinity. Therefore there are no local operators that carry nonzero electric charge. On the other hand there are clearly states of nonzero electric charge, such as a state with one electron in the center of space. These are created by acting on the vacuum with extended operators that create both the electron and its Coulomb field, and it is these extended operators which carry nonzero electric charge.[35]

Another important question about any internal symmetry in quantum field theory is whether or not the ground state $|\Omega\rangle$ is invariant. If it is not, then we say that the symmetry is **spontaneously broken**. Spontaneously broken global internal symmetries are very interesting in quantum field theory, for example being essential to our understanding of magnets, superfluidity, and nuclear physics. There is also a sense

[34] More formally local conservation is expressed as the requirement that we can continuously deform the slice on which U is defined without changing the operator. This is often described by saying that the symmetry operator U is a **topological surface operator**. In the continuous case this is a consequence of Noether's theorem: the charge $Q = \int d^{d-1}x J^0(t,\vec{x})$ can be written as $Q = \int_\Sigma n^\mu J_\mu$ where Σ is the surface $t = 0$ and n^μ is its normal vector, and then the fact that we can continuously deform Σ without changing Q is a consequence of the divergence theorem and the current conservation equation $\partial_\mu J^\mu = 0$. The basic idea in showing that the invariance of $T_{\mu\nu}$ implies this deformability in general is to use that the stress tensor is the functional derivative of the action with respect to the metric and that the action is invariant under arbitrary diffeomorphisms which act on both the dynamical fields and also the background spacetime metric.

[35] The distinction between gauge and global symmetry defined here is not the way this distinction is traditionally presented. The conventional definition is that in terms of the fundamental fields a global symmetry is one which acts the same way at all points in space while a gauge symmetry is one where the symmetry transformation can vary from point to point. This definition is problematic however, as most of the gauge transformations defined this way are mere redundancies of description and for discrete symmetries it isn't clear what the difference is. The algebraic definition I've given here isolates the physical distinction between the two without introducing confusing historical baggage.

in which gauge symmetries can be spontaneously broken, called the Anderson-Higgs mechanism, although the concept is somewhat less well-defined than for global symmetries. We will have more to say about spontaneous symmetry breaking next semester. If an internal global symmetry is **unbroken**, meaning that the ground state is invariant, then it implies a powerful constraint on the correlation functions of the theory. Indeed if we define

$$O'(x) = U^\dagger O(x) U, \tag{4.41}$$

then we must have

$$\langle O'_1(x_1) \dots O'_n(x_n) \rangle = \langle \Omega | U^\dagger O_1(x_1) U \dots U^\dagger O_n(x_n) U | \Omega \rangle = \langle O_1(x_1) \dots O_n(x_n) \rangle. \tag{4.42}$$

For example if we have a $U(1)$ global symmetry, this tells us that for all $\theta \in [0, 2\pi]$ we have

$$e^{i(q_1 + \dots q_n)\theta} \langle O_1(x_1) \dots O_n(x_n) \rangle = \langle O_1(x_1) \dots O_n(x_n) \rangle, \tag{4.43}$$

which shows that this correlation function obeys the selection rule that it must vanish unless the sum of the operator charges vanishes. For example this explains why you will find that $\langle \Phi(x)\Phi(y) \rangle = \langle \Phi^\dagger(x)\Phi^\dagger(y) \rangle = 0$ in the free complex scalar theory.

4.4 Spacetime symmetries in quantum field theory

We now turn to symmetries that act nontrivially on the spacetime coordinates, which are called spacetime symmetries. In terms of their action on operators, these are symmetries that move local operators around. In relativistic field theory the most familiar of these are Poincaré transformations, which for example act on a scalar field as

$$U(\Lambda, a)^\dagger \Phi(x) U(\Lambda, a) = \Phi(\Lambda^{-1}(x - a)). \tag{4.44}$$

There are three other kinds of spacetime symmetries that can show up in relativistic theories, which I'll mention here but not discuss further:

- **Conformal symmetry:** In quantum field theories which do not possess any dimensionful parameter, such as the massless free scalar theory, then in addition to Poincaré symmetry we also have a scaling symmetry $x^{\mu\prime} = \lambda x^\mu$ for any $\lambda > 0$. It is not obvious, but Poincaré symmetry plus scaling symmetry seems to imply the existence of a broader spacetime symmetry called conformal symmetry, which consists of arbitrary angle-preserving coordinate transformations. Field theories with this enhanced symmetry are called conformal field theories, and conformal field theories are very important to the logical structure of quantum field theory: any quantum field theory is supposed to asymptote to a conformal field theory in the limit of short or long distance.

- **Supersymmetry:** Supersymmetries are fermionic symmetries that exchange fermionic and bosonic fields. The spin-statistics theorem (which we will prove in section 6) shows that bosons must have integer spin and fermions must have half-integer spin, so a symmetry which exchanges them must transform nontrivially under rotations. Supersymmetries thus mix nontrivially with Poincaré transformations, and must thus be spacetime symmetries themselves. Supersymmetric field theories have many nice properties, and in particular many interesting quantities can be computed exactly. They are thus a source of interesting solvable examples of interesting field theory phenomena. There is also some hope that supersymmetry will be relevant in the real world, for example to address the hierarchy problem in particle physics (as we will discuss later), and also in string theory where supersymmetry seems to be necessary for the consistency of the theory.

- **Diffeomorphism symmetry:** There is a particularly simple kind of quantum field theory called a topological field theory, for which *arbitrary* coordinate transformations are symmetries. These theories arise in some interesting condensed matter systems such as those exhibiting the fractional quantum hall effect, and they also appear in some corners of string theory. One can think of topological field theory as a special kind of conformal field theory.

Returning to Poincaré symmetry, the full set of Poincaré transformations forms a group called the **Poincaré group** and it is useful to now make a few general comments about its global structure. Recall that this is defined to be the set of coordinate transformations

$$x^{\mu\prime} = \Lambda^{\mu}{}_{\nu} x^{\nu} + a^{\mu}, \tag{4.45}$$

with a^{μ} arbitrary and Λ obeying

$$\Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}\eta_{\mu\nu} = \eta_{\alpha\beta}. \tag{4.46}$$

The subgroup of the Poincaré group with $a = 0$ is called the **Lorentz** group, and it is denoted $O(d-1,1)$. Taking the determinant of (4.46) we see that

$$(\det \Lambda)^2 = 1, \tag{4.47}$$

and splitting the time and space terms of the 00 component of (4.46) we see that

$$(\Lambda^{0}{}_{0})^2 = 1 + \sum_i (\Lambda^{i}{}_{0})^2 \tag{4.48}$$

and thus

$$(\Lambda^{0}{}_{0})^2 \geq 1. \tag{4.49}$$

We therefore can split up the Lorentz group into four connected components labeled by the signs of $\det \Lambda$ and $\Lambda^{0}{}_{0}$. The simplest of these components is the one containing the identity transformation, which is called the identity component and denoted $SO^{+}(d-1,1)$ (here "S" indicates unit determinant and "+" indicates $\Lambda^{0}{}_{0} \geq 1$). Any element of the other components can be written as an element of $SO^{+}(d-1,1)$ multiplied by one of the following three Lorentz transformations:

$$\begin{aligned}
\mathcal{R} &: (t, x^1, x^2, \ldots x^{d-1}) \mapsto (t, -x^1, x^2, \ldots, x^{d-1}) \\
\mathcal{T} &: (t, x^1, x^2, \ldots x^{d-1}) \mapsto (-t, x^1, x^2, \ldots, x^{d-1}) \\
\mathcal{RT} &: (t, x^1, x^2, \ldots x^{d-1}) \mapsto (-t, -x^1, x^2, \ldots, x^{d-1}).
\end{aligned} \tag{4.50}$$

The transformation $\mathcal{R}$ reflects the spatial x^1 coordinate, the transformation $\mathcal{T}$ reverses time, and the transformation $\mathcal{RT}$ does both. Due to our general discussion above we should expect that $\mathcal{T}$ and $\mathcal{RT}$ are represented by antiunitary operators $\Theta_{\mathcal{T}}$ and $\Theta_{\mathcal{RT}}$, while $\mathcal{R}$ is represented by a unitary operator $U_{\mathcal{R}}$. Therefore two of the connected components of the Lorentz group are unitary and two are antiunitary. When d is even it is conventional to replace $\mathcal{R}$ by an operation $\mathcal{P}$, called parity, that reflects all spatial coordinates. When d is odd however $\mathcal{P}$ is in $SO^{+}(d-1,1)$, so in general it is best to stick with $\mathcal{R}$.

The fact that the Poincaré group has four connected components suggests the possibility that there could be relativistic field theories where only some of these components give genuine symmetries. We should always include the identity component $SO^{+}(d-1,1)$ (otherwise what would we mean by "relativistic field theory"), but there are indeed interesting theories where some of the other components are not symmetries. In fact this possibility is realized in the Standard Model of particle physics, which has neither parity nor time-reversal symmetry. On the other hand we will see in section 6 that there is a way of combining $\mathcal{RT}$ with an internal transformation $\mathcal{C}$, called **charge conjugation**, that gives a combined transformation $\mathcal{CRT}$ which is always a symmetry in any relativistic field theory (even if $\mathcal{C}$, $\mathcal{R}$, and $\mathcal{T}$ separately are not symmetries). Thus we always at least have a spacetime symmetry group $SO(d-1,1)$, where the absence of the + indicates that we have included the $\mathcal{RT}$ component of the Lorentz group but the S indicates that we have not included the $\mathcal{R}$ and $\mathcal{T}$ components.

The existence of a unitary representation of $SO^{+}(d-1,1)$ obeying $U(\Lambda,a)^{\dagger}\mathcal{A}[R]U(\Lambda,a) = \mathcal{A}[\Lambda^{-1}(\mathcal{R}-a)]$, the spectrum condition, nesting, and causality together form what are called the **Haag-Kastler axioms** for algebraic quantum field theory. It is widely agreed that these axioms are necessary for any reasonable definition of relativistic quantum field theory. There is less agreement on what else is needed, two things I personally would also include are duality and the existence of a conserved symmetric energy-momentum tensor that generates $SO^{+}(d-1,1)$.

4.5 Correlation functions of tensor fields

Just as in the case of internal symmetries, spacetime symmetries imply powerful constraints on correlation functions. First considering elements of the Poincare group with ${\Lambda^0}_0 \geq 1$, we can define

$$O'(x) = U^\dagger(\Lambda, a) O(x) U(\Lambda, a) \tag{4.51}$$

with $U(\Lambda, a)$ being unitary. Assuming the ground state is invariant under Poincare symmetry, we then have

$$\langle \Omega | O'_1(x_1) \dots O'_n(x_n) | \Omega \rangle = \langle \Omega | O_1(x_1) \dots O_n(x_n) | \Omega \rangle \tag{4.52}$$

just as in the internal case. In particular let's say that the operators $O(x)$ are **tensor fields**, meaning that they come with some number of raised and lowered indices such that their Poincare transformation is

$$O'^{\mu_1 \dots \mu_n}{}_{\nu_1 \dots \nu_m}(x) = {\Lambda^{\mu_1}}_{\alpha_1} \dots {\Lambda^{\mu_n}}_{\alpha_n} {\Lambda_{\nu_1}}^{\beta_1} \dots {\Lambda_{\nu_m}}^{\beta_m} O^{\alpha_1 \dots \alpha_n}{}_{\beta_1 \dots \beta_m}(\Lambda^{-1}(x-a)). \tag{4.53}$$

By taking Λ to be the identity we see that the correlation function must be invariant under translating all of the coordinates $x_1^\mu, \dots, x_n^\mu$ by an arbitrary vector a^μ, and thus that the correlation function can only depend on differences of these coordinates. When Λ is not the identity further constraints are imposed, for example the two-point function of a vector operator $V^\mu(x)$ must obey

$$\langle V^\mu(x_1) V^\nu(x_2) \rangle = {\Lambda^\mu}_\alpha {\Lambda^\nu}_\beta \langle V^\alpha(\Lambda^{-1} x_1) V^\beta(\Lambda^{-1} x_2) \rangle, \tag{4.54}$$

which determines the form of the two-point function to be

$$\langle V^\mu(x_1) V^\nu(x_2) \rangle = \eta^{\mu\nu} f\left((x_1 - x_2)^2\right) + (x_1^\mu - x_2^\mu)(x_1^\nu - x_2^\nu) g\left((x_1 - x_2)^2\right) \tag{4.55}$$

with f and g being functions of a single variable.

We can also consider Poincaré transformations with ${\Lambda^0}_0 \leq -1$, which are implemented by antiunitary operators $\Theta(\Lambda, a)$. The local operators transform as

$$O'(x) = \Theta^\dagger(\Lambda, a) O(x) \Theta(\Lambda, a) \tag{4.56}$$

as before, but the constraint on correlation functions is now a bit trickier to derive. Assuming that the ground state is invariant under $\Theta^\dagger$, we have

$$\begin{aligned} \langle O'_1(x_1) \dots O'_n(x_n) \rangle &= (\Omega, O'_1(x_1) \dots O'_n(x_n) \Omega) \\ &= (\Theta^\dagger \Omega, \Theta^\dagger O_1(x_1) \dots O_n(x_n) \Omega) \\ &= (O_1(x_1) \dots O_n(x_n) \Omega, \Omega) \\ &= (\Omega, (O_1(x_1) \dots O_n(x_n))^\dagger \Omega) \\ &= \langle O_n(x_n)^\dagger \dots O_1(x_1)^\dagger \rangle. \end{aligned} \tag{4.57}$$

Here we have switched to mathematician notation in the middle to handle the antiunitary operators. Thus we see that an antiunitary symmetry reverses the operator of the operators in a correlation function and takes their hermitian conjugates. This has the nice feature that it sends time-ordered correlation functions to time-ordered correlation functions.

4.6 Correlation functions involving conserved currents

We saw in section 2 that from the Lagrangian point of view, Noether's theorem tells us that any *continuous* symmetry in field theory leads to a conserved current $J^\mu(x)$.[36] The current conservation equation imposes

[36] This theorem has not quite been proven from the abstract point of view taken in this section (i.e. using only the Haag-Kastler axioms), and indeed in my long paper with Hirosi Ooguri we give some counterexamples. These counterexamples are in somewhat pathological theories however, and so far it seems likely that Noether's theorem is true for sufficiently well-behaved theories.

interesting constraints on correlation functions that contain such currents, since inserting $\partial_\mu J^\mu$ into any (Wightman) correlation function must give zero. For example you will show on the homework that imposing the conservation equation $\partial_\mu V^\mu = 0$ on the vector field appearing in (4.55) implies that the functions f and g obey the constraint

$$f'(x) + xg'(x) + \frac{d+1}{2} g(x) = 0. \tag{4.58}$$

There is also an interesting constraint on time-ordered correlation functions of a conserved current J^μ. We can illustrate the idea using a two-point function:

$$\begin{aligned} \partial_\mu \langle T J^\mu(x) O(y) \rangle &= \partial_\mu \left(\Theta(x^0 - y^0) \langle J^\mu(x) O(y) \rangle + \Theta(y^0 - x^0) \langle O(y) J^\mu(x) \rangle \right) \\ &= \delta(x^0 - y^0) \langle [J^0(x), O(y)] \rangle, \end{aligned} \tag{4.59}$$

where the term on the right-hand side comes from the derivative acting on the Heaviside Θ function. More generally we have

$$\partial_\mu \langle T J^\mu(x) O_1(y_1) \dots O_n(y_n) \rangle = \sum_{m=1}^{n} \delta(x^0 - y_m^0) \langle T O_1(y_1) \dots [J^\mu(x), O_m(y_m)] \dots O_n(y_n) \rangle. \tag{4.60}$$

Note that the commutators appearing on the right-hand side are at equal time due to the δ-function, and thus vanish when $x \neq y$. We can therefore expand them in the δ function and its derivatives:

$$[J^0(y^0, \vec{x}), O(y^0, \vec{y})] = A(y^0, \vec{y}) \delta^{d-1}(\vec{x} - \vec{y}) + B^i(y^0, \vec{y}) \partial_i \delta^{d-1}(\vec{x} - \vec{y}) + \dots \tag{4.61}$$

Integrating this equation over $\vec{x}$ we see that

$$A(y) = [Q, O(y)] = i \delta_S O(y), \tag{4.62}$$

and so we see that the divergence of a time-ordered correlation function involving a conserved current obeys the **Ward identity**:

$$\partial_\mu \langle T J^\mu(x) O_1(y_1) \dots O_n(y_n) \rangle = i \sum_{m=1}^{n} \delta^d(x - y_m) \langle T O_1(y_1) \dots \delta_S O(y_m) \dots O_n(y_n) \rangle + \dots, \tag{4.63}$$

where the "..." indicates terms proportional to derivatives of $\delta^d(x - y_m)$ with respect to x. In quantum field theory terms in correlation functions which vanish unless the operators are at the same point are called **contact terms**, and usually they have ambiguities depending on how the theory is regulated at short distance. The leading term in (4.61) is an exception, as we were able to determine it from the symmetry algebra.

4.7 Homework

1. Compute the two-point functions $\langle\Phi(x)\Phi(y)\rangle$ and $\langle\Phi^\dagger(x)\Phi(y)\rangle$ for a complex scalar field, giving each answer both as a covariant integral over spacetime momenta and also directly in position space in terms of a Bessel function. You are free to use our results for the real scalar field, so you shouldn't need to evaluate any new integrals.

2. Show that if R_1 and R_2 are open spatial regions (which recall for us means that each lies in a constant time slice in some Lorentz frame) obeying $R_1 \subset R_2$, then their domains of dependence obey $D[R_1] \subset D[R_2]$.

3. Show that $SU(N)$ is indeed a group, meaning that it is closed under matrix multiplication and matrix inverse.

4. Show that every Lorentz transformation is indeed a product of an element of $SO^+(d-1,1)$ with 1, $\mathcal{R}$, $\mathcal{T}$, or $\mathcal{RT}$. Hint: this shouldn't require any detailed calculation or explicit parameterization of the Lorentz group.

5. Argue that the vector two-point function indeed has the form (4.55), and also show that if $\partial_\mu V^\mu = 0$ then (4.58) follows.

6. Check that the two-point functions we computed for real and complex scalar fields are consistent with the time-reversal constraint (4.57).

7. Extra credit: Antiunitary operators may seem somewhat counter-intuitive, but there is an elegant characterization of any antiunitary operator due to Wigner that you will work out in this problem. First argue that if Θ is antiunitary then Θ^2 is unitary. There therefore must be a basis $|i\rangle$ in which we have $\Theta^2|i\rangle = e^{-2i\theta_i}|i\rangle$ and $(\Theta^\dagger)^2|i\rangle = (\Theta^2)^\dagger|i\rangle = e^{2i\theta_i}|i\rangle$ for some $\theta_i \in (-\pi/2, \pi/2]$. Work out how Θ and $\Theta^\dagger$ act in this basis, and then argue that their action on arbitrary superpositions follows from antilinearity. Hint: you want to show that up to phase redefinitions you can take this basis to consist of states which are invariant and pairs of states which are exchanged up to a phase by acting with Θ. You might start by showing that $\Theta|i\rangle$ is also an eigenstate of Θ^2.

5 Path integrals in quantum mechanics and quantum field theory

So far we have discussed quantum field theory in the Hamiltonian formalism. This formalism has many advantages, for example it is where the physical interpretation of a quantum system in terms of measurements and counting degrees of freedom is most clear, but it obscures the full symmetry of relativistic theories since one needs to pick a Lorentz frame to define the canonical momenta and the Hamiltonian.[37] Giving up on manifest Lorentz invariance makes it harder to demonstrate some of the deeper consequences of Lorentz invariance, such as the $\mathcal{CRT}$ and spin-statistics theorems, and it also makes practical calculations more difficult since each intermediate step seems to depend on the Lorentz frame but the end result doesn't. In classical mechanics there is a clear way to handle this problem: we can think more about the Lagrangian and less about the Hamiltonian. The goal of the path integral approach to quantum mechanics, first suggested by Dirac and then greatly expanded by Feynman, is to give an independent (but equivalent) formulation of quantum mechanics that based on the Lagrangian instead of the Hamiltonian. We will spend the rest of this section developing this approach.

5.1 Hamiltonian path integral in quantum mechanics

We will first discuss the path integral for a finite number M of quantum degrees of freedom, which we will refer to as Q^a with $a = 1, 2, \ldots, M$. They have canonical conjugate momenta P_a, and these obey the canonical commutation relations

$$\begin{aligned} [Q^a, P_b] &= i\delta^a_b \\ [Q^a, Q^b] &= 0 \\ [P_a, P_b] &= 0. \end{aligned} \tag{5.1}$$

We will take the Hamiltonian $H(Q, P)$ to be a polynomial in Q and P whose terms are ordered in such a way that all P's appear to the right of all Q's (using the canonical commutation relations we can always write any product of Ps and Qs as a sum of terms with this ordering), and we will work in the Heisenberg picture so that both Q and P are functions of time. For convenience we will take the Hamiltonian to be time-independent, but there is no real difficulty in repeating the argument for a time-dependent Hamiltonian. Let's say we are interested in computing the propagator $G(q_f, q_i; t_f, t_i)$ in the Q basis. In the Schrödinger picture this is given by

$$G(q_f, q_i; t_f, t_i) = \langle q_f | e^{-iH(t_f - t_i)} | q_i \rangle, \tag{5.2}$$

but since we are working in the Heisenberg picture we'll instead write it as

$$G(q_f, q_i; t_f, t_i) = \langle q_f, t_f | q_i, t_i \rangle, \tag{5.3}$$

where $|q, t\rangle$ is a simultaneous eigenstate of the $Q^a(t)$:

$$Q^a(t)|q, t\rangle = q^a |q, t\rangle. \tag{5.4}$$

Explicitly we have $|q, t\rangle = e^{iHt}|q, 0\rangle$.

The idea behind the path integral formalism is to break up the propagator into a repeated integral over propagators with smaller time separation by inserting complete sets of states:

$$\langle q_f, t_f | q_i, t_i \rangle = \prod_{m=1}^{N-1} \left(\int d^M q_m \right) \langle q_f, t_f | q_{N-1}, t_f - \epsilon \rangle \langle q_{N-1}, t_f - \epsilon | q_{N-2}, t_f - 2\epsilon \rangle \ldots \langle q_2, t_2 | q_1, t_1 \rangle \langle q_1, t_i + \epsilon | q_i, t_i \rangle. \tag{5.5}$$

[37] There actually is an approach to Hamiltonian mechanics that does not require a choice of Lorentz frame, which is called the **covariant phase space approach**. See my first paper with Jie-qiang for a review. Quantization in this approach is somewhat subtle, and we won't pursue the topic here.

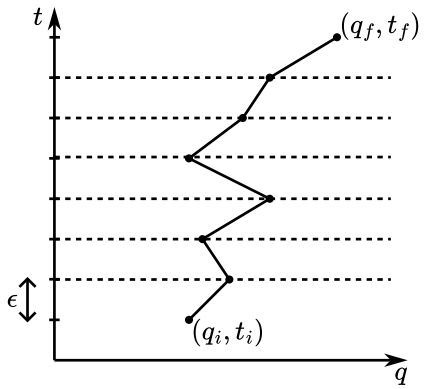


Figure 8: Discretizing a particle trajectory from (q_i, t_i) to (q_f, t_f). The dashed lines show the positions which are integrated over in the intermediate steps.

Here we have split the time interval $t_f - t_i$ into N pieces of size ϵ. We can think of the integration variables $q_1^a, \ldots q_{N-1}^a$ as giving a discretization of possible trajectories the system could follow from q_i^a at time t_i to q_f^a at time t_f, see figure 8 for an illustration in the case of a single particle moving in one dimension. The integral is therefore a sum over (discretized) intermediate trajectories; a **path integral**. The expression (5.5) however is not so useful: we need some way to compute the propagators. At finite ϵ this of course isn't any easier than computing the full propagator, but in the limit of small ϵ a simplification is possible:

$$\begin{aligned}
\langle q', t+\epsilon | q, t\rangle &= \langle q', t | e^{-iH\epsilon} | q, t\rangle \\
&\approx \langle q', t | (1 - i\epsilon H(Q(t), P(t))) | q, t\rangle \\
&= \int \frac{d^M p}{(2\pi)^M} \langle q', t | (1 - i\epsilon H(q', p)) | p, t\rangle \langle p, t | q, t\rangle \\
&= \int \frac{d^M p}{(2\pi)^M} (1 - i\epsilon H(q', p)) e^{i\sum_a p_a (q'^a - q^a)} \\
&\approx \int \frac{d^M p}{(2\pi)^M} e^{i\epsilon\left(\sum_a p_a \frac{q'^a - q^a}{\epsilon} - H(q', p)\right)}.
\end{aligned} \tag{5.6}$$

Here in going from the first to second and fourth to fifth lines we have neglected terms which are $O(\epsilon^2)$, in going from the second to the third line we have inserted a complete set of states and used that in H the momenta are ordered to the right, and in going from the third to the fourth line we have used the momentum-space wave function

$$\langle q, t | p, t\rangle = e^{i\sum_a p^a q_a}. \tag{5.7}$$

We can then use this repeatedly in (5.5) and take the limit $\epsilon \to 0$, which gives

$$\begin{aligned}
\langle q_f, t_f | q_i, t_i\rangle &= \lim_{\epsilon\to 0} \prod_{m=1}^{N-1} \left(\int d^M q_m\right) \prod_{n=0}^{N-1} \left(\int \frac{d^M p_n}{(2\pi)^M}\right) \exp\left[i\epsilon \sum_{\ell=0}^{N-1} \left(\sum_a p_{\ell,a} \frac{q_{\ell+1}^a - q_\ell^a}{\epsilon} - H\left(q_{\ell+1}, p_\ell\right)\right)\right] \\
&:= \int \mathcal{D}q|_{q_i}^{q_f} \int \mathcal{D}p \exp\left[i \int_{t_i}^{t_f} dt \left(\sum_a p_a(t) \dot{q}^a(t) - H(q(t), p(t))\right)\right].
\end{aligned} \tag{5.8}$$

Here we have defined $q_0 = q_i$ and $q_N = q_f$, and $\int \mathcal{D}q|_{q_i}^{q_f}$ indicates a **functional integral** over paths $q^a(t)$ obeying $q^a(t_i) = q_i^a$ and $q^a(t_f) = q_f^a$. $\int \mathcal{D}p$ indicates a functional integral over paths $p_a(t)$ in momentum space with no restrictions at t_i and t_f. Equation (5.8) is called a **Hamiltonian path integral** expression

for the propagator. The quantity appearing in the exponent is essentially i times the Lagrangian, except that p is treated as an independent variable instead of being related to q and $\dot{q}$.

In quantum field theory we are particularly interested in expectation values of products of Heisenberg operators, and these also have a useful path integral representation. Indeed we can consider the quantity

$$\langle q_f, t_f | O_M \left(Q(\bar{t}_M), P(\bar{t}_M)\right) \dots O_1 \left(Q(\bar{t}_1), P(\bar{t}_1)\right) | q_i, t_i \rangle, \tag{5.9}$$

where I've put a line over the times of the Heisenberg operators to distinguish them from the timesteps appearing in the path integral discretization. We will assume that the operators are **time-ordered**, meaning that

$$\bar{t}_1 \leq \bar{t}_2 \leq \dots \leq \bar{t}_M, \tag{5.10}$$

and we will also take these operators to be ordered so that all canonical momenta appear to the left (note that this is the opposite of the ordering we chose for the Hamiltonian). We can evaluate this quantity by inserting complete sets of states as before, except now we occasionally need to evaluate

$$\begin{aligned} \langle q', t+\epsilon | O(Q(t), P(t)) | q, t \rangle &= \int \frac{d^M p}{(2\pi)^M} \langle q', t | e^{-i\epsilon H(Q(t),P(t))} | p, t \rangle \langle p, t | O(Q(t), P(t)) | q, t \rangle \\ &\approx \int \frac{d^M p}{(2\pi)^M} e^{i\epsilon \left(\sum_a p_a \frac{q'^a - q^a}{\epsilon} - H(q',p) \right)} O(q(t), p(t)). \end{aligned} \tag{5.11}$$

Thus we see that the only effect of time-ordered operator insertions is to insert these operators evaluated as functions of q and p into the path integral:

$$\begin{aligned} \langle q_f, t_f | T O_1(Q(\bar{t}_1), P(\bar{t}_1)) \dots O_M(Q(\bar{t}_1), P(\bar{t}_M)) | q_i, t_i \rangle = \int \mathcal{D}q |_{q_i}^{q_f} \int \mathcal{D}p \, O_1(q(\bar{t}_1), p(\bar{t}_1)) \dots O_M(q(\bar{t}_M), p(\bar{t}_M)) \\ \times \exp \left[i \int_{t_i}^{t_f} dt \left(\sum_a p_a(t) \dot{q}^a(t) - H(q(t), p(t)) \right) \right] \end{aligned} \tag{5.12}$$

Here we have used the time-ordering symbol T on the left-hand side to ensure that operators are time-ordered, so we no longer need to impose (5.10).

5.2 Ground state preparation and the $i\epsilon$ prescription

In quantum field theory the canonical coordinates $Q^a(t)$ become Heisenberg fields, and it isn't so useful to consider expectation values in eigenstates of these fields. What we really want are *vacuum* expectation values, so for the path integral formulation to be useful in quantum field theory we need a path integral way to prepare the ground state. Fortunately there is a fairly simple way of doing this. Let's first recall that the eigenstates $|q, t\rangle$ of Q obey

$$|q, t\rangle = e^{iHt} |q, 0\rangle. \tag{5.13}$$

As with any state in the Hilbert space, we can expand $|q, 0\rangle$ in terms of energy eigenstates:

$$|q, 0\rangle = \sum_i C_i(q) |i\rangle, \tag{5.14}$$

with $H|i\rangle = E_i |i\rangle$. The idea is then to give t a small imaginary part via

$$t = e^{-i\epsilon} \tau, \tag{5.15}$$

with τ real and $0 < \epsilon \ll 1$, and then take τ to be large and negative. Working to leading order in ϵ we then have

$$|q, e^{-i\epsilon}\tau\rangle \approx |q, (1 - i\epsilon)\tau\rangle = e^{(i+\epsilon)\tau H} |q, 0\rangle = \sum_i C_i(q) e^{(i+\epsilon)E_i \tau} |i\rangle, \tag{5.16}$$

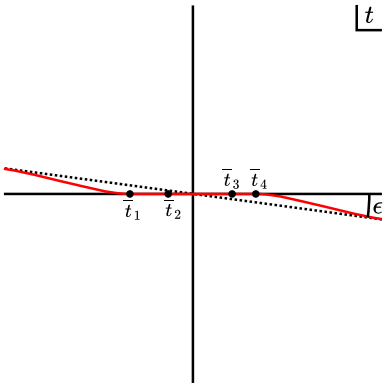


Figure 9: The $i\epsilon$ prescription for computing correlation functions in quantum field theory. Here $\bar{t}_1, \bar{t}_2, \ldots$ are the locations of the operators and the time contour is shown in red. In practice it simplifies calculations if we also analytically continue the operator times as $\bar{t}_m = e^{-i\epsilon}\bar{\tau}_m$, as then we can straighten the contour to the dashed one.

so if we take τ to $-\infty$ this gives us a state which is proportional to the ground state (which we renormalize to have zero energy):

$$|q_i, -(1-i\epsilon)\infty\rangle = C_0(q)|\Omega\rangle. \tag{5.17}$$

Therefore we can write a (Hamiltonian) path integral expression for the ground state wave function:

$$\langle q_f, 0|\Omega\rangle = \frac{1}{C_0(q_i)} \int \mathcal{D}q|_{q_i}^{q_f} \int \mathcal{D}p \, \exp\left[i \int_{-(1-i\epsilon)\infty}^{0} dt \left(\sum_a p_a(t)\dot{q}^a(t) - H(q(t), p(t)) \right) \right]. \tag{5.18}$$

We can also use (5.17) to give a path integral expression for the time-ordered correlation functions:

$$\begin{aligned} \langle \Omega | T O_1\left(Q(\bar{t}_1), P(\bar{t}_1)\right) \ldots O_M\left(Q(\bar{t}_M), P(\bar{t}_M)\right) |\Omega\rangle = & \frac{1}{|C_0(0)|^2} \int \mathcal{D}q|_0^0 \int \mathcal{D}p \, O_1(q(\bar{t}_1), p(\bar{t}_1)) \ldots O_M(q(\bar{t}_M), p(\bar{t}_M)) \\ & \times \exp\left[i \int_{-(1-i\epsilon)\infty}^{(1-i\epsilon)\infty} dt \left(\sum_a p_a(t)\dot{q}^a(t) - H(q(t), p(t)) \right) \right], \end{aligned} \tag{5.19}$$

where for convenience we have arbitrarily taken $q_i^a = q_f^a = 0$. The contour for the t integral is shown in figure 9. This contour prescription is the path integral version of the $i\epsilon$ prescription, and we will soon see that it gives rise to the same $i\epsilon$ prescription in the Feynman propagator that we found from the canonical approach in section 3. You may worry that this formula still requires us to know $|C_0(0)|$, but by removing the operator insertions we can also use it to give us a path integral formula for this,

$$|C_0(0)|^2 = \int \mathcal{D}q|_0^0 \int \mathcal{D}p \, \exp\left[i \int_{-(1-i\epsilon)\infty}^{(1-i\epsilon)\infty} dt \left(\sum_a p_a(t)\dot{q}^a(t) - H(q(t), p(t)) \right) \right], \tag{5.20}$$

so the correlation function is really a ratio of two path integrals. This is convenient because ambiguities in the normalization of the path integral measure cancel between the numerator and denominator.

5.3 An aside on Gaussian integrals

To proceed further, we now need to remember (or learn) a few things about Gaussian integrals. You hopefully haven't made it this far in your education without knowing that

$$\int_{-\infty}^{\infty} dx e^{-\frac{x^2}{2}} = \sqrt{2\pi}, \tag{5.21}$$

but just in case the proof is to look at the square of this integral and change to polar coordinates:

$$\left(\int_{-\infty}^{\infty} dx e^{-\frac{x^2}{2}}\right)^2 = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, e^{-\frac{x^2+y^2}{2}} = 2\pi \int_0^{\infty} dr\, r e^{-\frac{r^2}{2}} = 2\pi \int_0^{\infty} dr \frac{d}{dr}\left(-e^{-\frac{r^2}{2}}\right) = 2\pi. \tag{5.22}$$

Once we have this basic result we can derive others, for example for any $A > 0$ and any complex B we have

$$\int_{-\infty}^{\infty} dx\, e^{-\frac{A}{2}x^2+Bx} = \int_{-\infty}^{\infty} dx\, e^{-\frac{A}{2}\left(x-\frac{B}{A}\right)^2+\frac{B^2}{2A}} = \frac{1}{\sqrt{A}} e^{\frac{B^2}{2A}} \int_{-\infty}^{\infty} dz e^{-z^2} = \sqrt{\frac{2\pi}{A}} e^{\frac{B^2}{2A}}. \tag{5.23}$$

By differentiating this expression with respect to B we can compute all the moments of the Gaussian distribution, for example

$$\sqrt{\frac{A}{2\pi}} \int_{-\infty}^{\infty} dx x^2 e^{-\frac{A}{2}x^2} = \frac{d^2}{dB^2} e^{\frac{B^2}{2A}}\Big|_{B=0} = \frac{1}{A}. \tag{5.24}$$

We can also consider multiple integrals: given a symmetric matrix A which we will at first assume to be real and positive, and a complex vector B, we have the integral

$$Z[A,B] := \int dx e^{-\frac{1}{2}x^T Ax + B^T x} \tag{5.25}$$

where x is a real vector. We can diagonalize A as $A = O^T D O$, where O is orthogonal and D is diagonal with positive elements $d_1, d_2, \ldots$. We can then change variables to $\widetilde{x} = Ox$, giving

$$\begin{aligned} Z[A,B] &= \int d\widetilde{x} e^{-\frac{1}{2}\widetilde{x}^T D\widetilde{x} + (OB)^T\widetilde{x}} \\ &= \prod_i \left(\int d\widetilde{x}_i e^{-\frac{1}{2}\widetilde{x}_i d_i \widetilde{x}_i + \sum_j O_{ij}B_j\widetilde{x}_i} \right) \\ &= \prod_i \left(\sqrt{\frac{2\pi}{d_i}} e^{\frac{(\sum_j O_{ij}B_j)^2}{2d_i}} \right) \\ &= \frac{1}{\sqrt{\mathrm{Det}\left(\frac{A}{2\pi}\right)}} e^{\frac{1}{2}B^T A^{-1} B}. \end{aligned} \tag{5.26}$$

There is an easy way to remember this result: *up to a determinant factor, we can evaluate a Gaussian integral by evaluating its integrand on the value of x for which its exponent is stationary.* Indeed the exponent in (5.25) has a stationary point at

$$x = A^{-1}B, \tag{5.27}$$

and we then have

$$-\frac{1}{2}x^T Ax + Bx = \frac{1}{2}B^T A^{-1} B. \tag{5.28}$$

We can also use this result to compute correlation functions:

$$\frac{\int dx\, x_{i_1} \ldots x_{i_n} e^{-\frac{1}{2}x^T Ax}}{\int dx\, e^{-\frac{1}{2}x^T Ax}} = \frac{\partial}{\partial B_{i_1}} \cdots \frac{\partial}{\partial B_{i_n}} e^{\frac{1}{2}B^T A^{-1} B}\Big|_{B=0}, \tag{5.29}$$

in particular the two-point function is given by

$$\frac{\int dx\, x_i x_j e^{-\frac{1}{2}x^T Ax}}{\int dx\, e^{-\frac{1}{2}x^T Ax}} = (A^{-1})_{ij}. \tag{5.30}$$

In quantum mechanics we are not only interested in the situation where A is real and positive. We can extend our result (5.26) to more general A by analytic continuation; a minimal condition for the convergence of the integral (5.25) is that A has positive real part, meaning that $A + A^\dagger$ is positive, and (5.26) will apply for any such matrix provided that we are careful to define the sign of the square root by analytic continuation from real positive A.

5.4 Lagrangian path integral in quantum mechanics

So far the path integrals we have discussed have independent integrals over the trajectories $q(t)$ and $p(t)$. These manifestly rely on the Hamiltonian formalism, and thus are not manifestly covariant in relativistic theories. To get covariant expressions we need to get rid of $p(t)$. The best way to do this is to **integrate it out**, meaning to simply evaluate the functional integral over $p(t)$. In many theories of physical interest, including in particular the standard model of particle physics and also general relativity, the Hamiltonian is a quadratic function of the canonical momenta. The functional integral over $p(t)$ is therefore a Gaussian integral, and we can thus evaluate it using the methods of the previous subsection. Indeed the stationarity condition is simply Hamilton's equation

$$\dot{q}^a = \frac{\partial H}{\partial p_a}, \tag{5.31}$$

so evaluating the Gaussian integral over $p(t)$ has precisely the effect of converting the exponent in the path integral into the Lagrangian! More explicitly, considering expectation values of operators that depend only on Q (and not P) we have

$$\langle q_f, t_f | T O_1(Q(\bar{t}_1)) \dots O_M(Q(\bar{t}_M)) | q_i, t_i \rangle = \int \frac{\mathcal{D}q|_{q_i}^{q_f}}{\sqrt{\mathrm{Det}\,(2\pi A[q])}} O_1(q(\bar{t}_1)) \dots O_M(q(\bar{t}_M))\, e^{i\int_{t_i}^{t_f} L(q(t),\dot{q}(t))}. \tag{5.32}$$

Here $A[q]$ is the "matrix" appearing in the term in the Hamiltonian which is quadratic in P, as in equation (5.25). In simple theories (such as the harmonic oscillator or the standard model of particle physics) A is independent of q, in which case the determinant factor is a field-independent constant and can be absorbed into a rescaling of the measure.[38] Equation (5.32) is called the **Lagrangian path integral**, and unlike the Hamiltonian path integral it manifestly has (up to possible regularization issues) all the symmetries of the classical Lagrangian L. Using the $i\epsilon$ prescription we can also give a Lagrangian path integral expression for time-ordered correlation functions:

$$\langle \Omega | T O_1(Q(\bar{t}_1)) \dots O_M(Q(\bar{t}_M)) | \Omega \rangle = \frac{\int \frac{\mathcal{D}q|_0^0}{\sqrt{\mathrm{Det}(2\pi A[q])}} O_1(q(\bar{t}_1)) \dots O_M(q(\bar{t}_M)) e^{i\int_{-\infty(1-i\epsilon)}^{\infty(1-i\epsilon)} dt L(q(t),\dot{q}(t))}}{\int \frac{\mathcal{D}q|_0^0}{\sqrt{\mathrm{Det}(2\pi A[q])}} e^{i\int_{-\infty(1-i\epsilon)}^{\infty(1-i\epsilon)} dt L(q(t),\dot{q}(t))}}. \tag{5.34}$$

This expression, together with its Euclidean continuation we will introduce soon, is the starting point for many (most?) standard calculations in quantum field theory.

The restriction to operator insertions that don't depend on P is not so serious, as we can differentiate both sides of equation (5.32) with respect to the operator times $\bar{t}_1, \bar{t}_2, \dots$ to get path integral expressions for correlation functions involving time derivatives of q. The restriction to Hamiltonians which are quadratic in P is more concerning. In general the best that can be said is that by integrating out p we will always get *some* local Lagrangian which has whatever symmetries the theory has, but it won't in general be the Legendre transform of the Hamiltonian we started with. On the other hand in quantum field theory we usually end up writing down the most general local Lagrangian that is consistent with the symmetries in question (see our discussion of effective field theories in section 12), and the new Lagrangian resulting from integrating out p will differ from the one resulting from the Legendre transformation only by shifts of the values of the parameters in this Lagrangian. By starting with the Lagrangian approach we therefore land on the same class of theories as we did starting from the Hamiltonian approach, but now with a more complicated relationship between the two approaches. These comments also apply to the somewhat arbitrary choices we made for the operator ordering of H and O: other choices would just differ by shifting the coefficients of

[38] An example of a theory which is not "simple" in this regard is the "non-linear σ-model", which is a theory of multiple scalar fields ϕ^n with Lagrangian density

$$\mathcal{L} = -\frac{1}{2} g_{mn}(\phi) \partial_\mu \phi^m \partial^\mu \phi^n - V(\phi). \tag{5.33}$$

Here $g_{mn}(\phi)$ is a Euclidean metric on the target space of the fields $\phi^m(x)$. This theory shows up in the low-energy description of pions in nuclear physics.

the local terms appearing in H and O. In general shifts of this type are called **renormalizations**, and in defining path integrals we always give ourselves some leeway in how to renormalize both the Hamiltonian and the operators appearing in expectation values.

5.5 Path integral calculation of the harmonic oscillator ground state

As a first illustration of using a path integral for a practical calculation we can compute the ground state wave function of the simple harmonic oscillator. Up to normalization this is given by

$$\langle x_f|\Omega\rangle \propto \int \mathcal{D}x|_0^{x_f} e^{i\int_{-(1-i\epsilon)\infty}^0 dt\frac{1}{2}\left(\dot{x}^2-m^2x^2\right)}. \tag{5.35}$$

This integral is Gaussian, so we can evaluate it using our formula (5.26): we are supposed to find the saddle point of the exponent and then evaluate the integrand on it. The saddle point equation is

$$\frac{d^2x}{dt^2} = -m^2x, \tag{5.36}$$

but it is more convenient to rewrite this in terms of $\tau = (1+i\epsilon)t$:

$$\frac{d^2x}{d\tau^2} = -m^2(1-2i\epsilon)x. \tag{5.37}$$

We are interested in finding the saddle point which vanishes at $\tau = -\infty$ and is equal to x_f at $\tau = 0$; this is

$$x(\tau) = x_f e^{im(1-i\epsilon)\tau}. \tag{5.38}$$

Evaluating the exponent of the integrand we have

$$i(1-i\epsilon)\int_{-\infty}^0 d\tau\frac{x_f^2}{2}\left(-m^2(1-2i\epsilon)-m^2\right)e^{2im(1-i\epsilon)\tau} = -\frac{ix_f^2}{2}\frac{2m^2(1-i\epsilon)}{2im(1-i\epsilon)} = -\frac{mx_f^2}{2}, \tag{5.39}$$

which is indeed the exponent for the correct harmonic oscillator ground state.

5.6 Path integral calculation of the Feynman propagator in field theory

We can evaluate the Feynman propagator for a free scalar field along similar lines. Taking into account the $i\epsilon$ prescription, integrating by parts we can write the exponent as

$$\frac{i}{2}\int_{-\infty}^{\infty} d\tau\int d^{d-1}x\phi\left(-(1+2i\epsilon)\partial_\tau^2+\nabla^2-m^2\right)\phi. \tag{5.40}$$

Thuss the "matrix" A for this Gaussian integral is

$$A(x_1,x_2) = -i\left(-(1+2i\epsilon)\partial_\tau^2+\nabla^2-m^2\right), \tag{5.41}$$

which we can easily invert in momentum space:

$$\begin{aligned} A^{-1}(x_1,x_2) &= \int\frac{d^dp}{(2\pi)^d}\frac{ie^{ip(x_2-x_1)}}{(1+2i\epsilon)(p^0)^2-|p|^2-m^2} \\ &= \int\frac{d^dp}{(2\pi)^d}\frac{-ie^{ip(x_2-x_1)}}{p^2+m^2-i\epsilon}, \end{aligned} \tag{5.42}$$

where in the second line we rescaled ϵ by the positive quantity $(p^0)^2$. By equation (5.30) this should be equal to the Feynman propagator $G_F(x_1,x_2)$, and indeed it matches the expression we found in section 3 using the canonical formalism.

We can also use this approach to independently derive a position space expression for the Feynman propagator. Indeed from equation (5.41) the Feynman propagator should obey

$$\left(-(1+2i\epsilon)\partial_{\tau_2}^2+\nabla_2^2-m^2\right)G_F(x_1,x_2)=i\delta(x_1-x_2). \tag{5.43}$$

By Lorentz invariance G_F should really only be a function of

$$s=\sqrt{(\vec{x}_2-\vec{x}_1)^2-(t_2-t_1)^2}=\sqrt{(\vec{x}_2-\vec{x}_1)^2-(1-2i\epsilon)(\tau_2-\tau_1)^2}, \tag{5.44}$$

and substituting this into (5.43) we find that away from $s=0$ the Feynman propagator obeys (up to terms of order ϵ^2)

$$G_F''(s)+\frac{d-1}{s}G_F'(s)-m^2G_F(s)=0, \tag{5.45}$$

which is a standard ordinary differential equation whose solutions can be expressed in terms of Bessel functions (as you can easily check in mathematica). The solution which goes to zero at large positive s is

$$G_F\propto s^{-\frac{d-2}{2}}K_{\frac{d-2}{2}}(ms), \tag{5.46}$$

and we can fix the coefficient of proportionality either by requiring that this obeys

$$(\nabla_2^2-m^2)G_F(x_2,x_1)=i\delta^d(x_2-x_1) \tag{5.47}$$

or else by matching to the integral (5.42) in the massless limit that $ms\ll 1$ where the integral is easier to compute. This is the same position-space two-point function we quoted in section 3, except that now the $i\epsilon$ prescription we are using gives us the Feynman propagator instead of the two-point function.

5.7 Euclidean path integrals

We've seen that it is convenient to analytically continue t slightly into the complex plane via the $i\epsilon$ prescription $t=e^{-i\epsilon}\tau$. In fact it is an even better idea to continue all the way to $\epsilon=\pi/2$, i.e. to

$$t=-i\tau. \tag{5.48}$$

The path integral on this contour is called the **Euclidean path integral**, and for many questions the Euclidean path integral is the best way to think about it. Given its importance, it is worth repeating the deriviation we gave in Lorentzian signature directly in Euclidean signature. The idea is to define Euclidean Heisenberg operators by[39]

$$O(\tau)=e^{\tau H}O(0)e^{-\tau H}, \tag{5.49}$$

with eigenstates $|q,-i\tau\rangle=e^{H\tau}|q,0\rangle$. Proceeding as in the Lorentzian case, we can note that

$$\begin{aligned}\langle q',-i(\tau+\epsilon)|O(Q(\tau),P(\tau)|q,-i\tau\rangle&=\int\frac{d^Mp}{(2\pi)^M}\langle q',-i\tau|e^{-\epsilon H(Q(\tau),P(\tau))}|p,-i\tau\rangle\langle p,-i\tau|O(Q(\tau),P(\tau))|q,-i\tau\rangle\\&\approx\int\frac{d^Mp}{(2\pi)^M}e^{\epsilon\left(i\sum_a p_a\frac{q'^a-q^a}{\epsilon}-H(q',p)\right)}O(q(\tau),p(\tau))\end{aligned} \tag{5.50}$$

and therefore by inserting complete sets of states we have

$$\begin{aligned}\langle q_f,-i\tau_f|TO_1(Q(\overline{\tau}_1),P(\overline{\tau}_1))\dots O_M(Q(\overline{\tau}_M),P(\overline{\tau}_M))|q_i,-i\tau_i\rangle=\int\mathcal{D}q|_{q_i}^{q_f}\int\mathcal{D}p\,O_1(q(\overline{\tau}_1),p(\overline{\tau}_1))\dots O_M(q(\overline{\tau}_M),p(\overline{\tau}_M))\\\times\exp\left[\int_{\tau_i}^{\tau_f}d\tau\left(i\sum_a p_a(\tau)\dot{q}^a(\tau)-H(q(\tau),p(\tau))\right)\right].\end{aligned} \tag{5.51}$$

[39]These operators are somewhat delicate mathematically due to the presence of $e^{\tau H}$, which has a very limited domain. It is always ok to use them in time-ordered vacuum correlators however, which in the end is the only place we will use them.

Taking $\tau_i \to -\infty$ and $\tau_f \to \infty$ now automatically projects onto the ground state, so no analytic continuation is needed to convert this into a vacuum expectation value:

$$\begin{aligned}\langle\Omega|TO_1(Q(\overline{\tau}_1),P(\overline{\tau}_1))\dots O_M(Q(\overline{\tau}_M),P(\overline{\tau}_M))|\Omega\rangle =& \frac{1}{|C_0(0)|^2}\int \mathcal{D}q|_0^0 \int \mathcal{D}p\, O_1(q(\overline{\tau}_1),p(\overline{\tau}_1))\dots O_M(q(\overline{\tau}_M),p(\overline{\tau}_M)) \\ & \times \exp\left[\int_{-\infty}^{\infty} d\tau \left(i\sum_a p_a(\tau)\dot{q}^a(\tau) - H(q(\tau),p(\tau))\right)\right].\end{aligned} \tag{5.52}$$

Converting this into a Lagrangian path integral (with the same caveats as before), we end up with

$$\langle\Omega|TO_1(Q(\overline{\tau}_1))\dots O_M(Q(\overline{\tau}_M))|\Omega\rangle = \frac{\int \frac{\mathcal{D}q|_0^0}{\sqrt{\mathrm{Det}(2\pi A)}} O_1(q(\overline{\tau}_1))\dots O_M(q(\overline{\tau}_M)) e^{-\int_{-\infty}^{\infty} d\tau L_E(q,\dot{q})}}{\int \frac{\mathcal{D}q|_0^0}{\sqrt{\mathrm{Det}(2\pi A)}} e^{-\int_{-\infty}^{\infty} d\tau L_E(q,\dot{q})}}, \tag{5.53}$$

where L_E is the **Euclidean Lagrangian** defined in terms of the Lorentzian Lagrangian by

$$L_E\left(q,\frac{dq}{d\tau}\right) := -L\left(q,i\frac{dq}{d\tau}\right). \tag{5.54}$$

For example for the simple harmonic oscillator the Euclidean Lagrangian is

$$L_E = \frac{1}{2}\left(\dot{q}^2 + m^2q^2\right), \tag{5.55}$$

while for a free scalar field the Euclidean Lagrangian is the spatial integral of the Euclidean Lagrangian density

$$\mathcal{L}_E = \frac{1}{2}\left(\dot{\phi}^2 + \nabla\phi\cdot\nabla\phi + m^2\phi^2\right). \tag{5.56}$$

There are a few essential points to make about the Euclidean path integral:

- Mathematically it is much better behaved than the Lorentzian path integral. The Euclidean action $S_E = \int_{-\infty}^{\infty} d\tau L_E$ is often real and bounded from from below, as you can see from the harmonic oscillator and the free scalar, so the integrand e^{-S_E} exponentially suppresses field configurations which aren't near $\phi = 0$. This makes it possible to give it a mathematically rigorous formulation (at least in the case of a finite number of degrees of freedom), look up "Wiener measure" if you want to learn about it.

- In situations where S_E is real and bounded from below we can interpret the Euclidean path integral (5.53) as computing expectation values in a *classical* probability distribution. Many famous classical statistical systems arise in this way, for example the Euclidean path integral for a free particle is the classical theory of Brownian motion and the Euclidean path integral for a free scalar field with $d = 2$ is the classical theory of random surfaces. The critical point in the phase diagram of water is also described by a (interacting) Euclidean scalar field theory, as are the fluctuations of magnets at the Curie temperature. Euclidean path integrals also arise in quantitative finance: the prices of options as a function of time are fluctuating variables which can be characterized by a Euclidean path integral.

- In situations where the Euclidean path integral has a probabilistic interpretation it is amenable to explicit numerical evaluation. The standard approach to this is called the Monte Carlo method, which samples from the probability distribution and then assumes that the expectation value is dominated by its value on a typical instance. This is a very powerful method for evaluating high-dimensional integrals. For example in QCD, the theory of the strong nuclear force, my colleagues here in the Center for Theoretical Physics use this method to compute the masses of hadrons such as the proton

and neutron to quite good accuracy. The computational resources involved are somewhat terrifying, for example in a recent calculation my colleague Will Detmold used the fastest publicly-available supercomputer in the world, Frontier at Oak Ridge National Laboratory, to evaluate the Euclidean path integral of QCD on a Euclidean spacetime lattice with $72 \times 72 \times 72 \times 192$ sites, consuming of order 10^{11} Joules of energy in the process.

- In relativistic theories something particularly nice happens: if we have $SO^+(d-1,1)$ symmetry in Lorentzian signature then we have $SO(d)$ rotational symmetry in Euclidean signature. This Euclidean rotation invariance is at the heart of many famous results in quantum field theory, as we will see in the next section.

- We can also use the Euclidean path integral to compute Lorentzian correlation functions: to compute a time-ordered correlator of operators $O_1(\bar{t}_1), O_2(\bar{t}_2), \ldots$, we simply compute their Euclidean correlation function as a function of $\overline{\tau}_1, \overline{\tau}_2, \ldots$ and then analytically continue the time of each operator as $\overline{\tau} = i(1 - i\epsilon)\bar{t}$. This analytic continuation is called **Wick rotation**; essentially we are approaching the $i\epsilon$ contour shown in figure 9 from the Euclidean contour instead of the Lorentzian one. You will check in the homework that this continuation again gives the correct $i\epsilon$ prescription for the Feynman propagator.

- Euclidean path integrals arise naturally in the context of quantum statistical mechanics. For example to compute the partition function $Z(\beta) = \mathrm{Tr}\left(e^{-\beta H}\right)$, we evaluate the Euclidean path integral with periodic boundary conditions in time, with periodicity β.

5.8 Homework

1. Rewrite the operator $PQPQ$ as a sum of operators with all P to the right and all Q to the left.

2. Use the path integral to find the propagator $\langle q', t'|q, t\rangle$ of a free quantum particle moving on a line with Hamiltonian $H = \frac{P^2}{2m}$. Hint: use the discretized version of the Lagrangian path integral.

3. Use the path integral to find the propagator for the simple harmonic oscillator, with Hamiltonian $H = \frac{P^2}{2m} + \frac{k}{2}Q^2$. Hint: you should expand the function $q(t)$ you are integrating over as a classical solution q_{cl} plus a fluctuating piece δq, and then expand δq in Fourier modes and integrate over the coefficients of these modes. I recommend first doing the calculation neglecting any prefactors which are independent of k: you can find the k-independent prefactor at the end by comparing to your answer for the previous problem in the limit $k \to 0$.

4. Use the Lorentzian path integral with an $i\epsilon$ prescription to find the Feynman propagator of a free massive complex scalar field (remember that this is the time-ordered two-point function of Φ and $\Phi^\dagger$).

5. Use the Euclidean path integral followed by a Wick rotation to compute the Feynman propagator of a real free scalar field with mass m. Hint: you should find that the Euclidean Feynman propagator is a Greens function for the Euclidean Klein-Gordon operator, obeying $(\nabla_x^2 - m^2)G_F(x, y) = -\delta^d(x - y)$. It is ok to leave your expression for it in terms of a spacetime momentum integral, but you should make sure that after Wick rotation you get the right $i\epsilon$ prescription for the Feynman propagator in Lorentzian signature.

6 $\mathcal{CRT}$, spin-statistics, and all that

We've now developed two powerful formalisms for thinking about quantum field theory: the operator approach based on algebras acting on Hilbert spaces and the path integral approach, both in Lorentzian and Euclidean signature. In this section we will put the pieces together to prove some of the famous results in relativistic quantum field theory: the $\mathcal{CRT}$ theorem, the relation between spin and statistics, and the thermal nature of vacuum entanglement (the Unruh effect). All of these results are true non-perturbatively in any relativistic quantum field theory, as the arguments will hopefully make clear. The title of this section is shamelessly adapted from a famous book by Streater and Wightman, which discusses the first two of these from a rigorous (but somewhat out-dated) approach.

6.1 The $\mathcal{CRT}$ theorem

Let's first recall our Euclidean path integral expression for correlation functions in quantum field theory:

$$\langle \Omega | T O_1[\Phi] \dots O_M[\Phi] | \Omega \rangle = \frac{\int \mathcal{D}\phi \, O_1[\phi] \dots O_M[\phi] e^{-S_E[\phi]}}{\int \mathcal{D}\phi e^{-S_E[\phi]}} \tag{6.1}$$

Here I have switched from the particle notation we used in the last section to field notation, and also absorbed the determinant factor coming from integrating out the momenta into the measure $\mathcal{D}\phi$. In any relativistic quantum field theory this path integral is invariant under Euclidean rotation symmetry, in the sense that if F_Λ is a transformation of field space which implements a Euclidean rotation $\Lambda \in SO(d)$, i.e.

$$F_\Lambda \phi^a(x) = D_E(\Lambda)^a{}_b \phi^b(\Lambda^{-1} x) \tag{6.2}$$

on the dynamical fields, then the combination of the path integral measure and action are invariant:

$$\mathcal{D}(F_\Lambda \phi) e^{-S_E[F_\Lambda \phi]} = \mathcal{D}\phi e^{-S_E[\phi]}. \tag{6.3}$$

The invariance of the action is the classical statement of having a symmetry, while the invariance of the measure reflects the statement that the regularization of the theory implicit in the path integral does not destroy the symmetry (much later we will see examples of situations where this happens). Using this invariance we can derive a constraint on correlation functions:

$$\begin{aligned}
\langle \Omega | T O_1[\Phi] \dots O_M[\Phi] | \Omega \rangle &= \frac{\int \mathcal{D}\phi \, O_1[\phi] \dots O_M[\phi] e^{-S_E[\phi]}}{\int \mathcal{D}\phi e^{-S_E[\phi]}} \\
&= \frac{\int \mathcal{D}(F_\Lambda \phi) \, O_1[F_\Lambda \phi] \dots O_M[F_\Lambda \phi] e^{-S_E[F_\Lambda \phi]}}{\int \mathcal{D}\phi e^{-S_E[\phi]}} \\
&= \frac{\int \mathcal{D}\phi \, O_1[F_\Lambda \phi] \dots O_M[F_\Lambda \phi] e^{-S_E[\phi]}}{\int \mathcal{D}\phi e^{-S_E[\phi]}} \\
&= \langle \Omega | T O_1[F_\Lambda \Phi] \dots O_M[F_\Lambda \Phi] | \Omega \rangle.
\end{aligned} \tag{6.4}$$

In going from the first to the second line here we changed variables in the path integral, in going from the second to the third we used the symmetry condition (6.3), and in going from the third to the fourth we used (6.1).

To prove the $\mathcal{CRT}$ theorem we are interested in the Euclidean rotation $\Lambda = \mathcal{RT}$, which acts as[40]

$$\mathcal{RT} : (\tau, x^1, x_2, \dots, x^{d-1}) = (-\tau, -x^1, x^2, \dots, x^{d-1}) \tag{6.5}$$

[40] When d is even we can combine $\mathcal{RT}$ with spatial rotations to define an operation $\mathcal{PT}$ which simply acts as $\mathcal{PT} : x \mapsto -x$. This then leads to a symmetry called $\mathcal{CPT}$, which is a symmetry of any relativistic field theory when d is even. Historically the theorem discussed in this section has thus usually been called the $\mathcal{CPT}$ theorem, especially by particle physicists who only care about the case of $d = 4$, while the terminology $\mathcal{CRT}$ is of more recent origin. We have focused on $\mathcal{CRT}$ nonetheless because 1) it is the thing which works in any spacetime dimension and 2) it is what naturally arises from the proof of the theorem.

I emphasize that $\mathcal{RT}$ is indeed an element of $SO(d)$, it is a rotation by π in the plane of τ and x^1. This transformation reverses the direction of Euclidean time, so it also reverses the order of the operators in the Euclidean correlation function. To be more concrete, if O_1 lives at time τ_1, O_2 at time τ_2, etc, and for simplicity we assume that $\tau_1 \leq \tau_2 \leq \ldots$, then the Euclidean statement of this symmetry is that

$$\begin{aligned}\langle\Omega|O_M[\Phi]\ldots O_1[\Phi]|\Omega\rangle &= (-1)^{\frac{f(f-1)}{2}}\langle\Omega|O_1[F_{\mathcal{RT}}\Phi]\ldots O_M[F_{\mathcal{RT}}\Phi]|\Omega\rangle \\ &= (-1)^{f/2}\langle\Omega|O_1[F_{\mathcal{RT}}\Phi]\ldots O_M[F_{\mathcal{RT}}\Phi]|\Omega\rangle,\end{aligned} \tag{6.6}$$

where

$$f = f_{O_1} + \ldots + f_{O_M} \tag{6.7}$$

is the total number of fermionic operators appearing in $O_1 \ldots O_M$. This pesky minus sign arises from something we haven't discussed yet, which is that when you time-order fermionic operators the process is antisymmetric instead of symmetric. We'll see this in more detail when we discuss the fermionic path integral in a month or so, but for now the basic idea is that since fermionic fields anticommute instead of commute at spacelike separation it must be that the degrees of freedom which represent them in the path integral are also anticommuting. The second line of (6.6) follows from the first because correlation functions involving fermions vanish unless the total number of fermions is even, which is a consequence of the fact that the Lagrangian density is always bosonic (this is called **fermion parity** symmetry).

The $\mathcal{CRT}$ theorem is what we get when we analytically continue (6.6) to Lorentzian signature. We can formalize the analytic continuation be introducing a Wick rotation operation W, whose action on dynamical fields is defined to perform the analytic continuation $\tau = it$. On Euclidean scalar fields we have

$$W\Phi(t,\vec{x}) = \Phi(it,\vec{x}), \tag{6.8}$$

while for tensor fields each raised τ indices get a factor of i and each lowered τ indices get a factor of $-i$. So for example a vector field V^μ has

$$W\begin{pmatrix}V^0(t,\vec{x})\\ V^j(t,\vec{x})\end{pmatrix} = \begin{pmatrix}iV^0(it,\vec{x})\\ V^j(it,\vec{x})\end{pmatrix} \tag{6.9}$$

while a one-form field ω_μ has

$$W\begin{pmatrix}\omega_0(t,\vec{x})\\ \omega_j(t,\vec{x})\end{pmatrix} = \begin{pmatrix}-i\omega_0(it,\vec{x})\\ \omega_j(it,\vec{x})\end{pmatrix}. \tag{6.10}$$

These factors are necessary because we'd like to preserve e.g. the expressions $V = V^\mu\partial_\mu$ and $\omega = \omega_\mu dx^\mu$, so we should rotate V^0 in the same way as we rotate τ and ω_0 in the opposite way.[41] Analytic continuation of (6.6) thus gives

$$\langle\Omega|WO_M[\Phi]\ldots WO_1[\Phi]|\Omega\rangle = (-1)^{(f_{O_1}+\ldots f_{O_M})/2}\langle\Omega|WO_1[F_{\mathcal{RT}}\Phi]\ldots WO_M[F_{\mathcal{RT}}\Phi]|\Omega\rangle. \tag{6.11}$$

In order to give this a symmetry interpretation in Lorentzian signature, we can first observe that the symmetry must be antiunitary since it reverses time. To see what the antiunitary is, we need to first recall that for any antiunitary operator Θ that preserves the ground state we have the constraint

$$\langle\Omega|O_1\ldots O_M|\Omega\rangle = \langle\Omega|O_M'^\dagger\ldots O_1'^\dagger|\Omega\rangle \tag{6.12}$$

Thus we can interpret (6.11) as indicating that our Lorentzian theory has an antiunitary symmetry $\Theta_{\mathcal{CRT}}$ whose action on the dynamical fields is[42]

$$\Theta^\dagger_{\mathcal{CRT}}W\Phi^a(x)\Theta = i^{f_a}(WF_{\mathcal{RT}}\Phi^a(x))^\dagger, \tag{6.13}$$

[41] For future reference I'll mention that spinor fields do not pick up any phase factors under Wick rotation, but the γ-matrix γ^0 transforms as $W\gamma^\tau = i\gamma^t$. Don't worry about this if you don't yet know what it means.

[42] Note that this definition does not require or use independent definitions of $\mathcal{C}$, $\mathcal{R}$, and $\mathcal{T}$. In general these are not symmetries, and even when they are there is some freedom in how they are defined. The name $\mathcal{CRT}$ is thus in some sense a historical anachronism, the whole is better-defined than its parts.

where $f_a = 1$ if Φ^a is fermionic and $f_a = 0$ if Φ^a is bosonic. In particular note that we need to take the complex conjugate of the analytic continuation of the Euclidean rotation to match (6.12), this is the origin of the "$\mathcal{C}$" in $\mathcal{CRT}$. For example the action of $\mathcal{CRT}$ on a (Lorentzian) complex scalar Φ or complex vector V^μ is[43]

$$\begin{aligned}\Theta_{\mathcal{CRT}}^\dagger \Phi(x) \Theta_{\mathcal{CRT}} &= \Phi(\mathcal{RT}x)^\dagger \\ \Theta_{\mathcal{CRT}}^\dagger V^\mu(x) \Theta_{\mathcal{CRT}} &= (\mathcal{RT})^\mu{}_\nu V^\nu(\mathcal{RT}x)^\dagger.\end{aligned} \tag{6.14}$$

Once we have understood spinor fields we will also see that a Dirac spinor transforms as

$$\Theta_{\mathcal{CRT}}^\dagger \Psi(x) \Theta_{\mathcal{CRT}} = \gamma^{0*}\gamma^{1*}\Psi^*(\mathcal{RT}x). \tag{6.15}$$

In general we can write the $\mathcal{CRT}$ transformation in Lorentzian signature as

$$\Theta_{\mathcal{CRT}}^\dagger \Phi^a(x) \Theta_{\mathcal{CRT}} = i^{f_a} \left(D_E(\mathcal{RT})^a{}_b \Phi^b(\mathcal{RT}x)\right)^\dagger, \tag{6.16}$$

since any factors of i and $-i$ from the Wick rotation of any τ indices cancel between the two sides. In the homework you will show that this equation together with the spin-statistics theorem imply that

$$\Theta_{\mathcal{CRT}}^2 = 1, \tag{6.17}$$

in any quantum field theory.[44]

The $\mathcal{CRT}$ theorem is quite remarkable from the point of view of the topology Lorentz group. In Lorentzian signature $\mathcal{RT}$ lives in a component of $O(d-1,1)$ which is disconnected from the identity component $SO^+(d-1,1)$. If we just assume a relativistic theory has $SO^+(d-1,1)$ symmetry, there is no particular reason why we should expect any version of $\mathcal{RT}$ to be a symmetry. In Euclidean signature however $\mathcal{RT}$ is in the identity component $SO(d)$ of $O(d)$, and thus must be a symmetry. So far there does not seem to be any nice proof of the $\mathcal{CRT}$ theorem that doesn't involve analytic continuation away from Lorentzian signature. The only exception is a brute-force argument, given e.g. in Weinberg, that one simply can't make a polynomial Lagrangian out of tensor and spinor fields that isn't $\mathcal{CRT}$ invariant - the proof is just to check this for all possible terms. Any experimental observation of $\mathcal{CRT}$-violation would be a very big deal, as it would mean that we have to give up either locality or special relativity. And at least to the extent that quantum mechanics + special relativity implies locality, we'd really need to give up either on quantum mechanics or relativity![45]

6.2 Spin and statistics

In non-relativistic quantum mechanics we learn that each type of particle has a spin s which can take integer or half-integer values. We also learn that each type of particle should be a boson or a fermion, meaning that if we exchange two of them the wave function should be symmetric or antisymmetric. A priori there does not seem to be any reason why these two should be related, and indeed in non-relativistic quantum mechanics it is easy to write down theories of particles with arbitrary spin and statistics. On the other hand in relativistic quantum field theory there is a very simple rule:[46]

[43] Here we have slightly abused notation to use the same symbol $\mathcal{RT}$ for the Lorentzian map $\mathcal{RT} : (t, x^1, x^2, \ldots, x^{d-1}) \to (-t, -x^1, x^2, \ldots, x^{d-1})$.

[44] In theories with extra global symmetries people sometimes combine $\mathcal{CRT}$ with those symmetries to get something that doesn't square to one. The $\mathcal{CRT}$ we have constructed here however is the only unbreakable one, up to the possibility of multiplying it by the fermion parity operator $(-1)^F$ which acts as one on all bosonic states and minus one on all fermionic states.

[45] In quantum gravity there are good reasons to think that we can have quantum mechanics and (general) relativity without having locality, but as far as we can tell $\mathcal{CRT}$ continues to be a good symmetry even in quantum gravity. See my recent paper with Numasawa for more on this.

[46] The fully non-perturbative proof of the spin-statistics theorem we give here is due to Schwinger. In most quantum field theory books the theorem is proven in a more banal way that applies only to free fields. Essentially one tries to construct free fields for particles of various spin, and then finds that it only works if the fields commute for integer spin and anticommute for half-integer spin.

- **Spin-statistics relation:** Particles with integer spin are bosons, while particles with half-integer spin are fermions.

The idea behind this rule is quite easy to understand, but we first need to discuss the subtle fact (which hopefully you have seen before) that a rotation by 2π acts on objects of half-integer spin as -1. For example in the context of a spin $1/2$ particle in three spatial dimensions a rotation by θ about the z axis is implemented on the Hilbert space by

$$U(\theta) = e^{-i\theta\sigma_z/2}, \tag{6.18}$$

so $U(2\pi) = e^{-i\pi\sigma_z} = -1$. Mathematically we can express this by saying that the action of rotations on a spin $1/2$ particle does not give a genuine representation of the rotation group $SO(3)$ in the sense of a set of unitary operators $U(g)$ such that $U(g_1)U(g_2) = U(g_1g_2)$, since a rotation by 2π is equal to nothing in the rotation group but apparently it isn't equal to nothing acting on a spin $1/2$ particle. We will discuss this in more detail later in the semester, but the right way to understand this is that in a relativistic theory with half-integer spin particles the spacetime symmetry group isn't really $SO^+(d-1,1)$, but instead what is called its **double cover** $Spin^+(d-1,1)$. Locally $Spin^+(d-1,1)$ looks just like $SO^+(d-1,1)$, but globally it is different in that each element of $SO^+(d-1,1)$ corresponds to *two* elements of $Spin^+(d-1,1)$ which differ by a rotation by 2π. The rotation part of $Spin^+(3,1)$, which is the double cover of $SO(3)$, is precisely given by the set of matrices of the form $e^{-i\vec{\theta}\cdot\vec{\sigma}/2}$, which is nothing but the matrix group $SU(2)$. We will see how to extend this to a double cover of the full Lorentz group later in the semester when we discuss spinors.

Turning now to spin and statistics, the basic ingredient we will need is to understand in more detail how the Euclidean rotation matrix $D_E(\mathcal{RT})$ acts on the fields Φ^a and their complex conjugates. This is a bit tricky, so hold on tight! Let's first recall that in *Lorentzian* signature we have

$$U(\Lambda)^\dagger \Phi^a(x) U(\Lambda) = D(\Lambda)^a{}_b \Phi^b(\Lambda^{-1}x), \tag{6.19}$$

and thus

$$U(\Lambda)^\dagger \Phi^a(x)^\dagger U(\Lambda) = D^*(\Lambda)^a{}_b \Phi^b(\Lambda^{-1}x)^\dagger. \tag{6.20}$$

In particular when Λ is a boost of rapidity η in the x^1 direction we have

$$D(\Lambda) = e^{i\eta\mathcal{J}^{10}}, \tag{6.21}$$

where the matrix $\mathcal{J}^{10}$ is the boost generator in the representation D of the Lorentz group. To turn a boost into a Euclidean rotation, we want analytically continue $t = -i\tau$ and $\eta = -i\theta$ such that

$$\begin{aligned} t' &= \cosh(\eta)t + \sinh(\eta)x \\ x' &= \cosh(\eta)x + \sinh(\eta)t \end{aligned} \tag{6.22}$$

become

$$\begin{aligned} \tau' &= \cos(\theta)\tau + \sin(\theta)x \\ x' &= \cos(\theta)x - \sin(\theta)\tau. \end{aligned} \tag{6.23}$$

We therefore have

$$D_E(\Lambda) = e^{\theta\mathcal{J}^{10}} \tag{6.24}$$

for a Euclidean rotation by θ in the τ, x^1 plane. In Euclidean signature the rotation group $SO(d)$ is a compact group, and the finite-dimensional representations of such groups are always unitary. We therefore see that $\mathcal{J}^{10}$ must be anti-hermitian. $D^*(\Lambda)$ therefore analytically continues to

$$e^{-i\eta(\mathcal{J}^{10})^*} = e^{i\eta(\mathcal{J}^{10})^T} = e^{\theta(\mathcal{J}^{10})^T} = D_E(\Lambda)^T. \tag{6.25}$$

In particular this applies to $\mathcal{RT}$, which is just a Euclidean rotation by π.

We next need to understand how the hermitian conjugate of fields works in Euclidean signature. In Lorentzian signature we have the convenient fact that we can take the hermitian conjugate before or after time evolution and end up with the same thing:

$$O(t)^\dagger = (e^{iHt}Oe^{-iHt})^\dagger = e^{iHt}O^\dagger e^{-iHt} = O^\dagger(t). \tag{6.26}$$

In Euclidean signature we aren't so lucky, the hermitian conjugate now gives

$$O(\tau)^\dagger = (e^{H\tau}Oe^{-H\tau})^\dagger = e^{-H\tau}O^\dagger e^{H\tau} \neq e^{H\tau}O^\dagger e^{-H\tau}. \tag{6.27}$$

To deal with this it is conventional to define a **Euclidean adjoint**

$$O^*(\tau) = e^{H\tau}O^\dagger e^{-H\tau} = O(-\tau)^\dagger, \tag{6.28}$$

as this is the quantity which analytically continues to $O^\dagger(t)$ in Lorentzian signature. Therefore from the previous paragraph, in Euclidean signature we have the somewhat counter-intuitive symmetry transformations

$$\begin{aligned}\Phi'(x) &= D_E(\mathcal{RT})\Phi(x)\\ \Phi^{*\prime}(x) &= D_E(\mathcal{RT})^T\Phi^*(\mathcal{RT}x).\end{aligned} \tag{6.29}$$

To proceed further, we now change our basis of fields $\Phi^a(x)$ to diagonalize $D_E(\mathcal{RT})$. Recall that this is a unitary matrix, and since all fields have integer or half-integer spin it must obey $D_E(\mathcal{RT})^4 = 1$. Its eigenvalues are therefore ± 1 on fields of integer spin and $\pm i$ on fields of half-integer spin. We may then observe that

$$\begin{aligned}\langle\Omega|\Phi^*(\tau,\vec{0})\Phi(-\tau,\vec{0})|\Omega\rangle &= (-1)^{2j_\phi}\langle\Omega|T\Phi^*(-\tau,\vec{0})\Phi(\tau,\vec{0})|\Omega\rangle\\ &= (-1)^{2j_\phi+f_\phi}\langle\Omega|\Phi(\tau,\vec{0})\Phi^*(-\tau,\vec{0})|\Omega\rangle,\end{aligned} \tag{6.30}$$

where in the first line j_ϕ is the spin of Φ and we have used our Euclidean rotation rule (6.4) and also that if Φ has integer spin both rotations contribute ± 1 while if Φ has half-integer spin then they both contribute $\pm i$. Note that (6.29) here is crucial in getting the factor of $(-1)^{2j_\phi}$, as it ensures Φ and Φ^* contribute with the same sign in front of i in the fermionic case. The second line then follows from the antisymmetry of the time-ordered product for fermions, as explained below (6.6).

Finally we can complete the proof by showing that the correlation functions on both sides of (6.30) are strictly positive: this implies the theorem because then we need

$$(-1)^{2j_\phi+f_\phi} = 1, \tag{6.31}$$

which means that when j_ϕ is an integer we must have $f_\phi = 0$ while when j_ϕ is a half-integer we must have $f_\phi = 1$. It is easy to show that they are positive semidefinite, as they are the squared norms of states:

$$\begin{aligned}\langle\Omega|\Phi^*(\tau,\vec{0})\Phi(-\tau,\vec{0})|\Omega\rangle &= ||\Phi(-\tau,\vec{0})|\Omega\rangle||^2 \geq 0\\ \langle\Omega|\Phi(\tau,\vec{0})\Phi^*(-\tau,\vec{0})|\Omega\rangle &= ||\Phi^*(-\tau,\vec{0})|\Omega\rangle||^2 \geq 0,\end{aligned} \tag{6.32}$$

where here we have used (6.28). We will show later in the section that these norms cannot vanish, so provided that the theorem is proved![47]

It is instructive to consider how a naive version of this argument which doesn't use special relativity can fail. The idea of the naive argument is to do the same manipulation using a *spatial* rotation by π instead of $\mathcal{RT}$. We can derive the relation (6.30) just as before (except with the fields now being at $\pm x\hat{x}$), but the failure mode is that we can no longer show that the correlators aren't zero! Indeed in non-relativistic field theory you can have a field Φ that only has an annihilation part and such a field can indeed annihilate the vacuum.

[47] More precisely what we showed is that Φ and Φ^* commute/anticommute at spacelike separation if they have integer/half-integer spin. In the homework you will show that this implies the same for Φ with Φ and Φ^* with Φ^*.

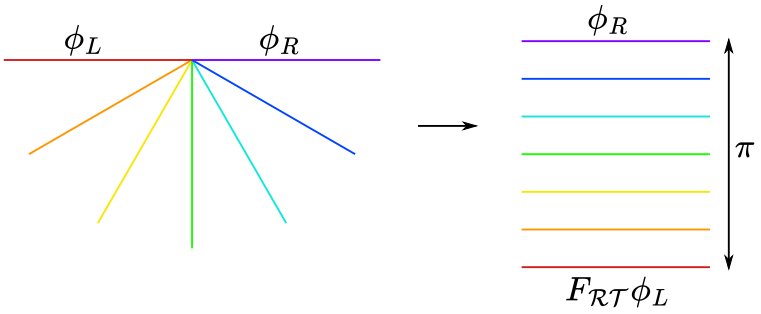


Figure 10: Re-interpreting the ground state wave function as a Euclidean transition amplitude in half of space.

It is also instructive to compare this argument to the more conventional one given e.g. in Weinberg. There one tries to construct free fields that create particles of arbitrary spin, finding out by brute force that it is impossible to choose the coefficient functions u_i and v_i such that the field both transforms in a valid representation of $Spin^+(d-1,1)$ and commutes/anticommutes at spacelike separation unless the spin-statistics relation is satisfied. The proof given here by contrast does not rely on free fields and is also more intuitive. As in the case of the $\mathcal{CRT}$ theorem, any experimental demonstration of a violation of the spin-statistics connection would be catastrophic for quantum mechanics and special relativity.

6.3 The structure of vacuum entanglement

There is a nice way to use the ideas we have been discussing to analyze the structure of the ground state wave function in relativistic quantum field theory. The idea is to decompose space into a "left" region with $x^1 < 0$ and a "right" region with $x^1 > 0$. To simplify our analysis we will restrict to bosonic theories where all fields commute at spacelike separation, in which case the fields in the L region are independent of the fields in the R region, so at least in the presence of a cutoff we can write the Hilbert space as a tensor product

$$\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_R. \tag{6.33}$$

The operators in the algebra $\mathcal{A}(L)$ are product operators of the form $O_L \otimes I_R$, while the operators in the algebra $\mathcal{A}(R) = \mathcal{A}'(L)$ are product operators of the form $I_L \otimes O_R$.[48] The ground state wave function is computed by the Euclidean path integral in the region $\tau < 0$:

$$\langle \phi_L \phi_R | \Omega \rangle \propto \int \mathcal{D}\phi|_0^{\phi_L,\phi_R} e^{-\int_{-\infty}^0 \int d^{d-1}x \mathcal{L}_E(\phi,\partial\phi)}. \tag{6.34}$$

The idea is to change our interpretation of this path integral from being split up on horizontal slices to being split up on radial slices, as shown in figure 10. We thus have

$$\begin{aligned} \langle \phi_L \phi_R | \Omega \rangle &\propto \langle \phi_R | e^{-\pi K_R} | F_{\mathcal{RT}} \phi_L \rangle \\ &= \sum_n e^{-\pi \omega_n} \langle \phi_R | n \rangle_R \langle n | F_{\mathcal{RT}} \phi_L \rangle \end{aligned} \tag{6.35}$$

where K_R is the right-sided boost operator

$$K_R = \int_{x^1>0} d^{d-1}x x^1 T_{00}(x), \tag{6.36}$$

[48] In fermionic theories this structure is more complicated because the fermionic fields in L and R need to anticommute instead of commute; we haven't introduced enough fermion technology to deal with this yet so for now we'll stick to bosons.

also sometimes called the **Rindler Hamiltonian**, and $|n\rangle$ is a complete basis of K_R eigenstates with eigenvalues ω_n. You can think of (6.35) as arising from applying our usual path integral derivation to Euclidean evolution by the Rindler Hamiltonian, which generates rotation in the τx^1 plane. To turn (6.35) into an expression for the ground state however we need to find a way to get ϕ_L into a bra instead of a ket. We can do this by introducing a "partial $\mathcal{CRT}$" operator $\Theta^R_{\mathcal{CRT}} : \mathcal{H}_R \to \mathcal{H}_L$ which acts as

$$\Theta^R_{\mathcal{CRT}}|\phi_R\rangle = |\phi'^*_L\rangle. \tag{6.37}$$

Here ϕ'_L indicates the $\mathcal{CRT}$ transformation of ϕ_L, which is indeed a function of ϕ_R. This operator implements $\mathcal{CRT}$ on operators in the *left* region, as we can check by noting that if x is in L we have:

$$\begin{aligned}\Phi(x)\Theta^R_{\mathcal{CRT}}|\phi_R\rangle &= \Theta^R_{\mathcal{CRT}}\left(\Theta^{R\dagger}_{\mathcal{CRT}}\Phi(x)\Theta^R_{\mathcal{CRT}}\right)|\phi_R\rangle \\ &= \Theta^R_{\mathcal{CRT}}\Phi'(x)|\phi_R\rangle \\ &= \Theta^R_{\mathcal{CRT}}\phi'_L(x)|\phi_R\rangle \\ &= \phi'^*_L(x)\Theta^R_{\mathcal{CRT}}|\phi_R\rangle. \end{aligned} \tag{6.38}$$

In the first line here we used that $\Theta^R_{\mathcal{CRT}}$ is antiunitary, the second line is just implementing $\mathcal{CRT}$ on Φ, in the third line we use that for bosonic theories Φ and $\Phi^\dagger$ are commuting so an eigenstates of Φ is also an eigenstates of Φ', and in the fourth line we used that $\Theta^R_{\mathcal{CRT}}$ is antilinear. From (6.13) we can rewrite (6.37) as

$$\Theta^R_{\mathcal{CRT}}|\phi_R\rangle = |F_{\mathcal{RT}}\phi_R\rangle, \tag{6.39}$$

and making the substitution $\phi_R = F_{\mathcal{RT}}\phi_L$ and using that $F^2_{\mathcal{RT}} = 1$ on bosons we have

$$\Theta^{R\dagger}_{\mathcal{CRT}}|\phi_L\rangle = |F_{\mathcal{RT}}\phi_L\rangle. \tag{6.40}$$

This then implies that

$$\langle n|F_{\mathcal{RT}}\phi_L\rangle = \langle n|\Theta^{R\dagger}_{\mathcal{CRT}}|\phi_L\rangle = \langle\phi_L|\Theta^R_{\mathcal{CRT}}|n\rangle, \tag{6.41}$$

and thus

$$\langle\phi_L\phi_R|\Omega\rangle \propto \sum_n e^{-\pi\omega_n}\langle\phi_L|\Theta^R_{\mathcal{CRT}}|n\rangle\langle\phi_R|n\rangle. \tag{6.42}$$

We therefore have shown that in any relativistic quantum field theory the ground state has the simple entangled form

$$|\Omega\rangle \propto \sum_n e^{-\pi\omega_n}\Theta^R_{\mathcal{CRT}}|n\rangle \otimes |n\rangle, \tag{6.43}$$

which is called the **Rindler decomposition**. Stated heuristically, the Rindler eigenstates in the right region are entangled with their $\mathcal{CRT}$ conjugates in the left region.[49]

6.4 Unruh Effect

The Rindler decomposition has two very important consequences, the first of which is the **Unruh effect**: an observer moving at constant acceleration a in the vacuum of a relativistic field theory feels a temperature

$$T_{Unruh} = \frac{\hbar a}{2\pi c k_B}, \tag{6.45}$$

[49] There were several points in this argument which need to be revisited if there are fermions. We don't yet have the tools to do so, but I'll mention that the result in that case becomes

$$|\Omega\rangle \propto \sum_n e^{-\pi\omega_n} i^{-F_L}\Theta^R_{\mathcal{CRT}}|n\rangle \otimes |n\rangle, \tag{6.44}$$

where F_L is the number of fermions in the left region.

where I have temporarily restored the unsightly dimensionful constants $\hbar$, c, and k_B. This is a quite remarkable statement, although not one which is easy to experience yourself. For example if we take a to be $9.8\,\mathrm{m/s^2}$ we get

$$T_{Unruh} \approx 4 \times 10^{-20}\mathrm{K}. \tag{6.46}$$

To derive this, we first note that an observer living in the right region can take the partial trace over the left region, leading to a vacuum density matrix

$$\rho_R \propto \sum_n e^{-2\pi\omega_n} |n\rangle\langle n| = e^{-2\pi K_R}. \tag{6.47}$$

This is nothing but a thermal density matrix, but with "Hamiltonian" K_R and "temperature" $T_K = \frac{1}{2\pi}$. The world should therefore look thermal to someone whose proper time is proportional to the boost rapidity η. From equation (6.22), we see that such a person should be moving on a trajectory

$$\begin{aligned} t(\eta) &= x_0 \sinh\eta \\ x(\eta) &= x_0 \cosh\eta. \end{aligned} \tag{6.48}$$

Note that this trajectory is the boost image of the point $(0, x_0)$. The proper time along this trajectory is related to η by

$$\tau = \eta x_0, \tag{6.49}$$

and we can compute the proper acceleration:

$$a = \sqrt{\left(\frac{d^2x}{d\tau^2}\right)^2 - \left(\frac{d^2t}{d\tau^2}\right)^2} = \frac{1}{x_0}. \tag{6.50}$$

Therefore the proper temperature seen by this observer is

$$T_{Unruh} = \frac{d\eta}{d\tau} T_K = \frac{1}{2\pi x_0} = \frac{a}{2\pi}. \tag{6.51}$$

This effect is the essence of Hawking's calculation showing that black holes evaporate into thermal radiation, and in fact Unruh discovered it by way of trying to come up with an intuitive interpretation of Hawking's paper.

6.5 Reeh-Schlieder property

There is a second important consequence of the Rindler decomposition, which we will call the Reeh-Schlieder property:

- In relativistic quantum field theory there are no nonzero local operators which annihilate the vacuum.

In free field theory this statement is quite intuitive: since any field is the sum of an annihilation and a creation part, to project onto the annihilation part we need to use a Fourier transform which is an integral over all of space. The proof for general field theories is quite simple: for any operator O localized in the right region R (or more carefully any element of $\mathcal{A}[R]$), using the Rindler decomposition the squared norm of the state created by acting with O on the vacuum is given by

$$\begin{aligned} \langle\Omega|O^\dagger O|\Omega\rangle &\propto \sum_n e^{-2\pi\omega_n} \langle n|O^\dagger O|n\rangle \\ &= \sum_{n,m} e^{-2\pi\omega_n} \langle n|O^\dagger|m\rangle\langle m|O|n\rangle \\ &= \sum_{n,m} e^{-2\pi\omega_n} |\langle m|O|n\rangle|^2. \end{aligned} \tag{6.52}$$

The final expression here is a sum of positive semi-definite terms, so it can vanish only if each term vanishes. Therefore if O annihilates the vacuum, all of its matrix elements must vanish - in other words O must itself be zero. We note in passing that the Reeh-Schlieder property is precisely what we needed to complete our proof of the spin-statistics theorem, so that theorem is now proved as well. We also note that this argument actually proves something stronger: it shows that no operator which is an element of $\mathcal{A}[R]$ in some Lorentz frame can annihilate the vacuum. For example any nonzero product of a finite number of local operators at arbitrary points also cannot annihilate the vacuum, since by an appropriate spacetime translation we can put all the operators into the domain of dependence of the right region R and the vacuum is translation-invariant.

The Reeh-Schlieder property has a rather surprising consequence: it implies that *any* state in the Hilbert space can be obtained by acting on the vacuum with an operator which is supported only in the left region L (by symmetry the same is of course true for the right region R, or more generally for the left or right region in any Lorentz frame). The proof goes like this: suppose by contradiction that there is a nontrivial subspace $S \subset \mathcal{H}$ which is orthogonal to all the states which can be written as $O|\Omega\rangle$ for some $O \in \mathcal{A}[L]$. We will argue that the projection P_S is a nonzero element of $\mathcal{A}[R]$ which annihilates $|\Omega\rangle$. By the Reeh Schlieder property this is not allowed, and so the subspace S must be zero-dimensional. The idea is to first consider $P_{S^\perp} = 1 - P_S$, which is the projection onto the subspace of states which *can* be created by acting on $|\Omega\rangle$ with elements of $\mathcal{A}[L]$. For any O in $\mathcal{A}[L]$ we have

$$OP_{S^\perp} = P_{S^\perp} O P_{S^\perp} \tag{6.53}$$

and

$$O^\dagger P_{S^\perp} = P_{S^\perp} O^\dagger P_{S^\perp}, \tag{6.54}$$

where in both cases the argument is that both sides of the equation act as zero on S and as $O/O^\dagger$ on $S^\perp$. Taking the dagger of the second equation and combining them, we see that

$$OP_{S^\perp} = P_{S^\perp} O, \tag{6.55}$$

and thus that $P_{S^\perp}$ is in the commutant of $\mathcal{A}[L]$. By Haag duality this is equal to $\mathcal{A}[R]$, and so we have $P_S = 1 - P_{S^\perp} \in \mathcal{A}[R]$. Moreover P_S clearly annihilates $|\Omega\rangle$ since $|\Omega\rangle \in S^\perp$.

We only proved the Reeh-Schlieder property for half-space regions, but in fact it is true for *any* region which is not a complete time slice.[50] In other words any operator which annihilates the vacuum cannot be in $\mathcal{A}[R]$ for *any* region R that is not a complete time slice. The argument just given then implies an even more shocking consequence: for *any* open spatial region R and *any* quantum state $|\psi\rangle$, we can find an element O of $\mathcal{A}[R]$ such that[51]

$$|\psi\rangle = O|\Omega\rangle. \tag{6.56}$$

For example we can instantaneously create the moon by acting on the vacuum with an operator that has support only in this classroom! This is a rather extreme example of what is called quantum teleportation.[52]

It is worth briefly mentioning some standard mathematical terminology which is used in discussing the Reeh-Schlieder property. In von Neumann algebra a state $|\Omega\rangle$ with the property that it is not annihilated by any nonzero element of a von Neumann algebra $\mathcal{A}$ is said to be **separating** for that algebra. Similarly a state $|\Omega\rangle$ with the property that $\mathcal{A}|\Omega\rangle$ is a dense set of states in the Hilbert space $\mathcal{H}$ is said to be a **cyclic** state for $\mathcal{A}$. What the Reeh Schlieder property says is that in quantum field theory the vacuum is both cyclic and separating for the algebra $\mathcal{A}[R]$ associated to any spatial region which isn't a complete time slice.

[50] Unfortunately I'm not aware of a simple proof of this generalization, except in the special case of conformal field theories.

[51] This statement isn't actually quite true: if we are careful about infinite-dimensional Hilbert spaces, what we find from the proof in the previous paragraph is that we can create a state which as close to $|\psi\rangle$ as we like in the Hilbert space norm. A mathematician would describe this situation by saying that the set $\mathcal{A}[R]|\Omega\rangle$ is **dense** in the Hilbert space $\mathcal{H}$.

[52] To be clear, the operator which does this is not unitary so we can't use it to communicate faster than light. This seeming non-locality is of the EPR variety, rather than the worse non-locality we found in the first section by trying to quantize a relativistic quantum particle.

6.6 Homework

1. Using (6.16) and also the spin-statistics theorem, show that $\Theta^2_{\mathcal{CRT}} = 1$.

2. Check that the complex scalar action is invariant under $\mathcal{CRT}$.

3. Check that the massive (real) vector action with Lagrangian

$$\mathcal{L} = -\frac{1}{4}(\partial_\mu V_\nu - \partial_\nu V_\mu)(\partial^\mu V^\nu - \partial^\nu V^\mu) - \frac{m^2}{2} V^\mu V_\mu \tag{6.57}$$

   is also invariant under $\mathcal{CRT}$.

4. Let's model the hydrogen atom by a classical electron orbiting the proton in a circle whose radius is the Bohr radius $a_0 = 5 \times 10^{-11}\,\mathrm{m}$. What is the Unruh temperature experienced by the electron? How does it compare to the binding energy of hydrogen?

5. Show that if $\Phi(x)\Phi^\dagger(y) \pm \Phi^\dagger(y)\Phi(x) = 0$ at spacelike separation, then we also have $\Phi(x)\Phi(y) \pm \Phi(y)\Phi(x) = 0$ and $\Phi^\dagger(x)\Phi^\dagger(y) \pm \Phi^\dagger(y)\Phi^\dagger(x) = 0$ at spacelike separation. Hint: you should assume that $\Phi(x)\Phi(y) + s\Phi(y)\Phi(x) = 0$ with either $s = 1$ or $s = -1$, and then show that s needs to be the same sign as appears in $\Phi(x)\Phi^\dagger(y) \pm \Phi^\dagger(y)\Phi(x) = 0$. I recommend considering the norm of the state $\Phi(x)\Phi(y)|\Omega\rangle$, and you will need to use the Reeh-Schlieder property and also that as $(x-y)^2 \to +\infty$ we have

$$\langle \Phi^\dagger(x)\Phi(x)\Phi^\dagger(y)\Phi(y)\rangle \to \langle \Phi^\dagger(x)\Phi(x)\rangle\langle \Phi^\dagger(y)\Phi(y)\rangle, \tag{6.58}$$

   which is an example of what is called **cluster decomposition**. In general cluster decomposition says that the connected correlation functions of local operators should always decay at large separation, in this case the connected two-point function of the composite operator $\Phi^\dagger\Phi$.

7 Perturbative calculation of correlation functions in interacting theories

So far our results in this class have fallen into two categories:

- Explicit calculations in free field theory
- General formal results (such as the $\mathcal{CRT}$ and spin-statistics theorems) which are valid in any relativistic quantum field theory.

Free field theory is quite useful for getting an initial picture of how quantum field theory works, and formal results are of course important for understanding the general structure of quantum field theory, but in the end of the day most field theories are not free and formal results won't get us to detailed predictions that can be quantivatively compared to experiment. It is time for us to learn how to do some explicit calculations in field theories that are not free.

The simplest interacting field theory is called ϕ^4 theory, and its Lagrangian density is given by

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{m^2}{2}\phi^2 - \frac{\lambda}{4!}\phi^4. \tag{7.1}$$

It must be acknowledged from the outset that no analytic solution of this theory is known. It is not difficult to see why it cannot be solved using the methods we have discussed so far: the Heisenberg equation of motion

$$(\nabla^2 - m^2)\Phi = \frac{\lambda}{6}\Phi^3 \tag{7.2}$$

is non-linear, and thus cannot be solved using the Fourier transform, and the path integral

$$\int \mathcal{D}\phi e^{i\int d^dx\mathcal{L}} \tag{7.3}$$

is not Gaussian so we can't compute it using our Gaussian tricks. In fact there is an even more severe problem: for $d \geq 4$ this model is widely expected to not even have a continuum limit: it can only be defined precisely in the presence of a finite UV cutoff such as a lattice. Nonetheless there is much to be gained by studying this model, and the key idea that will allow us to make progress is **perturbation theory**: we treat the parameter λ, called the coupling constant, as small, and then we compute interacting correlation functions as power series in λ about their free field values. There is a beautiful diagrammatic way of organizing such calculations, called **Feynman diagrams**, which we will meet for the first time in this section. Perturbative calculations using Feynman diagrams are the central focus of a large fraction of the practicing quantum field theorists in the world, especially those working in particle physics, and developing a good intuition for them is essential for any aspiring theoretical physicist (or any aspiring particle experimentalist).

7.1 Perturbation series for an integral

As a first illustration of the perturbative method, we'll consider the integral

$$f(\lambda) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dx\, e^{-\frac{1}{2}x^2 - \frac{\lambda}{4!}x^4}. \tag{7.4}$$

with $\lambda > 0$. This integral can be evaluated in closed form, according to Mathematica we have

$$f(\lambda) = \sqrt{\frac{3}{2\pi\lambda}}e^{\frac{3}{4\lambda}}K_{1/4}(\frac{3}{4\lambda}), \tag{7.5}$$

but our approach here will be to ignore this and try to approximate $f(\lambda)$ when $\lambda \ll 1$. The idea is to Taylor expand the "interaction" term, which allows us to rewrite the integral as a sum over Gaussian moments:

$$\begin{aligned} f(\lambda) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx e^{-\frac{1}{2}x^2} \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{\lambda x^4}{4!}\right)^n \\ \text{“}&= \text{”} \frac{1}{\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-\lambda}{4!}\right)^n \int_{-\infty}^{\infty} dx x^{4n} e^{-\frac{1}{2}x^2} \end{aligned} \tag{7.6}$$

I've put the equality in quotes in the second line since we have recklessly exchanged the order of summation and integration, a sin for which we will shortly pay a price. Proceeding boldly ahead in the meantime, we can be encouraged by the fact that the terms in the sum are suppressed by higher powers of λ as n increases, and so we can hope that truncating this sum to the first few terms gives a good approximation to $f(\lambda)$ when λ is small. The easiest way to evaluate these Gaussian moments is to remember the integral definition

$$\Gamma(y) = \int_0^{\infty} ds s^{y-1} e^{-s} \tag{7.7}$$

of the Euler Γ-function and change variables $x^2 = 2s$, which gives

$$\int_{-\infty}^{\infty} dx x^{4n} e^{-\frac{1}{2}x^2} = 2^{2n+1/2} \Gamma(2n + 1/2), \tag{7.8}$$

so we can write the perturbative expansion as

$$f(\lambda) \text{“}= \text{”} \frac{1}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{1}{n!} \Gamma(2n + 1/2) \left(-\frac{\lambda}{6}\right)^n. \tag{7.9}$$

The first few terms in the sum are given by

$$f(\lambda) \text{“}= \text{”} 1 - \frac{\lambda}{8} + \frac{35\lambda^2}{384} - \frac{385\lambda^3}{3072} + \frac{25025\lambda^4}{98304} + \ldots. \tag{7.10}$$

In figure 11 we show how this approximation does against the exact expression (7.5): at least in the range $0 < \lambda < .3$ including higher order terms indeed seems to give us a better and better approximation to $f(\lambda)$.

Unfortunately things are not so simple as this plot might suggest. Recalling Stirling's approximation that at large x we have

$$\Gamma(x) = \exp\left[x \log x - x + O(\log x)\right], \tag{7.11}$$

we see that the coefficients of λ^n in the series (7.9) eventually grow like $e^{2n \log n}$ at large n, which is faster than $\left(\frac{\lambda}{6}\right)^n$ is decreasing no matter the size of λ: the perturbation series (7.9) is divergent! This is the price we pay for our earlier illegal exchange of an integral and an infinite sum. Another way to anticipate this trouble is that the integral for $f(\lambda)$ is badly divergent for $\lambda < 0$, so asking for a convergent power series at $\lambda = 0$ is asking for too much. You may be desperately hoping that this problem is special to this particular example, but I assure that it isn't: almost any perturbation series in quantum field theory (or even in non-relativistic quantum mechanics) is divergent. We therefore need to decide what to do. Discarding the method altogether is too drastic given the impressive success shown in figure 11, but we'd like to get a better sense of when it succeeds and when it doesn't. The key idea to remember is that perturbation theory is an **asymptotic series**, which means that if we sum the first N terms in the series we get an approximation to the function whose error is of order λ^N for sufficiently small λ. The reason the series doesn't converge is that as N gets larger, we need to go to smaller λ for this approximation to be good. Asymptotic series are written using the "$\sim$" symbol, so we can rewrite (7.9) as

$$f(\lambda) \sim \frac{1}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{1}{n!} \Gamma(2n + 1/2) \left(-\frac{\lambda}{6}\right)^n. \tag{7.12}$$

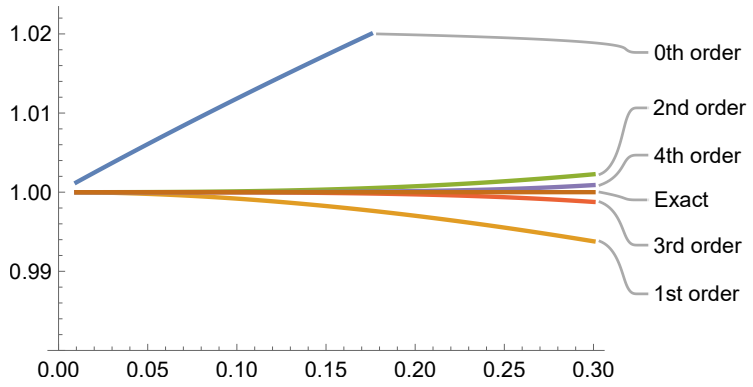


Figure 11: Comparing the first few terms in perturbation theory to the exact answer. What is plotted here is the ratio of the partial sum of the first few terms to the exact answer; for $\lambda < .3$ the first order result already brings us within a percent of right answer, and including higher order terms gets us even closer.

To show that this series is indeed asymptotic, note that we can legally move a finite number of the terms in the sum past the integral to get

$$f(\lambda) = \frac{1}{\sqrt{\pi}} \sum_{n=0}^{N-1} \frac{1}{n!} \Gamma(2n+1/2) \left(-\frac{\lambda}{6}\right)^n + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx e^{-\frac{1}{2}x^2} \sum_{n=N}^{\infty} \frac{1}{n!} \left(-\frac{\lambda x^4}{4!}\right)^n, \tag{7.13}$$

and therefore

$$f(\lambda) - \frac{1}{\sqrt{\pi}} \sum_{n=0}^{N-1} \frac{1}{n!} \Gamma(2n+1/2) \left(-\frac{\lambda}{6}\right)^n = \left(-\frac{\lambda}{4!}\right)^N \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx e^{-\frac{1}{2}x^2} \sum_{m=0}^{\infty} \frac{1}{(m+N)!} \left(-\frac{\lambda}{4!}\right)^m x^{4(m+N)}. \tag{7.14}$$

In the second line we relabeled the sum to pull out an overall factor of $\left(-\frac{\lambda}{4!}\right)^N$. The thing it multiplies approaches a constant as $\lambda \to 0$, so the error of the series is indeed of order λ^N at small λ.

We can understand the implications of the asymptotic nature of this series as follows: the series will not begin to diverge until we get to large enough n that

$$e^{\log n} \lambda \sim 1, \tag{7.15}$$

or in other words

$$n \sim \frac{1}{\lambda}. \tag{7.16}$$

At this point the terms are of order

$$\epsilon_{min} = c^{\frac{1}{\lambda}} = e^{-\frac{\log(1/c)}{\lambda}}, \tag{7.17}$$

where c is some $O(1)$ constant which is less than one. ϵ_{min} is the most accurate that the perturbation series can be, after this including more terms only causes the error to get larger. We illustrate this qualitative behavior in figure 12. Effects which are of order ϵ_{min} or smaller are typically referred to as **non-perturbative effects**, and in situations where they are of interest we need to use methods that go beyond perturbation theory. For reasonable values of λ however this minimal error can be quite small, for example in quantum electrodynamics we have $\lambda \approx \frac{1}{137}$ so the QED perturbation series should be good up to an unrecoverable error which is of order

$$\epsilon_{min} \sim e^{-137}. \tag{7.18}$$

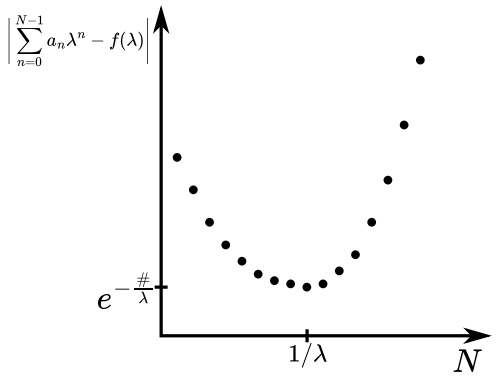


Figure 12: The qualitative behavior of perturbation theory: adding more terms to the series increases the accuracy until we get to $N \sim \frac{1}{\lambda}$ terms, at which point the error of the series is of order $e^{\frac{-\#}{\lambda}}$. After this the series begins to diverge and the approximation gets worse and worse. In the plot label a_n indicates the coefficient of λ^n in the perturbative expansion for $f(\lambda)$.

I'd say this is close enough for most practical purposes! From now on we will therefore use perturbation theory without further handwringing about its validity, except in non-perturbative situations where we are indeed interested in effects of order ϵ_{min}.[53]

7.2 Feynman diagrams for Gaussian integrals

In the previous section we took advantage of our knowledge of Γ-functions to immediately compute the coefficients in the perturbation series. In more general examples this is not possible, so we need another method. The idea which always works is to compute the integral (7.8) using our Gaussian integral technology. Indeed recall that we have

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dx e^{-\frac{1}{2}x^2+Bx} = e^{\frac{B^2}{2}}, \tag{7.19}$$

and therefore

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dx x^m e^{-\frac{1}{2}x^2} = \left(\frac{d}{dB}\right)^m e^{\frac{B^2}{2}}|_{B=0}. \tag{7.20}$$

At first the combinatorics of computing these derivatives is somewhat intimidating, after all it has to give the series (7.10) whose coefficients are not so simple-looking, but life is simple once we realize that each derivative can only do one of two things: bring down a factor of B from the exponent or compute the derivative of the existing prefactor. We therefore have

$$\left(\frac{d}{dB}\right)^m e^{\frac{B^2}{2}} = e^{\frac{B^2}{2}}\left(B+\frac{d}{dB}\right)\left(B+\frac{d}{dB}\right)\ldots\left(B+\frac{d}{dB}\right)\times 1, \tag{7.21}$$

where are there m copies of $\left(B+\frac{d}{dB}\right)$. In order to get a term which survives when we set $B=0$, there must be an equal number of Bs and $\frac{d}{dB}$s, with each derivative appearing to the left of the B that it acts on. There are no such terms when m is odd, so we see that the integral (7.20) vanishes unless n is even. When m is even, the number of terms is equal to the number of pairings of m objects since each derivative needs to be

[53]Of course if λ is not small then neither is ϵ_{min}, so in that case we are obviously interested in effects of order ϵ_{min}! More interesting are situations where λ is small but we nonetheless still care about some non-perturbatively small effect. For example there could be some process whose rate is zero to all orders in perturbation theory but not zero, in which case non-perturbative effects give the leading contribution. The possible decay of the Higgs vacuum is an example of such an effect.

paired with the B it acts on. The number of such pairings is

$$N_m := \frac{m!}{2^{m/2}\,(m/2)!}, \tag{7.22}$$

since we can chose the first element of the first pair, the second element of the first pair, and so on down to the 2nd element of the $m/2$nd pair, and then we need to divide by a factor of two for each pair since the order doesn't matter and also divide by the number of permutations of the pairs. Therefore we have

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dx x^m e^{-\frac{1}{2}x^2} = \begin{cases} N_m & m \text{ even} \\ 0 & m \text{ odd} \end{cases}. \tag{7.23}$$

This of course is equal to what we found using the Γ function (with the replacement $m = 4n$), as you will check on the homework.

In quantum field theory we are really interested in multi-dimensional Gaussian integrals, which we found obey

$$\sqrt{\mathrm{Det}\left(\frac{A}{2\pi}\right)}\int dx \exp\left[-\frac{1}{2}x^T A x + B^T x\right] = \exp\left[\frac{1}{2}B^T A^{-1} B\right] \tag{7.24}$$

and thus

$$\sqrt{\mathrm{Det}\left(\frac{A}{2\pi}\right)}\int dx x_{i_1}\ldots x_{i_m} e^{-\frac{1}{2}x^T A x} = \frac{\partial}{\partial B_{i_1}}\cdots\frac{\partial}{\partial B_{i_m}} e^{\frac{1}{2}B^T A^{-1} B}\bigg|_{B=0}. \tag{7.25}$$

We can think of this as the m-point correlation function in the Gaussian distribution. To compute it we again can observe that each derivative again does one of two things, which now are to bring down a factor of $A^{-1}B$ or to take the derivative of the existing prefactor, so we have

$$\frac{\partial}{\partial B_{i_1}}\cdots\frac{\partial}{\partial B_{i_m}} e^{\frac{1}{2}B^T A^{-1} B} = e^{\frac{1}{2}B^T A^{-1} B}\left(\sum_{j_1} A^{-1}_{i_1 j_1} B_{j_1} + \frac{\partial}{\partial B_{i_1}}\right)\cdots\left(\sum_{j_m} A^{-1}_{i_m j_m} B_{j_m} + \frac{\partial}{\partial B_{i_m}}\right)\times 1. \tag{7.26}$$

As before we can only get a term that survives taking $B = 0$ if each partial derivative is paired with a B to its right, so the integral again vanishes for odd m while for even m we have

$$\sqrt{\mathrm{Det}\left(\frac{A}{2\pi}\right)}\int dx x_{i_1}\ldots x_{i_m} e^{-\frac{1}{2}x^T A x} = \sum_P \prod_{(j,k)\in P} A^{-1}_{i_j i_k}. \tag{7.27}$$

Here P indicates pairings of $1,\ldots,m$. As before there are N_m such pairings, but now they can make different contributions to the integral. For example for $m = 4$ we have

$$\sqrt{\mathrm{Det}\left(\frac{A}{2\pi}\right)}\int dx x_{i_1} x_{i_2} x_{i_3} x_{i_4} e^{-\frac{1}{2}x^T A x} = A^{-1}_{i_1 i_2} A^{-1}_{i_3 i_4} + A^{-1}_{i_1 i_3} A^{-1}_{i_2 i_4} + A^{-1}_{i_1 i_4} A^{-1}_{i_2 i_3}. \tag{7.28}$$

We are now ready for our first meeting with Feynman diagrams. These are simply a graphical way of representing the different pairings appearing on the right side of equation 7.27. The idea is quite trivial: we draw a dot for each x_i appearing in the correlation function, and then we draw lines connecting them to indicate the pairing. Each pairing contributes a "propagator" A^{-1}. The $m = 4$ case is shown in figure 13.

7.3 Feynman diagrams for an "interacting" integral

We can explore the idea of Feynman diagrams further by considering an "interacting" integral

$$f(\lambda) = \sqrt{\mathrm{Det}\left(\frac{A}{2\pi}\right)}\int dx e^{-\frac{1}{2}x^T A x - \frac{\lambda}{4!}\sum_i x_i^4}, \tag{7.29}$$

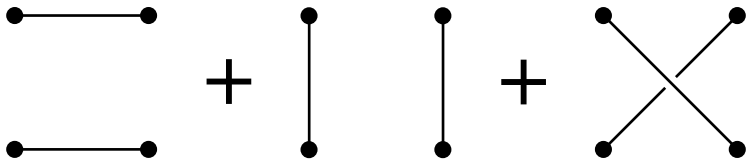


Figure 13: Feynman diagrams for the four-point function in the Gaussian distribution.

which you can think of as a simple model of the interacting ϕ^4 theory we began the section with. The perturbative expansion for this integral is

$$f(\lambda) \sim \sqrt{\mathrm{Det}\left(\frac{A}{2\pi}\right)} \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{\lambda}{4!}\right)^n \sum_{i_1 \dots i_n} \int dx x_{i_1}^4 \dots x_{i_n}^4 e^{-\frac{1}{2}x^T A x}, \tag{7.30}$$

and we can evaluate these integrals using our pairing formula (7.27). We now meet a new phenomenon however, which is that many of the pairings give the same answer due to the repeated indices in the interaction. For example the first order $n = 1$ contribution to the series is

$$\begin{aligned} -\frac{\lambda}{4!} \sum_i \sqrt{\mathrm{Det}\left(\frac{A}{2\pi}\right)} \int dx x_i^4 e^{-\frac{1}{2}x^T A x} &= -\frac{\lambda}{4!} \times 3 \times \sum_i (A_{ii}^{-1})^2 \\ &= -\frac{\lambda}{8} \sum_i (A_{ii}^{-1})^2, \end{aligned} \tag{7.31}$$

where all three pairings appearing in (7.28) contribute equally. The second order contribution has three distinct kinds of pairings: those where each interaction has two self-pairings, those where each interaction has one self-pairing, and those where there are no self-pairings. These lead to

$$\begin{aligned} \frac{\lambda^2}{2 \times (4!)^2} \sum_{ij} \sqrt{\mathrm{Det}\left(\frac{A}{2\pi}\right)} \int dx x_i^4 x_j^4 e^{-\frac{1}{2}x^T A x} &= \frac{\lambda^2}{2 \times (4!)^2} \sum_{ij} \left(9(A_{ii}^{-1})^2 (A_{jj}^{-1})^2 + 72 A_{ii}^{-1} A_{jj}^{-1} (A_{ij}^{-1})^2 + 24 (A_{ij}^{-1})^4\right) \\ &= \lambda^2 \sum_{i,j} \left(\frac{1}{128}(A_{ii}^{-1})^2 (A_{jj}^{-1})^2 + \frac{1}{16} A_{ii}^{-1} A_{jj}^{-1} (A_{ij}^{-1})^2 + \frac{1}{48}(A_{ij}^{-1})^4\right), \end{aligned} \tag{7.32}$$

where the factors of 9, 72, and 24 count how many pairings there are of each type. Counting these pairings takes a bit of practice to get used to, we illustrate the idea in figure 14.

The diagrams in figure 14 are useful for counting pairings, but it is also useful to have a simpler set of diagrams which are designed so that the same diagram automatically represents all the pairings in each equivalence class. Following Feynman, the idea is to combine all the dots appearing in each factor of the interaction $\sum_i x_i^4$ to a single **interaction vertex**, giving us the Feynman diagram expansion. See figure 15 for the set of Feynman diagrams contributing to $f(\lambda)$ up through order n^2. In terms of these diagrams we can rewrite our asymptotic series for $f(\lambda)$ as

$$f(\lambda) \sim 1 + \sum_D (-\lambda)^{n_D} \frac{1}{s_D} \sum_{i_1 \dots i_{n_D}} \prod_{(m,\ell) \in L_D} A_{i_m i_\ell}^{-1}, \tag{7.33}$$

where m and ℓ label the interaction vertices of the diagram, n_D indicates the number of interaction vertices in D, L_D indicates the set of (unoriented) links in D, and s_D is called the **symmetry factor** of the diagram and is given by

$$s_D = \frac{n_D!(4!)^{n_D}}{p_D}, \tag{7.34}$$

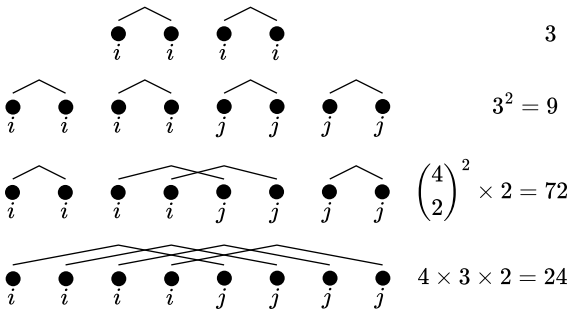


Figure 14: Counting pairings at first and second order. For the $n = 1$ pairings, we need to pick which of three other is to pair the first i with. For the $n = 2$ pairings where each interaction has two self-pairings, we need to make this choice independently for each interaction. For the $n = 2$ pairings where both interactions have a single self-pairing, for each interaction we need to pick which two of the four is are self-paired, and then there are two ways to do the remaining pairings. For the $n = 2$ pairings with no self-pairings, we need to pick which of the four js pairs with the first i, which of the remaining three js pairs with the second i, and which of the remaining two js pairs with the third i.

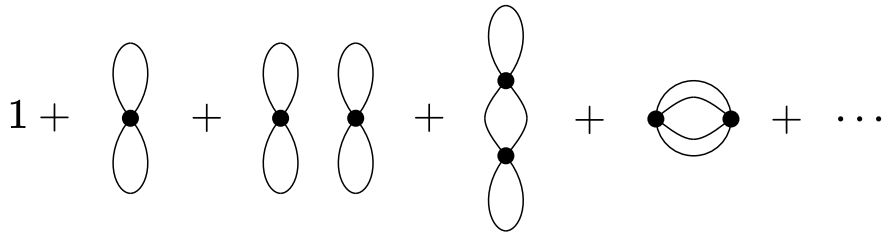


Figure 15: Feynman diagrams contributing to $f(\lambda)$ up through order λ^2. As we saw above, the symmetry factors for these diagrams are $s_D = 1, 8, 128, 16$, and 48.

with p_D the number of pairings which give rise to this diagram as in figure 14. Except for s_D all factors in (7.33) are easy to read off by visual inspection of D, so Feynman diagrams give a powerful way of immediately seeing what is going on at each order in perturbation theory. There is actually also a way to compute s_D directly from the diagram, it is the size of the automorphism group of the diagram, but as long as you do not intend to become a high-order amplitudes expert it is easy enough (and perhaps safer) to just use the method of figure 14 to compute p_D.[54] In more realistic theories where the interaction vertices are less symmetric we conveniently often have $S_D = 1$.

7.4 Exponentiation of connected diagrams

You may have already noticed in figure 15 that at second order we started getting diagrams which are topologically disconnected. This makes the computation of the perturbation series for $f(\lambda)$ somewhat redundant, as diagrams from lower orders are constantly reappearing at higher orders. In fact there is a beautiful combinatoric simplification: the sum (7.33) of all disconnected Feynman diagrams is actually the *exponential* of the sum of connected diagrams only! In other words we have

$$\log f(\lambda) \sim \sum_C (-\lambda)^{n_C} \frac{1}{s_C} \sum_{i_1 \dots i_{n_C}} \prod_{(m,\ell)\in L_C} A^{-1}_{i_m i_\ell}, \tag{7.35}$$

where C indicates the set of **connected** Feynman diagrams.

To derive (7.35), we need to understand how to evaluate a disconnected diagram in terms of its connected components. We will indicate by

$$V_D = (-\lambda)^{n_D} \frac{1}{s_D} \sum_{i_1 \dots i_{n_D}} \prod_{(m,\ell)\in L_D} A^{-1}_{i_m i_\ell}. \tag{7.36}$$

the "value" of a Feynman diagram. If D is disconnected then most of these terms are just products of the analogous terms for its connected components, but we need to be careful about the symmetry factor. Indeed let's say that a disconnected diagram D has connected components $C_1, C_2, \dots, C_M$, which we will momentarily take to be all distinct from each other. We can write the pairing number p_D of the full disconnected diagram as

$$\begin{aligned} p_D &= \binom{n_D}{n_{C_1}} \binom{n_D - n_{C_1}}{n_{C_2}} \dots \binom{n_D - n_{C_1} - \dots - n_{C_{M-1}}}{n_{C_M}} \times p_{C_1} \dots p_{C_M} \\ &= \frac{n_D!}{n_{C_1}! \dots n_{C_M}!} \times p_{C_1} \dots p_{C_M}, \end{aligned} \tag{7.37}$$

where the combinatoric factors account for the number of ways we can choose which interaction vertices get assigned to which connected components, and we then multiply by the number of pairings we can do within each component. If the diagrams appear with repetitions, say m_a repetitions of C_a, then we need to divide by additional factors of $m_a!$ since exchanging identitical connected components of a pairing gives the same pairing. We thus in general have

$$p_D = \frac{n_D!}{(n_{C_1}!)^{m_1} \dots (n_{C_M}!)^{m_M}} \times \frac{1}{m_1! \dots m_M!} \times (p_{C_1})^{m_1} \dots (p_{C_M})^{m_M}, \tag{7.38}$$

[54] This interpretation of S_D is actually the reason we included the factor of $1/4!$ in the interaction vertex. The basic idea is that $n_D!(4!)^{n_D}$ gives an "estimate" of how many pairings there are with a given diagram topology, since permuting the n_D vertices and permuting which of the four dots at each vertex get attached to other dots can't change the graph topology. This sometimes is an overestimate however, as whenever the graph has an automorphism then acting on a pairing with it gives the same pairing. Therefore S_D is precisely counting the number of such automorphisms. For example in the second diagram of figure 15 there are three $\mathbb{Z}_2$ automorphisms: one that reflects the top lobe, one that reflects the bottom lobe, and one that exchanges the two lobes. We therefore have $S_D = 8$. Similarly for the fifth diagram there is a four-fold permutation symmetry of the links, as well as a $\mathbb{Z}_2$ symmetry that exchanges the two vertices, so we have $S_D = 4! \times 2 = 48$.

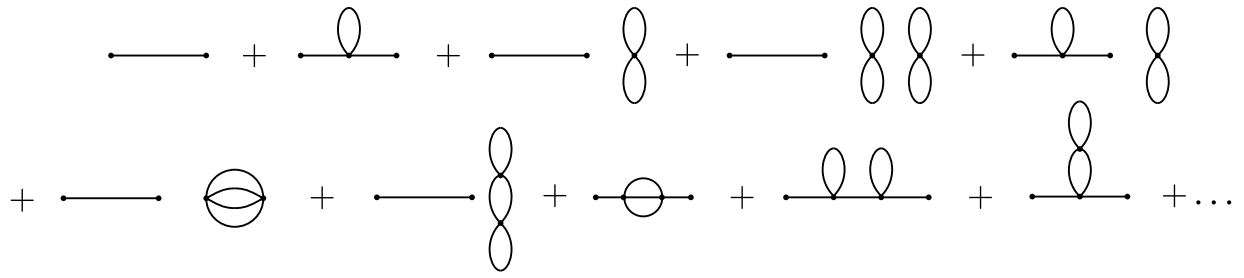


Figure 16: Feynman diagrams for computing the numerator of (7.42) with two external points. Dividing by the denominator of (7.42) removes all disconnected diagrams. The symmetry factors of the connected diagrams here are $s_D = 1, 2, 6, 4,$ and 4.

and therefore

$$\frac{1}{S_D} = \frac{p_D}{n_D!(4!)^{n_D}} = \frac{1}{S_{C_1}^{m_1} \dots S_{C_M}^{m_M}} \times \frac{1}{m_1! \dots m_M!}. \tag{7.39}$$

We can therefore write the value of D as

$$V_D = \prod_C \frac{(V_C)^{m_C}}{m_C!}. \tag{7.40}$$

Finally we can observe that these are precisely the coefficients that these values appear with in

$$e^{\sum_C V_C} = \prod_C e^{V_C} = \prod_C \left(\sum_{m_C} \frac{(V_C)^{m_C}}{m_C!} \right), \tag{7.41}$$

which computes the proof of (7.35).

7.5 Perturbative computation of correlation functions

Let's now see how to use Feynman diagrams to compute perturbative corrections to the Gaussian correlation functions (7.27). We want to evaluate

$$\langle x_{i_1} \dots x_{i_M} \rangle = \frac{\int dx x_{i_1} \dots x_{i_M} e^{-\frac{1}{2} x^T A x - \frac{\lambda}{4!} \sum_i x_i^4}}{\int dx e^{-\frac{1}{2} x^T A x - \frac{\lambda}{4!} \sum_i x_i^4}}. \tag{7.42}$$

We already know how to compute the denominator perturbatively: it is the exponential of the sum of connected Feynman diagrams with only interaction vertices (divided by a factor of $\sqrt{\mathrm{Det}\left(\frac{A}{2\pi}\right)}$ that will cancel with the same factor in the numerator). Let's think about how to compute the numerator. The perturbation series for the numerator is

$$\begin{aligned} \sqrt{\mathrm{Det}\left(\frac{A}{2\pi}\right)} \int dx x_{i_1} \dots x_{i_M} e^{-\frac{1}{2} x^T A x - \frac{\lambda}{4!} \sum_i x_i^4} &\sim \sqrt{\mathrm{Det}\left(\frac{A}{2\pi}\right)} \sum_n \frac{1}{n!} \left(-\frac{\lambda}{4!}\right)^n \int dx x_{i_1} \dots x_{i_M} \left(\sum_i x_i^4\right)^n e^{-\frac{1}{2} x^T A x} \\ &= \sum_n \frac{1}{n!} \left(-\frac{\lambda}{4!}\right)^n \sum_{i_{M+1} \dots i_{M+n}} \sum_P \prod_{(j,k) \in P} A^{-1}_{i_j, i_k}, \end{aligned} \tag{7.43}$$

where in the second line I've labeled the n interaction vertices as $i_{M+1}, \dots, i_{M+n}$. In such calculations the x_{i_a} with $a \in (1, M)$ are referred to as "external" and the i_a with $a \in (M+1, M+n)$ are referred to as "internal" or "interaction".[55]

[55] In this terminology the denominator of (7.42) (multiplied by $\sqrt{\mathrm{Det}\left(\frac{A}{2\pi}\right)}$) is the exponential of the sum of connected diagrams with no external legs, also sometimes called the sum over "vacuum bubbles".

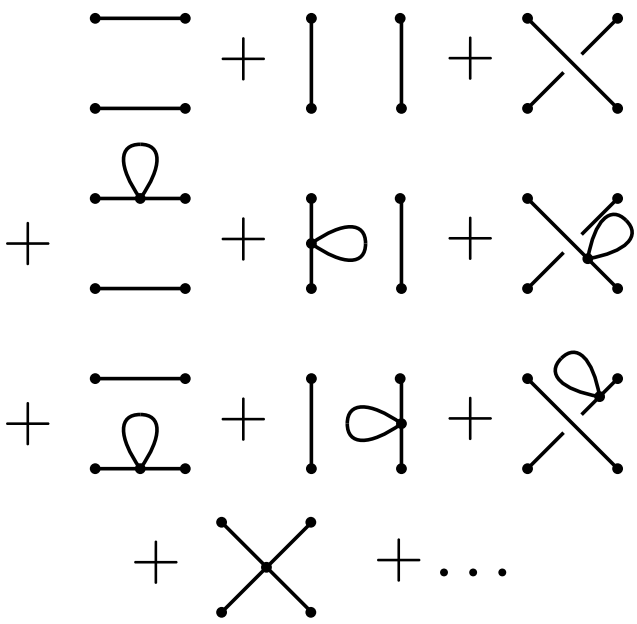


Figure 17: Feynman diagrams contributing to the four-point function up through $O(\lambda)$. Note however that the second two rows are all really just incorporating corrections to the two point functions in the first row; it is only the fourth row that is a "genuinely four-point" contribution.

As in the previous subsection we can group this sum over pairings into equivalence classes labeled by Feynman diagrams, with the diagrams contributing to the two-point function through second order shown in figure 16. In general in terms of diagrams we have

$$\sqrt{\mathrm{Det}\left(\frac{A}{2\pi}\right)}\int dx x_{i_1}\dots x_{i_M} e^{-\frac{1}{2}x^T Ax - \frac{\lambda}{4!}\sum_i x_i^4} \sim \sum_D (-\lambda)^{n_D}\frac{1}{S_D}\sum_{i_{M+1}\dots i_{M+n_D}}\prod_{m,\ell\in L_D} A^{-1}_{i_m i_\ell}, \tag{7.44}$$

where now n_D is the number of interaction vertices, m and ℓ run over the links of the diagram including links to external points, and S_D is again the symmetry factor

$$S_D = \frac{n_D!(4!)^{n_D}}{p_D} \tag{7.45}$$

with p_D the number of pairings of the $m+N$ points that give rise to the diagram D. As before we can interpret S_D as counting the automorphisms of the diagram, now restricting to those automorphisms which keep the external points fixed. We also have an exponentiation result: the numerator of (7.42) is equal to the sum over diagrams where all interaction vertices are connected to at least one external point times the exponential of the sum over connected vacuum bubbles. The second factor just cancels the denominator (7.42), so we then have

$$\langle x_{i_1}\dots x_{i_M}\rangle \sim \sum_{\hat{C}}(-\lambda)^{n_{\hat{C}}}\frac{1}{S_{\hat{C}}}\sum_{i_{M+1}\dots i_{M+n_{\hat{C}}}}\prod_{m,\ell\in L_{\hat{C}}} A^{-1}_{i_m i_\ell} \tag{7.46}$$

where $\hat{C}$ runs over the set of diagrams where each interaction vertex is connected to at least one external point. The first few such diagrams for the four-point function are shown in figure 17. Note that these diagrams still are not all connected, essentially because there are diagrams which amount to just correcting

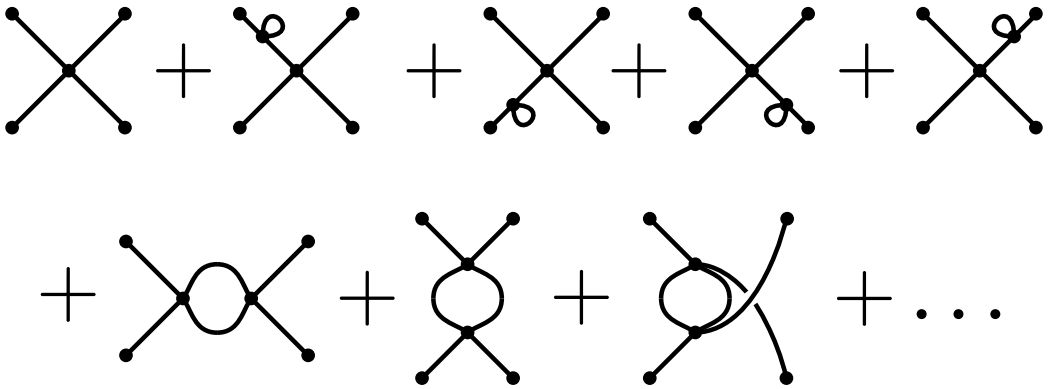


Figure 18: Feynman diagrams for computing the connected four-point function, including all contributions up through λ^2. The symmetry factor of the first diagram is one and the symmetry factors for the others are all two.

the two-point functions appearing in figure 13 rather than giving "genuinely four-point" contributions. To focus on the latter, we should look at the **connected four-point function**, which is defined by

$$\langle x_{i_1} \dots x_{i_4} \rangle_c = \langle x_{i_1} \dots x_{i_4} \rangle - \langle x_{i_1} x_{i_2} \rangle \langle x_{i_3} x_{i_4} \rangle - \langle x_{i_1} x_{i_3} \rangle \langle x_{i_2} x_{i_4} \rangle - \langle x_{i_1} x_{i_4} \rangle \langle x_{i_2} x_{i_3} \rangle. \tag{7.47}$$

More generally the connected M-point function $\langle x_{i_1} \dots x_{i_M} \rangle_c$ is defined recursively by[56]

$$\langle x_{i_1} \dots x_{i_M} \rangle = \sum_S \langle \prod_{j \in S_1} x_{i_j} \rangle_c \dots \langle \prod_{j \in S_L} x_{i_j} \rangle_c, \tag{7.48}$$

where the sum is over partitions S of M into parts $S_1, \dots, S_L$. This defines $\langle x_{i_1} \dots x_{i_M} \rangle_c$ in terms of lower-point connected correlation functions and the full correlation function $\langle x_{i_1} \dots x_{i_M} \rangle$. Recursing down to the lowest level, we take $\langle x_i \rangle_c = \langle x_i \rangle$. Forgetting for a moment that in this theory the odd moments of x_i vanish, the first few explicit solutions of this definition are

$$\begin{aligned} \langle x_i \rangle_c &= \langle x_i \rangle \\ \langle x_i x_j \rangle_c &= \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle \\ \langle x_i x_j x_k \rangle_c &= \langle x_i x_j x_k \rangle - \langle x_i x_j \rangle \langle x_k \rangle - \langle x_i x_k \rangle \langle x_j \rangle - \langle x_j x_k \rangle \langle x_i \rangle + 2 \langle x_i \rangle \langle x_j \rangle \langle x_k \rangle. \end{aligned} \tag{7.49}$$

In practice however the definition (7.48) is more useful, as it shows that what the connected correlation function is really doing is removing all parts of the full correlation function which are mere products of lower-point correlation functions. Said differently, it builds up the full correlation function out of connected components in precisely the same way as Feynman diagrams do. We therefore can express the connected correlation function as a sum over connected diagrams only:

$$\langle x_{i_1} \dots x_{i_M} \rangle_c \sim \sum_C (-\lambda)^{n_C} \frac{1}{S_C} \sum_{i_{M+1} \dots i_{M+n_C}} \prod_{m,\ell \in L_C} A^{-1}_{i_m i_\ell}, \tag{7.50}$$

where now the sum is over genuinely connected diagrams C. We show the first few diagrams contributing to the connected four-point function in figure 18.

Already some patterns may be apparent in the diagrams we have discussed. Let's emphasize two of them:

- The number of diagrams grows quite rapidly as we go to higher orders in λ. Roughly speaking grows like some power of $n_D!$, since the total number of pairings grows like this and the symmetry factors grow

[56] In other contexts the connected correlation functions are called "cumulants" or "Ursell functions".

too slowly to make up for it (after all generic diagrams should have few symmetries). This growth is consistent with the idea that the series should be divergent, since a power of $n_D!$ will always eventually beat λ^{n_D}. It also means that computing higher-order Feynman diagrams is a rather laborious process, requiring many clever tricks to make progress.

- For a fixed number of external legs, as we go to higher order the number of loops in the diagram increases by one for each power of λ. Diagrams are thus often classified by the number of loops rather than the number of interaction vertices, as it is really the number of loops that determines the complexity of evaluating individual diagrams. Connected diagrams with interaction vertices but no loops are called **tree diagrams**, while higher loop diagrams are referred to as **one-loop diagrams**, **two-loop diagrams**, and so on. Most theoretical physicists these days never need to evaluate a diagram with more than one loop, so in this class our focus will be on computing tree and one-loop diagrams rather than developing machinery for higher loop computations.

7.6 Feynman diagrams for perturbative correlation functions in ϕ^4 theory

Having now set up all of our machinery, it is quite easy (at least formally) to generalize from perturbative correlation functions for the integral (7.42): we just replace $\int dx$ by the Euclidean path integral $\int \mathcal{D}\phi$ leading to

$$\langle T\phi(x_1)\dots\phi(x_M)\rangle = \frac{\int \mathcal{D}\phi\phi(x_1)\dots\phi(x_M)e^{-S_0-\frac{\lambda}{4!}\int d^dx\phi^4}}{\int \mathcal{D}\phi e^{-S_0-\frac{\lambda}{4!}\int d^dx\phi^4}}. \tag{7.51}$$

Here S_0 is the Euclidean action

$$S_0 = \frac{1}{2}\int d^dx \left(\dot{\phi}^2 + \vec{\nabla}\phi\cdot\vec{\nabla}\phi + m^2\phi^2\right) \tag{7.52}$$

of the free massive scalar field ($\dot{\phi}$ indicates the derivative with respect to the Euclidean time τ). The perturbative evaluation of this path integral is precisely the same as in the previous section, leading to

$$\langle T\phi(x_1)\dots\phi(x_M)\rangle_c \sim \sum_C (-\lambda)^{n_C}\frac{1}{S_C}\int d^dx_{M+1}\dots\int d^dx_{M+n_C}\prod_{m,\ell\in L_C} G_E(x_m-x_\ell), \tag{7.53}$$

where

$$G_E(x) = \int \frac{d^dp}{(2\pi)^d}\frac{e^{ip\cdot x}}{p^2+m^2} \tag{7.54}$$

is the Euclidean propagator. The two-point function is thus computed by the sum of the connected subset of the diagrams appearing in figure 16, while the four-point function is computed by the sum of connected diagrams appearing in figure 18. So for example the first few terms in the expansion for the Euclidean two-point function are

$$\begin{aligned}\langle T\phi(x_1)\phi(x_2)\rangle \sim & G_E(x_2-x_1) - \frac{\lambda}{2}\int d^dx_3 G_E(x_1-x_3)G_E(x_2-x_3)G_E(0)\\ & + \lambda^2\int d^dx_3 d^dx_4\left(\frac{1}{6}G_E(x_1-x_3)G_E(x_2-x_4)G_E(x_3-x_4)^3\right.\\ & \left. + \frac{1}{4}G_E(x_1-x_3)G_E(x_3-x_4)G_E(x_2-x_4)G_E(0)^2 + \frac{1}{4}G_E(x_1-x_3)G_E(x_2-x_3)G_E(x_3-x_4)^2G_E(0)\right)\\ & + \dots.\end{aligned} \tag{7.55}$$

You may be alarmed by the factors of $G_E(0)=\infty$ in the one-loop and two-loop contributions to this formula. These are further "UV divergences" of the type we met already in computing the Hamiltonian in free

field theory. There we saw the divergence could be absorbed into a redefinition of the cosmological constant via a process we called **renormalization**. We'll eventually see that we can also absorb the divergences here into a redefinition of the particle mass m and a rescaling of the field ϕ. To get a first sense of the former we can observe that a change δm^2 in the mass squared corrects the Euclidean propagator as

$$\delta G_E(x) = -\int \frac{d^d p}{(2\pi)^d} \frac{\delta m^2}{(p^2+m^2)^2} e^{ip\cdot x}, \tag{7.56}$$

while the one-loop contribution to the propagator is

$$-\frac{\lambda}{2}\int \frac{d^d p_1}{(2\pi)^d}\frac{d^d p_2}{(2\pi)^d}\frac{G_E(0)}{(p_1^2+m^2)(p_2^2+m^2)}\int d^d x_3 e^{ip_1\cdot(x_1-x_3)+ip_2(x_2-x_3)} = -\frac{\lambda G_E(0)}{2}\int \frac{d^d p}{(2\pi)^d}\frac{e^{ip\cdot(x_2-x_1)}}{(p^2+m^2)^2}, \tag{7.57}$$

so the only effect of this diagram is to shift the mass-squared of the free propagator by

$$\delta m^2 = \frac{\lambda G_E(0)}{2}. \tag{7.58}$$

Said differently, if we call the mass in the Lagrangian m_{bare} then the "true" mass squared is

$$m^2_{\text{true}} = m^2_{\text{bare}} + \frac{\lambda G_E(0)}{2} + O(\lambda^2). \tag{7.59}$$

In particular if we want m_{true} to be finite, then we should tune m^2_{bare} to cancel this divergent contribution and leave a leftover finite piece. This is another example of the process of renormalization. It is important to emphasize that the fact that there is a nontrivial relationship between m_{true} and m_{bare} has nothing to do with the fact that the relationship involves UV divergent objects. Even in a theory where all momentum integrals were convergent, such as a lattice theory or a theory in low spacetime dimensions, it would still be true that in interacting theories $m_{\text{true}} \neq m_{\text{bare}}$. This may make you wonder what we even mean by m_{true}, we will answer this question in the next section.

The fine-tuning of m^2_{bare} which is needed to cancel the mass of a scalar particle is sometimes viewed with suspicion, and the fact that we need to do it for the mass of the Higgs boson (a scalar) in the standard model of particle physics is often called the **hierarchy problem**. We will discuss this more in later sections.

We can also compute higher-point correlation functions, for example the leading contribution to the connected Euclidean four-point function is

$$\langle T\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\rangle_c \sim -\lambda\int d^d x_5 G_E(x_1-x_5)G_E(x_2-x_5)G_E(x_3-x_5)G_E(x_4-x_5)+\dots \tag{7.60}$$

To get perturbative expressions for time-ordered Lorentzian correlation functions, we should rotate $\tau = i(1-i\epsilon)t$ in all external locations $x_1,\dots,x_M$ and also in all interaction locations $x_{M+1},\dots,x_{M+n_C}$. This has two effects: it converts all Euclidean propagators to Feynman propagators

$$G_F(x) = \int \frac{d^d p}{(2\pi)^d}\frac{-ie^{ip\cdot x}}{p^2+m^2-i\epsilon}, \tag{7.61}$$

and provides an extra factor of i^{n_C} from the $d\tau$ factors in the integrals over interaction locations. Thus we have the Lorentzian formula

$$\langle T\phi(x_1)\dots\phi(x_M)\rangle_c \sim \sum_C (-i\lambda)^{n_C}\frac{1}{S_C}\int d^d x_{M+1}\dots\int d^d x_{M+n_C}\prod_{m,\ell\in L_C} G_F(x_m-x_\ell). \tag{7.62}$$

In these calculations the exponentiated sum over connected bubble diagrams canceled between the numerator and denominator of (7.51). It is worth mentioning however that this sum does have a physical interpretation: it renormalizes the cosmological constant. To see this, note that any connected bubble diagram will be proportional to the volume of spacetime since there is a symmetry of translating all interaction vertices by the same amount. We can therefore view each connected diagram with no external legs as giving a contribution to the cosmological constant.

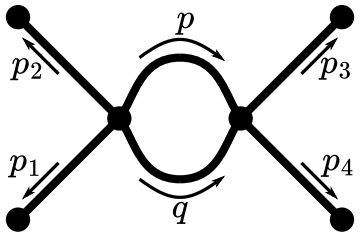


Figure 19: Momentum labels for a one-loop contribution to the four-point function.

7.7 Feynman rules in momentum space for correlation functions in ϕ^4 theory

Earlier in the semester we met the QFT mantra that one should think in position space but compute in momentum space. This mantra certainly applies to perturbative correlation functions. Working now in Lorentzian signature, the basic idea is to replace G_F in (7.62) by its momentum representation (7.61) and then evaluate all of the integrals over interaction vertex locations to get momentum-conserving delta functions. The only subtlety in doing this is that although $G_F(x_2 - x_1)$ is symmetric under exchanging x_1 and x_2, the momentum representation isn't manifestly symmetric so in assigning a momentum to a link in the graph we need to pick a direction. This is typically indicated on the diagram by drawing a small arrow next to the link which is labeled by the momentum, as in figure 19. For this diagram the momentum representation is

$$\begin{aligned}\langle T\phi(x_1)\dots\phi(x_4)\rangle_c \supset & \frac{(-i\lambda)^2}{2}\int \frac{d^dp_1}{(2\pi)^d}\cdots\frac{d^dp_4}{(2\pi)^d}\frac{d^dp}{(2\pi)^d}\frac{d^dq}{(2\pi)^d}e^{ip_1\cdot x_1+\dots ip_4\cdot x_4}\\ & \times\frac{-i}{p_1^2+m^2-i\epsilon}\frac{-i}{p_2^2+m^2-i\epsilon}\frac{-i}{p_3^2+m^2-i\epsilon}\frac{-i}{p_4^2+m^2-i\epsilon}\frac{-i}{p^2+m^2-i\epsilon}\frac{-i}{q^2+m^2-i\epsilon}\\ & \times(2\pi)^d\delta^d(p_1+p_2+p+q)(2\pi)^d\delta^d(p_3+p_4-p-q).\end{aligned} \tag{7.63}$$

This looks a bit nicer if we take the Fourier transform and use one of the δ-functions to evaluate the q integral, giving us

$$\begin{aligned}\langle T\phi(p_1)\dots\phi(p_4)\rangle_c \supset & \frac{(-i\lambda)^2}{2}(2\pi)^d\delta^d(p_1+p_2+p_3+p_4)\frac{-i}{p_1^2+m^2-i\epsilon}\cdots\frac{-i}{p_4^2+m^2-i\epsilon}\\ & \times\int\frac{d^dp}{(2\pi)^d}\frac{-i}{p^2+m^2-i\epsilon}\frac{-i}{(p-p_3-p_4)^2+m^2-i\epsilon}.\end{aligned} \tag{7.64}$$

Here the "time-ordered correlator in momentum space" just means the Fourier transform of the time-ordered position space correlator, and the δ-function in front is called a **momentum-conserving δ-function**. Such a δ-function appears in every momentum-space correlation function, and is a consequence of the fact that the correlation functions in position space only depend on relative positions due to spacetime translation invariance. The "hard" part of computing this diagram is evaluating the integral over the loop momentum on the second line, we will learn how to evaluate such integrals in a few weeks.

The procedure employed in the previous paragraph is easily formalized into an algorithm for evaluating any Feynman diagram for a momentum-space correlation function. This algorithm is called the **Feynman rules**, and given a connected Feynman diagram C contributing to the connected M-point function it goes like this:

1. Write a factor of $(-i\lambda)^{n_C}$, where n_C is the number of interaction vertices in C.
2. Divide by the symmetry factor S_C.
3. Label the momenta of all propagators, with external momenta pointed outwards.

4. Multiply by an overall momentum-conserving δ-function $(2\pi)^d \delta^d(p_1 + \ldots + p_M)$.

5. Multiply by a factor of $\frac{-i}{p^2+m^2-i\epsilon}$ for each propagator, both internal and external, imposing momentum conservation at all interaction vertices (e.g. imposing $q = p_3 + p_4 - p$) in the previous diagram).

6. Integrate over any internal momenta which are not determined by the δ-functions.

This algorithm is the daily routine of every perturbative quantum field theorist, although as we already saw in the previous section some care is needed to deal with UV (and eventually IR) divergences and at higher loops one has to be rather organized.

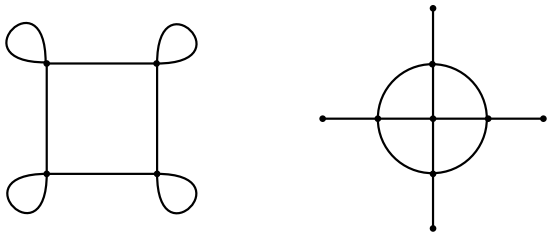


Figure 20: Diagrams for problem 3.

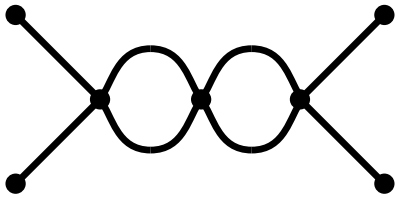


Figure 21: Diagram for problem 4.

7.8 Homework

1. Using Mathematica (or your favorite competitor), for $f(\lambda)$ given by (7.5) make plots of $\log|\sum_{n=0}^{N-1} a_n\lambda^n - f(\lambda)|$ as a function of N for $\lambda = .5, .2, .1,$ and $.05$. Here a_n are the coefficients in the perturbation series (7.12). Are your plots consistent with the qualitative story in figure 12? In particular note the maximal accuracy and the value of N at which the series begins to diverge.

2. Starting from the definition (7.7), show that the Γ-function obeys $\Gamma(x+1) = x\Gamma(x)$ (to use (7.7) you can assume $\mathrm{Re}\, x > 0$, but if you are comfortable with analytic continuation then you should also argue that this identity holds for all complex x). Also show that $\Gamma(1/2) = \sqrt{\pi}$. Using these results, show that (7.8) and (7.23) are compatible.

3. Check the symmetry factors quoted in the captions of figures 16 and 18, and also compute the symmetry factors for the two diagrams shown in figure 20.

4. Using the momentum-space Feynman rules, write down an expression for the contribution of the two-loop diagram in figure 21 to the Fourier transform of the Lorentzian time-ordered four-point function in $\lambda\phi^4$ theory. You should evaluate all momentum integrals which can be evaluated using the momentum-conserving δ-functions at the vertices, but you can leave any remaining integrals unevaluated. Make sure to label the directions of the momenta on your diagram.

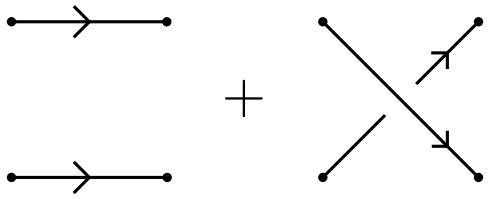


Figure 22: Feynman diagrams for the four-point function in a Gaussian integral over complex degrees of freedom, note that there is one fewer diagram than in the real case.

5. Show that

$$\int dxdx^* e^{-x^\dagger Ax + B^T x + \widetilde{B}^T x^*} = \frac{1}{\mathrm{Det}\left(\frac{iA}{2\pi}\right)} e^{\widetilde{B}^T A^{-1} B}, \tag{7.65}$$

where x is a vector with complex components, A is a positive symmetric matrix, and the measure is defined by $dxdx^* = -2id\mathrm{Re}(x)d\mathrm{Im}(x)$. Use this to show that correlation functions of the form

$$\langle x_{i_1} \dots x_{i_M} x^*_{j_1} \dots x^*_{j_N} \rangle = \mathrm{Det}\left(\frac{iA}{2\pi}\right) \int dxdx^* x_{i_1} \dots x_{i_M} x^*_{j_1} \dots x^*_{j_M} e^{-x^\dagger Ax} \tag{7.66}$$

are given by a sum over pairings as in (7.27), but where now each pair must contain one x and one x^* (so in particular they vanish unless $M = N$). Feynman diagrams for a complex degree of freedom therefore include an arrow on each propagator that points from x to x^*, as in figure 22.

6. Consider now an interacting complex scalar field with Lagrangian[57]

$$\mathcal{L} = -\partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi - \frac{\lambda}{4} |\phi|^4. \tag{7.67}$$

Using the results of the previous problem, what are the Feynman rules for time-ordered connected correlation functions in momentum space? Using your rules, write out an expression in momentum space for the contribution of the diagram in figure 19 to the four-point function where the two left dots are ϕs and the two right dots are ϕ^*s (you will need to add arrows to the propagators). You don't need to evaluate the final momentum integral.

[57] Note that the denominator in the interaction term is 4, not 4!. This is because the symmetry of the vertex is now only exchanging two ϕs and two ϕ^*s instead of permuting four ϕs.

8 Particles and Scattering

So far we have organized our discussion of quantum field theory in terms of *correlation functions* - vacuum expectation values of products of Heisenberg fields. At least in free field theory however we saw that we could also interpret correlation functions in terms of **particles**: we found a basis of eigenstates of the Hamiltonian whose elements each have some definite number of non-interacting bosons each carrying a definite spacetime momentum. It is natural to ask if interacting field theories also have a description in terms of particles. In general they don't, which is why we have focused on correlation functions so far, but many of the interacting quantum field theories of most interest in physics do indeed give rise to particles and it is therefore worthwhile for us to spend some time understanding when this happens and how to relate it to our knowledge of correlation functions.

8.1 One-particle states

What do we mean by a particle? Our first approach to answering this question is to say what we mean by a one-particle state. We have already met particles of spin zero in our free scalar theory, but now it is time to understand the general case. To motivate the problem I'll mention a question that bothered me as a student: it is often said that photons have spin one, but if so then why do they only have two polarization states instead of three? We'll answer this question at the end of this section.

The definition we propose is the following:[58]

- In any relativistic quantum system, a **one-particle state** is an eigenstate of the total spacetime momentum operator P^μ which is part of an irreducible representation of the (identity component of the) Poincaré group whose basis element labels are discrete and finite except for the eigenvalue of P^μ.

In other words we can expand the set of one-particle states in a basis of P^μ eigenstates

$$P^\mu|p,\sigma\rangle = p^\mu|p,\sigma\rangle, \tag{8.1}$$

where σ runs over a finite set and Poincaré transformations $U(\Lambda,a)$ act within the subspace of the full Hilbert space of the theory that is spanned by this basis. As in previous sections we will normalize these states so that

$$\langle p',\sigma'|p,\sigma\rangle = (2\pi)^{d-1}\delta^{d-1}(\vec{p}'-\vec{p})\delta_{\sigma'\sigma}, \tag{8.2}$$

where ensuring diagonality in the σ indices may require some Gram-Schmidt procedure. The action of spacetime translations in this basis is simple, we have

$$e^{-ia_\mu P^\mu}|p,\sigma\rangle = e^{-ia_\mu p^\mu}|p,\sigma\rangle, \tag{8.3}$$

so what we need to understand is the action of Lorentz transformations $U(\Lambda)$. These obey

$$\begin{aligned} P^\mu U(\Lambda)|p,\sigma\rangle &= U(\Lambda)U(\Lambda^{-1})P^\mu U(\Lambda)|p,\sigma\rangle \\ &= U(\Lambda)\Lambda^\mu{}_\nu P^\nu|p,\sigma\rangle \\ &= \Lambda^\mu{}_\nu p^\nu U(\Lambda)|p,\sigma\rangle, \end{aligned} \tag{8.4}$$

so we see that we need to have

$$U(\Lambda)|p,\sigma\rangle = \sum_{\sigma'} C_{\sigma',\sigma}(\Lambda,p)|\Lambda p,\sigma'\rangle \tag{8.5}$$

for some matrices $C_{\sigma',\sigma}(\Lambda,p)$.

[58] Recall that a **representation** π of a group G is a set of linear maps $D_\pi(G)$ on a vector space V obeying $D_\pi(g_1)D_\pi(g_2) = D_\pi(g_1g_2)$. A representation is **irreducible** if the only subspaces $S \subset V$ which are preserved by all elements of π are V itself and the empty subspace. It is unfortunately standard to also refer to V itself as the representation, as we have already done in the above definition of one-particle states.

To work out the structure of the $C_{\sigma',\sigma}(\Lambda, p)$, it is useful to first consider the special case where ${\Lambda^\mu}_\nu p^\nu = p^\mu$. Given a spacetime momentum p^μ, the subgroup of the Lorentz group which preserves p^μ is called the **little group** for p^μ. Given a Lorentz transformation W which is in the little group for some spacetime momentum k^μ, acting on $|k, \sigma\rangle$ the transformation (8.5) simplifies to

$$U(W)|k, \sigma\rangle = \sum_{\sigma'} D_{\sigma',\sigma}(W)|k, \sigma'\rangle, \tag{8.6}$$

where we've defined

$$D_{\sigma',\sigma}(W) = C_{\sigma',\sigma}(W, k). \tag{8.7}$$

These D-matrices give a finite-dimensional representation of the little group for k^μ:

$$\sum_{\sigma'} D_{\sigma,\sigma'}(W_1) D_{\sigma',\sigma''}(W_2) = D_{\sigma,\sigma''}(W_1 W_2). \tag{8.8}$$

A warning:

- The D-matrices appearing here are operators which represent the little group on the Hilbert space of quantum mechanics. They are NOT the same as the $D(\Lambda)$ matrices we met in earlier sections, which represented the full Lorentz group acting on the components of the fields in the theory via $U(\Lambda^{-1})\Phi^a(x)U(\Lambda) = \sum_b D^a{}_b(\Lambda)\Phi^b(\Lambda^{-1}x)$. Many people have wasted a lot of time being confused about the difference between the (typically infinite-dimensional) representation $U(\Lambda)$ of Lorentz symmetry acting on Hilbert space and the (typically finite-dimensional) representation $D(\Lambda)$ of Lorentz symmetry acting on fields. The D-matrices we are introducing now are involved in the part of the former which acts on one-particle states, *not* the latter.

The key point is then that once we have decided on a representation for the little group, the representation of the full Lorentz group is determined as well. The idea is that the set of possible spacetime momenta p^μ for a particle are all related by Lorentz transformations, so we can write each p^μ as a Lorentz transformation L_p of some fixed reference momentum k^μ. The detailed form of k^μ depends on whether the particle is massive or massless, and will be considered in the next paragraph. So far we have not said anything about how the σ indices at different momenta are related, we can determine this by simply adopting a convention where the state $|p, \sigma\rangle$ is related to the state $|k, \sigma\rangle$ by

$$|p, \sigma\rangle = N(p)U(L_p)|k, \sigma\rangle. \tag{8.9}$$

Here $N(p)$ is a normalization factor that we include to maintain the normalization (8.2). We showed back in section 3 that this requires

$$N(p) = \sqrt{\frac{k^0}{p^0}}, \tag{8.10}$$

with the idea being that the object $\frac{d^d p}{(2\pi)^d} 2\pi\delta(p^2+m^2)\Theta(p^0) = \frac{d^{d-1}p}{(2\pi)^{d-1}}\frac{1}{2\omega_{\vec{p}}}$ defines a Lorentz-invariant measure on spatial momenta and this implies a Lorentz transformation $\delta^{d-1}(\vec{p}'_\Lambda - \vec{p}_\Lambda) = \frac{\omega_{\vec{p}}}{\omega_{\vec{p}_\Lambda}}\delta^{d-1}(\vec{p}' - \vec{p})$. For general Λ we then have

$$\begin{aligned} U(\Lambda)|p, \sigma\rangle &= N(p)U(\Lambda)U(L_p)|k, \sigma\rangle \\ &= N(p)U(L_{\Lambda p})U(L_{\Lambda p}^{-1}\Lambda L_p)|k, \sigma\rangle \\ &= N(p)\sum_{\sigma'} D_{\sigma',\sigma}(L_{\Lambda p}^{-1}\Lambda L_p)U(L_{\Lambda p})|k, \sigma'\rangle \\ &= \frac{N(p)}{N(\Lambda p)}\sum_{\sigma'} D_{\sigma',\sigma}(L_{\Lambda p}^{-1}\Lambda L_p)|\Lambda p, \sigma'\rangle, \end{aligned} \tag{8.11}$$

where we have observed that $L_{\Lambda p}^{-1}\Lambda L_p$ is in the little group with respect to k^μ, and thus that

$$C_{\sigma',\sigma}(\Lambda,p) = \sqrt{\frac{(\Lambda p)^0}{p^0}} D_{\sigma',\sigma}(L_{\Lambda p}^{-1}\Lambda L_p). \tag{8.12}$$

This way of building a representation of a group out of a representation for one of its subgroups is called the method of induced representations.

To discuss the structure of the little group in more detail, we need to be more explicit about which Lorentz group we are considering: do we include its non-identity components, and do we go to the double cover where a rotation by 2π acts as -1 on particles of half-integer spin? Let's first consider the more familiar case where we take 2π to be the identity, in which case we are interested in the identity component $SO^+(d-1,1)$ of the Lorentz group. The little group depends on whether p^μ is timelike, null, or spacelike, and in the timelike and null cases it also depends on whether it is future or past pointing. There are no known particles whose momentum is spacelike (this would be called "tachyons" and would allow causality violation), nor are there any known particles with $p^0 < 0$ (these would have negative energy and destabilize the vacuum). We will thus focus on the cases where p^μ is timelike or null with $p^0 > 0$, which describe massive and massless particles respectively.

In the massive case we have $p \cdot p = -m^2$ for some $m > 0$, and by going to the rest from of the particle we can choose our reference momentum to be

$$k^\mu = (m, 0, \ldots 0). \tag{8.13}$$

The little group thus consists of those elements of $SO^+(d-1,1)$ which preserve the vector $(m,0,\ldots,0)$. No Lorentz transformation which involves a boost can do this, so the little group of a massive particle is just the spatial rotation group $SO(d-1)$. Therefore each massive particle is characterized by an irreducible representation of the spatial rotation group, which of course is what we call the **spin** of the particle. In particular for $d = 4$ the irreducible representations of $SO(3)$ are labeled by integers $j \geq 0$, with the spin-j representation having dimension $2j+1$ as you hopefully know. If we generalize $SO^+(d-1,1)$ to its double cover $Spin^+(d-1,1)$, then the little group becomes $Spin(d-1)$ so more representations are allowed. In particular for $d = 4$ the little group becomes $Spin(3) = SU(2)$, which allows for half-integer j.

The massless case is perhaps more novel. A massless particle has no rest frame, so the best we can do is choose the reference momentum k^μ to point in the positive x^1 direction:

$$k^\mu = (\kappa, \kappa, 0, \ldots 0) \tag{8.14}$$

with $\kappa > 0$. To find the little group the easiest way to proceed is to find the set of Lorentz generators which annihilate k^μ. We can write a general Lorentz generator as

$$J = -i \begin{pmatrix} 0 & b_1 & b_2 & \ldots & b_{d-1} \\ b_1 & 0 & -c_2 & \ldots & -c_{d-1} \\ b_2 & c_2 & & & \\ \vdots & \vdots & & A & \\ b_{d-1} & c_{d-1} & & & \end{pmatrix}, \tag{8.15}$$

where the b_i multiply boost generators in the i direction and the c_i multiply rotation generators in the $1-i$ plane. The matrix A is an arbitrary real antisymmetric matrix, which we can think of as generating a rotation that doesn't involve the x^1 direction. Demanding that this annihilates (8.14) tells us that we need $b_1 = 0$ and $b_i = -c_i$ for all $i \geq 2$, so we see that the little group of a massless particle moving in the x^1 direction is generated by rotations involving the $x^2, \ldots, x^{d-1}$ directions and combinations of boosts and rotations whose generators have the form

$$A^i = J^{i0} + J^{i1}. \tag{8.16}$$

These generators are mutually commuting, and are rotated into each other by rotations involving the $x^2, \ldots, x^{d-1}$ directions, so the little group of a massless particle is isomorphic to the group $ISO(d-2)$

of Euclidean rotations and translations of $\mathbb{R}^{d-2}$. Since the A^i are mutually commuting we can simultaneously diagonalize them, but this leads to a problem: since their eigenvalues give a vector in $\mathbb{R}^{d-2}$ which can be continuously rotated, if we do not have $A^i = 0$ for all $i \geq 2$ then the index σ cannot run only over a finite (or even discrete) range. Such representations are called "continuous spin particles", and they are typically viewed as pathological. For example such a particle could never be in thermal equilibrium, and seems quite hard to reconcile with quantum gravity. There has also never been any evidence for a continuous spin particle. More pedantically, continuous spin particles don't actually obey our definition of particle state since σ is not finite. For any of these reasons, we will from now on restrict to representations with $A^i = 0$. This reduces the little group to $SO(d-2)$, and the choice of an irreducible $SO(d-2)$ representation associated to a particle is called its **helicity**. In particular for $d = 4$ helicity is determined by an irreducible representation of $SO(2)$, and these are one-dimensional and labeled by an integer j. Indeed the representation is simply

$$D_j(\theta) = e^{ij\theta}, \tag{8.17}$$

where θ is the rotation angle in the $2-3$ plane. If we generalize $SO^+(d-1,1)$ to $Spin^+(d-1,1)$, then the little group of a massless particle (again with $A^i = 0$) is $Spin(d-2)$. For $d = 4$ this is again isomorphic to $SO(2)$, but now with half-integer j allowed (so that we have to rotate by 4π to get back to where we started).

Now we can return to the question of why photons do not have three spin states. The reason is that they don't really have spin one, spins are for massive particles. What they have is helicity one! Photons with right circular polarization have helicity $j = 1$, while photons with left circular polarization have helicity $j = -1$. These states transform in distinct irreducible representations of the Lorentz group, and in particular they are not mixed together by Lorentz transformations (which is quite different from the situation for a massive particle of spin one).

There is a somewhat confusing terminology issue with helicity that is related to the idea of spatial reflection symmetry $\mathcal{R}$ (or parity $\mathcal{P}$ in even dimensions). Photons with helicity one and helicity minus one are not mixed by the action of the connected Lorentz group $SO^+(d-1,1)$, so strictly speaking according to our definition we should view them as different types of particle. On the other hand all interactions which involve photons also preserve $\mathcal{R}$ symmetry, and $\mathcal{R}$ symmetry *does* mix photons of opposite helicity. It is therefore conventional to refer to both helicities as photons. The same is true for gravitons, which have helicity $j = \pm 2$. In the standard model of particle physics however there are particles called neutrinos, which are involved in nuclear reactions, and which are treated as massless with helicity $j = -1/2$. The standard model also has massless particles with helicity $j = 1/2$, which are called antineutrinos. The interactions of these particles do not respect $\mathcal{R}$ symmetry, and so they are given different names.[59]

8.2 Multiparticle states in non-interacting theories

We now consider states with more than one particle. In non-interacting theories this is a fairly straightforward matter: a multiparticle state has the form

$$|p_1, \sigma_1, n_1; p_2, \sigma_2, n_2; \ldots; p_M, \sigma_M, n_M\rangle, \tag{8.18}$$

where the new label n_i tells us the type of the ith particle (i.e. is it a photon, an electron, etc). The Poincaré transformation of such a state is just the product transformation

$$U(\Lambda, a)|p_1, \sigma_1, n_1; \ldots; p_M, \sigma_M, n_M\rangle = \prod_{i=1}^{M} \left(e^{-ia\cdot\Lambda p} \sqrt{\frac{(\Lambda p_i)^0}{p_i^0}} \sum_{\sigma_i'} D^{n_i}_{\sigma_i', \sigma_i} \left(L^{-1}_{\Lambda p_i} \Lambda L_{p_i} \right) \right) |\Lambda p_1, \sigma_1', n_1; \ldots; \Lambda p_M, \sigma_M', n_M\rangle. \tag{8.19}$$

[59] There is clear evidence that at least two of the three known types of neutrino is massive. Most particle physicists expect that in fact they all are, but so far this has not been confirmed experimentally. Understanding the nature of the neutrino mass matrix is one of the main goals of current particle physics research.

The normalization of these states is a bit trickier since we need to account for identical particles. For example in our free scalar theory we have

$$\begin{aligned}\langle\Omega|a_{\vec{p}_2'}a_{\vec{p}_1'}a^\dagger_{\vec{p}_1}a^\dagger_{\vec{p}_2}|\Omega\rangle =&(2\pi)^{d-1}\delta^{d-1}(\vec{p}_1'-\vec{p}_1)\times(2\pi)^{d-1}\delta^{d-1}(\vec{p}_2'-\vec{p}_2)\\ &+(2\pi)^{d-1}\delta^{d-1}(\vec{p}_1'-\vec{p}_2)\times(2\pi)^{d-1}\delta^{d-1}(\vec{p}_2'-\vec{p}_1).\end{aligned}\tag{8.20}$$

In general we impose

$$\langle p_1',\sigma_1',n_1';\ldots;p_M',\sigma_M',n_M'|p_1,\sigma_1,n_1;\ldots;p_M,\sigma_M,n_M\rangle=\sum_\pi(-1)^{f_\pi}\prod_{i=1}^M(2\pi)^{d-1}\delta^{d-1}(\vec{p}'_{\pi(i)}-\vec{p}_i)\delta_{\sigma'_{\pi(i)}\sigma_i}\delta_{n'_{\pi(i)}n_i},\tag{8.21}$$

where the sum is over permutations π of M objects and f_π indicates the number of fermions which are exchanged by the permutation. In free field theory this sum is automatically generated by the algebra of creation/annihilation operators, as we saw in (8.20). It is also useful to introduce the idea of a complete set of multiparticle states, which we can write as

$$I=\sum_{M=0}^\infty\prod_{i=1}^M\left(\int\frac{d^{d-1}p_i}{(2\pi)^{d-1}}\sum_{\sigma_i}\sum_{n_i}\right)\frac{1}{S(n)}|p_1,\sigma_1,n_1;\ldots;p_M,\sigma_M,n_M\rangle\langle p_1,\sigma_1,n_1;\ldots;p_M,\sigma_M,n_M|,\tag{8.22}$$

where the "symmetry factor" $S(n)$ counts the number of possible permutations of identitical particles in $|p_1,\sigma_1,n_1\ldots p_M,\sigma_M,n_M\rangle$ (so in particular if none of the n_i are equal then $S(n)=1$). Here by convention the state with zero particles is of course the vacuum $|\Omega\rangle$.

As you can already see the notation for multiparticle states is somewhat tedious, so following Weinberg we'll adopt an abbreviated notation where a multiparticle state is simply called $|\alpha\rangle$, the inner product (8.21) is written as

$$\langle\alpha|\beta\rangle=\delta(\alpha-\beta),\tag{8.23}$$

and the resolution of the identity (8.22) is written as

$$I=\int d\alpha|\alpha\rangle\langle\alpha|.\tag{8.24}$$

8.3 Multiparticle states in interacting theories

Now we come to the crucial question: to what extent do interacting quantum field theories have particles? It is clear that asking for a complete basis of states of the form (8.18) with Poincaré transformation (8.19) is asking for too much: in particular taking $\Lambda=1$ and a to be a time translation, (8.19) would imply that the energy of a multiparticle state is just the sum of the single-particle energies while we know that this isn't true e.g. because of potential energy between particles. (8.19) would also imply that the number of particles is conserved with time, while we know clearly from experiment that it is not. On the other hand there is no such objection to interacting theories having *one-particle* states obeying (8.5). For example the hydrogen atom is a completely stable one-particle state in quantum electrodynamics, as is the electron and the proton. I emphasize that from the point of view of scattering theory there is no difference between "fundamental" particles such as the electron, that correspond to fields in the Lagrangian, and "bound state" particles such as the hydrogen atom. After all one can never be sure that there isn't a "more fundamental" theory in which electrons are also bound states, as indeed is true for protons (which are bound states of quarks and gluons).

Given the existence of one-particle states it is natural to hope for the existence of multiparticle states, at least in the limit that the particles live in wave packets that are localized far away from each other. Moreover in this limit we can hope for the energy and momentum of the particles to indeed be additive. In order to realize these hopes however, we need the interactions between the particles to fall off sufficiently fast with distance. This is always true in theories where all particles are massive, and that is the regime

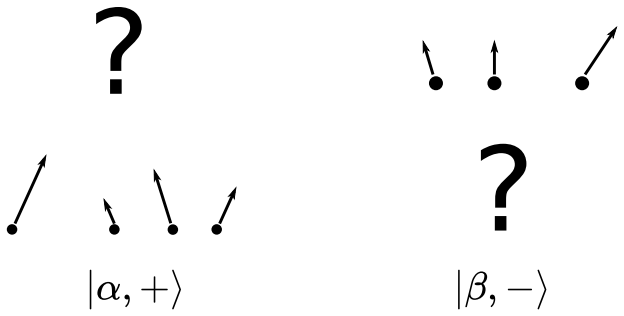


Figure 23: In and out states in scattering: in the in state $|\alpha, +\rangle$ we have a definite particle configuration at $t \to -\infty$, while in an out state $|\beta, +\rangle$ we have a definite particle configuration at $t \to \infty$. In general an in state evolves to a complicated superposition of out states, with the coefficients being given by the S-matrix.

where scattering theory is most clearly established. In theories with massless particles one can still try to use scattering theory, but one often encounters "infrared divergences" when doing so and these typically need to be dealt with on a somewhat case-by-case basis. Those of you who have studied scattering in non-relativistic quantum mechanics should already be familiar with this problem, as attempts to treat scattering off of a Coulomb potential using standard methods break down due to logarithmic divergences. Our approach will be simply to proceed with the assumption that additive multiparticle states exist, with the understanding that this will sometimes lead to trouble with massless particles that we will need to address when it arises. We formalize this as follows:

- A quantum mechanical theory with Hamiltonian H has a **scattering description** if H has a complete set of "in state" eigenstates, denoted $|\alpha, +\rangle$, and also a complete set of "out state" eigenstates $|\alpha, -\rangle$, both with eigenvalues
$$H|\alpha, \pm\rangle = E_\alpha |\alpha, \pm\rangle, \tag{8.25}$$
where E_α are the eigenvalues of a non-interacting multiparticle Hamiltonian H_0 with eigenstates $|\alpha\rangle$, such that we have
$$\lim_{t \to \mp\infty} \int d\alpha\, g(\alpha) e^{-iE_\alpha t} |\alpha, \pm\rangle = \lim_{t \to \mp\infty} \int d\alpha\, g(\alpha) e^{-iE_\alpha t} |\alpha\rangle \tag{8.26}$$
for arbitrary smooth (and integrable) wave packets $g(\alpha)$ which respect the exchange symmetry of any identical particles.

What this definition says is that wave packets of the in states look like non-interacting multiparticle eigenstates at early times, while wave packets of the out states look like non-interacting multiparticle eigenstates at late times. The basic idea is illustrated in figure 23.

One immediate consequence of this definition is that the inner product of in states with in states and out states with out states is the same as for non-interacting particles:

$$\langle \beta, \pm | \alpha, \pm \rangle = \delta(\beta - \alpha), \tag{8.27}$$

which follows because the inner product is time-independent so we can compute at early/late times for in/out states where they coincide with the non-interacting eigenstates. More interesting is the overlap between in and out states, which by definition is called the **S-matrix**:

$$S_{\beta\alpha} := \langle \beta, - | \alpha, + \rangle. \tag{8.28}$$

The S-matrix is the primary object of interest in scattering theory; it tells us the quantum amplitude to find the system in an out state $|\beta, -\rangle$ given that it started in an in state $|\alpha, +\rangle$. More earthily, the S-matrix

provides the answer to a question well known to children everywhere: if you take some stuff and slam it together, what comes out? Many physicists, who after all have much in common with children, spend their days studying precisely this question.

A very important property of this S-matrix is that it is *unitary*, which follows immediately from its definition since it is a change of basis between two complete sets of orthonormal states. We can also check this explicitly:

$$\int d\beta\, S^*_{\beta\alpha} S_{\beta\gamma} = \int d\beta\, \langle \alpha, +|\beta, -\rangle \langle \beta, -|\gamma, +\rangle = \langle \alpha, +|\gamma, +\rangle = \delta(\alpha - \gamma). \tag{8.29}$$

One reason why the unitarity of the S-matrix is interesting is that if we have a perturbative expansion for S then the unitarity constraint mixes different orders in perturbation theory, which sometimes lets us determine higher-order contributions from lower-order ones.

It will be useful in what follows to write down the Lorentz transformation of the S-matrix: from (8.19) we have (in more explicit notation)

$$\begin{aligned}\langle p'_1, \sigma'_1, n'_1; \ldots; p'_N, \sigma'_N, n'_N, -|p_1, \sigma_1, n_1; \ldots p_M, \sigma_M, n_M, +\rangle = &\prod_{i=1}^{N} \left(\sqrt{\frac{(\Lambda p'_i)^0}{p'^0_i}} \sum_{\overline{\sigma}'_i} D^{n'_i *}_{\overline{\sigma}'_i, \sigma'_i}(W_p) \right) \prod_{j=1}^{M} \left(\sqrt{\frac{(\Lambda p_j)^0}{p^0_j}} \sum_{\overline{\sigma}_j} D^{n_j}_{\overline{\sigma}_j, \sigma_j}(W_p) \right) \\ &\times \langle \Lambda p'_1, \overline{\sigma}'_1, n'_1; \ldots; \Lambda p'_N, \overline{\sigma}'_N, n'_N, -|\Lambda p_1, \overline{\sigma}_1, n_1; \ldots \Lambda p_M, \overline{\sigma}_M, n_M, +\rangle,\end{aligned} \tag{8.30}$$

with

$$W_p = L^{-1}_{\Lambda p_i} \Lambda L_{p_i}. \tag{8.31}$$

Due to the unitarity of the little group representations this formula simplifies if we take the absolute value squared and sum over all initial and final spins/helicities:

$$\begin{aligned}\sum_{\sigma'_1 \ldots \sigma'_N, \sigma_1 \ldots \sigma_M} |\langle p'_1, \sigma'_1, n'_1; \ldots; p'_N, \sigma'_N, n'_N, -|p_1, \sigma_1, n_1; \ldots p_M, \sigma_M, n_M, +\rangle|^2 = &\prod_{i=1}^{N} \left(\frac{(\Lambda p'_i)^0}{p'^0_i} \right) \prod_{j=1}^{M} \left(\frac{(\Lambda p_j)^0}{p^0_j} \right) \\ \times \sum_{\sigma'_1 \ldots \sigma'_N, \sigma_1 \ldots \sigma_M} &|\langle \Lambda p'_1, \sigma'_1, n'_1; \ldots; \Lambda p'_N, \sigma'_N, n'_N, -|\Lambda p_1, \sigma_1, n_1; \ldots \Lambda p_M, \sigma_M, n_M, +\rangle|^2.\end{aligned} \tag{8.32}$$

In other words in our condensed notation if we define

$$\widetilde{S}_{\beta\alpha} = \left(\prod_{i=1}^{N_\beta} \sqrt{2E_{\beta,i}} \right) \left(\prod_{j=1}^{N_\alpha} \sqrt{2E_{\alpha,j}} \right) S_{\beta\alpha}, \tag{8.33}$$

where $E_{\alpha,i}$ is the energy of the ith ingoing particle and $E_{\beta,j}$ is the energy of the jth outgoing particle, then the quantity

$$\sum_{\text{spin/helicity}} |\widetilde{S}_{\beta\alpha}|^2 \tag{8.34}$$

is Lorentz invariant.

8.4 Cross sections and decay rates

Before considering how to compute the S-matrix in quantum field theory, we will first make an aside to explain how to use it. Experimentalists typically do not report the outcomes of scattering experiments directly in terms of the S-matrix, but rather in terms of cross sections and decay rates. In this section we will work out how to relate these to the S-matrix.[60]

The first thing we need to understand is why experimentalists do not directly measure the S-matrix. The basic issue is that the S-matrix is always proportional to a momentum-conserving δ-function, so the quantity

[60] Cross sections will not be much of a focus in this class, but it is still good to know what they are!

$|S_{\beta\alpha}|^2$, which naively is the probability to find an out state $|\beta,-\rangle$ given that we start in an in state $|\alpha,+\rangle$, is infinite. To see the necessity of this δ-function, we can observe that

$$S_{\beta\alpha} = \langle\beta,-|e^{iP\cdot a}e^{-iP\cdot a}|\alpha,+\rangle = e^{i(p_\beta - p_\alpha)\cdot a}S_{\beta\alpha}, \tag{8.35}$$

for all spacetime vectors a, which is only possible if $S_{\beta\alpha}$ vanishes if $p_\alpha \neq p_\beta$. Here p_α and p_β are the total spacetime momenta of the in and out states (α and β label the states, they are NOT Lorentz indices). Moreover if we integrate $S_{\beta\alpha}$ against generic normalized wave packets we expect a finite and nonzero answer since we are computing an overlap of normalized states which have no reason to be orthogonal in general, so whatever support $S_{\beta\alpha}$ has when $p_\alpha = p_\beta$ must be strong enough to integrate to a nonzero result - in other words there must be a δ-function. It is convenient to extract this δ-function, and also the non-interacting contribution $\delta(\beta-\alpha)$, in hopes that the remaining part of $S_{\beta\alpha}$ is nonsingular:

$$S_{\beta\alpha} = \delta(\beta-\alpha) + i\times(2\pi)^d\delta^d(p_\beta - p_\alpha)\mathcal{M}_{\beta\alpha}. \tag{8.36}$$

The factor of i here is conventional, I offer no explanation for it. $\mathcal{M}_{\beta\alpha}$ can still have further δ-function singularities, but only when a subset of the particles in α has exactly the same spacetime momenta as a subset of the particles in β. These additional singularities can be removed by defining a "connected" S-matrix in exactly the same way we did for correlation functions, to avoid this we will just restrict to studying the S-matrix away from these special kinematic points.

To understand what to do about the divergence of $|S_{\beta\alpha}|^2$, we first need to realize that it is an *infrared* divergence: the quantity $\delta^{d-1}(0)$ is infinity in momentum space because we are working in infinite volume. In finite volume the inner product of our one-particle states is

$$\langle p',\sigma'|p,\sigma\rangle = \delta_{\sigma\sigma'}\int d^{d-1}x e^{i(\vec{p}-\vec{p}')\cdot\vec{x}} = V\delta_{\sigma\sigma'}\delta_{\vec{p}\vec{p}'}, \tag{8.37}$$

so we can formally interpret $\delta^{d-1}(0)$ as the volume of space. The square of the S-matrix is therefore diverging because we defined it as an overlap of states whose norms are order $\sqrt{V}$ instead of states whose norms are one. The principled way to fix this is to only consider scattering of normalizeable wave packets. In particular the quantity

$$\Big|\int d\alpha g(\alpha)S_{\beta\alpha}\Big|^2 \tag{8.38}$$

with

$$\int d\alpha |g(\alpha)|^2 = 1 \tag{8.39}$$

should be finite. A somewhat lazier approach, but which also works and leads more quickly to the same answer, is to just work in finite volume and stick to momentum eigenstates. Indeed in finite volume we can define properly normalized in and out states

$$|\alpha,\pm\rangle_V = \frac{1}{V^{N_\alpha/2}}|\alpha,\pm\rangle, \tag{8.40}$$

in terms of which the differential transition probability from α to β is

$$dP(\alpha\to\beta) = |\langle\beta,-|\alpha,+\rangle_V|^2 dN_\beta = \frac{|S_{\beta\alpha}|^2}{V^{N_\alpha}}d\beta, \tag{8.41}$$

where

$$dN_\beta = V^{N_\beta}d\beta \tag{8.42}$$

is the number of states in the infinesimal phase space window $d\beta$. To derive this, recall that for a single particle momentum in finite volume we have

$$dN = V\frac{d^{d-1}p}{(2\pi)^{d-1}} \tag{8.43}$$

since momenta in the box are quantized (for example if we take it to be square torus of length L) as

$$\vec{p} = \frac{2\pi\vec{n}}{L} \tag{8.44}$$

with $\vec{n}$ a spatial vector of integers. If we avoid choices of α and β where $\mathcal{M}_{\beta\alpha}$ has additional δ-functions (or just focus on connected scattering), then we can write this as

$$dP(\alpha \to \beta) = V^{-N_\alpha} \left((2\pi)^d \delta^d(p_\beta - p_\alpha)\right)^2 |\mathcal{M}_{\beta\alpha}|^2 d\beta. \tag{8.45}$$

To deal with the square of the δ-function, we notice that in finite volume we can write it as

$$\begin{aligned} \left((2\pi)^d \delta^d(p_\beta - p_\alpha)\right)^2 &= \left(V \delta_{\vec{p}_\alpha \vec{p}_\beta} T \delta_{E_\alpha E_\beta}\right)^2 \\ &= VT \times \left(VT \delta_{\vec{p}_\alpha \vec{p}_\beta} \delta_{E_\alpha E_\beta}\right) \\ &= VT (2\pi)^d \delta^d(p_\beta - p_\alpha), \end{aligned} \tag{8.46}$$

where T is the total time elapsed in the scattering process (i.e. we work in a "time box" as well as a spatial box). The transition rate, which is the transition probability per unit time, is therefore given by

$$d\Gamma(\alpha \to \beta) = V^{1-N_\alpha} (2\pi)^d \delta^d(p_\beta - p_\alpha) |\mathcal{M}_{\beta\alpha}|^2 d\beta. \tag{8.47}$$

We've now succeeding in pushing all infrared divergences into an overall power of the volume, as everything else appearing here is sensible in the large volume limit.

To proceed further, we need to think a bit about how to connect this setup to what experimentalists actually do. The easiest case is $N_\alpha = 1$, for which the power of V just cancels. We are then studying the decay of an unstable particle, whose differential decay rate into a final state β is apparently given by

$$d\Gamma(\alpha \to \beta) = (2\pi)^d \delta^d(p_\beta - p_\alpha) |\mathcal{M}_{\beta\alpha}|^2 d\beta. \tag{8.48}$$

I must confess however that this formula (although correct when interpreted properly) is a bit of a cheat: an unstable particle isn't really a one-particle state of the theory in infinite volume, so we can't really interpret $\mathcal{M}_{\beta\alpha}$ as part of the S-matrix. On the other hand if the total decay rate

$$\Gamma = \int d\beta (2\pi)^d \delta^d(p_\beta - p_\alpha) |\mathcal{M}_{\beta\alpha}|^2 \tag{8.49}$$

is small compared to the inverse of our time interval T then we can effectively treat the particle as stable in our box setup. This formula should therefore be correct as long as the lifetime of the particle is long compared to all other scales in the problem.[61]

A somewhat more complicated (but also better defined) case is $N_\alpha = 2$, which is the scattering of two particles to many. The classic setup for this experiment is shown in figure 24: we have a beam of incident identical particles with momentum p_1 aimed at a target particle with momentum $p_2 = (m_2, \vec{0})$. The natural thing to measure is the differential rate for scattering into the out state β divided by the incident flux, which is called the **differential cross section**:

$$d\sigma(\alpha \to \beta) = \frac{d\Gamma(\alpha \to \beta)}{f_\alpha}. \tag{8.50}$$

Since we are working in the rest frame of the target particle, the incident flux f_α is given by

$$f_\alpha = |v_1| \rho_1 \tag{8.51}$$

[61] There is a more rigorous justification for this based on defining an unstable particle as a resonance in the scattering of stable particles and $M_{\alpha\beta}$ in terms of the residue of this resonance, but we won't explore it here. If you just compute $M_{\alpha\beta}$ using the Feynman rules we'll find below extrapolated to the case of one ingoing particle then you will get the right answer.

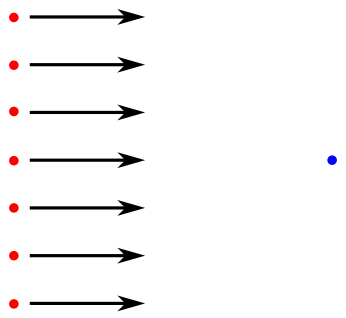


Figure 24: Fixed-target scattering: an incident beam of identical particles (shown in red) with identical momenta are scattered off of a single target particle (shown in blue). What is the rate at which each out state β is produced per unit incident flux? The answer to this question is the differential cross section $\frac{d\sigma}{d\beta}$.

where $\vec{v}_1$ is the velocity of incident particles and ρ_1 is their density. In a general Lorentz frame (which after all we had better include since the target particle could be massless) the flux is instead defined to be

$$f_\alpha = u_\alpha \rho_1, \tag{8.52}$$

where

$$u_\alpha = \frac{\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}}{E_1 E_2} \tag{8.53}$$

is called the **relative velocity**. You will show in the homework that when $\vec{p}_1$ and $\vec{p}_2$ are collinear (i.e. proportional to each other) then we have

$$u_\alpha = |\vec{v}_1 - \vec{v}_2|. \tag{8.54}$$

For general p_1, p_2 the motivation for this definition of flux is that it makes the spin-summed differential cross section Lorentz invariant, as we will see in a moment. Returning now to our box setup with $N_\alpha = 2$, our one-particle box states $|\alpha, \mp\rangle_V$ are properly normalized so the number of particles in the box in such a state is one. The particle density is therefore $1/V$, and we can think of the transition rate (8.47) as arising from a beam of particle one with density $\rho_1 = \frac{1}{V}$ and flux u_α/V scattering off of the other particle in the box (wherever it is) just as in figure 24. The differential cross section is therefore given by

$$d\sigma(\alpha \to \beta) = u_\alpha^{-1} (2\pi)^d \delta^d(p_\beta - p_\alpha) |\mathcal{M}_{\beta\alpha}|^2 d\beta. \tag{8.55}$$

This formula is used anytime someone wants to compare a theoretical calculation of two-particle scattering to experiment!

Let's briefly consider the Lorentz transformation properties of the differential cross section. $\mathcal{M}_{\beta\alpha}$ has the same Lorentz transformation properties as $S_{\beta\alpha}$, and we saw in equation (8.32) that the Lorentz transformations of $|S_{\beta\alpha}|^2$ is simple once we sum over spins/helicities and multiply by the product of initial and final energies of each particle. More concretely, if we define

$$\widetilde{\mathcal{M}}_{\beta\alpha} = \left(\prod_{i=1}^{N_\beta} \sqrt{2E_{\beta,i}} \right) \left(\prod_{j=1}^{N_\alpha} \sqrt{2E_{\alpha,j}} \right) \mathcal{M}_{\beta\alpha}, \tag{8.56}$$

then

$$\sum_{\text{spin/helicity}} |\widetilde{\mathcal{M}}_{\beta\alpha}|^2 \tag{8.57}$$

is Lorentz invariant. We also saw in section 3 (and mentioned below equation (8.10)) that the quantity

$$\widetilde{d\beta} = \frac{d\beta}{\prod_{i=1}^{N} 2E_{\beta,i}} \tag{8.58}$$

is Lorentz invariant, where again $E_{\beta,i}$ is the energy of the ith particle in the final state. We therefore are motivated to sum (8.55) over initial and final spins/helicities and then rewrite it as

$$\sum_{\text{spin/helicity}} d\sigma(\alpha \to \beta) = \frac{1}{4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}} \times (2\pi)^d \delta^d(p_\beta - p_\alpha) \times \sum_{\text{spin/helicity}} |\widetilde{M}_{\beta\alpha}|^2 \times \widetilde{d\beta}, \tag{8.59}$$

where the right-hand side is now a product of manifestly Lorentz-invariant quantities (in particular because of our definition (8.53) of the relative velocity). Thus we see that the spin-summed differential cross section is Lorentz invariant!

You perhaps are wondering about the physical motivation for summing over initial and final spins/helicities. In fact what we really should do is sum over final spins/helicities and *average* over initial spins/helicities, for the following reasons:

- Typically the method for preparing a beam of particles does not preferentially treat one spin/helicity state over another. We therefore should expect the initial state in a scattering process to be a mixed quantum state where all spin/helicity configurations are equally likely, in which case we should average over initial spins/helicities in the transition rate.
- Typically in measuring the final state we do not get a good measurement of the spins/helicities of the particles. We should therefore sum over these to compute the transition rate which does not distinguish between different spins/helicities.

In situations where either of these statements is not the case, then we need to deal with the full differential cross section.

8.5 Unitarity and the optical theorem

The unitarity of the S-matrix has some nice consequences for cross sections and decay rates. We can first observe that

$$\begin{aligned}
\delta(\gamma - \alpha) &= \int d\beta S^*_{\beta\gamma} S_{\beta\alpha} \\
&= \int d\beta \left(\delta(\beta - \gamma) - i(2\pi)^d \delta^d(p_\beta - p_\gamma)\mathcal{M}^*_{\beta\gamma}\right)\left(\delta(\beta - \alpha) + i(2\pi)^d \delta^d(p_\beta - p_\alpha)\mathcal{M}_{\beta\alpha}\right) \\
&= \delta(\gamma - \alpha) + (2\pi)^d \delta^d(p_\gamma - p_\alpha)\left(i\mathcal{M}_{\gamma\alpha} - i\mathcal{M}^*_{\alpha\gamma} + \int d\beta (2\pi)^d \delta^d(p_\beta - p_\alpha)\mathcal{M}^*_{\beta\gamma}\mathcal{M}_{\beta\alpha}\right).
\end{aligned} \tag{8.60}$$

The $\delta(\gamma - \alpha)$ terms cancel on both sides, so the remaining equality tells this that for any states α and γ such that $p_\gamma = p_\alpha$ we should have

$$i\mathcal{M}_{\gamma\alpha} - i\mathcal{M}^*_{\alpha\gamma} + \int d\beta (2\pi)^d \delta^d(p_\beta - p_\alpha)\mathcal{M}^*_{\beta\gamma}\mathcal{M}_{\beta\alpha} = 0. \tag{8.61}$$

In particular this is true if $\alpha = \gamma$, in which case we find

$$\operatorname{Im}\mathcal{M}_{\alpha\alpha} = \frac{1}{2}\int d\beta (2\pi)^d \delta^d(p_\beta - p_\alpha)|\mathcal{M}_{\beta\alpha}|^2. \tag{8.62}$$

From equation (8.47) we can rewrite the right-hand side of this equation in terms of the total transition rate

$$\Gamma(\alpha) = \int d\beta \frac{d\Gamma(\alpha \to \beta)}{d\beta}, \tag{8.63}$$

which gives

$$\Gamma(\alpha) = 2V^{1-N_\alpha} \operatorname{Im}\mathcal{M}_{\alpha\alpha}. \tag{8.64}$$

In particular for $N_\alpha = 1$ we have

$$\Gamma(\alpha) = 2\mathrm{Im}\,\mathcal{M}_{\alpha\alpha}, \tag{8.65}$$

so the lifetime of an unstable particle is just two times the imaginary part of its forward scattering amplitude. For $N_\alpha = 2$ we can rewrite things in terms of the total cross section

$$\sigma(\alpha) = \int d\beta \frac{d\sigma(\alpha \to \beta)}{d\beta} = \frac{V\Gamma(\alpha)}{u_\alpha}, \tag{8.66}$$

which gives

$$\sigma(\alpha) = \frac{2}{u_\alpha}\mathrm{Im}\,\mathcal{M}_{\alpha\alpha}. \tag{8.67}$$

This last result is called the **optical theorem**. Both (8.65) and (8.67) express the idea that by unitary any decay or scattering which is possible must decrease the probability that no scattering happens, which is certainly a reasonable thing to expect!

8.6 Homework

1. The helicity of a photon in general dimensions is that of the vector representation of $SO(d-2)$, so a photon in d dimensions has $d-2$ independent values of σ (i.e. independent polarization states). How would you interpret this in $1+1$ and $2+1$ dimensions? Hint: think about how the classical polarization of an EM wave should work in these dimensions.

2. The helicity of a graviton in general dimensions is that of the representation of $SO(d-2)$ furnished by a symmetric traceless two-tensor h_{ij}. How many independent polarizations does a graviton have in d spacetime dimensions?

3. Check that our resolution (8.22) of the identity indeed acts as the identity on free scalar two-particle states of the form $a^\dagger_{\vec{p}} a^\dagger_{\vec{p}'} |\Omega\rangle$.

4. Consider the scattering of a non-relativistic quantum particle off of a δ-function potential, with Hamiltonian
$$H = \frac{p^2}{2m} + V_0 \delta(x). \tag{8.68}$$
You can assume $V_0 > 0$. Give explicit formulas for the In and Out states of this theory, and compute the S-matrix. Hint: to get a complete basis you need to consider incident waves from both the left and the right, and you need to make sure your states are eigenstates of the full Hamiltonian. This theory arises from the more general scattering theory we've considered in the limit where one of the incident particles is infinitely massive and its interaction with the other particle has infinitely-short range.

5. Confirm that (8.32) follows from (8.30) by the unitarity of the little group representations.

6. Show that the relative velocity (8.53) becomes (8.54) in the "collinear" situation where the spatial momenta are proportional to each other (possibly with opposite sign).

9 Scattering from correlation functions in quantum field theory

In the last section we studied scattering theory in quantum mechanics. In particular we encountered the idea of "in" and "out" states $|\alpha, \pm\rangle$ and the S-matrix

$$S_{\beta\alpha} = \langle \beta, -|\alpha, +\rangle = \delta(\beta - \alpha) + i(2\pi)^d \delta^d(p_\beta - p_\alpha) M_{\beta\alpha}. \tag{9.1}$$

We also learned how to convert the S-matrix into observable transition rates such as the differential cross-section

$$d\sigma(\alpha \to \beta) = u_\alpha^{-1} (2\pi)^d \delta^d(p_\beta - p_\alpha) |M_{\beta\alpha}|^2 d\beta \tag{9.2}$$

for two-particle scattering. In this section we return to quantum field theory, looking to answer two questions:

1. How can we tell when an interacting quantum field theory has a scattering description in terms of particles?
2. In quantum field theories which do have a scattering description, how can we compute the S-matrix starting from the correlation functions?

We will see that the answer to the first of these questions is that the existence of particles in a quantum field theory leads to poles in the Fourier transform of its two-point functions, and the answer to the second is given by the **LSZ reduction formula**. Once we establish these tools, we will finally be in a position to compute genuine observables in interacting quantum field theories for comparison with experiment![62]

9.1 Exact two-point function in interacting quantum field theory

As a warmup to our discussion of scattering, we'll begin with a general discussion of the structure of the two-point function of Heisenberg operators in interacting quantum field theory. The idea is to consider

$$\langle T O_2^{a_2}(k_2) O_1^{a_1}(k_1) \rangle_\epsilon = \int dx_1 \int dx_2 e^{-i(k_2 \cdot x_2 + k_1 \cdot x_1)} e^{-\epsilon |t_2 - t_1|} \langle \Omega | T O_2^{a_2}(x_2) O_1^{a_1}(x_1) | \Omega \rangle, \tag{9.3}$$

where $O_1^{a_1}$ and $O_2^{a_2}$ are local operators that transform in irreducible representations of the Lorentz group:

$$U^\dagger(\Lambda) O_i^{a_i}(x) U(\Lambda) = \sum_{b_i} D_i^{a_i b_i}(\Lambda) O_i^{b_i}(\Lambda^{-1} x), \tag{9.4}$$

with $i = 1, 2$. We will assume that

$$\langle \Omega | O_i^{a_i}(x) | \Omega \rangle = 0, \tag{9.5}$$

which follows from Lorentz symmetry unless O_i is a scalar and which we can achieve in the scalar case by redefining O_i by an additive shift. The convergence factor $e^{-\epsilon|t_2 - t_1|}$ is useful to include, as we will soon see.

In a theory with a scattering description we can evaluate this two-point function by inserting a complete set of in/out states:

$$\begin{aligned}\langle T O_2^{a_2}(k_2) O_1^{a_1}(k_1) \rangle_\epsilon = \int d\alpha \int dx_1 \int dx_2 e^{-i(k_2 \cdot x_2 + k_1 \cdot x_1)} e^{-\epsilon|t_2 - t_1|} \Big(&\Theta(t_2 - t_1) \langle \Omega | O_2^{a_2}(x_2) | \alpha, \pm \rangle \langle \alpha, \pm | O_1^{a_1}(x_1) | \Omega \rangle \\ &+ (-1)^{f_O} \Theta(t_1 - t_2) \langle \Omega | O_1^{a_1}(x_1) | \alpha, \pm \rangle \langle \alpha, \pm | O_2^{a_2}(x_2) | \Omega \rangle \Big),\end{aligned} \tag{9.6}$$

where $f_O = 1$ if O_1 and O_2 are fermionic and $f_O = 0$ if they are bosonic. We can simplify the matrix elements appearing here by noting that

$$O_i^{a_i}(x_i)) = e^{-iP \cdot x_i} O_i^{a_i}(0) e^{iP \cdot x_i}, \tag{9.7}$$

[62] My apologies that it is taking such a long time to get there. The basic curse of quantum field theory pedagogy is that if you want to get to practical applications quickly you won't understand what you are doing, and in this class we've decided to take our time and learn things properly.

and thus

$$\begin{aligned}\langle\Omega|O_i^{a_i}(x_i)|\alpha,\pm\rangle &= e^{ip_\alpha\cdot x_i}\langle\Omega|O_i^{a_i}(0)|\alpha,\pm\rangle\\ \langle\alpha,\pm|O_i^{a_i}(x_i)|\Omega\rangle &= e^{-ip_\alpha\cdot x_i}\langle\alpha,\pm|O_i^{a_i}(x_i)|\Omega\rangle,\end{aligned}\tag{9.8}$$

from which we have

$$\begin{aligned}\langle TO_2^{a_2}(k_2)O_1^{a_1}(k_1)\rangle_\epsilon = \int d\alpha\Bigg[&\int_{t_2>t_1} dx_1dx_2e^{-i(k_1+p_\alpha)\cdot x_1-i(k_2-p_\alpha)\cdot x_2-\epsilon(t_2-t_1)}\langle\Omega|O_2^{a_2}(0)|\alpha,\pm\rangle\langle\alpha,\pm|O_1^{a_1}(0)|\Omega\rangle\\ &+(-1)^{f_O}\int_{t_2<t_1} dx_1dx_2e^{-i(k_1-p_\alpha)\cdot x_1-i(k_2+p_\alpha)\cdot x_2+\epsilon(t_2-t_1)}\langle\Omega|O_1^{a_1}(0)|\alpha,\pm\rangle\langle\alpha,\pm|O_2^{a_2}(0)|\Omega\rangle\Bigg].\end{aligned}\tag{9.9}$$

The spatial integrals here give simple δ-functions, but the integrals over t_1 and t_2 are a bit trickier:

$$\begin{aligned}\int_{-\infty}^{\infty}dt_1\int_{t_1}^{\infty}dt_2\,e^{i(k_1^0+p_\alpha^0-i\epsilon)t_1+i(k_2^0-p_\alpha^0+i\epsilon)t_2} &= \int_{-\infty}^{\infty}dt_1\,e^{i(k_1^0+k_2^0)t_1}\frac{i}{k_2^0-p_\alpha^0+i\epsilon}\\ &= 2\pi\delta(k_1^0+k_2^0)\times\frac{i}{k_2^0-p_\alpha^0+i\epsilon}\end{aligned}\tag{9.10}$$

and

$$\begin{aligned}\int_{-\infty}^{\infty}dt_2\int_{t_2}^{\infty}dt_1\,e^{i(k_1^0-p_\alpha^0+i\epsilon)t_1+i(k_2^0+p_\alpha^0-i\epsilon)t_2} &= \int_{-\infty}^{\infty}dt_2\,e^{i(k_1^0+k_2^0)t_2}\frac{i}{k_1^0-p_\alpha^0+i\epsilon}\\ &= 2\pi\delta(k_1^0+k_2^0)\times\frac{i}{k_1^0-p_\alpha^0+i\epsilon}.\end{aligned}\tag{9.11}$$

We thus have

$$\begin{aligned}\langle TO_2^{a_2}(k_2)O_1^{a_1}(k_1)\rangle_\epsilon = \int d\alpha\Bigg[&(2\pi)^{d-1}\delta^{d-1}(\vec{k_2}-\vec{p}_\alpha)(2\pi)^{d-1}\delta^{d-1}(\vec{k_1}+\vec{p}_\alpha)2\pi\delta(k_1^0+k_2^0)\frac{i}{k_2^0-p_\alpha^0+i\epsilon}\langle\Omega|O_2^{a_2}(0)|\alpha,\pm\rangle\langle\alpha,\pm|O_1^{a_1}(0)|\Omega\rangle\\ &+(-1)^{f_O}(2\pi)^{d-1}\delta^{d-1}(\vec{k_2}+\vec{p}_\alpha)(2\pi)^{d-1}\delta^{d-1}(\vec{k_1}-\vec{p}_\alpha)2\pi\delta(k_1^0+k_2^0)\frac{i}{k_1^0-p_\alpha^0+i\epsilon}\langle\Omega|O_1^{a_1}(0)|\alpha,\pm\rangle\langle\alpha,\pm|O_2^{a_2}(0)|\Omega\rangle\Bigg]\end{aligned}\tag{9.12}$$

The key point to notice here are the pole factors $\frac{i}{k_2^0-p_\alpha^0+i\epsilon}$ and $\frac{i}{k_1^0-p_\alpha^0+i\epsilon}$: what we will now show is that the contribution to the α integral coming from one-particle states turns these poles into poles of the momentum-space correlation function $\langle TO_2^{a_2}(k_2)O_1^{a_1}(k_1)\rangle_\epsilon$. For one-particle states we simply have

$$\int d\alpha = \sum_{\sigma,n}\int\frac{d^{d-1}p}{(2\pi)^{d-1}}\tag{9.13}$$

and

$$p_\alpha = (\omega_{n,\vec{p}},\vec{p}),\tag{9.14}$$

with

$$\omega_{n,\vec{p}} = \sqrt{|p|^2+m_n^2}.\tag{9.15}$$

Evaluating the momentum integrals we thus find

$$\begin{aligned}\langle TO_2^{a_2}(k_2)O_1^{a_1}(k_1)\rangle_\epsilon \supset (2\pi)^d\delta(k_2+k_1)\sum_{n,\sigma}\Bigg(&\frac{i}{k_2^0-\omega_{n,\vec{k}_2}+i\epsilon}\langle\Omega|O_2^{a_2}(0)|k_2,\sigma,n\rangle\langle k_2,\sigma,n|O_1^{a_1}(0)|\Omega\rangle\\ &+(-1)^{f_O}\frac{i}{k_1^0-\omega_{n,\vec{k}_1}+i\epsilon}\langle\Omega|O_1^{a_1}(0)|k_1,\sigma,n\rangle\langle k_1,\sigma,n|O_2^{a_2}(0)|\Omega\rangle\Bigg).\end{aligned}\tag{9.16}$$

The two-point function therefore has a pole whenever the external momenta go "on-shell" for any particle species n for which the matrix elements do not vanish. This may seem technical, but in fact it is profound:

- The way we tell if a quantum field theory has particles is we look for on-shell poles in the Fourier transform of the time-ordered two-point function: they exist if and only if the theory has one-particle states, and we can determine the masses of the particles from the locations of the poles.

In particular I want to emphasize that nothing in this derivation assumed that the particles are "fundamental" in the sense of being associated with fields in the Lagrangian, for example in QED the hydrogen atom contributes poles to two-point functions and the same is true for protons in QCD.

What about other contributions to the α integral? The vacuum contribution vanishes because of (9.5). We will not try to show it systematically, but the multi-particle states only contribute branch cuts which in massive theories are away from the "on-shell" poles at $k_2^0 = \pm\omega_{\vec{k}_2}$ so the pole contributions comes only from one-particle states. The basic idea is that the integral

$$\int_{z_{min}}^{z_{max}} \frac{dz}{k^0 - z + i\epsilon} = \log\left(\frac{k^0 - z_{min} + i\epsilon}{k^0 - z_{max} + i\epsilon}\right) \tag{9.17}$$

has only logarithmic branch singularities, where here z stands for the integral over the additional momenta in a multi-particle state and we'd have $z_{min} = m_n + \omega_{n,\vec{k}}$ so the branch point is at $k^0 = m_n + \omega_{n,\vec{k}}$ which is different from $\omega_{n,\vec{k}}$ unless $m_n = 0$. When $m_n = 0$ some more care is needed, but in essence we can still distinguish a pole from a branch point even if they are right on top of each other.

9.2 Matrix elements

Let's now say something about the matrix elements $\langle\Omega|O^a(0)|\vec{k},\sigma,n\rangle$ and $\langle\vec{k},\sigma,n|O^a(0)|\Omega\rangle$ appearing in the residues of the one-particle poles we just found (to make life simpler we have dropped the i indices for now since we are only considering the Os one at a time). These are strongly constrained by Lorentz symmetry, as we will now show.[63] Taking Λ to be an element of the little group for k, i.e. a Lorentz transformation for which $\Lambda k = k$, we have

$$\langle\Omega|O^a(0)|k,\sigma,n\rangle = \langle\Omega|U(\Lambda)O^a(0)U^\dagger(\Lambda)U(\Lambda)|k,\sigma,n\rangle = \sum_{a',\sigma'} D^{aa'}(\Lambda^{-1})\hat{D}^n_{\sigma'\sigma}(\Lambda)\langle\Omega|O^{a'}(0)|k,\sigma',n\rangle. \tag{9.18}$$

Here I've put a hat on the little group representation matrices $\hat{D}^n$ to distinguish them from the operator representation matrices D appearing in (9.4), and used the little group transformation

$$U(\Lambda)|k,\sigma,n\rangle = \sum_{\sigma'} D^n_{\sigma'\sigma}(\Lambda)|k,\sigma',n\rangle. \tag{9.19}$$

Viewing $\langle\Omega|O^a(0)|k,\sigma,n\rangle$ as a matrix $T^a{}_\sigma$, we can write this as

$$T = D(\Lambda^{-1})T\hat{D}^n(\Lambda), \tag{9.20}$$

or equivalently

$$D(\Lambda)T = T\hat{D}^n(\Lambda). \tag{9.21}$$

Similarly we have

$$\langle k,\sigma,n|O^a(0)|\Omega\rangle = \sum_{a',\sigma'} D^{aa'}(\Lambda^{-1})\hat{D}^{n*}_{\sigma'\sigma}(\Lambda)\langle k,\sigma',n|O^{a'}(0)|\Omega\rangle, \tag{9.22}$$

and so the matrix

$$\widetilde{T}^a{}_\sigma = \langle k,\sigma,n|O^a(0)|\Omega\rangle \tag{9.23}$$

[63] This argument is somewhat mathematical, the answer is in equation (9.41) if you want to just trust me.

obeys

$$D(\Lambda)\widetilde{T} = \widetilde{T}\hat{D}^{n*}(\Lambda). \tag{9.24}$$

We have taken both D and $\hat{D}^n$ to form irreducible representations of the little group, and equations (9.21) and (9.24) can be interpreted in terms of group theory as saying that the matrices T and $\widetilde{T}$ are **intertwiners** between these irreducible representations. More precisely, T is an intertwiner from the $\hat{D}$ representation to the D representation and $\widetilde{T}$ is an intertwiner from the conjugate of the $\hat{D}$ representation to the D representation. It is a theorem in group theory that nonzero intertwiners between finite-dimensional irreducible representations exist only if the representations are equivalent by a similarity transformation, and moreover that even in this case the intertwiner is unique up to an overall constant factor.[64]

In fact we have already discussed these intertwiners, in the context of free field theory. Given a particle type n, we argued in the first section that to make a relativistic quantum theory we should begin by constructing a free field which annihilates that particle with the form

$$\Phi^a(x) = \sum_\sigma \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{1}{\sqrt{2\omega_{n,\vec{p}}}} \left[u^a(\vec{p},\sigma,n) e^{ip\cdot x} a_{\vec{p}\sigma n} + v^a(\vec{p},\sigma,n^c) e^{-ip\cdot x} a^\dagger_{\vec{p}\sigma n^c} \right]. \tag{9.25}$$

Here n^c is the antiparticle of n (which may coincide with n or not), and the functions u^a and v^a are chosen so that the field commutes/anticommutes at spacelike separation and we have the Lorentz transformation

$$U^\dagger(\Lambda)\Phi^a(x)U(\Lambda) = \sum_{a'} D^{aa'}(\Lambda)\Phi^{a'}(\Lambda^{-1}x). \tag{9.26}$$

Due to our little group transformation (9.19) we see that the creation and annihilation operators in this field must transform as

$$\begin{aligned} U(\Lambda)a^\dagger_{\vec{p}\sigma n^c}U(\Lambda^{-1}) &= \sum_{\sigma'} \hat{D}^{n^c}_{\sigma'\sigma}(\Lambda) a^\dagger_{\vec{p}\sigma' n^c} \\ U(\Lambda)a_{\vec{p}\sigma n}U(\Lambda^{-1}) &= \sum_{\sigma'} \hat{D}^{n*}_{\sigma'\sigma}(\Lambda) a_{\vec{p}\sigma' n}, \end{aligned} \tag{9.27}$$

where Λ is in the little group for p. In order for the field to have the Lorentz transformation (9.26), these transformations must combine with u^a and v^a to give

$$\begin{aligned} \sum_{\sigma'} u^a(\vec{p},\sigma',n)\hat{D}^n_{\sigma',\sigma}(\Lambda) &= \sum_{a'} D^{aa'}(\Lambda) u^{a'}(\vec{p},\sigma,n) \\ \sum_{\sigma'} v^a(\vec{p},\sigma',n)\hat{D}^{n^c*}_{\sigma'\sigma}(\Lambda) &= \sum_{a'} D^{aa'}(\Lambda) v^{a'}(\vec{p},\sigma,n^c). \end{aligned} \tag{9.28}$$

In other words u^a and v^a are intertwiners, so by the uniqueness of intertwiners they must be proportional to T and $\widetilde{T}$ respectively:

$$\begin{aligned} \langle \Omega | O^a(0) | k,\sigma,n\rangle &= A_n(\vec{k})\, u^a(\vec{k},\sigma,n) \\ \langle k,\sigma,n^c | O^a(0) | \Omega\rangle &= \widetilde{A}_{n^c}(\vec{k})\, v^a(\vec{k},\sigma,n^c). \end{aligned} \tag{9.29}$$

We can learn more about the proportionality functions A_n and $\widetilde{A}_{n^c}$ by considering general Lorentz transformations.[65] For general Λ instead of (9.18) we have

$$\langle \Omega | O^a(0) | k,\sigma,n\rangle = \sqrt{\frac{\omega_{n,\vec{k}_\Lambda}}{\omega_{\vec{k}}}} \sum_{a',\sigma'} D^{aa'}(\Lambda^{-1})\hat{D}^n_{\sigma'\sigma}(L^{-1}_{\Lambda k}\Lambda L_k) \langle \Omega | O^{a'}(0) | \Lambda k,\sigma',n\rangle, \tag{9.30}$$

[64] If you know a little representation theory the proof is a fairly straightforward application of Schur's lemmas, see theorem A.5 in my long paper with Ooguri.

[65] You may wonder why we restricted to the little group above if we need the general transformation anyways here: the reason is that the uniqueness result for intertwiners only applies to finite-dimensional irreducible representations, and it is only the little group which acts in a finite-dimensional representation on particle states.

which from (9.29) tells us that

$$A_n(\vec{k})u^a(\vec{k},\sigma,n) = \sqrt{\frac{\omega_{n,\vec{k}_\Lambda}}{\omega_{\vec{k}}}}A_n(\vec{k}_\Lambda)\sum_{a',\sigma'} D^{aa'}(\Lambda^{-1})\hat{D}^n_{\sigma'\sigma}(L^{-1}_{\Lambda k}\Lambda L_k)u^{a'}(\vec{k}_\Lambda,\sigma',n). \tag{9.31}$$

Here we are using the notation that $\Lambda(\omega_{n,\vec{k}},\vec{k}) = (\omega_{n,\vec{k}_\Lambda},\vec{k}_\Lambda)$, and as in the previous section L_p is the Lorentz transformation which maps a reference momentum to p. Similarly from the transformation of $\langle k,\sigma,n^c|O^a(0)|\Omega\rangle$ we have

$$\widetilde{A}_{n^c}(\vec{k})v^a(\vec{k},\sigma,n^c) = \sqrt{\frac{\omega_{n,\vec{k}_\Lambda}}{\omega_{\vec{k}}}}\widetilde{A}_{n^c}(\vec{k}_\Lambda)\sum_{a',\sigma'} D^{aa'}(\Lambda^{-1})\hat{D}^{n^c*}_{\sigma'\sigma}(L^{-1}_{\Lambda k}\Lambda L_k)v^{a'}(\vec{k}_\Lambda,\sigma',n^c). \tag{9.32}$$

We can simplify these by noting that in general the transformations of the creation and annihilation operators are

$$\begin{aligned} U(\Lambda)a^\dagger_{\vec{p}\sigma n^c}U(\Lambda^{-1}) &= \sqrt{\frac{\omega_{n,\vec{p}_\Lambda}}{\omega_{\vec{p}}}}\sum_{\sigma'}\hat{D}^{n^c}_{\sigma'\sigma}(L^{-1}_{\Lambda p}\Lambda L_p)a^\dagger_{\vec{p}_\Lambda\sigma' n^c} \\ U(\Lambda)a_{\vec{p}\sigma n}U(\Lambda^{-1}) &= \sqrt{\frac{\omega_{n,\vec{p}_\Lambda}}{\omega_{\vec{p}}}}\sum_{\sigma'}\hat{D}^{n*}_{\sigma'\sigma}(L^{-1}_{\Lambda p}\Lambda L_p)a_{\vec{p}_\Lambda\sigma' n}, \end{aligned} \tag{9.33}$$

so the Lorentz transformation (9.26) for the free field implies (extra credit homework) that we have

$$\begin{aligned} \sum_{\sigma'} u^a(\vec{k}_\Lambda,\sigma',n)\hat{D}^n_{\sigma'\sigma}(L^{-1}_{\Lambda k}\Lambda L_k) &= \sum_{a'} D^{aa'}(\Lambda)u^{a'}(\vec{k},\sigma,n) \\ \sum_{\sigma'} v^a(\vec{k}_\Lambda,\sigma',n^c)\hat{D}^{n^c*}_{\sigma'\sigma}(L^{-1}_{\Lambda k}\Lambda L_k) &= \sum_{a'} D^{aa'}(\Lambda)v^{a'}(\vec{k},\sigma,n^c). \end{aligned} \tag{9.34}$$

Using these in (9.31), (9.32) we see that we must have

$$\begin{aligned} \sqrt{\omega_{n,\vec{k}}}A_n(\vec{k}) &= \sqrt{\omega_{n,\vec{k}_\Lambda}}A_n(\vec{k}_\Lambda) \\ \sqrt{\omega_{n,\vec{k}}}\widetilde{A}_{n^c}(\vec{k}) &= \sqrt{\omega_{n,\vec{k}_\Lambda}}\widetilde{A}_{n^c}(\vec{k}_\Lambda), \end{aligned} \tag{9.35}$$

and thus

$$\begin{aligned} A_n(\vec{k}) &= \frac{Z_n}{\sqrt{2\omega_{n,\vec{k}}}} \\ \widetilde{A}_{n^c}(\vec{k}) &= \frac{\widetilde{Z}_{n^c}}{\sqrt{2\omega_{n,\vec{k}}}} \end{aligned} \tag{9.36}$$

where Z_n and $\widetilde{Z}_{n^c}$ are pure numbers. Moreover Z_n and $\widetilde{Z}_{n^c}$ are related by $\mathcal{CRT}$ symmetry: by $\mathcal{CRT}$ we have

$$\langle\Omega|O^a(0)|k,\sigma,n\rangle = \langle\Theta^\dagger_{\mathcal{CRT}}k\sigma n|(\Theta^\dagger_{\mathcal{CRT}}O^a(0)\Theta_{\mathcal{CRT}})^\dagger|\Omega\rangle, \tag{9.37}$$

where $\langle\Theta^\dagger_{\mathcal{CRT}}k\sigma n|$ is the bra dual to the ket $\Theta^\dagger_{\mathcal{CRT}}|k\sigma n\rangle$ and we have used that $\Theta_{\mathcal{CRT}}$ is antiunitary and leaves the vacuum unchanged. Recalling that for $\mathcal{CRT}$ we have

$$(\Theta^\dagger_{\mathcal{CRT}}O^a(0)\Theta_{\mathcal{CRT}})^\dagger = i^{-f_O}D_E(\mathcal{RT})^a{}_bO^b(0), \tag{9.38}$$

and also that $\mathcal{CRT}$ maps particles to antiparticles, we see that (9.37) gives a proportionality relation between Z_n and $\widetilde{Z}_{n^c}$. Working out the proportionality coefficient this way is a bit tricky (we'd need to sort out how

$\mathcal{CRT}$ acts on one-particle states), but once we know such a relation exists we can instead just determine the coefficient by comparing to free field theory. There we have

$$\begin{aligned}\langle\Omega|\Phi^a(0)|k,\sigma,n\rangle &= \frac{1}{\sqrt{2\omega_{n,\vec{k}}}}u^a(\vec{k},\sigma,n)\\ \langle k,\sigma,n^c|\Phi^a(0)|\Omega\rangle &= \frac{1}{\sqrt{2\omega_{n,\vec{k}}}}v^a(\vec{k},\sigma,n^c),\end{aligned} \tag{9.39}$$

so the coefficient of proportionality is one:

$$Z_n = \widetilde{Z}_{n^c}. \tag{9.40}$$

Thus we at last have

$$\begin{aligned}\langle\Omega|O^a(0)|k,\sigma,n\rangle &= \frac{Z_n}{\sqrt{2\omega_{n,\vec{k}}}}\,u^a(\vec{k},\sigma,n)\\ \langle k,\sigma,n^c|O^a(0)|\Omega\rangle &= \frac{Z_n}{\sqrt{2\omega_{n,\vec{k}}}}\,v^a(\vec{k},\sigma,n^c),\end{aligned} \tag{9.41}$$

which up to the overall factor of Z_n are rather remarkably the same as we would have obtained simply from replacing O^a by Φ^a and using free field theory!

9.3 Back to the two-point function

We can now at last go back to our expression (9.16) for the two-point function. Given our new knowledge about the matrix elements, and now taking $O_1^\dagger = O_2 = O$, we have

$$\begin{aligned}\langle TO^{a_2}(k_2)O^{a_1\dagger}(k_1)\rangle_\epsilon \supset (2\pi)^d\delta(k_2+k_1)\sum_n |Z_n|^2\Bigg(&\frac{i}{k_2^0-\omega_{n,\vec{k}_2}+i\epsilon}\frac{1}{2\omega_{n,\vec{k}_2}}\sum_\sigma u^{a_2}(\vec{k_2},\sigma,n)u^{a_1*}(\vec{k_2},\sigma,n)\\ &+(-1)^{f_O}\frac{i}{k_1^0-\omega_{n,\vec{k}_1}+i\epsilon}\frac{1}{2\omega_{n,\vec{k}_1}}\sum_\sigma v^{a_2}(\vec{k_1},\sigma,n^c)v^{a_1*}(\vec{k_1},\sigma,n^c)\Bigg),\end{aligned} \tag{9.42}$$

where I again remind you that we are focusing on the contribution of the on-shell pole. You will show in the homework that up to the factor of $|Z_n|^2$, this is precisely the Fourier transform of the time-ordered two point function of the free field (9.25):[66]

$$\langle TO^{a_2}(k_2)O^{a_1\dagger}(k_1)\rangle_\epsilon \xrightarrow[k_i^0\to\omega_{n,\vec{k}_i}]{} |Z_n|^2\langle T\Phi^{a_2}(k_2)\Phi^{a_1\dagger}(k_1)\rangle_\epsilon. \tag{9.43}$$

Thus we see that in the vicinity of an on-shell pole, the exact two-point function in any quantum field theory with a particle description just becomes that of free field theory up to an overall factor! It is important to make several comments about this however:

- In general the particles appearing here have nothing to do with the fields appearing in the Lagrangian. The free fields we are discussing here may thus look nothing like the "true" fields appearing in the Lagrangian.

[66] If there are multiple types of particle with exactly the same mass and spin/helicity (besides just n and n^c) then O^a could create a superposition of them, in which case there could still be a sum over some subset of n here. In this case however we can just redefine our basis of particle types to treat this superposition as its own type of particle. Typically this situation only arises when there is a global symmetry to enforce the degeneracy, in which case this redefinition will just be a global symmetry transformation.

- On the other hand in situations when the interactions are weak and we are interested in particles which *do* correspond to fundamental fields, we indeed can (and will) chose O^a to just be the fundamental field for the particle in question. It cannot be emphasized enough however that mass m_n appearing in this formula is the genuine particle mass, *not* the mass parameter appearing in the Lagrangian. We saw already at one loop in ϕ^4 theory that these are not the same.

- Moreover the two-point function of a field whose kinetic term in the Lagrangian is normalized as $-\frac{1}{2}\partial_\mu\phi\partial^\mu\phi$ (in the scalar case) will NOT have $Z_n = 1$ in the interacting theory. We have not yet computed enough loop diagrams to see this happen, but we eventually will (unfortunately in ϕ^4 theory one needs to go to two loops to see it). Rescaling the fundamental field to give us something whose two-point function doesn't have a factor of $|Z_n|^2$ is (for historical reasons) called **wave function renormalization**.

- In a situation where the particle we want to create is not fundamental, it may not seem so clear which operator O^a we should use to get a two-point function with a non-vanishing Z_n. In fact it is easy: we simply look for any local operator with the same symmetry charges as a free field which annihilates that particle should have. So for example in QCD it is easy to construct a local operator out of quark and gluon fields with the same symmetry transformations as a field that would annihilate the proton, and we can just use that one even though it undoubtedly will create all sorts of mess in the multiparticle states which do not contribute to the pole.

9.4 The LSZ reduction formula

Having worked at some length with the exact two-point function in interacting theories, we are now at last ready to give a general algorithm for extracting the S-matrix from correlation functions in quantum field theories with a particle description. The argument is somewhat delicate, and I must confess that some steps will be motivated by analogy to the above discussion rather than justified in detail, but the final answer is quite simple and forms the basis for all scattering computations in quantum field theory. The idea is that the S-matrix element for M particles going to N particles is the coefficient of a multi-dimensional pole in the Fourier transform of a time-ordered $M+N$-point correlation function.

We'll begin with the object

$$\begin{aligned}\langle\Omega|TO_N'^{a_N}(k_N')\dots O_1'^{a_1}(k_1')O_1^{b_1\dagger}(k_1)\dots O_M^{b_M\dagger}(k_M)|\Omega\rangle_\epsilon := \int dx_1\dots\int dx_{N+M}e^{-i\left(k_1\cdot x_1+\dots+k_M\cdot x_M+k_1'\cdot x_{M+1}+\dots+k_N'\cdot x_{M+N}\right)}\\ \times\langle\Omega|TO_N'^{a_N}(x_{M+N})\dots O_1'^{a_1}(x_{M+1})O_1^{b_1\dagger}(x_1)\dots O_M^{b_M\dagger}(x_M)|\Omega\rangle\\ \times e^{-\epsilon(t_{max}-t_{min})}\end{aligned} \tag{9.44}$$

where $O_1^{b_1},\dots,O_M^{b_M}$ and $O_1'^{a_1},\dots,O_N'^{a_N}$ are Heisenberg operators in some quantum field theory with a scattering description that transform in irreducible representations of the Lorentz group. In the ϵ-regulator, t_{min} and t_{max} are the least and greatest of the times $t_1,\dots,t_{M+N}$. We will eventually arrange so that $k_1,\dots,k_M$ are the spacetime momenta of the ingoing particles and $k_1',\dots,k_N'$ are the spacetime momenta of the outgoing particles. Very roughly we'll see that you can think of the $O_i^\dagger(k_i)$ as creation operators for particles in an "in" state and the $O_i'(k_i)$ as annihilation operators for particles in an "out" state. Our goal is to show that this object has a multi-dimensional pole as we take the external momenta to be on-shell, and in particular when we do this so that

$$\begin{aligned}k_i^0 &\to -\omega_{n_i,\vec{k}_i}\\ k_i'^0 &\to \omega_{n_i',\vec{k}_i'}.\end{aligned} \tag{9.45}$$

This pole arises from the region of integration where we have

$$t_1 \le t_2 \le \dots \le t_{M+N}, \tag{9.46}$$

and is a generalization of the first term in (9.42) (other regions of integration give poles where some of the k_i^0 are equal to $\omega_{n_i,\vec{k}_i}$ and some of the $k_i'^0$ are equal to $-\omega_{n_i',\vec{k}_i'}$). Focusing on this region of the integral, and defining

$$G_\epsilon := \langle\Omega|TO_N'^{a_N}(k_N')\dots O_1'^{a_1}(k_1')O_1^{b_1\dagger}(k_1)\dots O_M^{b_M\dagger}(k_M)|\Omega\rangle_\epsilon \tag{9.47}$$

to save space, we can insert complete sets of scattering states to get

$$\begin{aligned}G_\epsilon \supset \int_{t_1\leq t_2\dots\leq t_{M+N}} dx_1\dots dx_{N+M}\int d\alpha_1\dots d\alpha_M d\beta_1\dots d\beta_N e^{-i\left(k_1\cdot x_1+\dots+k_M\cdot x_M+k_1'\cdot x_{M+1}+\dots+k_N'\cdot x_{M+N}\right)-\epsilon(t_{M+N}-t_1)}\\ \times\langle\Omega|O_N'^{a_N}(x_{M+N})|\beta_1\rangle\dots\langle\beta_{N-1}|O_1'^{a_1}(x_{M+1})|\beta_N\rangle\langle\beta_N|\alpha_M\rangle\langle\alpha_M|O_1^{b_1\dagger}(x_1)|\alpha_{M-1}\rangle\dots\langle\alpha_1|O_M^{b_M\dagger}(x_M)|\Omega\rangle,\end{aligned} \tag{9.48}$$

where to save more space I've here adopted a convention that α states are "in" states and β states are "out" states. Note the appearance of the M-particle to N-particle S-matrix $\langle\beta_N|\alpha_M\rangle$; our goal is now to show that this can be extracted by isolating the on-shell pole.

As before we can extract the position dependence of the matrix elements in scattering states as

$$\langle\gamma|O(x)|\delta\rangle = e^{-i(p_\gamma-p_\delta)\cdot x}, \tag{9.49}$$

so we can rewrite things as

$$\begin{aligned}G_\epsilon \supset \int d\alpha_1\dots d\alpha_M d\beta_1\dots d\beta_N\,\langle\beta_N|\alpha_M\rangle\int_{t_1\leq t_2\dots\leq t_{M+N}} dx_1\dots dx_{N+M}\\ \times e^{-i\left[(k_1+p_{\alpha_M}-p_{\alpha_{M-1}})\cdot x_1+\dots+(k_M+p_{\alpha_1})\cdot x_M+(k_1'-p_{\beta_N}+p_{\beta_{N-1}})\cdot x_{M+1}+\dots+\left(k_N'-p_{\beta_1}\right)\cdot x_{M+N}\right]}e^{-\epsilon(t_{M+N}-t_1)}\\ \times\langle\Omega|O_N'^{a_N}(0)|\beta_1\rangle\dots\langle\beta_{N-1}|O_1'^{a_1}(0)|\beta_N\rangle\langle\alpha_M|O_1^{b_1\dagger}(0)|\alpha_{M-1}\rangle\dots\langle\alpha_1|O_M^{b_M\dagger}(0)|\Omega\rangle.\end{aligned} \tag{9.50}$$

The integrals over spatial positions now give simple δ-functions, but the integrals over time require a bit more work. Defining

$$\begin{aligned}\mathcal{T} := \int_{-\infty}^{\infty} dt_1 e^{i\left(k_1^0+p^0_{\alpha_M}-p^0_{\alpha_{M-1}}-i\epsilon\right)t_1}\int_{t_1}^{\infty} dt_2 e^{i\left(k_2^0+p^0_{\alpha_{M-1}}-p^0_{\alpha_{M-2}}\right)t_2}\dots\int_{t_{M-1}}^{\infty} dt_M e^{i\left(k_M^0+p^0_{\alpha_1}\right)t_M}\\ \times\int_{t_M}^{\infty} dt_{M+1} e^{i\left(k_1'^0-p^0_{\beta_N}+p^0_{\beta_{N-1}}\right)t_{M+1}}\dots\int_{t_{M+N-2}}^{\infty} dt_{M+N-1} e^{i\left(k_{N-1}'^0-p^0_{\beta_2}+p^0_{\beta_1}\right)t_{M+N-1}}\int_{t_{M+N-1}}^{\infty} dt_{M+N} e^{i\left(k_N'^0-p^0_{\beta_1}+i\epsilon\right)t_{M+N}},\end{aligned} \tag{9.51}$$

we can evaluate the integrals from right to left to find

$$\begin{aligned}\mathcal{T} =& 2\pi\delta(k_{tot}^0+k_{tot}'^0+p^0_{\alpha_M}-p^0_{\beta_N})\\ &\times\frac{-i}{k_1^0+p^0_{\alpha_M}-p^0_{\alpha_{M-1}}-i\epsilon}\frac{-i}{k_1^0+k_2^0+p^0_{\alpha_M}-p^0_{\alpha_{M-2}}-i\epsilon}\dots\frac{-i}{k_{tot}^0-k_M^0+p^0_{\alpha_M}-p^0_{\alpha_1}-i\epsilon}\\ &\times\frac{i}{k_{tot}'^0-p^0_{\beta_N}+i\epsilon}\dots\frac{i}{k_{N-1}'^0+k_N'^0-p^0_{\beta_2}+i\epsilon}\frac{i}{k_N'^0-p^0_{\beta_1}+i\epsilon}.\end{aligned} \tag{9.52}$$

Here we've defined

$$\begin{aligned}k_{tot} &= k_1+\dots+k_M\\ k_{tot}' &= k_1'+\dots+k_N'.\end{aligned} \tag{9.53}$$

Evaluating the spatial integrals, we thus have

$$\begin{aligned} G_\epsilon \supset & \int d\alpha_1 \dots d\alpha_M d\beta_1 \dots d\beta_N \, \langle \beta_N | \alpha_M \rangle \, \mathcal{T} \\ & \times (2\pi)^{d-1} \delta^{d-1} \left(\vec{k}_1 + \vec{p}_{\alpha_M} - \vec{p}_{\alpha_{M-1}} \right) \dots (2\pi)^{d-1} \delta^{d-1} \left(\vec{k}_M + \vec{p}_{\alpha_1} \right) \\ & \times (2\pi)^{d-1} \delta^{d-1} \left(\vec{k}'_1 - \vec{p}_{\beta_N} + \vec{p}_{\beta_{N-1}} \right) \dots (2\pi)^{d-1} \delta^{d-1} \left(\vec{k}'_N - \vec{p}_{\beta_1} \right) \\ & \times \langle \Omega | O'^{a_N}_N(0) | \beta_1 \rangle \dots \langle \beta_{N-1} | O'^{a_1}_1(0) | \beta_N \rangle \langle \alpha_M | O^{b_1 \dagger}_1(0) | \alpha_{M-1} \rangle \dots \langle \alpha_1 | O^{b_M \dagger}_M(0) | \Omega \rangle. \end{aligned} \tag{9.54}$$

The effect of these δ-functions is simply to impose that

$$\begin{aligned} \vec{p}_{\alpha_1} &= -\vec{k}_M \\ \vec{p}_{\alpha_2} &= -(\vec{k}_M + \vec{k}_{M-1}) \\ &\vdots \\ \vec{p}_{\alpha_M} &= -\vec{k}_{tot}, \end{aligned} \tag{9.55}$$

and also that

$$\begin{aligned} \vec{p}_{\beta_1} &= \vec{k}'_N \\ \vec{p}_{\beta_2} &= (\vec{k}'_N + \vec{k}'_{N-1}) \\ &\vdots \\ \vec{p}_{\beta_N} &= \vec{k}'_{tot}. \end{aligned} \tag{9.56}$$

To make sure we get a pole of maximum strength we should choose the multiparticle states appearing in (9.54) to ensure that the answer has no remaining momentum integrals. The way to do this to take α_1 and β_1 to be one-particle states, α_2 and β_2 to be two-particle states, and so on.

To proceed further we need to say something about the matrix elements of the O operators. For each O matrix element the bra has one fewer particle than the ket, so we are interested in the part of O which is proportional to an annihilation operator. In the previous section we saw that we can write this as

$$O^a(0) = \sum_{n,\sigma} Z_n \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{u^a(\vec{p},\sigma,n)}{\sqrt{2\omega_{n,\vec{p}}}} a_{\vec{p},\sigma,n} \tag{9.57}$$

Similarly for each of the $O^\dagger$s the bra has one more particle than the ket, so are interested in the part of $O^\dagger$ which is proportional to a creation operator, which is given by

$$O^{b\dagger}(0) = \sum_{n,\sigma} Z^*_n \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{u^{b*}(\vec{p},\sigma,n)}{\sqrt{2\omega_{n,\vec{p}}}} a^\dagger_{\vec{p},\sigma,n}. \tag{9.58}$$

To simplify life we'll assume that we've chosen either our particle basis or our operators O such that Z_n is nonzero for only one n with a given mass and spin, in which case we can drop the sum on n in these expressions. There is then only one way to satisfy all of the spatial δ-functions: $O^{B_M \dagger}_M$ must create a particle of spatial momentum $-\vec{k}_M$, $O^{b_{M-1}\dagger}_{M-1}$ must create a particle of spatial momentum $-\vec{k}_{M-1}$, and so on, and similarly $O'^{a_1}_1$ must annihilate a particle of momentum $\vec{k}_1$, $O'^{b_2}_2$ must annihilate a particle of momentum $\vec{k}_2$, and so on. We can therefore simplify the quantity $\mathcal{T}$ from the time integrals:

$$\mathcal{T} = 2\pi\delta(k^0_M + \omega_{n_M,\vec{k}_M}) \times \frac{-i}{k^0_1 + \omega_{n_1,\vec{k}_1} - i\epsilon} \cdots \frac{-i}{k^0_{M-1} + \omega_{n_{M-1},\vec{k}_{M-1}} - i\epsilon} \frac{i}{k'^0_1 - \omega_{n'_1,\vec{k}'_1} + i\epsilon} \cdots \frac{i}{k'^0_N - \omega_{n'_N \vec{k}'_N} + i\epsilon}. \tag{9.59}$$

It may seem strange that we have treated k_M^0 differently than all the other energies, this is because momentum conservation doesn't let us really vary all the momenta independently. We can restore the symmetry by using

$$2\pi\delta(k_M^0+\omega_{n_M,\vec{k}_M})=\frac{i}{k_M^0+\omega_{n_M,\vec{k}_M}+i\epsilon}-\frac{i}{k_M^0+\omega_{n_M,\vec{k}_M}-i\epsilon}, \tag{9.60}$$

with the understanding that we are interested in the pole where $k_M^0=-\omega_{n_M,\vec{k}_M}-i\epsilon$, in which case we can simply write

$$\mathcal{T}\supset\frac{-i}{k_1^0+\omega_{n_1,\vec{k}_1}-i\epsilon}\cdots\frac{-i}{k_M^0+\omega_{n_M,\vec{k}_M}-i\epsilon}\frac{i}{k_1'^0-\omega_{n_1',\vec{k}_1'}+i\epsilon}\cdots\frac{i}{k_N'^0-\omega_{n_N',\vec{k}_N'}+i\epsilon}. \tag{9.61}$$

Having now done all of the momentum integrals, we arrive at last at the **LSZ formula**:[67]

$$\begin{aligned}G_\epsilon\xrightarrow[\substack{k_i^0\to-\omega_{n_i,\vec{k}_i}\\ k_i'^0\to\omega_{n_i',\vec{k}_i'}}]{}&\prod_{j=1}^N\left(Z_{n_j'}\sum_{\sigma_j'}\frac{u^{a_j}(\vec{k}_j',\sigma_j',n_j')}{\sqrt{2\omega_{n_j',\vec{k}_j'}}}\times\frac{i}{k_j'^0-\omega_{n_j',\vec{k}_j'}+i\epsilon}\right)\prod_{i=1}^M\left(Z_{n_i}^*\sum_{\sigma_i}\frac{u^{b_i*}(-\vec{k}_i,\sigma_i,n_i)}{\sqrt{2\omega_{n_i,\vec{k}_i}}}\times\frac{-i}{k_i^0+\omega_{n_i,\vec{k}_i}-i\epsilon}\right)\\&\times\langle k_1',\sigma_1',n_1';\ldots;k_N',\sigma_N',n_N',-|-k_1,\sigma_1,n_1;\ldots;-k_M,\sigma_M,n_M,+\rangle.\end{aligned} \tag{9.62}$$

The third line here is just the S-matrix with arbitrary external particles, so by stripping off the factors in the first two lines we can directly extract it! The only remaining difficulty is how to "undo" the sums over spin/helicity; there is a standard way to do this but we will postpone discussion of it until we discuss fields with nonzero spin more explicitly.

The LSZ formula is often written in a slightly more covariant way by noting that near the poles we have

$$\frac{i}{k^0-\omega_{\vec{k}}+i\epsilon}=\frac{-i(k^0+\omega_{\vec{k}}-i\epsilon)}{(\omega_{\vec{k}}-i\epsilon)^2-(k^0)^2}\approx\frac{-i2\omega_{\vec{k}}}{k^2+m^2-i\epsilon} \tag{9.63}$$

and

$$\frac{-i}{k^0+\omega_{\vec{k}}-i\epsilon}=\frac{-i(\omega_{\vec{k}}-i\epsilon-k^0)}{(\omega_{\vec{k}}-i\epsilon)^2-(k^0)^2}\approx\frac{-i2\omega_{\vec{k}}}{k^2+m^2-i\epsilon}, \tag{9.64}$$

and thus[68]

$$\begin{aligned}G_\epsilon\xrightarrow[\substack{k_i^0\to-\omega_{n_i,\vec{k}_i}\\ k_i'^0\to\omega_{n_i',\vec{k}_i'}}]{}&\prod_{j=1}^N\left(Z_{n_j'}\sum_{\sigma_j'}u^{a_j}(\vec{k}_j',\sigma_j',n_j')\times\frac{-i\sqrt{2\omega_{n_j',\vec{k}_j'}}}{k_i^2+m_{n_i}^2-i\epsilon}\right)\times\prod_{i=1}^M\left(Z_{n_i}^*\sum_{\sigma_i}u^{b_i*}(-\vec{k}_i,\sigma_i,n_i)\times\frac{-i\sqrt{2\omega_{n_i,\vec{k}_i}}}{k_i^2+m_{n_i}^2-i\epsilon}\right)\\&\times\langle k_1',\sigma_1',n_1';\ldots;k_N',\sigma_N',n_N',-|-k_1,\sigma_1,n_1;\ldots;-k_M,\sigma_M,n_M,+\rangle.\end{aligned} \tag{9.65}$$

Due to the pesky signs in the momenta of the "in" state here, one sometimes defines G_ϵ to have the opposite sign for k_i (anyways the signs in the Fourier transform are a matter of convention).

There are a few points which are worth making about this formula:

- The LSZ formula is completely non-perturbative, computing the exact S-matrix of the true asymptotic states of the theory. Choosing the operators O and O' in general requires you to know enough about the theory to be able to find a local operator that creates/annihilates each particle type n with nonzero Z_n. As mentioned above this is usually not difficult however: you just find an operator that has the right symmetry charges and then a nonzero Z_n is generic.

[67] LSZ stands for Lehmann, Symanzik, and Zimmerman. Their original paper from 1954 only treats the scalar case and assumes weak coupling. It is written in German, Ein Prosit if you can read it!

[68] In many textbooks the "in" and "out" states are normalized in a more covariant way by absorbing the factors of $\sqrt{2\omega_{n_j',\vec{k}_j'}}$ and $\sqrt{2\omega_{n_i,\vec{k}_i}}$ into their definitions, which makes this formula look even more covariant.

- Comparison to (9.42) may have you worried that we are only computing the S-matrix for particles and not antiparticles, but of course it is arbitrary which particles we view as an antiparticles. Given a free field $\Phi^a(x)$ as in equation (9.25), we can exchange the role of particles and antiparticles by taking the adjoint of the field. We have implicitly done this in our presentation of the LSZ formula since we have only the u^a intertwiners appearing, so in particular if the amplitude involves both a particle and its anti-particle then we have used an O analogous to Φ for the former and an O analogous to $\Phi^\dagger$ for the latter. If we instead want to adhere to some pre-existing convention for which particles are antiparticles (for example if we want positrons to be antiparticles), and we want to take all Os to be analogous to Φ, then in the LSZ formula we should make the replacements $Z^* u^{a*} \to Z v^a$ for each antiparticle in the initial state and $Z u^a \to Z^* v^{a*}$ for each antiparticle in the final state.

- Since we have related S-matrix elements to correlation functions, all symmetry constraints on correlation functions must imply symmetry constraints on S-matrix elements. For example if there is a $U(1)$ global symmetry under which the incoming particles have charges $q_1, q_1, \ldots$ and the outgoing particles have charges $q'_1, q'_2, \ldots$, then we must have

$$q_1 + q_2 + \ldots + q_M = q'_1 + q'_2 + \ldots + q'_N. \tag{9.66}$$

- The masses appearing in the poles are again the physical masses, *not* bare masses that appear in the Lagrangian. After all the latter do not even make sense for composite particles. The factors of $Z_{n'}$ and Z_n are again called wave function renormalization.

In the next section we will learn how to use the LSZ formula to compute the S-matrix in weakly interacting theories using Feynman diagrams.

9.5 Homework

1. Show that the expression (9.16) is compatible with our expression $(2\pi)^d\delta^d(k_2+k_1)\frac{-i}{k_2^2+m^2-i\epsilon}$ for the momentum-space Feynman propagator in free scalar field theory. You should take $O_1(x) = O_2(x) = \Phi(x)$, where Φ is a real free scalar field.

2. Consider the derivative $\partial^\mu\Phi(x)$ of a free scalar field. From the point of view of this section this is just as good of a candidate for a field that creates a free scalar particle as $\Phi(x)$ itself is. What are u^μ and v^μ for the free field $\partial^\mu\Phi$? Show that these u^μ and v^μ obey the intertwiner equations (9.28) and (9.34).

3. Evaluate the Fourier transform of the time-ordered two-point function $\langle\Omega|T\Phi^{a_2}(x_2)\Phi^{a_1\dagger}(x_1)|\Omega\rangle$ of a general free field as in equation (9.25), and show that it gives the right hand side of (9.42) but without the $\sum_n |Z_n|^2$.

4. Evaluate the integrals in $\mathcal{T}$ and show that they lead to (9.52). Make sure to go from right to left, and be prepared to use the δ-function at the end to rewrite the poles involving k_i^0.

5. Extra credit: starting from the general Lorentz transformation properties of one-particle states, the free-field expression (9.25), and the field transformation (9.26), derive the creation and annihilation operator transformations (9.33) and the u^a and v^a transformations (9.34).

6. Extra extra credit: derive $Z_n = \widetilde{Z}_{n^c}$ directly from (9.37) and (9.38), without using free field theory. You will need to figure out how one-particle states transform under $\mathcal{CRT}$, which requires you to think about how to analytically continue the machinery of the little group to Euclidean signature.

10 Scattering in perturbation theory

In the last section we met the LSZ formula relating the S-matrix to the Fourier transform of time-ordered correlation functions:

$$\begin{aligned}\langle\Omega|TO'^{a_N}_N(k'_N)\dots O'^{a_1}_1(k'_1)O^{b_1\dagger}_1(k_1)\dots O^{b_M\dagger}_M(k_M)|\Omega\rangle \xrightarrow[\substack{k_i^0\to-\omega_{n_i,\vec{k}_i}\\ k_i'^0\to\omega_{n'_i,\vec{k}'_i}}]{} &\prod_{j=1}^N\left(Z_{n'_j}\sum_{\sigma'_j}u^{a_j}(\vec{k}'_j,\sigma'_j,n'_j)\times\frac{-i\sqrt{2\omega_{n'_j,\vec{k}'_j}}}{k_j'^2+m^2_{n'_j}-i\epsilon}\right)\\ &\times\prod_{i=1}^M\left(Z^*_{n_i}\sum_{\sigma_i}u^{b_i*}(-\vec{k}_i,\sigma_i,n_i)\times\frac{-i\sqrt{2\omega_{n_i,\vec{k}_i}}}{k_i^2+m^2_{n_i}-i\epsilon}\right)\\ &\times\langle k'_1,\sigma'_1,n'_1;\dots;k'_N,\sigma'_N,n'_N,-|-k_1,\sigma_1,n_1;\dots;-k_M,\sigma_M,n_M,+\rangle.\end{aligned} \tag{10.1}$$

In this section we will learn how to use this formula in perturbation theory to compute the S-matrix. To simplify expressions we will restrict to particles with zero spin/helicity and take the operators O and O' to be scalars, in which case the formula simplifies to[69]

$$\begin{aligned}\langle\Omega|TO'_N(k'_N)\dots O'_1(k'_1)O^\dagger_1(k_1)\dots O^\dagger_M(k_M)|\Omega\rangle \xrightarrow[\substack{k_i^0\to-\omega_{n_i,\vec{k}_i}\\ k_i'^0\to\omega_{n'_i,\vec{k}'_i}}]{} &\prod_{j=1}^N\left(\frac{-iZ_{n'_j}\sqrt{2\omega_{n'_j,\vec{k}'_j}}}{k_j'^2+m^2_{n'_j}-i\epsilon}\right)\times\prod_{i=1}^M\left(\frac{-iZ^*_{n_i}\sqrt{2\omega_{n_i,\vec{k}_i}}}{k_i^2+m^2_{n_i}-i\epsilon}\right)\\ &\times\langle k'_1,n'_1;\dots;k'_N,n'_N,-|-k_1,n_1;\dots;-k_M,n_M,+\rangle.\end{aligned} \tag{10.2}$$

We can invert this to get a formula directly for the S-matrix:

$$\begin{aligned}\langle k'_1,n'_1;\dots;k'_N,n'_N,-|k_1,n_1;\dots;k_M,n_M,+\rangle_c=&\prod_{j=1}^N\left(\frac{k_j'^2+m^2_{n'_j}-i\epsilon}{-iZ_{n'_j}\sqrt{2\omega_{n'_j,\vec{k}'_j}}}\right)\prod_{i=1}^M\left(\frac{k_i^2+m^2_{n_i}-i\epsilon}{-iZ^*_{n_i}\sqrt{2\omega_{n_i,\vec{k}_i}}}\right)\\ &\times\langle\Omega|TO'_N(k'_N)\dots O'_1(k'_1)O^\dagger_1(-k_1)\dots O^\dagger_M(-k_M)|\Omega\rangle_c\Big|_{\substack{k_i^0\to\omega_{n_i,\vec{k}_i}\\ k_i'^0\to\omega_{n'_i,\vec{k}'_i}}},\end{aligned} \tag{10.3}$$

where I've taken the liberty of flipping the sign of the ingoing momenta in the Fourier transform. I've also taken the connected part of the S-matrix, which is defined in just the same way as the connected part of the correlation functions and therefore can be computed by using the connected correlation function on the right-hand side. Let's study this formula specifically in the context of our interacting ϕ^4 theory, with Lagrangian density

$$\mathcal{L}=-\frac{1}{2}\partial_\mu\phi\partial^\mu\phi-\frac{m_0^2}{2}\phi^2-\frac{\lambda}{4!}\phi^4. \tag{10.4}$$

We will take all of the Os and O's to just be Φ. Since this theory has only one kind of particle and Φ is hermitian, we can further simplify (10.3) to

$$\langle k'_1;\dots;k'_N,-|k_1;\dots;k_M,+\rangle_c=\prod_{j=1}^N\left(\frac{k_j'^2+m^2-i\epsilon}{-iZ\sqrt{2\omega_{\vec{k}'_j}}}\right)\prod_{i=1}^M\left(\frac{k_i^2+m^2-i\epsilon}{-iZ\sqrt{2\omega_{\vec{k}_i}}}\right)\langle\Omega|T\Phi(k'_N)\dots\Phi(k'_1)\Phi(-k_1)\dots\Phi(-k_M)|\Omega\rangle_c\Big|_{\substack{k_i^0\to\omega_{\vec{k}_i}\\ k_i'^0\to\omega_{\vec{k}'_i}}}. \tag{10.5}$$

[69] It is easy to put back the spin, we will do it later when we consider particles of spin/helicity 1/2 and 1.

We've written m_0 for the "bare" mass in the Lagrangian to distinguish it from the genuine physical mass m which appears in equation (10.5). Recall that Z here is defined by

$$\langle \Omega | \Phi(0) | k \rangle = \langle k | \Phi(0) | \Omega \rangle = \frac{Z}{\sqrt{2\omega_{\vec{k}}}}, \tag{10.6}$$

with

$$\omega_{\vec{k}} = \sqrt{|k|^2 + m^2}. \tag{10.7}$$

Z must be real by the hermiticity of Φ.

We learned in section 7 that in perturbation theory we can compute the Fourier transform of the connected time-ordered correlation functions of Φ using the momentum-space Feynman rules:

1. Write down a connected Feynman diagram C with $N + M$ external points.
2. Write a factor of $-i\lambda$ for each interaction vertex
3. Divide by the symmetry factor s_C.
4. Label all internal and external momenta (with external momenta outgoing), enforcing momentum conservation at all interaction vertices.
5. Write down an overall momentum-conserving δ-function for the external momenta.
6. For each propagator supply a factor of $\frac{-i}{p^2+m_0^2-i\epsilon}$
7. Integrate over all remaining loop momenta.
8. Sum over connected diagrams C.

What we need to figure out now is how these rules are modified by the pole factors in (10.5). Naively these just remove the propagators attached to the external points and multiply by factors of Z^{-1}, but we need to be careful about the difference between m and m_0.

10.1 Self-energy and the two-point function

To clarify the difference between m and m_0 we can be a little more organized about the perturbative calculation of the two-point function. Let's first define the **self-energy** $\Sigma(p^2)$ of the scalar particle by

$$\langle T\Phi(p)\Phi(p') \rangle = (2\pi)^d \delta^d(p + p') \frac{-i}{p^2 + m_0^2 + \Sigma(p^2) - i\epsilon}. \tag{10.8}$$

The δ-function here is a consequence of translation invariance, and the fact that it depends only on p^2 is a consequence of Lorentz invariance. We saw in the previous section that this correlation function has a pole at $p^2 = -m^2$, so the relationship between the bare and physical masses is determined by solving

$$m^2 = m_0^2 + \Sigma(-m^2). \tag{10.9}$$

We also saw that the residue of this pole is $-i(2\pi)^d \delta^d(p + p')Z^2$, so apparently we have

$$Z^2 = \frac{1}{1 + \Sigma'(-m^2)}. \tag{10.10}$$

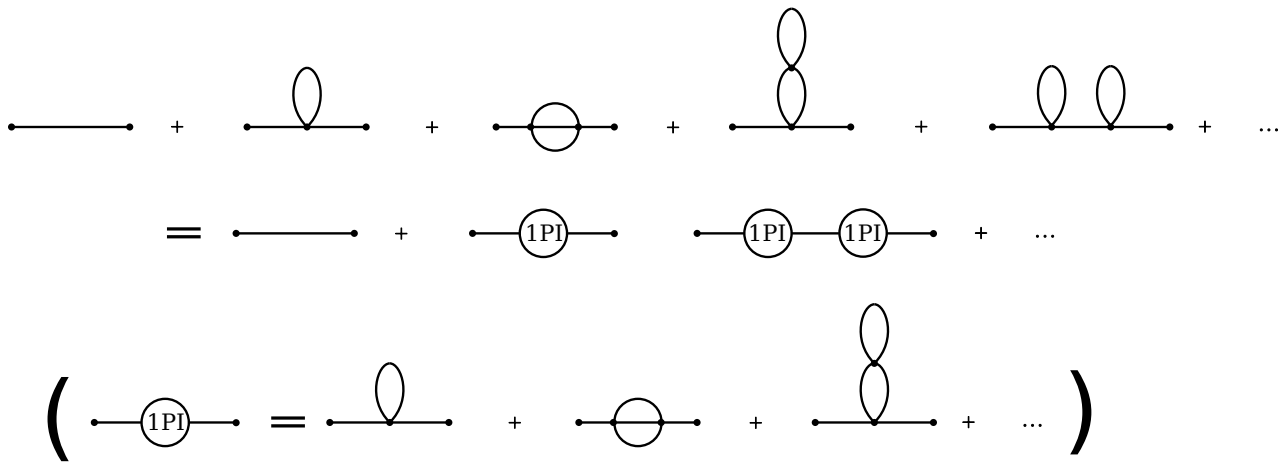


Figure 25: The one-particle-irreducible (1PI) decomposition of the two-point function. The full two-point function is built out of a sum of increasing numbers of 1PI bubbles chained together by propagators, leading to a geometric series.

We can then consider how to compute the self-energy perturbatively. In the free theory with $\lambda = 0$ we have $\Sigma = 0$, so we can rewrite the momentum space propagator perturbatively:

$$\begin{aligned}\frac{-i}{p^2 + m_0^2 + \Sigma(p^2) - i\epsilon} &= \frac{-i}{p^2 + m_0^2 - i\epsilon} \times \frac{p^2 + m_0^2 - i\epsilon}{p^2 + m_0^2 + \Sigma(p^2) - i\epsilon} \\ &= \frac{-i}{p^2 + m_0^2 - i\epsilon} \times \frac{1}{1 + \frac{\Sigma(p^2)}{p^2 + m_0^2 - i\epsilon}} \\ &= \frac{-i}{p^2 + m_0^2 - i\epsilon} \times \left(1 + \left(-i\Sigma(p^2)\frac{-i}{p^2 + m_0^2 - i\epsilon}\right) + \left(-i\Sigma(p^2)\frac{-i}{p^2 + m_0^2 - i\epsilon}\right)^2 + \dots\right).\end{aligned} \tag{10.11}$$

In the last line here we have "unsummed" a geometric series to get a Taylor expansion in Σ, which we should think of as being $O(\lambda)$. We can then compare this expression to what we get from the Feynman diagram expansion for the two-point function, for which the first few diagrams are shown in the first line of figure 25. To isolate the contribution of Σ, we note that we can organize this series as a geometric sum by splitting it up into "one-particle irreducible" (1PI) pieces. By definition a 1PI Feynman diagram is a connected diagram with at least one interaction vertex and also the property that there is no internal link such that removing that link splits the diagram into two connected components. The second, third, and fourth diagrams in the first line of figure 25 are 1PI but the first and fifth are not. Comparing the 1PI decomposition to the last line of equation (10.11), we see the following rule:

$$\begin{gathered}-i\Sigma(p^2) = \text{sum over two-point 1PI diagrams with external propagators} \\ \text{and the momentum-conserving } \delta\text{-function removed.}\end{gathered}$$

In particular at one loop the only contribution to Σ is the "snail" diagram (the first diagram in the third line of figure 25), so we have

$$\begin{aligned}-i\Sigma(p^2) &= -\frac{i\lambda}{2}\int \frac{d^d q}{(2\pi)^d}\frac{-i}{q^2 + m_0^2 - i\epsilon} + O(\lambda^2) \\ &= -\frac{i\lambda}{2}G_F(0) + O(\lambda^2).\end{aligned} \tag{10.12}$$

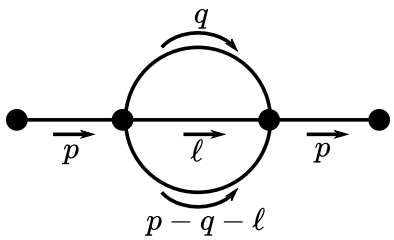


Figure 26: Momentum labels for the sunset diagram.

Equation(10.9) then shows us that

$$m^2 = m_0^2 + \lambda G_F(0)/2 + O(\lambda^2), \tag{10.13}$$

just as we found back in section 7. We have now done better than we did then however, as we have seen that by resumming an infinite sum of diagrams this mass shaft indeed is a shift of the pole location to all orders in perturbation theory.

You may have already noticed that at one loop $\Sigma(p^2)$ is actually independent of p^2, so from (10.10) we see that

$$Z = 1 + O(\lambda^2). \tag{10.14}$$

To get a nonzero contribution to Z we need a 1PI diagram that has nontrivial p^2 dependence once the external propagators are removed, and the first diagram with this property is the two-loop "sunset" diagram (the second diagram in the third line of figure 25). Choosing momentum labels as in figure 26, the contribution of this diagram to $\Sigma(p^2)$ is

$$-i\Sigma(p^2) \supset -\frac{\lambda^2}{6}\int \frac{d^dq}{(2\pi)^2}\int \frac{d^d\ell}{(2\pi)^d}\frac{-i}{q^2+m_0^2-i\epsilon}\frac{-i}{\ell^2+m_0^2-i\epsilon}\frac{-i}{(p-\ell-q)^2+m_0^2-i\epsilon}. \tag{10.15}$$

You can see the explicit p^2 dependence here in the third propagator. For now we won't try to evaluate the loop integrals. In the homework you'll meet another scalar field theory which already has a nonzero contribution to Z at one loop.

10.2 Perturbative calculation of the S-matrix

Returning now to the perturbative computation of the connected S-matrix, from the LSZ formula (10.5) we want to start with the Fourier transform of the time-ordered $N+M$-point function and send all of the external momenta on-shell, dividing by an exact momentum-space two-point function for each external leg as we do, and then multiply by Z for each external factor since the residue of the two-point function pole is Z^2 while we only have Z appearing in (10.5). To get a sense of what this means at the level of Feynman diagrams, we show the first contributions to the four-point function (appropriate for computing $2 \to 2$ scattering) in figure 27. In this diagram we should view the second row of diagrams as simply providing the one-loop corrections to the external propagators in the single diagram in the first row, and so these diagrams will be canceled when we divide out by the exact two-point functions on the legs. We can make this cancellation automatic by only summing over "pruned" connected Feynman diagrams, meaning connected diagrams with at least one interaction vertex and also the property that there is no internal link such that removing that link would result in one external point being in different connected component from all of the others external points.[70] This pruning removes all diagrams which would give corrections to the external propagators, so to finish

[70] 1PI diagrams are always pruned, but the converse is not true since we could have an internal line whose removal splits the diagram into pieces each of which contains multiple external points. A simple example is the six-point tree level diagram where three external legs are connected to one interaction vertex and three external legs are connected to another interaction vertex, with one link connecting the two vertices.

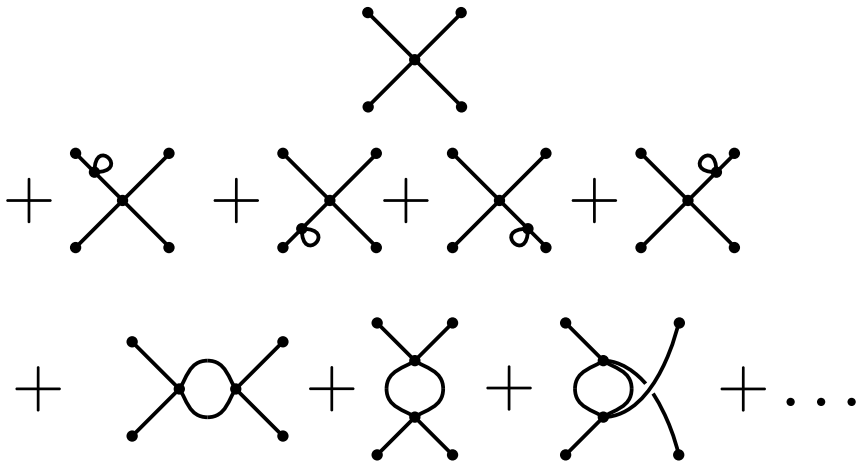


Figure 27: Contributions to the four-point function up to $O(\lambda^2)$. The diagrams in the second row are not pruned, so we should remove them in computing the S-matrix.

removing the exact two-point functions on the external legs we just need to divide by the free propagators. We thus have the following rule:[71]

$$\sqrt{2\omega_{\vec{k}'_1}}\cdots\sqrt{2\omega_{\vec{k}'_N}}\sqrt{2\omega_{\vec{k}_1}}\cdots\sqrt{2\omega_{\vec{k}_M}}\langle k'_1;\ldots;k'_N,-|k_1;\ldots k_M,+\rangle_c = Z^{N+M}\times \text{ sum over all pruned connected Feynman diagrams with external propagators removed.}$$

In our discussion of scattering theory it was convenient to remove the overall momentum conserving δ-function, defining the quantity $M_{\beta\alpha}$ by

$$S_{\beta\alpha} = \delta(\beta-\alpha) + i\times(2\pi)^d\delta^d(p_\beta - p_\alpha)\mathcal{M}_{\beta\alpha}, \tag{10.16}$$

and we can also get rid of the annoying factors of the energy in our formula by defining (as we did in section 8)

$$\widetilde{\mathcal{M}}_{\beta\alpha} = \sqrt{2\omega_{\vec{k}'_1}}\cdots\sqrt{2\omega_{\vec{k}'_N}}\sqrt{2\omega_{\vec{k}_1}}\cdots\sqrt{2\omega_{\vec{k}_M}}\mathcal{M}_{\beta\alpha}. \tag{10.17}$$

To give us more room to describe the states we'll also write this as $\widetilde{M}(\alpha\to\beta)$, in which case we have

$$i\widetilde{\mathcal{M}}_c(k_1,\ldots,k_M\to k'_1,\ldots,k'_N) = Z^{N+M}\times\text{sum over all pruned connected Feynman diagrams with external propagators and the momentum-conserving }\delta\text{-function removed.}$$

As a first example we can consider the contributions to the $2\to 2$ S-matrix. At tree level we just have the single diagram in the first row of figure 28, which simply gives

$$i\widetilde{\mathcal{M}}_c(k_1,k_2\to k'_1,k'_2) = -i\lambda + O(\lambda^2). \tag{10.18}$$

At one loop we then have the three diagrams in the third row, whose momenta we label as in figure 28, which add up to

$$i\widetilde{\mathcal{M}}_c(k_1,k_2\to k'_1,k'_2) \supset -\frac{\lambda^2}{2}\int\frac{d^d\ell}{(2\pi)^d}\frac{-i}{\ell^2+m_0^2-i\epsilon}\Big[\frac{-i}{(\ell+k'_1-k_1)^2+m_0^2-i\epsilon} + \frac{-i}{(\ell-k_1-k_2)^2+m_0^2-i\epsilon} + \frac{-i}{(\ell+k_2-k'_1)^2+m_0^2-i\epsilon}\Big]. \tag{10.19}$$

We will learn how to evaluate this integral in the next section.

[71] Pruned diagrams are usually instead called "amputated", but pruning feels less gruesome to me.

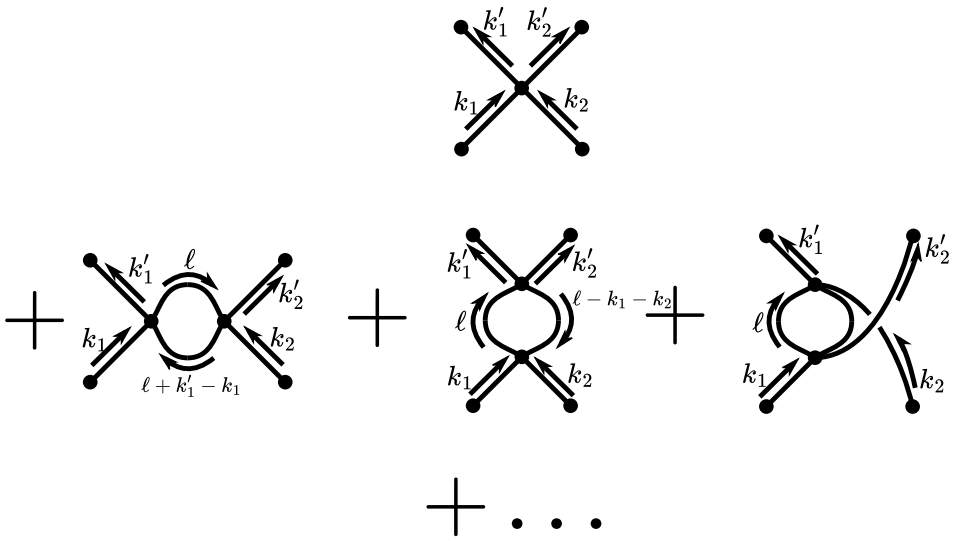


Figure 28: Momentum labels for the pruned diagrams contributing to the $2 \to 2$ S-matrix at tree level and one loop.

10.3 Computing the cross section

We'll conclude this section by computing the tree level differential cross section for $2 \to 2$ scattering in ϕ^4 theory. Let's first consider the general differential cross section with two particles in the final state and potentially different masses for all four particles, which we saw earlier is given by

$$d\sigma(\alpha \to \beta) = u_\alpha^{-1}(2\pi)^d \delta^d(p_\beta - p_\alpha)|\mathcal{M}(\alpha \to \beta)|^2 d\beta \tag{10.20}$$

with

$$u_\alpha = \frac{\sqrt{(k_1 \cdot k_2)^2 - m_{n_1}^2 m_{n_2}^2}}{\omega_{n_1,\vec{k}_1}\omega_{n_2,\vec{k}_2}}. \tag{10.21}$$

Writing out these factors more explicitly with two particles in the final state and restricting to the situation where the individual ingoing momenta are not equal to the individual outgoing momenta we have

$$\begin{aligned} d\sigma(k_1,\sigma_1,n_1;k_2,\sigma_2,n_2 \to k_1',\sigma_1',n_1';k_2,\sigma_2',n_2') = {} & \frac{1}{4\sqrt{(k_1\cdot k_2)^2 - m_{n_1}^2 m_{n_2}^2}}(2\pi)^d\delta^d(k_1'+k_2'-k_1-k_2) \\ & \times \frac{|\widetilde{\mathcal{M}}_c(\alpha\to\beta)|^2}{4\omega_{n_1',\vec{k}_1'}\omega_{n_2',\vec{k}_2'}} \times \frac{1}{2^{I_{final}}}\frac{d^{d-1}k_1'}{(2\pi)^{d-1}}\frac{d^{d-1}k_2'}{(2\pi)^{d-1}}. \end{aligned} \tag{10.22}$$

Here I_{final} is equal to zero if the final state particles are distinguishable (i.e. if $n_1' \neq n_2'$) and equal to one if they are indistinguishable (i.e. if $n_1' = n_2'$) (such a factor was part of the definition of $d\beta$). We can use the spatial momentum-conserving δ-function to evaluate the integral over $\vec{k_1}'$, so we are left with

$$\begin{aligned} d\sigma(k_1,\sigma_1,n_1;k_2,\sigma_2,n_2 \to k_1+k_2-k_2',\sigma_1',n_1';k_2',\sigma_2',n_2') = {} & \frac{1}{2^{I_{final}}}\frac{1}{4\sqrt{(k_1\cdot k_2)^2 - m_{n_1}^2 m_{n_2}^2}} \\ & \times 2\pi\delta(-\omega_{n_1,\vec{k}_1} - \omega_{n_2,\vec{k}_2} + \omega_{n_1',\vec{k}_1'} + \omega_{n_2',\vec{k}_2'})\frac{|\widetilde{M}_c|^2}{4\omega_{n_1',\vec{k}_1'}\omega_{n_2',\vec{k}_2'}}\frac{d^{d-1}k_2'}{(2\pi)^{d-1}}, \end{aligned} \tag{10.23}$$

where now we set

$$\vec{k}_1' = \vec{k}_1 + \vec{k}_2 - \vec{k}_2' \tag{10.24}$$

and the differential cross section is to be integrated over only k_2'. We will study this in the "center of mass frame" where

$$\vec{k}_2 = -\vec{k_1} := \vec{k} \tag{10.25}$$

and

$$\vec{k}_2' = -\vec{k}_1' := \vec{k}'. \tag{10.26}$$

In this frame the spacetime momenta are

$$\begin{aligned} k_1 &= \left(\sqrt{|k|^2 + m_{n_1}^2}, -\vec{k}\right) \\ k_2 &= \left(\sqrt{|k|^2 + m_{n_2}^2}, \vec{k}\right) \\ k_1' &= \left(\sqrt{|k'|^2 + m_{n_1'}^2}, -\vec{k}'\right) \\ k_2' &= \left(\sqrt{|k'|^2 + m_{n_2'}^2}, \vec{k}'\right), \end{aligned}$$

with total center of mass energy

$$E_{tot} := \sqrt{|k|^2 + m_{n_1}^2} + \sqrt{|k|^2 + m_{n_2}^2}. \tag{10.27}$$

and

$$\sqrt{(k_1 \cdot k_2)^2 - m_{n_1}^2 m_{n_2}^2} = |k| E_{tot}. \tag{10.28}$$

The energy conserving δ-function sets

$$E_{tot} = \sqrt{|k'|^2 + m_{n_1'}^2} + \sqrt{|k'|^2 + m_{n_2'}^2}, \tag{10.29}$$

which has no solution if $E_{tot} < m_{n_1'} + m_{n_2'}$ and has solution

$$|k'| = \frac{\sqrt{(E_{tot}^2 - m_{n_1'}^2 - m_{n_2'}^2)^2 - 4 m_{n_1'}^2 m_{n_2'}^2}}{2E_{tot}} \tag{10.30}$$

if $E_{tot} \geq m_{n_1'} + m_{n_2'}$. In order to use the δ function to simplify the differential cross section we need to rewrite by noting that

$$\frac{d}{d|k'|}\left(\sqrt{|k'|^2 + m_{n_1'}^2} + \sqrt{|k'|^2 + m_{n_2'}^2}\right) = \frac{|k'|}{\sqrt{|k'|^2 + m_{n_1'}^2}} + \frac{|k'|}{\sqrt{|k'|^2 + m_{n_2'}^2}} = \frac{|k'| E_{tot}}{\sqrt{|k'|^2 + m_{n_1'}^2}\sqrt{|k'|^2 + m_{n_2'}^2}} \tag{10.31}$$

and thus

$$2\pi\delta(-E_{tot} + \sqrt{|k'|^2 + m_{n_1'}^2} + \sqrt{|k'|^2 + m_{n_2'}^2}) = 2\pi \frac{\sqrt{|k'|^2 + m_{n_1'}^2}\sqrt{|k'|^2 + m_{n_2'}^2}}{|k'| E_{tot}} \delta\left(|k|' - \frac{\sqrt{(E_{tot}^2 - m_{n_1'}^2 - m_{n_2'}^2)^2 - 4 m_{n_1'}^2 m_{n_2'}^2}}{2E_{tot}}\right). \tag{10.32}$$

The integration measure is

$$\frac{d^{d-1}k'}{(2\pi)^{d-1}} = (2\pi)^{-(d-1)} |k'|^{d-2} d|k'| d\Omega_{d-2}, \tag{10.33}$$

where $d\Omega$ is the volume measure on a unit $\mathbb{S}^{d-2}$. Putting this all together we see that the differential cross section (now only to be integrated over the angular coordinates on $\mathbb{S}^{d-2}$) is

$$\frac{d\sigma}{d\Omega_{d-2}} = \frac{1}{2^{I_{final}}} \frac{1}{(2\pi)^{d-2}} \frac{|k'|^{d-3}}{16|k| E_{tot}^2} |\widetilde{M}_c|^2, \tag{10.34}$$

with $|k'|$ and E_{tot} being given in terms of $|k|$ by equations (10.30) and (10.27).

Returning now to our ϕ^4 theory, since all external masses are equal equation (10.30) simplifies to

$$|k'| = |k| \tag{10.35}$$

and so we have

$$\frac{d\sigma}{d\Omega_{d-2}} = \frac{1}{2}\frac{1}{(2\pi)^{d-2}}\frac{|k|^{d-4}}{16E_{tot}^2}|\widetilde{M}_c|^2. \tag{10.36}$$

At tree level we simply have $|\widetilde{M}_c|^2 = \lambda^2$, so the differential cross section is therefore given by

$$\frac{d\sigma}{d\Omega_{d-2}} = \frac{1}{2}\frac{1}{(2\pi)^{d-2}}\frac{|k|^{d-4}}{16E_{tot}^2}\left(\lambda^2 + O(\lambda^3)\right). \tag{10.37}$$

This answer is independent of angle, so the outgoing particles are equally likely to come out in any direction. The total cross section σ is therefore just the differential cross section times the volume[72]

$$\Omega_{d-2} = \frac{2\pi^{(d-1)/2}}{\Gamma(\frac{d-1}{2})} \tag{10.38}$$

of a unit $\mathbb{S}^{d-2}$:

$$\sigma = \frac{\Omega_{d-2}}{32(2\pi)^{d-2}}\frac{|k|^{d-4}}{E_{tot}^2}\left(\lambda^2 + O(\lambda^3)\right). \tag{10.39}$$

At one loop the differential cross section becomes angle-dependent, leading to a more interesting differential cross section. In the homework you will consider another scalar field theory which already at tree level has an angle-dependent differential cross section.

It is interesting to consider the high- and low-energy limits of the tree-level cross section as a function of incident energy. This dependence is given by

$$\sigma \propto \frac{|k|^{d-4}}{|k|^2 + m^2}\lambda^2. \tag{10.40}$$

To get a sense of the real strength of the interactions we should compare σ to some other quantity with units of area, and at high energies the only such quantity available is $|k|^{-(d-2)}$. We thus can get a rough estimate of the interaction strength at high energy by

$$\sigma|k|^{d-2} \sim \lambda^2|k|^{2(d-4)}. \tag{10.41}$$

Thus for $d > 4$ the interaction strength grows with energy, and one might worry whether the theory really makes sense at short distances (it probably doesn't). In the massless limit this scaling also controls the theory at low energies, and so when $d < 4$ the interaction strength grows at low energies in the massless case. Thus for $d < 4$ perturbation theory will not be valid at low energy and we will need to use some more exotic technique. In particular this is true for $d = 3$, and the strongly-interacting theory one reaches at low energy in that case governs the behavior of classical Ising magnets in three spatial dimensions. In the case of $d = 4$ the interaction strength is constant, and then we need to go to higher order in perturbation theory to see what happens. We will soon see that at one-loop the interactions grow logarithmically with energy in $d = 4$. This kind of argument is made more precisely using the idea of the **renormalization group**, which we will return to soon.

[72] You will derive this volume in the homework.

10.4 Homework

1. Derive equation (10.38) for the volume of a sphere in general dimensions. Hint: The easiest way to do this is to evaluate the multi-dimensional Gaussian integral $\int d^{d-1}x e^{-|x|^2}$ in both cartesian and spherical coordinates and then compare the answers.

2. Another simple scalar field theory we can study is

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{m_0^2}{2}\phi^2 - \frac{g_0}{3!}\phi^3 \tag{10.42}$$

with $g > 0$. Non-perturbatively this theory is rather sick, as the naive ground state near $\phi = 0$ can decay by tunneling through the barrier and then rolling down to $\phi = 0 - \infty$. It is also rather fine-tuned, as we arbitrarily didn't write down a linear term proportional to ϕ in the potential that would have been consistent with all the symmetries of the theory. Nonetheless it is a useful model for playing with Feynman diagrams, which are not sophisticated enough to see the non-perturbative instability. In fact some textbooks (such as Srednicki) use this theory as their primary example of an interacting field theory, as the Feynman diagrams are more similar to those of QED.

   (a) Make a sketch of the potential for the field in this theory.

   (b) Draw the Feynman diagrams which contribute to the self energy $\Sigma(p^2)$ up through two loops.

   (c) Draw the Feynman diagrams which contribute to the $2 \to 2$ scattering amplitude up through one loop.

   (d) Evaluate the tree-level $2 \to 2$ scattering amplitude and differential cross section. In what space-time dimension is the cross section measured in units of the wavelength roughly constant at large energies?

11 Loop diagrams

We've now learned how to compute the perturbative S-matrix and perturbative correlation functions in quantum field theory. In particular we wrote down several one- and two- loop Feynman integrals, but so far we have not attempted to actually integrate over any of the loop momenta. The goal of this section is to rectify that.

11.1 Self-energy at one loop

Let's first return to our one-loop expression for the self-energy of the scalar field with interacting Lagrangian[73]

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{m_0^2}{2}\phi^2 - \frac{\lambda_0}{4!}\phi^4, \tag{11.1}$$

which we found in the previous section to be

$$\Sigma(p^2) = \frac{\lambda_0}{2}\int \frac{d^dq}{(2\pi)^d}\frac{-i}{q^2+m_0^2-i\epsilon}. \tag{11.2}$$

The first thing we will do with this integral is rotate the q^0 contour to Euclidean signature by substituting $q_L^0 = iq_E^0$, leading to

$$\Sigma(p^2) = \frac{\lambda_0}{2}\int \frac{d^dq}{(2\pi)^d}\frac{1}{q^2+m_0^2}, \tag{11.3}$$

where we can drop the $i\epsilon$ since the denominator of the propagator is now positive-definite. We can rewrite this in radial coordinates as

$$\Sigma(p^2) = \frac{\lambda_0}{2}\frac{\Omega_{d-1}}{(2\pi)^d}\int_0^\infty \frac{q^{d-1}dq}{q^2+m_0^2}, \tag{11.4}$$

which is an integral that diverges at large q for $d \geq 2$. To make sense of the integral we therefore need to regulate it in some way. We will consider four regulators in turn, understanding the advantages and disadvantages of each.

11.1.1 Lattice regulator

The most physical way to regulate a quantum field theory at short distance is by using a spacetime lattice. For example in Euclidean signature we could say that the field $\phi(x)$ only lives at the points

$$(x_E^0, x^1, \ldots x^{d-1}) = (an^0, an^1, \ldots, an^{d-1}), \tag{11.5}$$

where a is some short distance scale called the **lattice spacing**. This is called a **cubic spacetime lattice**. When we introduce the Fourier transform on a lattice, the momenta which can appear are restricted in an interesting way. This is because when $x \in a\mathbb{Z}$ we have

$$e^{i(p+\frac{2\pi m}{a})x} = e^{ipx} \tag{11.6}$$

for any integer m. To get a genuinely independent set of Fourier modes, we should therefore restrict to p in the range

$$p \in (-\frac{\pi}{a}, \frac{\pi}{a}). \tag{11.7}$$

What this does in loop integrals is that it restricts each component of q_E^μ to lie in this range. Since at finite a this range is finite, this rule assigns a finite value to all loop integrals.

[73] I've relabeled the bare coupling constant to λ_0, anticipating that some renormalization will be necessary to convert this to a "physical" coupling λ.

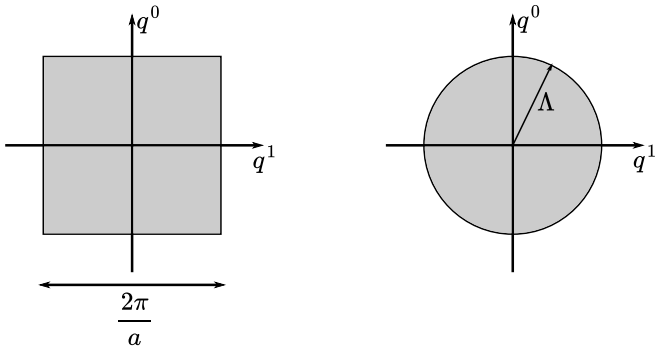


Figure 29: Momentum integration regions for lattice (on the left) and hard momentum cutoff (on the right) regulators in $1+1$ dimensions.

In most cases lattice regularization is by far the best way to do non-perturbative calculations in interacting quantum field theories. You "simply" put the Euclidean path integral on a lattice and then evaluate it with a big computer using the Monte Carlo method. It is also the best way to *think* about regularization in quantum field theory, as there is a clear physical picture of what is going on. Unfortunately however the lattice regulator is rather awkward for concrete calculations in perturbation theory, as the region over which the momentum integral is evaluated breaks most of the Euclidean symmetry of the problem (see figure 29). Lattice regulators are therefore rarely used in perturbative calculations. For example even the $d = 2$ lattice-regulated self-energy integral

$$\Sigma(p^2) = \frac{\lambda_0}{2} \int_{-\pi/a}^{\pi/a} \frac{dk_E^0}{2\pi} \int_{-\pi/a}^{\pi/a} \frac{dk^1}{2\pi} \frac{1}{(k_E^0)^2 + (k^1)^2 + m_0^2} \tag{11.8}$$

is too hard for me to evaluate.[74]

11.1.2 Hard momentum cutoff regulator

There is an obvious way to "improve" the lattice regulation to be more symmetric: we can simply cut off the momentum integral in a spherically-symmetric way,

$$q^2 \leq \Lambda^2, \tag{11.9}$$

where Λ is some fixed large energy scale. This makes the integral much easier, as we can now go to radial coordinates:

$$\Sigma(p^2) = \frac{\lambda_0}{2} \frac{\Omega_{d-1}}{(2\pi)^d} \int_0^\Lambda \frac{q^{d-1} dq}{q^2 + m_0^2}. \tag{11.10}$$

This integral can be done for general $d \geq 0$ in terms of a hypergeometric function, but it is perhaps more instructive to just give the answers for $d = 1, 2, 3, 4$:

$$\Sigma(p^2) = \frac{\lambda_0}{2} \frac{\Omega_{d-1}}{(2\pi)^d} \times \begin{cases} \frac{1}{m_0} \arctan\left(\frac{\Lambda}{m_0}\right) & d = 1 \\ \log\left(\frac{\Lambda}{m_0}\right) + \frac{1}{2} \log\left(1 + \frac{m_0^2}{\Lambda^2}\right) & d = 2 \\ \Lambda - m_0 \arctan\left(\frac{\Lambda}{m_0}\right) & d = 3 \\ \frac{\Lambda^2}{2} + \frac{m_0^2}{2} \log \frac{m_0^2}{\Lambda^2 + m_0^2} & d = 4 \end{cases}. \tag{11.11}$$

[74] Mathematica did give me some terrifying expression, but when I asked it to expand this answer for small a it gave me 1.5mb of garbage.

Expanding these at large Λ we have

$$\Sigma(p^2) = \frac{\lambda_0}{2}\frac{\Omega_{d-1}}{(2\pi)^d} \times \begin{cases} \frac{\pi}{2m_0} + \dots & d = 1 \\ \log\left(\frac{\Lambda}{m_0}\right) + \dots & d = 2 \\ \Lambda - \frac{m_0\pi}{2} + \dots & d = 3 \\ \frac{\Lambda^2}{2} - m_0^2 \log\frac{\Lambda}{m_0} + \dots & d = 4 \end{cases}, \tag{11.12}$$

where in each case "..." indicates terms which vanish as $\Lambda \to \infty$. The $d = 1$ case gives a (finite) quantum correction to the frequency of the quartic anharmonic oscillator, while for $d = 2, 3, 4$ we see that we have increasingly divergent corrections to mass.

11.1.3 Pauli-Villars regularization

The hard momentum cutoff has made our life somewhat easier, but for more complicated loop integrals it still leads to several problems. In particular if our loop diagram has several propagators where the same loop momentum appears shifted in different ways, such as we had for the one-loop contribution to $2 \to 2$ scattering at the end of the previous section, then a hard momentum regulator still isn't rotationally invariant. It also leads to difficulties once we introduce gauge fields such as the electromagnetic field. It turns out to be nicer to "gradually" turn off the contributions of large momenta rather than brutally discarding them completely. This is called **Pauli-Villars regularization**, with the canonical approach being to modify all propagators via the replacement

$$\frac{-i}{p^2 + m_0^2 - i\epsilon} \to \frac{-i}{p^2 + m_0^2 - i\epsilon} + \frac{i}{p^2 + \Lambda^2 - i\epsilon}. \tag{11.13}$$

In Euclidean signature this becomes

$$\frac{1}{p^2 + m_0^2} \to \frac{1}{p^2 + m_0^2 - i\epsilon} - \frac{1}{p^2 + \Lambda^2}. \tag{11.14}$$

The new term in the propagator here is small when $p^2 \ll \Lambda^2$, but for $p^2 \gg \Lambda^2$ it improves the high-momentum behavior of the propagator from $1/p^2$ to $1/p^4$. This improves the convergence of loop integrals.

Unfortunately the canonical Pauli-Villars regulation (11.13) doesn't always render loop integrals finite. For example our self-energy integral in $d = 4$ is still logarithmically divergent at high momentum, and in higher dimensions things are only worse. To deal with this we will instead consider an "improved" Pauli-Villars regulator, which in Euclidean signature modifies the propagator as

$$\frac{1}{p^2 + m_0^2} \to \frac{e^{-\frac{p^2}{\Lambda^2}}}{p^2 + m_0^2}. \tag{11.15}$$

The exponential factor is close to one when $p^2 \ll \Lambda^2$ just as before, but now for $p^2 \gg \Lambda^2$ the exponential suppression ensures that all loop integrals will be finite in any dimension and for any number of propagators. Making use of Mathematica, with this regulator the self-energy of our scalar field theory becomes

$$\Sigma(p^2) = \frac{\lambda_0}{2}\frac{\Omega_{d-1}}{(2\pi)^d} \times \begin{cases} \frac{\pi}{2m_0} + \dots & d = 1 \\ \log\left(\frac{\Lambda}{m_0}\right) - \frac{\gamma}{2} + \dots & d = 2 \\ \frac{\sqrt{\pi}}{2}\Lambda - \frac{m_0\pi}{2} + \dots & d = 3 \\ \frac{\Lambda^2}{2} - m_0^2 \log\frac{\Lambda}{m_0} + \frac{\gamma}{2}m_0^2 + \dots & d = 4 \end{cases}, \tag{11.16}$$

where again "..." indicates terms which vanish as $\Lambda \to \infty$ and

$$\gamma := \lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{1}{k} - \log n\right) \approx .577 \tag{11.17}$$

is called the Euler-Mascheroni constant.

Already a pattern is hopefully apparent: the power-law divergent contributions to $\Sigma(p^2)$ are different for different choices of regulator, but when there is a logarithmic divergence its coefficient is universal and when there is no logarithmic divergence the finite term is also universal. The finite piece cannot be universal when there is a logarithmic divergence since we can always rescale the cutoff:

$$a \log \frac{b\Lambda}{m_0} = a \log \frac{\Lambda}{m_0} + a \log b. \tag{11.18}$$

As you can see this changes the finite piece but doesn't change the coefficient of the logarithm. Such rescalings of course also change the coefficients of any power-law divergences.

Something else is also hopefully clear: the relationship between the bare and physical mass is *regularization-dependent*: for example for $d = 4$ we at one loop have

$$m^2 = m_0^2 + \frac{\lambda_0}{16\pi^2} \begin{cases} \frac{\Lambda^2}{2} - m_0^2 \log \frac{\Lambda}{m_0} & \text{hard momentum cutoff} \\ \frac{\Lambda^2}{2} - m_0^2 \log \frac{\Lambda}{m_0} + \frac{\gamma}{2} m_0^2 & \text{improved Pauli-Villars} \end{cases}. \tag{11.19}$$

Thus the value of m_0 we should choose to match the observed value of m depends on which regularization scheme we use. This is sometimes described by saying that bare masses are **scheme-dependent**.

These expressions also illustrate what is called the **hierarchy problem**: in order for m^2 to be small, we need to carefully tune m_0 to cancel some power of Λ. Why should m_0, which after all is the fundamental parameter in the Lagrangian, be tuned in this way? Shouldn't we generically just expect that the physical mass will be of order the cutoff scale, say at the Planck scale of order 10^{18} GeV? This argument has sometimes been used to suggest that fundamental scalar fields should not appear in nature, and in particular to argue that the LHC shouldn't have found the Higgs boson (which is a massive scalar field). The LHC did find the Higgs boson of course, and actually we saw back when we studied free fields that there is an even more severe version of this problem for the cosmological constant. Most likely the issue is simply that we do not have a good understanding of which values of bare parameters are "typical", my own hope is that this question will be answered through a deeper understanding of quantum gravity.

11.1.4 Dimensional regularization

There is one more regularization scheme we will consider. It is simultaneously the most useful and the least physical - the **dimensional regularization** method of 't Hooft and Veltman. The basic problem with choosing a regulator is that it is usually hard to choose a regulator which preserves all symmetries of your theory. What 't Hooft and Veltman proposed to get around this is the following hack:

- Evaluate all loop integrals in low enough spacetime dimension d that they are convergent. Once you have the result, analytically continue back to the dimension you are actually interested in.

It is hopefully clear that this is only a hack. For example as far as I am aware it doesn't actually make sense to think about quantum field theory in 3.5 spacetime dimensions. On the other hand this procedure clearly will give some kind of well-defined answer, and as long as we only assign physical interpretations to universal quantities we can hope that this answer is the same as what we'd have gotten from some more physical regularization such as a lattice.

Let's see how this works in practice for the self-energy. The integral we need to evaluate is

$$\int_0^\infty \frac{q^{d-1} dq}{q^2 + m_0^2}, \tag{11.20}$$

which is convergent for $0 < d < 2$ and gives[75]

$$\int_0^\infty \frac{q^{d-1} dq}{q^2 + m_0^2} = \frac{\pi m_0^{d-2}}{2 \sin\left(\frac{d\pi}{2}\right)}. \tag{11.21}$$

[75] This is a special case of the integral (11.47) below, as you can check using the Γ function reflection formula $\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}$.

Combining this with the expression

$$\Omega_{d-1} = \frac{2\pi^{d/2}}{\Gamma(d/2)} \tag{11.22}$$

that we derived last time, we have

$$\Sigma(p^2) = \frac{\lambda_0}{2} \frac{2\pi^{d/2}}{\Gamma(d/2)} \frac{1}{(2\pi)^d} \frac{\pi m_0^{d-2}}{2\sin\frac{d\pi}{2}}. \tag{11.23}$$

Let's now think a bit about the analytic structure of this expression. Most of the factors are well-behaved for positive d, but the $\sin\frac{d\pi}{2}$ in the denominator leads to poles at each even value of d. Let's first therefore consider the odd values - in particular we have

$$\frac{\pi m_0^{d-2}}{2\sin\left(\frac{d\pi}{2}\right)} = \begin{cases} \frac{\pi}{2m_0} & d=1 \\ -\frac{\pi m_0}{2} & d=3 \end{cases}. \tag{11.24}$$

These are precisely the universal finite contributions we found using hard momentum cutoff and improved Pauli-Villars above! For $d = 1$ this is no mystery, since anyways the integral is convergent so the regulator can't matter, but for $d = 3$ the dimensional regularization method has automatically removed the linear divergence but kept the correct finite piece.

In even dimensions we need to be more careful due to the poles. The basic idea is to work in $d = 2(n-\epsilon)$ dimensions, in which case the pole at $d = 2n$ will show up as a factor of $1/\epsilon$. To expand (11.23) near $d = 2n$ we need two pieces of information. The first is the behavior of the sin factor near $d = 2n$, which is easily shown to be

$$\frac{1}{\sin(\pi(n-\epsilon))} = \frac{(-1)^{n+1}}{\pi\epsilon}\left(1 + O(\epsilon^2)\right). \tag{11.25}$$

We also need to know how to deal with the $\Gamma(n+\epsilon)$ in the denominator of (11.23). This is a bit trickier: from the Taylor expansion we have

$$\Gamma(n-\epsilon) = \Gamma(n)\left(1 - \psi(n)\epsilon + O(\epsilon^2)\right), \tag{11.26}$$

where

$$\psi(z) = \frac{\Gamma'(z)}{\Gamma(z)} \tag{11.27}$$

is called the **digamma function**. For our purposes it is useful to know that by taking the logarithm of the Γ-function recursion relation $\Gamma(z+1) = z\Gamma(z)$ and then differentiating we see that the digamma function obeys

$$\psi(z+1) = \psi(z) + \frac{1}{z}, \tag{11.28}$$

and thus

$$\psi(n) = \sum_{k=1}^{n-1} \frac{1}{k} + \psi(1) \tag{11.29}$$

for any positive integer n. Computing $\psi(1) = \Gamma'(1)$ is a bit tricky, but after a nasty integral evaluation (or more elegantly by using the Weierstrass product representation of Γ) one finds

$$\psi(1) = -\gamma \tag{11.30}$$

and thus

$$\psi(n) = \sum_{k=1}^{n-1} \frac{1}{k} - \gamma. \tag{11.31}$$

We will also at several points need to use the fact that for $a > 0$ we have

$$a^\epsilon = e^{\epsilon \log a} = 1 + \epsilon \log a + O(\epsilon^2). \tag{11.32}$$

Using all these we have

$$\Sigma(p^2) = \frac{\lambda_0}{2} \frac{\Omega_{2n-1}}{(2\pi)^{2n}} \times \frac{(-1)^{n+1} m_0^{2n-2}}{2} \left(\frac{1}{\epsilon} - \log(4\pi^3) - \gamma + \sum_{k=1}^{n-1} \frac{1}{k} - \log m_0^2 + O(\epsilon) \right), \tag{11.33}$$

so again we see that the logarithmic term in m_0 matches the logarithmic term we got from the hard cutoff and improved Pauli-Villars regulators.

You may be puzzled about how we were able to get a dimensionful quantity (m_0^2) inside of a logarithm in (11.33). To understand this we need to account for the dimensions of the bare coupling constant λ_0. In d spacetime dimensions a scalar field needs to have units of energy to the $(d-2)/2$, since we need the kinetic term $\partial_\mu \phi \partial^\mu \phi$ to have energy dimension d so that integrating it against $d^d x$ gives a dimensionless quantity. The interaction term $\lambda_0 \phi^4$ must also have energy dimension d, which means that the energy dimension of λ_0 is $(4-d)$. Energy dimensions are a very useful notion in quantum field theory, so there is a special notation for them: if a quantity O has units of energy to the Δ, then we write

$$[O] = \Delta. \tag{11.34}$$

For example in our free scalar theory we have

$$\begin{aligned} [\mathcal{L}] &= d \\ [\phi] &= \frac{d-2}{2} \\ [m_0^2] &= 2 \\ [\lambda_0] &= 4 - d. \end{aligned} \tag{11.35}$$

When doing dimensional regularization we wish to expand things around $d = 2n$, so we can write the bare coupling "near" $d = 2n$ in terms of the bare coupling "at" $d = 2n$ as

$$\lambda_0^{\{2(n-\epsilon)\}} = \mu^{2\epsilon} \lambda_0^{\{2n\}} = \lambda_0^{\{2n\}} \left(1 + \epsilon \log \mu^2\right), \tag{11.36}$$

where μ is an arbitrary quantity with energy dimension one. Substituting this into (11.33) we get

$$\Sigma(p^2) = \frac{\lambda_0^{\{2n\}}}{2} \frac{\Omega_{2n-1}}{(2\pi)^{2n}} \times \frac{(-1)^{n+1} m_0^{2n-2}}{2} \left(\frac{1}{\epsilon} - \log(4\pi^3) - \gamma + \sum_{k=1}^{n-1} \frac{1}{k} + \log \frac{\mu^2}{m_0^2} + O(\epsilon) \right), \tag{11.37}$$

which looks more sensible dimensionally. The scale μ is called the **renormalization scale**, we will discuss its physical interpretation below. Putting everything together we have the expressions

$$m^2 = m_0^2 + \begin{cases} \frac{\lambda_0}{4m_0} & d = 1 \\ \frac{\lambda_0}{8\pi} \left(\frac{1}{\epsilon} - \gamma + \log\left(\frac{\mu^2}{4\pi^3 m_0^2} \right) \right) & d = 2 \\ -\frac{\lambda_0 m_0}{8\pi} & d = 3 \\ -\frac{\lambda_0 m_0^2}{32\pi^2} \left(\frac{1}{\epsilon} - \gamma + 1 + \log\left(\frac{\mu^2}{4\pi^3 m_0^2} \right) \right) & d = 4 \end{cases} \tag{11.38}$$

for determining the physical mass at one-loop in terms of the bare mass and coupling in dimensional regularization. In practice the way this expression is usually used is the opposite however: the physical mass m is measured and then we use this formula to determine m_0.

11.2 Two-to-two scattering at one loop

Let's now see how we can use these ideas to study the one-loop contribution to $2 \to 2$ scattering in ϕ^4 theory that we computed last time:

$$i\widetilde{\mathcal{M}}_c(k_1,k_2\to k_1',k_2') \supset -\frac{\lambda_0^2}{2}\int\frac{d^d\ell}{(2\pi)^d}\frac{-i}{\ell^2+m_0^2-i\epsilon}\Big[\frac{-i}{(\ell+k_1'-k_1)^2+m_0^2-i\epsilon} + \frac{-i}{(\ell-k_1-k_2)^2+m_0^2-i\epsilon}+\frac{-i}{(\ell+k_2-k_1')^2+m_0^2-i\epsilon}\Big]. \tag{11.39}$$

The external momenta k_i, k_i' here should be taken on-shell to give a genuine scattering matrix element, but for now it is convenient to allow them to take general values so that we can analytically continue to Euclidean signature:

$$i\widetilde{\mathcal{M}}_c(k_1,k_2\to k_1',k_2') \supset \frac{i\lambda_0^2}{2}\int\frac{d^d\ell}{(2\pi)^d}\frac{1}{\ell^2+m_0^2}\Big[\frac{1}{(\ell+k_1'-k_1)^2+m_0^2} + \frac{1}{(\ell-k_1-k_2)^2+m_0^2}+\frac{1}{(\ell+k_2-k_1')^2+m_0^2}\Big]. \tag{11.40}$$

This integral is convergent for $0 < d < 4$ and logarithmically divergent for $d = 4$. We could study it using any of the regulators we discussed above, but we will stick to dimensional regularization so for now we are assuming that d is in the convergent range. The three terms give the same integral three times, so we just need to figure out how to compute

$$\mathcal{I}(q) = \int\frac{d^d\ell}{(2\pi)^d}\frac{1}{\ell^2+m_0^2}\frac{1}{(\ell+q)^2+m_0^2} \tag{11.41}$$

for general Euclidean q (we will eventually analytically continue back to on-shell Lorentzian q). This integral may look difficult, but there is a clever trick due to Feynman which makes it tractable: we use the identity (that you will derive in the homework)

$$\frac{1}{AB} = \int_0^1\frac{dx}{(xA+(1-x)B)^2}, \tag{11.42}$$

which is valid for $A, B > 0$, to combine the denominators:

$$\begin{aligned} x(\ell^2+m_0^2)+(1-x)((\ell+q)^2+m_0^2) &= \ell^2+2(1-x)\ell\cdot q+(1-x)q^2+m_0^2 \\ &= (\ell+(1-x)q)^2+x(1-x)q^2+m_0^2, \end{aligned} \tag{11.43}$$

and thus

$$\begin{aligned} \mathcal{I}(q) &= \int_0^1 dx\int\frac{d^d\ell}{(2\pi)^d}\frac{1}{\left((\ell+(1-x)q)^2+x(1-x)q^2+m_0^2\right)^2} \\ &= \int_0^1 dx\int\frac{d^d\ell}{(2\pi)^d}\frac{1}{\left(\ell^2+x(1-x)q^2+m_0^2\right)^2} \qquad (11.44) \\ &= \frac{\Omega_{d-1}}{(2\pi)^d}\int_0^1 dx\int_0^\infty\frac{\ell^{d-1}d\ell}{\left(\ell^2+x(1-x)q^2+m_0^2\right)^2} \qquad (11.45) \end{aligned}$$

In going to the second line we have made an additive shift of the integration variable, and in going to the third we changed to radial coordinates. x here is called a **Feynman parameter**. The propagators in more general loop diagrams can be combined using multiple Feynman parameters, for example we have

$$\frac{1}{ABC} = \int_0^1 dx\int_0^{1-x} dy\,\frac{2}{(xA+yB+(1-x-y)C)^3}. \tag{11.46}$$

The remaining momentum integral in (11.45) can then be evaluated using the general formula (that you will derive in the homework)

$$\int_0^\infty d\ell \frac{\ell^{a-1}}{(\ell^2+\sigma^2)^b} = \frac{\sigma^{a-2b}}{2}\frac{\Gamma(a/2)\Gamma(b-a/2)}{\Gamma(b)}, \tag{11.47}$$

which is valid for $\sigma > 0$ and $0 < a < 2b$, giving us

$$\begin{aligned}\mathcal{I}(q) &= \frac{\Omega_{d-1}}{(2\pi)^d}\frac{1}{2}\int_0^1 dx \left(m_0^2 + x(1-x)q^2\right)^{\frac{d-4}{2}} \frac{\Gamma(d/2)\Gamma(2-d/2)}{\Gamma(2)} \\ &= \frac{\Gamma(2-d/2)}{(4\pi)^{d/2}}\int_0^1 dx \left(m_0^2 + x(1-x)q^2\right)^{\frac{d-4}{2}}. \end{aligned} \tag{11.48}$$

Defining the **Mandelstam variables**

$$\begin{aligned} s &= -(k_1+k_2)^2 \\ t &= -(k_1'-k_1)^2 \\ u &= -(k_1'-k_2)^2, \end{aligned} \tag{11.49}$$

we can then write the one-loop contribution to the scattering amplitude as

$$i\widetilde{M}_c(k_1,k_2\to k_1',k_2') \supset \frac{i\lambda_0^2}{2}\frac{\Gamma(2-d/2)}{(4\pi)^{d/2}}\int_0^1 dx \left[\left(m_0^2 - x(1-x)s\right)^{\frac{d-4}{2}} + \left(m_0^2 - x(1-x)t\right)^{\frac{d-4}{2}} + \left(m_0^2 - x(1-x)u\right)^{\frac{d-4}{2}}\right]. \tag{11.50}$$

We will now focus on the cases of $d = 3$ and $d = 4$.

For $d = 3$ we are in the convergent region, so we simply have

$$\mathcal{I}(q) = \frac{1}{8\pi}\int_0^1 \frac{dx}{\sqrt{m_0^2 + x(1-x)q^2}}. \tag{11.51}$$

For $q^2 > 0$ this integral gives

$$\mathcal{I}(q) = \frac{1}{8\pi}\frac{\pi - 2\arctan\left(\frac{2m_0}{\sqrt{q^2}}\right)}{\sqrt{q^2}}, \tag{11.52}$$

which is the regime we need for the terms involving t and u since on shell we always have $t < 0$ and $u < 0$. To compute the integral involving s we need to restore the $i\epsilon$ by taking $q^2 = -(s+i\epsilon)$, which leads to

$$\mathcal{I}(q) = \frac{1}{8\pi}\frac{i\pi + \log\left(\frac{s+2m_0\sqrt{s}}{s-2m_0\sqrt{s}}\right)}{\sqrt{s}} \tag{11.53}$$

where we have used that $s \geq 4m_0^2$. We won't have too much to say about these results, but one comment is that at this order we can replace $m_0 \to m$ since the difference between the two is higher-order in λ_0, so the one loop contribution to $2 \to 2$ scattering is finite without any further renormalization. It also decays with energy since you will show in the homework that s is essentially just the center of mass energy squared.

For $d = 4$ we need to be more careful due to the pole in $\Gamma(2-d/2)$. Setting $d = 4-2\epsilon$, we have

$$\Gamma(\epsilon) = \frac{\Gamma(1+\epsilon)}{\epsilon} = \frac{1}{\epsilon}\left(1 + \psi(1)\epsilon + \ldots\right) = \frac{1}{\epsilon} - \gamma + O(\epsilon). \tag{11.54}$$

Using this in (11.48), together with (11.32), we have

$$\mathcal{I}(q) = \frac{1}{16\pi^2}\left(\frac{1}{\epsilon} - \gamma + \log(4\pi) - \int_0^1 dx \log\left(m_0^2 + x(1-x)q^2\right)\right). \tag{11.55}$$

To turn this into an expression for $\widetilde{M}_c$ we need to again be careful about units near 4 dimensions: noting that

$$[\widetilde{M}_c] = [\lambda_0] = 4 - d, \tag{11.56}$$

with factors of μ^ϵ inserted so that $\widetilde{M}$ and λ_0 have their $d = 4$ units we have

$$i\widetilde{M}_c = -i\lambda_0 + \frac{i\lambda_0^2}{32\pi^2}\left[\frac{3}{\epsilon} - 3\gamma + 3\log(4\pi) + \int_0^1 dx \left(\log\frac{\mu^2}{m_0^2 - sx(1-x)} + \log\frac{\mu^2}{m_0^2 - tx(1-x)} + \log\frac{\mu^2}{m_0^2 - ux(1-x)}\right)\right]. \tag{11.57}$$

This expression is indeed UV divergent as $\epsilon \to 0$. Before discussing how to fix this, I'll mention that in a more physical regularization scheme (e.g. such as a lattice or Pauli-Villars) we'd instead have found

$$i\widetilde{M}_c = -i\lambda_0 + \frac{i\lambda_0^2}{32\pi^2}\left[\int_0^1 dx \left(\log\frac{\Lambda^2}{m_0^2 - sx(1-x)} + \log\frac{\Lambda^2}{m_0^2 - tx(1-x)} + \log\frac{\Lambda^2}{m_0^2 - ux(1-x)}\right)\right] \tag{11.58}$$

where Λ is the UV cutoff. Either way we are now ready for the key question: what are we supposed to do about this UV divergence?

There is only one sensible thing to do: *we absorb this divergence into a redefinition of the bare coupling constant* λ_0. In the context of the bare mass m_0 we had a physical motivation for doing this: we wanted to write things in terms of the physical mass m instead of the scheme-dependent bare mass m_0. Is there a similar justification here? Indeed there is - the bare coupling λ_0 is no more directly measurable than the bare mass m_0. What *is* measurable is the $2 \to 2$ S-matrix, so the simplest thing we can do is simply *define* a physical coupling λ so that the *exact* $2 \to 2$ S-matrix is equal to its tree-level value at some preferred choice for the initial momenta. More concretely we will impose that

$$i\widetilde{M}_c|_{s=4m^2,t=0,u=0} = -i\lambda. \tag{11.59}$$

Looking at our above expressions, this requires that

$$\lambda = \lambda_0 - \frac{\lambda_0^2}{32\pi^2}\left[\frac{3}{\epsilon} - 3\gamma + 3\log(4\pi) + 2\log\frac{\mu^2}{m_0^2} + \int_0^1 dx \log\frac{\mu^2}{m_0^2 - 4m^2x(1-x)}\right] \tag{11.60}$$

in dimensional regularization and

$$\lambda = \lambda_0 - \frac{\lambda_0^2}{32\pi^2}\left[2\log\frac{\Lambda^2}{m_0^2} + \int_0^1 dx \log\frac{\Lambda^2}{m_0^2 - 4m^2x(1-x)}\right] \tag{11.61}$$

in a more physical regularization scheme. At the order we are working in λ_0 we can easily rewrite these as expressions for the bare coupling in terms of the physical mass and the physical coupling:

$$\begin{aligned}\lambda_0 =& \lambda + \frac{\lambda^2}{32\pi^2}\left[\frac{3}{\epsilon} - 3\gamma + 3\log(4\pi) + 2\log\frac{\mu^2}{m^2} + \int_0^1 dx \log\frac{\mu^2}{m^2 - 4m^2x(1-x)}\right] \\ =& \lambda + \frac{\lambda^2}{32\pi^2}\left[\frac{3}{\epsilon} - 3\gamma + 2 + 3\log(4\pi) + 3\log\frac{\mu^2}{m^2}\right]\end{aligned} \tag{11.62}$$

and

$$\begin{aligned}\lambda_0 =& \lambda + \frac{\lambda^2}{32\pi^2}\left[2\log\frac{\Lambda^2}{m^2} + \int_0^1 dx \log\frac{\Lambda^2}{m^2 - 4m^2x(1-x)}\right] \\ =& \lambda + \frac{\lambda^2}{32\pi^2}\left[3\log\frac{\Lambda^2}{m^2} + 2\right]\end{aligned} \tag{11.63}$$

Now the moment of truth: using either (11.62) or (11.63) we can rewrite the scattering amplitude $\widetilde{M}_c$ for general s, t, u in terms of the physical mass and coupling:

$$i\widetilde{M}_c = -i\lambda + \frac{i\lambda^2}{32\pi^2}\int_0^1 dx \left[\log\left(\frac{m^2 - 4m^2x(1-x)}{m^2 - sx(1-x)}\right) + \log\left(\frac{m^2}{m^2 - tx(1-x)}\right) + \log\left(\frac{m^2}{m^2 - ux(1-x)}\right)\right]. \tag{11.64}$$

All UV divergences are gone, and the answer is now independent of which regularization scheme we used! We have to fit two parameters (λ and m) to experiment, but this expression gives a function's worth of predictions in exchange. The integrals can again be evaluated in terms of inverse trig functions but I won't bother.

This argument leading to the finite and scheme-independent scattering amplitude (11.64) may have seemed a bit like magic. Why did this happen? Does it continue to happen at higher loops and for more complicated scattering amplitudes? Are there more parameters we need to tune, or is it just λ_0 and m_0? It is far from obvious, but the answers to the latter two questions are "yes it continues" and "no it is just λ_0 and m_0". Understanding why is our next order of business.

As a first indication that things may not be so mysterious, I'll mention that the derivative of λ_0 with respect to the logarithm of either the renormalization scale μ (in dim reg) or the cutoff Λ (in a more physical scheme) holding the physical coupling λ fixed is a very useful quantity, usually called the β-function. Here it is given by

$$\beta(\lambda) := \Lambda \frac{d\lambda_0}{d\Lambda} = \frac{3\lambda^2}{16\pi^2}. \tag{11.65}$$

Note in particular that λ_0 grows with energy, so when we reach a regime where $\frac{3\lambda}{16\pi^2} \log \frac{\Lambda}{m} \sim 1$, or in other words the cutoff reaches

$$\Lambda_{strong} \sim m e^{\frac{16\pi^2}{3\lambda}}, \tag{11.66}$$

then the theory becomes strongly coupled and our perturbative approach breaks down. This is usually viewed as evidence that the continuum limit does not really exist for ϕ^4 theory in $d = 4$. Fortunately if λ is small this is a rather high energy scale, for example in the standard model of particle physics the Higgs boson is a scalar field theory whose mass is 125GeV and whose quartic coupling is of order $\lambda \sim .1$, so the scale where the Higgs becomes strongly coupled is

$$\Lambda_{strong} \sim (125\,\mathrm{GeV}) \times e^{525}, \tag{11.67}$$

which is a far higher energy scale than the Planck scale of $M_p \sim 10^{18}$GeV where quantum gravity effects are expected to become important. If we view our theory as having a genuine cutoff Λ at some scale which is large compared to where we do experiments but small compared to Λ_{strong}, then these UV divergences start looking less scary and perhaps we will be able to tame them more systematically. Doing so is the goal of the next section.

11.3 Homework

1. Check the Feynman parameter identity (11.42) for $A, B > 0$.

2. Evaluate the general loop integral (11.47). One strategy is the following: 1) rescale ℓ to extract the overall power of σ, 2) rewrite the integral in terms of the Euler β-function

$$\beta(z_1, z_2) = \int_0^1 t^{z_1-1}(1-t)^{z_2-1}. \tag{11.68}$$

   using the change of variables $t = \frac{1}{1+\ell^2}$, and 3) use a famous expression for the β function,

$$\beta(z_1, z_2) = \frac{\Gamma(z_1)\Gamma(z_2)}{\Gamma(z_1+z_2)}. \tag{11.69}$$

   To derive this last expression, start with $\Gamma(z_1)\Gamma(z_2) = \int_0^\infty ds_1 \int_0^\infty ds_2 s_1^{z_1-1} s_2^{z_2-1} e^{-s_1-s_2}$ and then use the change of variables $s_1 = st$ and $s_2 = s(1-t)$.

3. Show that on shell the Mandelstam variables (11.49) obey $s + t + u = 4m^2$ and $s = E_{tot}^2$, where E_{tot} is the total energy in the center-of-mass frame. Also show that $t, u \leq 0$.

12 Renormalizability and the Renormalization Group

In the previous section we saw that once we expressed the one-loop $2 \to 2$ scattering amplitude of ϕ^4 theory in terms of the physical mass and coupling parameters m and λ, the amplitude was independent of the short-distance cutoff Λ (or the renormalization scale μ in dimensional regularization). In this section we will sketch a general understanding of why this is true, starting with a more "old-fashioned" approach based on showing that the divergences cancel in certain "renomalizeable" theories such as the ϕ^4 theory and then moving to a more modern "Wilsonian" approach based on viewing the cutoff Λ as being physical and then seeking to understand physics at energy scales which are low compared to Λ.

12.1 Power counting and renormalizability

Let's now consider in a rather general way the possible divergences of Feynman diagrams. We'll consider a general quantum field theory with a set of fields Φ^a and a set of interaction vertices labeled by i, with each interaction vertex involving N_i^a powers of Φ^a and d_i derivatives. We'll write the interaction Lagrangian as

$$\mathcal{L} = \sum_i \lambda_i O_i, \tag{12.1}$$

where O_i is some power of the fields and their derivatives and λ_i is a coupling constant. We will discuss below how to normalize the fields so that O_i and λ_i are separately well-defined. In this section we will study the convergence of a general one-particle irreducible diagram with E^a external Φ^a legs, I^a internal Φ^a legs, and V_i vertices of type i. We will focus on the particular region of the loop integration space where all loop momenta become large at the same rate. This is not the only region a divergence can come from, but our results will be indicative of the general case.

Before beginning we need to think a bit about the large-momentum behavior of the propagator. For a scalar field this is easy, it just goes like $1/k^2$. In section 9 you showed on the homework that the momentum-space Feynman propagator of a field of general spin is

$$G_F(k) = \frac{i}{k^0 - \omega_{n,\vec{k}} + i\epsilon} \frac{1}{2\omega_{n,\vec{k}}} \sum_\sigma u^a(\vec{k},\sigma,n) u^{b*}(\vec{k},\sigma,n) - (-1)^F \frac{i}{k^0 + \omega_{n,\vec{k}} - i\epsilon} \frac{1}{2\omega_{n,\vec{k}}} \sum_\sigma v^a(-\vec{k},\sigma,n^c) v^{b*}(-\vec{k},\sigma,n^c). \tag{12.2}$$

We haven't discussed the spin sums of the intertwiners yet, but they always give polynomials of k so at large k this propagator will go as

$$G_F \sim k^{2s_a - 2}, \tag{12.3}$$

where $2s_a$ is highest power of k appearing in the spin sums. Roughly speaking s_a is the spin of the ath field, for example for a spin 1/2 field we will see that $s_a = 1/2$ and for a massive vector field we'll have $s_a = 1$. In the massless case however sometimes s_a is lower than expected due to gauge symmetry, for example $s_f = 0$ for photons and gravitons. I will adopt a convention where the field is normalized so that the highest power of k in G_F has a coefficient of order one (perhaps multiplied by some dimensionless tensor such as $\eta^{\mu\nu}$ to make up the a, b^* indices),[76] in which case (12.3) tells us that the energy dimension of the field obeys

$$2[\Phi^a] - d = 2s_a - 2 \tag{12.6}$$

[76] This convention sometimes leads to surprising normalizations. For example for a massive vector field with Lagrangian

$$\mathcal{L} = -\frac{1}{4}(\partial_\mu V_\nu - \partial_\nu V_\mu)(\partial^\mu V^\nu - \partial^\nu V^\mu) - \frac{m^2}{2} V^\mu V_\mu \tag{12.4}$$

the Feynman propagator is

$$\langle T V^\mu(x) V^\nu(y) \rangle = \int \frac{d^d k}{(2\pi)^d} \frac{-i\left(\eta^{\mu\nu} - \frac{k^\mu k^\nu}{m^2}\right)}{k^2 + m^2 - i\epsilon} e^{ik\cdot(x-y)}, \tag{12.5}$$

so the field with dimension $s_a + \frac{d-2}{2}$ (with $s_a = 1$) is mV^μ. In particular this means that interactions such as $V^\mu V_\mu \phi^2$ in $d = 4$ are non-renormalizable in the classification introduced in a moment, even though when expressed in terms of V^μ (as opposed to mV^μ) they naively appear to be renormalizable.

and thus

$$[\Phi^a] = s_a + \frac{d-2}{2}. \tag{12.7}$$

Turning now to the question of the divergence of the loop integrals, each internal propagator contributes an integral $\frac{d^d k}{(2\pi)^d}$. The total number of loop integrals is $d\sum_a I_a$, but each vertex contributes a momentum-conserving δ function so the total number of loop integrals is

$$d\left(\sum_a I_a - \sum_i V_i + 1\right) \tag{12.8}$$

since there will always be a single momentum-conserving δ function left over which doesn't constrain the loop momenta. Going to spherical coordinates in this full space of loop integrals thus gives a radial momentum integral of the form

$$\int dk k^{d\left(\sum_a I_a - \sum_i V_i + 1\right) + 2\sum_a I_a(s_a - 1) + \sum V_i d_i - 1} \tag{12.9}$$

which will be divergent if the **degree of divergence**

$$D := \sum_a I_a \left(d + 2s_a - 2\right) + \sum_i V_i(d_i - d) + d \tag{12.10}$$

is greater than or equal to zero. We can simplify this expression by observing that since each internal a line connects two vertices and each external line is connected to one vertex we have

$$\sum_i V_i N_i^a = 2I_a + E_a, \tag{12.11}$$

and thus

$$D = d - \sum_a E_a \left(s_a + \frac{d-2}{2}\right) - \sum_i V_i \left(d - d_i - \sum_a N_i^a \left(s_a + \frac{d-2}{2}\right)\right). \tag{12.12}$$

We can write this more simply by observing that the quantity multiplying V_i in the sum over i is just d minus the energy dimension

$$\Delta_i := d_i + \sum_a N_i^a [\Phi^a] \tag{12.13}$$

of the operator appearing in the ith interaction vertex. We therefore have

$$D = d - \sum_a E_a [\Phi^a] - \sum_i V_i (d - \Delta_i). \tag{12.14}$$

The qualitative behavior of this formula depends very strongly on Δ_i: if all interactions obey $\Delta_i \leq d$, then adding additional interaction vertices cannot increase the degree of divergence. In this case the theory is said to be **renormalizable**. More generally we can classify interaction vertices into three groups:

- Vertices with $\Delta_i < d$ are called **super-renormalizable**. For example the ϕ^4 interaction in $d = 3$ obeys

  $$[\phi^4] = 2 < 3, \tag{12.15}$$

  and is thus super-renormalizable.

- Vertices with $\Delta_i \leq d$ are called **renormalizable**. For example the ϕ^4 interaction in $d = 4$ obeys

  $$[\phi^4] = 4, \tag{12.16}$$

  and is thus renormalizable but not super-renormalizable.

- Vertices with $\Delta_i > d$ are called **non-renormalizable**. For example the ϕ^4 interaction in $d = 5$ obeys

$$[\phi^4] = 6 > 5, \tag{12.17}$$

  and is thus non-renormalizable.

A theory with at least one non-renormalizable interaction is said to be non-renormalizable, as it has the property that diagrams become more and more divergent as we go to higher and higher orders in perturbation theory. On the other hand in a renormalizable theory there are only a finite number of amplitudes which have $D \geq 0$, namely those with

$$\sum_a E_a[\Phi^a] \leq d. \tag{12.18}$$

So for example a real scalar field has

$$[\Phi] = \frac{d-2}{2}, \tag{12.19}$$

so a scattering amplitude with E external particles can have $D \geq 0$ only if

$$E \leq \frac{2d}{d-2}. \tag{12.20}$$

For $d = 4$ this is $E \leq 4$, while for $d = 3$ this is $E \leq 6$. We will now argue that this translates into the statement that in a renormalizable theory UV divergences can be removed by absorbing them into a finite number of coupling constants. This is just what we found at one-loop in ϕ^4 theory in $d \leq 4$. In a theory where all interactions are super-renormalizable something even stronger is true: there are only a finite number of diagrams which are UV divergent. The UV divergences in a super-renormalizable theory can thus be completely removed by coupling constant shifts which are polynomials in the coupling, and that can be computed at some fixed order in perturbation theory. For example in the ϕ^4 theory in $d = 3$ we found no divergence at one loop in $2 \to 2$ scattering.

Before continuing it is worth mentioning that there is a simple interpretation of the condition for a vertex to be non-renormalizable: the quantity $d - \Delta_i$ is precisely the energy dimension of the coupling λ_i which appears in front of the interaction operator O_I. Non-renormalizable interactions are thus those with coupling constants that have negative energy dimension, while super-renormalizable interactions are those with positive energy dimension.

12.2 Cancellation of divergences in renormalizable theories

Continuing with our study of the divergence of a general 1PI diagram in the integration region where all momenta go to infinity together, we can consider what happens when we differentiate the diagram with respect to some external momentum p. Each time we do this it decreases the degree of divergence of the diagram by one since the derivative acts on the propagators, for example we have

$$\frac{d}{dp}\int \frac{d\ell}{2\pi}\frac{p+\ell+a}{(p+\ell)^2+b} = -\int \frac{d\ell}{2\pi}\frac{(p+\ell)^2+2a(p+\ell)-b}{((p+\ell)^2+b)^2}, \tag{12.21}$$

where the integral on the left is logarithmically divergent at large ℓ but the integral on the right is convergent. More generally if we differentiate a diagram with $D \geq 0$ a total number of $(D+1)$ times it becomes convergent. The divergent part of the diagram must therefore consist of a polynomial in the external momenta, heuristically of the form

$$\Lambda^D p^{M_D} + \Lambda^{D-1} p^{M_{D-1}} + \ldots \log \Lambda p^{M_0}, \tag{12.22}$$

where we have written p^{M_n} to represent any product of M_n components of the external momenta. These powers will also be multiplied by various mass scales from the coupling constants of the theory to make sure they have the right units (in a massless theory the units will need to work out without this). These however

are precisely the form of divergence which can be removed by adding local terms to the Lagrangian! More concretely, to remove a divergence of the heuristic form $\Lambda^{D-n}p^{M_{D-n}}$ in a diagram with E_a external Φ^a legs, we introduce a term with E_a factors of each field and M_{D-n} derivatives acting on those fields with the same index structure as in the divergence. For example let's say we are computing the self-energy of a scalar in $d = 4$ and we find the divergences

$$\Sigma(p^2) \supset a\Lambda^2 + (b + cp^2)\log\frac{\Lambda}{m}. \tag{12.23}$$

We can absorb $a\Lambda^2 + b\log\frac{\Lambda}{m}$ divergence into a shift of the bare mass term, and we can absorb the $cp^2\log\frac{\Lambda}{m}$ divergence into a shift of the kinetic term $\partial_\mu\Phi\partial^\mu\Phi$, i.e. into a wave function renormalization. In the previous section we computed a and b at one loop, and we pointed out that c is also nonzero at two loops. More generally we only need to do this subtraction for diagrams with $D \geq 0$, and thus we only need to include shifts of interaction terms with $\Delta_i \leq d$.

In the traditional way of describing this process one rewrites the bare Lagrangian in terms of the physical mass and coupling m and λ, and also defines a rescaled field

$$\Phi_R = \frac{\Phi}{Z} \tag{12.24}$$

which has a finite two-point function and in particular whose on-shell residue is the same as that of a free field. We thus have

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\Phi\partial^\mu\Phi - \frac{m_0^2}{2}\Phi^2 - \frac{\lambda_0}{4!}\Phi^4 \tag{12.25}$$

$$= -\frac{Z^2}{2}\partial_\mu\Phi_R\partial^\mu\Phi_R - \frac{Z^2m_0^2}{2}\Phi_R^2 - \frac{\lambda_0 Z^4}{4!}\Phi_R^4 \tag{12.26}$$

$$= -\frac{1}{2}\partial_\mu\Phi_R\partial^\mu\Phi_R - \frac{m}{2}\Phi_R^2 - \frac{\lambda}{4!}\Phi_R^4 + \mathcal{L}_{ct}, \tag{12.27}$$

where

$$\mathcal{L}_{ct} = -\frac{Z^2-1}{2}\partial_\mu\Phi_R\partial^\mu\Phi_R - \frac{Z^2m_0^2 - m^2}{2}\Phi_R^2 - \frac{\lambda_0 Z^4 - \lambda}{4!}\Phi_R^4 \tag{12.28}$$

is called the **counterterm Lagrangian** and its individual terms are called **counterterms**. In the old-fashioned approach to renormalization one views these counterterms as "corrections" to the original theory which are included to cancel the infinities. They are treated as additional interaction vertices, providing corrections to a free theory whose mass is now the physical mass and whose interaction vertex is now $-i\lambda$ instead of $-i\lambda_0$. This is not actually different from what we did in the previous section, where we followed the Wilsonian approach (to be developed further in the next section) of tuning the bare couplings so that the physical couplings have their observed values: the counterterms are just an alternative way of describing this tuning.

Before proceeding to the Wilsonian approach, we need to confront the fact that so far we have only considered the region of momentum integration where all loop momenta go to infinity together. This of course is an important contribution to the integrals, but we also need to consider the possibility of divergences that arise when only a subset of the momenta go to infinity. This is a subtle and difficult problem, whose traditional solution we won't describe in detail since the Wilsonian approach deals with it in a much cleaner (but more abstract) manner. We will instead content ourselves with a few remarks about the ingredients which go into the traditional proof that the same renormalization which removes the divergences in the integration region we have considered so far also removes them for the full range of momentum integration.

- The first step in proving renormalizability is **Weinberg's theorem**, which says that a multi-loop integral will be convergent if and only if its degree of divergence is negative as we take any linear combination of the loop momenta to infinity. This means that we can show convergence using a generalization of the method employed so far.

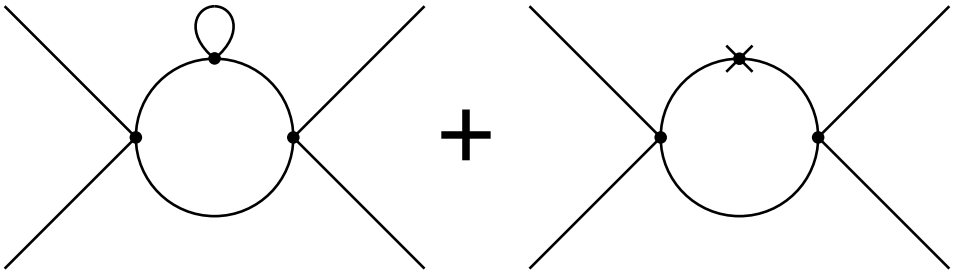


Figure 30: Canceling a divergent subdiagram with a counterterm. For $d = 2, 3$ the full diagram has $D < 0$, but the subdiagram is still divergent. Here the dot with an x through it indicates an insertion of the mass renormalization term $-\frac{Z^2 m_0^2 - m^2}{2}\Phi_R^2$.

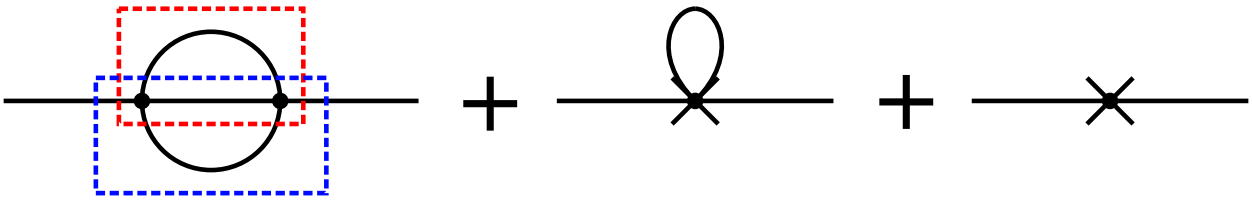


Figure 31: A Feynman diagram with overlapping divergences. The red and blue dashed lines surround four-point subdiagrams which each are logarithmically divergent in $d = 4$, but we can only use a four-point counterterm to cancel the divergence from one of them. The remaining divergence must be canceled by an additional two-point counterterm.

- One can think of the various options for which momenta go to infinity together in terms of **subdiagrams** of the full Feynman diagram. For example a diagram whose degree of divergence as defined above is negative can still diverge due to a subdiagram whose degree of divergence is positive. See figure 30 for an example.
- In simple cases there is a simple fix to the presence of a divergent subdiagram: we can simply ignore the rest of the diagram, in which case we have already seen that the divergence can be canceled by including an appropriate counterterm. At least to the extent that the propagators involved in the divergent subdiagram are not involved in other divergent subdiagrams, this cancellation works also in the full diagram (see figure 30).
- The key technical problem with this approach however is the possibility of **overlapping divergences**, meaning situations where we have multiple divergent subdiagrams with propagators in common. See figure 31 for an example. In such a case it isn't so clear that we can cancel both divergences with counterterms, as once we replace one of the subdiagrams by a counterterm we have lost part of the other subdiagram. The systematic approach to dealing with this goes under the name "BPHZ", for Bogoliubov, Parasiuk, Hepp, and Zimmerman, and it requires a detailed analysis of the structure of the diagrams using the infamous "forest formula". In the end everything does work though, and the renormalization which fixes the divergences in the region where all momenta scale together indeed removes the divergences from subdiagrams as well.

12.3 The Wilsonian approach

At least to my taste, the above discussion of renormalizability leaves something to be desired. We started with very simple ideas, essentially based on dimensional analysis, but then to turn these into a full proof of renormalizability we found that we needed to address some annoying technical problems. Shouldn't there be a better way that makes the physical meaning of renormalization obvious? Fortunately there is: the **Wilsonian approach** to renormalization.[77]

The first essential idea for the Wilsonian approach is to **view the cutoff as physical**. In a condensed matter system this is self-explanatory: at the atomic scale in a solid there is a genuine lattice of ions, with electrons constrained be near the ions, and at shorter distances there is nothing. In high-energy physics it is less obvious that there needs to be a genuine cutoff at short distances (or equivalently high energies), but the quantization of gravity seems to require major modifications of the laws of physics at the (absurdly small) **Planck length**:

$$\ell_p = \sqrt{\frac{\hbar G}{c^3}} \approx 10^{-35}\,\mathrm{m}. \tag{12.29}$$

Moreover there are several indications from particle physics (such as the mass of neutrinos, the existence of dark matter, and the small baryon-to-photon ration of the universe) that some kind of modification of the standard model of particle physics is necessary at sufficiently short distances.

The second essential idea for the Wilsonian approach is **decoupling**. This means that the details of what is going on at large energies/short distances do not affect what is going on low energies/long distances. For example if we regulate a scalar field theory by putting it on a lattice, when we look at the low-energy physics of the system we cannot tell whether the lattice has a cubic structure or a hexagonal structure. We also cannot detect the existence of very heavy particles by doing low-energy experiments.

The third essential idea for the Wilsonian approach is **integrating out**. The idea here is that since low-energy physics does not depend on the details of high-energy physics, rather than carrying around all that high-energy physics for no reason we can simply sum over it in the path integral once and for all. This produces a "low-energy effective field theory", where all effects of the high-energy modes are repackaged into the values of the low-energy coupling constants.

Indeed let's consider a rather general-looking quantum field theory with an explicit cutoff Λ, with action

$$S_\Lambda = \sum_i \int d^d x g_i(\Lambda) \Lambda^{d-\Delta_i} O_i. \tag{12.30}$$

Here O_i are some basis for all the scalar local operators in the theory, and Δ_i are their energy dimensions. In general there are infinitely many such operators, so the sum over i here needs to be viewed somewhat heuristically. We have chosen to extract a power of the cutoff Λ from the coupling constants, which is chosen so that the quantities $g_i(\Lambda)$ are dimensionless. The idea of the Wilsonian approach is that if we lower the cutoff from Λ_0 to Λ (with $\Lambda < \Lambda_0$), we should tune the Λ-dependence of the couplings so that the low-energy physics is not affected. You may worry whether or not we can do this, but in fact there is a simple path integral method: we split all fields into a "high-energy" part Φ_H, consisting of the modes which exist for cutoff Λ_0 but not for cutoff Λ, and a "low-energy" part Φ_L, consisting of the modes which exist for both cutoffs. For any observable $O_L[\Phi_L]$ built only out of the low-energy modes we then have

$$\begin{aligned}\langle O_L[\Phi_L]\rangle &= \frac{\int \mathcal{D}\phi_L \mathcal{D}\phi_H O_L[\phi_L] e^{-S_{\Lambda_0}[\phi_L,\phi_H]}}{\int \mathcal{D}\phi_L \mathcal{D}\phi_H e^{-S_{\Lambda_0}[\phi_L,\phi_H]}} \\ &= \frac{\int \mathcal{D}\phi_L O_L[\phi_L] e^{-S_\Lambda[\phi_L]}}{\int \mathcal{D}\phi_L e^{-S_\Lambda[\phi_L]}},\end{aligned} \tag{12.31}$$

[77] Kenneth G. Wilson's academic biography is an interesting one: during his PhD and also for eight years after he wrote almost no papers (in particular he wrote zero papers as a graduate student and his 1961 thesis still has zero citations). Somehow he managed to get a faculty position anyways, and also tenure at Cornell. He then proceeded to revolutionize physics, explaining the real meaning of renormalization in the process, and ended up with a Nobel Prize. I do not recommend trying to replicate this trajectory.

where

$$e^{-S_\Lambda[\phi_L]} := \int \mathcal{D}\phi_H e^{-S_{\Lambda_0}[\phi_L,\phi_H]}. \tag{12.32}$$

In other words, the low-energy effective action is obtained by starting with the full action and then integrating out the high-energy modes. This process gives a flow in the space of actions (or equivalently a flow in the space of coupling constants) which is called **renormalization group flow**.[78]

The operation (12.32) has an important defect: in general there is no reason for the action S_Λ to be local even if we start with a local action S_{Λ_0}. On the other hand since we only integrated out modes whose wavelengths are at most of order $\frac{1}{\Lambda}$, any non-localities we generate should be constrained to this scale. We therefore can Taylor expand them to express S_Λ as a local action order by order in $\frac{1}{\Lambda}$. This suppression is already built into our expression (12.30), as each derivative increase the dimension of O_i and thus costs a power of Λ. As a simple example of this, we can consider a non-local term

$$\int d^d x d^d y K(x-y)\phi(x)\phi(y). \tag{12.33}$$

Since this came from integrating out short-distance modes with momenta roughly between Λ and Λ_0, the Fourier transform of $K(x-y)$ should be a reasonably smooth function with compact support in k. K will therefore be an analytic function that decays rapidly at separations which are large compared to $\frac{1}{\Lambda}$. For ϕ configurations which vary only on scales which are large compared to $\frac{1}{\Lambda}$, we can therefore approximate K as a sum of δ-functions and their derivatives. In this way given a local action S_{Λ_0} with couplings

$$g_i(\Lambda_0) := g_i^0 \tag{12.34}$$

we can construct a local action S_Λ with couplings $g_i(\Lambda)$ that gives the same low-energy physics. At first order in $\Lambda_0 - \Lambda$ the new couplings are functions of the old couplings only, so they must obey Wilson's renormalization group (RG) equation

$$\Lambda \frac{dg_i}{d\Lambda} = \beta_i(g(\Lambda)). \tag{12.35}$$

In other words we can think of the renormalization group flow as being the integral curves generated by a vector field β_i on the space of couplings. In the last section we computed the β function for the scalar coupling $\frac{\lambda}{4!}\phi^4$ in $d = 4$ at one loop, finding

$$\beta_\lambda(\lambda) = \frac{3\lambda^2}{16\pi^2}. \tag{12.36}$$

It must be emphasized that this flow generically generates all possible couplings which are allowed by the symmetries of the theory - it does *NOT* only generate renormalizable couplings with $\Delta_i \leq d$.[79]

12.4 Polchinski's theorem

Using Wilson's ideas, there is a beautiful argument due to Polchinski which explains in a deep way the renormalizability results sketched in the previous subsections of this section. The idea is that in theories which start out weakly-coupled at the initial cutoff scale Λ_0, the RG equation (12.35) has a powerful focusing behavior that suppresses any information about what is going on at the scale Λ_0. More precisely, there is a finite-dimensional manifold in the space of coupling constants, whose dimensionality is equal to the number of renormalizable couplings, which is an attractor for the renormalization group. Wherever we start out, we eventually end up near this manifold (at least as long as we stay within the region of validity for perturbation

[78] The name is misleading, as renormalization group flow is not invertible (how would you "un-integrate"?) so there isn't really a group structure. A more accurate name would be "renormalization semigroup", but unfortunately we are stuck with this one.

[79] The only exceptions I know of to this rule are free theories, conformal field theories for which all β_i vanish, and supersymmetric theories.

theory). The only remaining high-energy information about where we started is where on this manifold we end up. The space of low-energy theories therefore can be parametrized by the renormalizable couplings only.

To see this focusing behavior, we can study how the renormalization group equation behaves under a small change δg_i in the trajectory. Working to first order in δg_i, we have

$$\Lambda \frac{d\delta g_i}{d\Lambda} = \sum_j M_{ij} \delta g_j, \tag{12.37}$$

with

$$M_{ij} = \frac{\partial \beta_i}{\partial g_j}. \tag{12.38}$$

Introducing matrix notation and also using $'$ to indicate a derivative with respect to $\log \Lambda$, we can rewrite this equation as

$$\delta g' = M \delta g. \tag{12.39}$$

So far this equation does not distinguish between renormalizable and non-renormalizable couplings. To distinguish them, it is useful to introduce a projection matrix

$$P_{ij} = \begin{cases} \delta_{ij} & i \text{ renormalizable} \\ 0 & \text{otherwise,} \end{cases} \tag{12.40}$$

and also a matrix

$$D_{ij} = \frac{\partial g_i}{\partial g_j^0}, \tag{12.41}$$

where in D_{ij} the derivative is computed for the particular trajectory $g_i(\Lambda)$ which obeys $g_i(\Lambda_0) = g_i^0$. Following Polchinski we can then introduce a clever second projection

$$\Pi = 1 - DP(PDP)^{-1}P, \tag{12.42}$$

which is designed to decouple the renormalizable and non-renormalizable couplings in the RG equation.[80] Π is indeed a projection in the linear algebra sense of obeying

$$\Pi^2 = \Pi, \tag{12.43}$$

but it is *not* orthogonal in the sense that it doesn't obey $\Pi^\dagger = \Pi$. P and Π are related by the equations[81]

$$\begin{aligned} P\Pi &= 0 \\ \Pi(1-P) &= (1-P). \end{aligned} \tag{12.44}$$

We can then define a projected coupling variation

$$\xi = \Pi \delta g, \tag{12.45}$$

which by construction obeys

$$P\xi = 0 \tag{12.46}$$

[80] The inverse matrix $(PDP)^{-1}$ here should only be used on vectors which are in the image of P. Otherwise the inverse does not exist.

[81] The relationship between P and Π is interesting from a linear algebra point of view. If Π were hermitian then equations (12.44) would imply that $\Pi = 1 - P$. Since Π is not hermitian, we can only conclude that $(1-P)v = v \Leftrightarrow \Pi v = v$. We will see im a moment however that what we are really interested in the null space of Π, and this need not coincide with the null space of $1 - P$.

so all renormalizable couplings have been removed. To compute the derivative of ξ with respect to scale we first need to understand the scale dependence of Π. Differentiating both sides of the RG equation (12.35) with respect to g_j^0 and using that partials of g_i with respect to Λ and g_j^0 commute, we have

$$\frac{\partial g_i'}{\partial g_j^0} = \sum_k \frac{\partial \beta_i}{\partial g_k} \frac{\partial g_k}{\partial g_j^0} \tag{12.47}$$

and thus

$$D' = MD. \tag{12.48}$$

Moreover since for any matrix N we have

$$(N^{-1})' = -N^{-1}N'N^{-1}, \tag{12.49}$$

we also have

$$\begin{aligned} \Pi' &= 1 - MDP(PDP)^{-1}P + DP(PDP)^{-1}PMDP(PDP)^{-1}P \\ &= -\Pi MDP(PDP)^{-1}P. \end{aligned} \tag{12.50}$$

We thus straightforwardly have

$$\begin{aligned} \xi' &= \Pi \delta g' + \Pi' \delta g \\ &= \Pi M \delta g - \Pi MDP(PDP)^{-1}P\delta g \\ &= \Pi M \xi, \end{aligned} \tag{12.51}$$

so the projection Π has succeeded in decoupling the RG equation. Rewriting this in terms of P and D we have

$$\xi' = \left(M - DP(PDP)^{-1}PM\right)\xi. \tag{12.52}$$

So far our discussion has been non-perturbative. In a situation where perturbation theory is valid, we can usefully approximate the matrix M using free field theory. In free field theory the action should not have any cutoff dependence since there are no loop diagrams, so we need the quantities

$$g_i(\Lambda)\Lambda^{d-\Delta_i} \tag{12.53}$$

to be cutoff-independent as the interactions go to zero. We therefore have

$$g_i' \approx (\Delta_i - d)g_i, \tag{12.54}$$

and thus

$$\beta_i \approx (\Delta_i - d)g_i \tag{12.55}$$

and

$$M_{ij} = \partial_j \beta_i \approx (\Delta_i - d)\delta_{ij}. \tag{12.56}$$

The key point is then the following. The renormalizable components of ξ are zero by construction, and to the extent that M is diagonal in the same basis as P we can ignore the second term in (12.52) since then

$$PM\xi \approx MP\xi = 0. \tag{12.57}$$

We therefore have

$$\xi_i' \approx \begin{cases} 0 & i \text{ renormalizable} \\ (\Delta_i - d)\xi_i & i \text{ non-renormalizable.} \end{cases} \tag{12.58}$$

The non-renormalizable couplings are precisely those for which $\Delta_i - d > 0$, so we thus see that the entire vector ξ vanishes like a power of $\frac{\Lambda}{\Lambda_0}$ as we flow to $\Lambda \ll \Lambda_0$! And moreover this conclusion is preserved under perturbative corrections as long as these are small compared to $\Delta_i - d$ (which they always will be for small enough coupling). Once this suppression is complete, the full set of coupling variations needs to obey

$$\Pi \delta g = 0, \tag{12.59}$$

or more explicitly

$$\delta g = DP(PDP)^{-1} P \delta g. \tag{12.60}$$

In other words we can determine the change in all of the infinitely many non-renormalizable couplings by looking at the change in the renormalizable couplings alone. Said differently, if we know the values of all of the renormalizable couplings in the low-energy action then the non-renormalizable couplings are all determined. This, in essence, is the statement of renormalizability!

To see more closely the connection between (12.60) and renormalizability, we first should note that although the matrix D, which depends on the cutoff Λ_0 and initial couplings g_i^0, appears in equation (12.60), the relationship determining the non-renormalizable couplings in terms of the renormalizable ones actually can't depend on these. This is because the focusing behavior of the RG equation (12.35) is a purely local affair in the space of couplings: we are solving a first-order differential equation, and we only need to know what is going on in the vicinity of where we are solving it. The finite-dimensional attractor manifold therefore cannot depend on where the flows started. This is the essential point: all low-energy observables can be computed using only the low-energy action $S_\Lambda[\phi_L]$, and thus expressed entirely in terms of where we are on the attractor manifold. We can parametrize where we are on this manifold using the low-energy renormalizable couplings, in which case all results will depend only on these low-energy couplings and the (low) cutoff scale Λ, *NOT* on the initial cutoff Λ_0 or initial couplings g_i^0. But this is precisely the statement of renormalizability: all observables can be expressed as functions of the low-energy couplings and kinematic variables without any dependence on the cutoff or the bare couplings.

12.5 Why renormalizability?

In the traditional approach to renormalizability there is a preferred set of theories, the renormalizable theories, where only renormalizable terms in the action have nonzero coupling constants. This is a powerful constraint on theories, as it forbids most of the potential terms one could write in the Lagrangian. At any point you could have asked me why I did not add terms like ϕ^{42} or $(\partial_\mu \phi \partial^\mu \phi)^2$ to the Lagrangian of our interacting scalar theory, and in the end renormalizability is the reason. For example the standard model of particle physics is a renormalizable theory (at least until we include neutrino masses and gravity), and this seems essential for it to be predictive. In particular in the first section we mentioned as a great success of quantum field theory that we can compute the anomalous magnetic moment of the electron to many significant figures and it agrees with experiment. On the other hand there is a simple non-renormalizable term we could add to the Lagrangian which would allow us to tune this quantity to be whatever we want, and if this were allowed then we could no longer say that the standard model predicts a definite value.

In the Wilsonian approach however, we generically study actions with nonzero coupling constants for all possible terms in the action. Does this mean that we have given up on the spectacular predictivity of renormalizable theories? In fact we have not: what Polchinski's theorem shows is that even if we allow non-renormalizeable terms to be turned on, at low energies we still only have a finite number of parameters for the theory; in fact we have precisely the same number of parameters as we'd have had by restricting to renormalizeable terms alone. From the Wilsonian point of view, the point is not that we need to restrict to theories with renormalizable terms only: instead the right statement is that **even if we allow non-renormalizable terms in the bare action, at low energies the theory will still look like a renormalizable theory!** In particular we do not lose any low-energy information by setting the non-renormalizeable couplings to zero in the bare theory, so we might as well do so: we are back to the old-fashioned renormalized perturbation theory we constructed above.

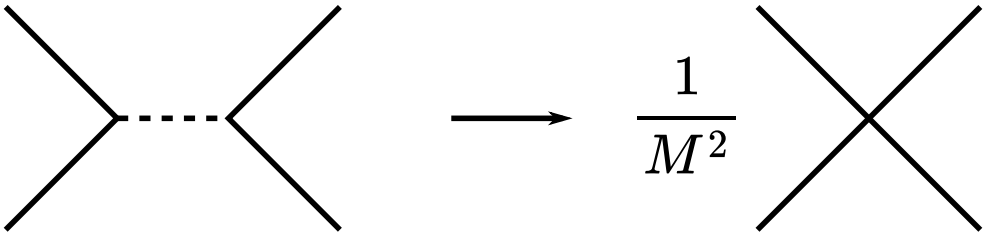


Figure 32: Integrating out a heavy particle of mass M creates new interactions for the light fields which are suppressed by the mass of the particle.

How did non-renormalizeable terms become so much less threatening than they seemed in the traditional approach? The reason is that our formula (12.14) for the degree of divergence of a Feynman diagram assumed that the coupling constants λ_i are independent of the cutoff scale Λ, while in the Wilsonian approach we take $\lambda_i = g_i \Lambda^{d-\Delta_i}$. The increasing degree of divergence as we bring down more powers of a non-renormalizeable vertex is thus offset by the factor of $\Lambda^{d-\Delta_i}$ which multiplies g_i. For this reason Wilson invented a new terminology for classifying operators to replace the old one of renormalizable vs. non-renormalizable:

- An operator O_i with dimension $\Delta_i < d$ is said to be **relevant**
- An operator O_i with dimension $\Delta_i = d$ is said to be **marginal**
- An operator O_i with dimension $\Delta_i > d$ is said to be **irrelevant**.

In particular note the demotion of operators with $\Delta_i > d$ from "non-renormalizable" to "irrelevant": if we change the dimensionless coupling for an irrelevant operator by an $O(1)$ amount at short distance, the only effect at low energies is a shift of where we are on the attractor manifold that could just as well have been achieved by changing the coefficients of the relevant and marginal operators alone. These days most non-ancient theoretical physicists prefer this terminology for classifying operators to the old one, and in fact it has been something of a chore for me to not use it thus far. From now on I will switch to using it.

12.6 Effective field theory

We've now seen that the RG flow equation tends to suppress information about high-energy physics. If E is the energy scale where we are doing experiments, then high-energy details at some cutoff scale Λ are suppressed by powers of the dimensionless ratio $\frac{E}{\Lambda}$. This is both a blessing and a curse: it means that we can figure out a theory of low-energy physics without needing to know what is happening at high-energy, but it also makes it hard to figure out what is going on at high energy!

Fortunately for us there are two situations where this suppression of high-energy physics is not complete. The obvious situation is when $\frac{E}{\Lambda}$ isn't that small - there is some new physics at an energy scale Λ which is high but not too high. For example there could be a heavy massive particle with mass M interacting with a massless particle at energy scale $E \ll M$, in which case the massless particle would have irrelevant interactions suppressed by powers of M arising from Feynman diagrams where the massive particle was exchanged (see figure 32). As long as we work only up to a fixed order in $\frac{E}{M}$ we only need to include a finite number of diagrams, since including more propagators of the heavy particle gives us higher and higher inverse powers of M.

The less obvious, but still very important, situation where irrelevant interactions can't be neglected is when they are the *only* kind of interactions. The canonical example of this situation is in general relativity, which is Einstein's theory of gravity. Its Lagrangian density (with zero cosmological constant) is

$$\mathcal{L} = \frac{1}{16\pi G} R, \tag{12.61}$$

where G is Newton's constant and R is the Ricci scalar. To make this look more like quantum field theory we define

$$g_{\mu\nu} = \eta_{\mu\nu} + \sqrt{G} h_{\mu\nu}, \tag{12.62}$$

in terms of which we heuristically have

$$\mathcal{L} = \partial h \partial h \left(1 + \sqrt{G} h + G h^2 + \dots\right). \tag{12.63}$$

Here we have not indicated how the indices are contracted or attempted to compute $O(1)$ factors. The key point however is that all interaction terms are suppressed by powers of the Planck mass

$$M_p = \sqrt{\frac{\hbar c}{G}} \approx 2.2 \times 10^{-8}\,\text{kg}. \tag{12.64}$$

This may not seem like a large mass compared to your own mass, but it is a gigantic energy scale for an elementary particle. For example it is about 10^{19} times the mass of a proton, which is about 1 GeV. Nonetheless gravity is a part of our every day experience, since the tiny gravitational force of each proton in the earth on each atom in our bodies adds up and there is no competing force to overwhelm it.

Another example of a field with only irrelevant interactions is a real scalar field with a shift symmetry $\phi' = \phi + a$. The Lagrangian for this theory can only be made out of derivatives of ϕ, and the only relevant or marginal term of this type is the massless kinetic term $-\frac{1}{2}\partial_\mu\phi\partial^\mu\phi$. We therefore need to include irrelevant operators such as

$$\mathcal{L} \supset -\frac{g}{\Lambda^d}(\partial_\mu\phi\partial^\mu\phi)^2 \tag{12.65}$$

to get nontrivial scattering. This example has great physical relevance, as it arises whenever there is a "spontaneously-broken continuous global symmetry". For example this happens in nuclear physics, where the pion fields are the scalars, and also in condensed matter physics systems such as liquid helium at the critical point. We will learn more about these next semester.

The unifying theme of these examples is that we can parametrize the effects of unknown high-energy physics by including irrelevant operators in the low-energy theory suppressed by powers of the energy scale Λ of that unknown physics. The theory including these terms is only valid when viewed as computing an expansion in $\frac{E}{\Lambda}$. Such a theory is called an **effective field theory**. Our current best understanding of the laws of physics is an effective field theory, as it includes irrelevant operators to explain gravity and also the observed nonzero values of neutrino masses. Effective field theories inevitably break down when we consider energies of order Λ, and to understand what happens then we need to know the real high-energy physics. For pions and liquid helium we already know this, while finding it for gravity is one of the biggest problems in physics. We will meet effective field theories again in the next semester, and in fact there is an entire class about them taught by Iain Stewart here at MIT.[82]

12.7 Fixed points and conformal symmetry

One of the most important aspects of the renormalization group is the possibility of **fixed points**, meaning points g_i^* in the space of couplings where

$$\beta_i(g^*) = 0 \tag{12.66}$$

for all i. The theories which live at these couplings are necessarily scale-invariant, meaning that in addition to Poincaré symmetry they also have **dilation symmetry**

$$x^{\mu'} = \lambda x^\mu. \tag{12.67}$$

This is because at a fixed point the theory has no dimensionful parameters, so any dimensionless observable must be invariant under a change of units.

[82] You can find this class on MIT open courseware!

It is widely expected that in relativistic field theories a fixed point must also have a larger spacetime symmetry group called **conformal symmetry**, which is the set of diffeomorphisms which send the spacetime metric to a scalar multiple of itself.[83] Infinitesimally conformal transformations are generated by **conformal Killing vectors**, which in Minkowski space obey

$$\partial_\mu \xi_\nu + \partial_\nu \xi_\mu = A \eta_{\mu\nu}. \tag{12.68}$$

We can determine A by contracting both sides of this equation with $\eta^{\mu\nu}$, leading to

$$\partial_\mu \xi_\nu + \partial_\nu \xi_\mu = \frac{2}{d} \partial_\alpha \xi^\alpha \eta_{\mu\nu}. \tag{12.69}$$

The general solution of this equation is

$$\xi_\mu(x) = a_\mu + \omega_{\mu\nu} x^\nu + b x_\mu + 2 c_\alpha x^\alpha x_\mu - c_\mu x^2, \tag{12.70}$$

where $\omega_{\mu\nu} = -\omega_{\nu\mu}$. The first two terms here are infinitesimal Poincaré transformations, while b parametrizes an infinitesimal dilation. The vector c^α parametrizes the infinitesimal version of what is called a **special conformal transformation**. A quantum field theory with conformal symmetry is called a **conformal field theory**, so in relativistic field theory a fixed point of the renormalization group likely always corresponds to a conformal field theory. In what follows we will not need to use conformal symmetry however, so we will stick to the language of fixed points.

Fixed points are natural "starting" and "ending" points for the renormalization group flow. The typical situation is that we begin with a "UV" fixed point, deform by a relevant operator with a small coefficient, and then flow off in the space of couplings until we reach some other "IR" fixed point. As a simple example we can consider our old friend the free massive scalar theory:

$$\mathcal{L} = -\frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{m^2}{2} \phi^2. \tag{12.71}$$

In this theory the only nontrivial coupling is the mass m^2, which we will parametrize as a function of the cutoff by

$$m^2 = g_2 \Lambda^2 \tag{12.72}$$

as usual. The renormalization group equation is quite simple:

$$g_2' = -2 g_2, \tag{12.73}$$

so the β-function is

$$\beta_2(g_2) = -2 g_2. \tag{12.74}$$

Thus we see that to get a fixed point we need $g_2 = 0$, or equivalently $m^2 = 0$. This is quite sensible: if m^2 is not zero then the theory has a dimensionful parameter and cannot be scale invariant. This fixed point is the simplest conformal field theory: the massless free scalar. If we now deform the action by turning on a small nonzero value g_2^0 for g_2 at some cutoff scale Λ_0, then we now have

$$g_2(\Lambda) = \left(\frac{\Lambda_0}{\Lambda} \right)^2 g_2^0. \tag{12.75}$$

When $\Lambda \sim \Lambda_0$ this is a small contribution, but as we lower Λ it grows and once we get to the regime where

$$\Lambda \sim \Lambda_0 \sqrt{g_2^0} \tag{12.76}$$

this deformation has a large effect on the theory. Of course in this case the right-hand side of (12.76) is just the mass m, so it is hardly news that the mass becomes important for energies $E \lesssim m$. Indeed below this

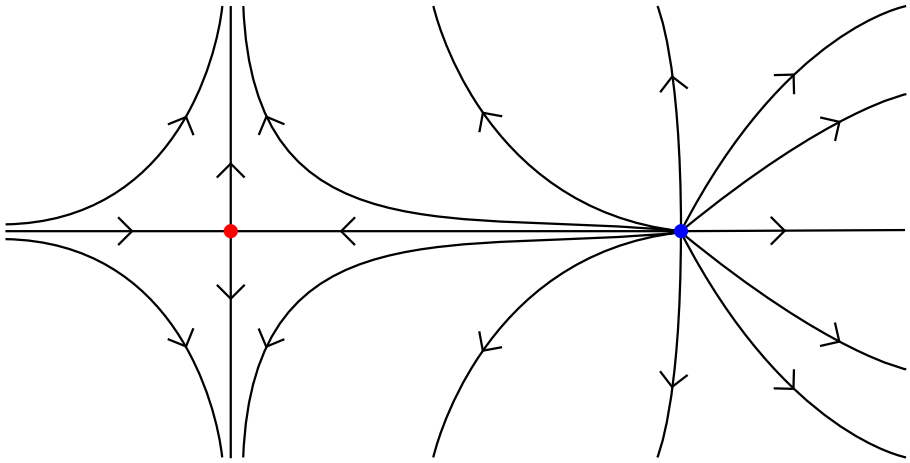


Figure 33: Renormalization group flows in the vicinity of a UV fixed point (shown in blue) with two relevant operators and an IR fixed point (shown in red) with one relevant operator. To hit the IR fixed point, we need to tune the initial flow direction from the blue point, otherwise we flow off to what is likely a trivial theory.

scale there are no states except for the ground state, so in this particular renormalization group flow the IR fixed point is the trivial conformal field theory with zero degrees of freedom.

This last example may have you worried that the IR fixed point in quantum field theory is often the trivial CFT. Indeed this is generically the case, for a simple reason: as long as the candidate IR CFT has at least one scalar relevant operator which is invariant under all symmetries of the UV CFT, then generically the RG flow is repelled from the fixed point in this direction so we need to tune a continuous parameter in the initial conditions to hit it. See figure 33 for an illustration. If the candidate IR CFT has more than one invariant relevant operator, then we need to tune a continuous parameter for each such operator. Sometimes however you get lucky: the IR CFT can have no invariant relevant operators or there may be some reason why they cannot be turned on. We will meet examples of this type next semester.

12.8 Critical phenomena

Let's now use all this technology to predict something for a real experiment. The system we will study is a classical Ising magnet in three spatial dimensions. This has a phase transition as we vary the temperature, exhibiting spontaneous magnetization when $T < T_c$ where T_c is either called the Curie temperature or the critical temperature. In particular for $T \approx T_c$ this system is described by the Euclidean version of our old friend the massive scalar field ϕ, with Lagrangian

$$\mathcal{L}_E = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi + \frac{m^2}{2}\phi^2 + \frac{\lambda}{4!}\phi^4. \tag{12.77}$$

Here ϕ is essentially the average magnetization, the Ising spin-flip symmetry is represented as $\phi' = -\phi$, and[84]

$$m^2 \propto T - T_c \tag{12.78}$$

so the phase transition happens at $m = 0$. This tuning to $m = 0$ is precisely the tuning mentioned at the end of the previous section, so at low energy this theory at $m = 0$ should be described by a nontrivial conformal

[83]This expectation is a theorem for $d = 2$, proven by our friend Polchinski, and so far no counterexamples have been found for $d > 2$.

[84]The argument for this that m^2 should vanish at $T = T_c$ by scale invariance and the effective Lagrangian should be analytic in T, so generically it should vanish linearly.

field theory with one relevant operator that is invariant under the spin-flip symmetry. This theory is not so easy to compute in, as for $d = 3$ the operator ϕ^4 in the free theory (the UV fixed point) is relevant so at low energy its dimensionless coupling becomes strong. Indeed finding a reliable way to do computations in the IR fixed point of the critical Ising model in $d = 3$ is one of the most famous problems in theoretical physics.[85] We will now see that by using a clever trick due to Wilson and Fisher we can compute some aspects of this theory surprisingly reliably using results we have already obtained.

We first need to get a sense of what kind of quantity we would like to compute. The first thing to note is that as we flow to the IR fixed point some renormalization of the operator ϕ^2 will typically be necessary. In other words the operator which has cutoff-independent correlation functions will have the form

$$\phi^2 = [\phi^2]_0 \Lambda^{\gamma_{\phi^2}}, \tag{12.79}$$

where $[\phi^2]_0$ is the "bare" ϕ^2 operator at the UV fixed point and γ_{ϕ^2} is called the **anomalous dimension** of ϕ^2. The full energy dimension of ϕ^2 at the IR fixed point (working for the moment in d spacetime dimensions) is thus

$$\Delta_{\phi^2} = d - 2 + \gamma_{\phi^2}. \tag{12.80}$$

We can read off the anomalous dimension of ϕ^2 from its Euclidean two-point function at the critical point, since by dimensional analysis this must be given by

$$\langle \phi^2(x)\phi^2(y)\rangle = \frac{C}{|x-y|^{2\Delta_{\phi^2}}} \tag{12.81}$$

with C a dimensionless constant. It is convenient to take the Fourier transform of this, which by dimensional analysis must be

$$\langle \phi^2(k_1)\phi^2(k_2)\rangle = (2\pi)^d \delta^d(k_1+k_2) D |k_1|^{2\Delta_{\phi_2}-d}, \tag{12.82}$$

with D again a dimensionless constant. Since the quantity $m^2\phi^2$ must have dimension d, we must have

$$[m^2] = d - \Delta_{\phi^2} = 2 - \gamma_{\phi^2}, \tag{12.83}$$

so we can write

$$m^2 = \xi^{\gamma_{\phi^2}-2} \tag{12.84}$$

where ξ has units of length and is called the **correlation length** of the system. Combining this with (12.78), we see that we must have

$$\frac{1}{\xi} \sim (T - T_c)^\nu \tag{12.85}$$

with

$$\nu = \frac{1}{2 - \gamma_{\phi^2}}. \tag{12.86}$$

ν here is an example of what is called a **critical exponent**, and the relation (12.85) is easily measurable in a real magnet. There are other critical exponents for other thermodynamic quantities, and all of them can be related to the energy dimensions of relevant operators in the IR CFT. Computing these dimensions is thus the central problem in understanding the Ising phase transition.[86]

Now, following Wilson and Fisher, let's see how to compute the anomalous dimension γ_{ϕ^2}. The method we will use is called the "ϵ-expansion", and if this is the first time you are hearing it you may think I am

[85] For $d \geq 4$ the ϕ^4 coupling is marginal or irrelevant (and in the marginal $d = 4$ case it still flows to zero in the IR since the one-loop β function is positive), so the IR CFT is just the massless free scalar theory. For $d = 2$ the scalar description breaks down due to infrared divergences and other methods are needed; we will show next semester that the IR fixed point for $d = 2$ is actually a free fermion theory.

[86] In fact the same IR CFT governs many other physical systems, including the critical point of the phase diagram of water. All of these systems have the same critical exponents, which is a rather remarkable convergence due to the great differences in the underlying physics of these system. This "universality" is a beautiful illustration of the focusing power of the renormalization group.

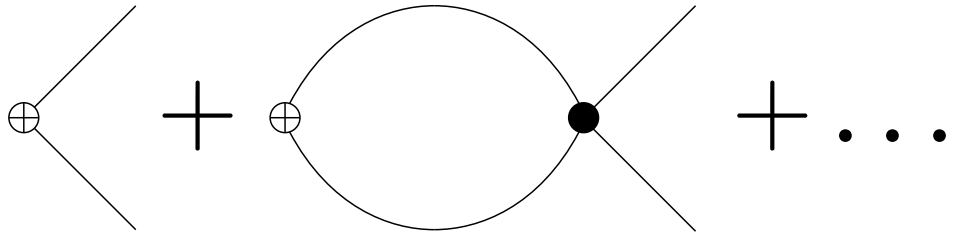


Figure 34: Leading diagrams contributing to an insertion of $[\phi^2]_0$ into a correlation function.

crazy. The idea is to continuously connect the nontrivial IR CFT in $d = 3$ to the free scalar CFT in $d = 4$ by taking $d = 4 - 2\epsilon$, expanding perturbatively in ϵ, and then setting $\epsilon = 1/2$. It is not clear a priori that this is a good thing to do, but it turns out that the $O(1)$ coefficients in this expansion work out in such a way that $\epsilon = 1/2$ is small enough to get a decent approximation.[87] Let's first recall that in $d = 4$ we found the expression

$$\beta = \frac{3\lambda^2}{16\pi^2} \tag{12.87}$$

for the β-function of the quartic coupling in $\lambda\phi^4$ theory. For $d = 4 - 2\epsilon$ this coupling becomes dimensionful, so following the Wilsonian approach we should introduce a dimensionless coupling g_4 via

$$\lambda = g_4 \Lambda^{2\epsilon}. \tag{12.88}$$

The coupling g_4 thus has nontrivial scale dependence even in the free theory, scaling like $g_4(\Lambda) \sim \Lambda^{-2\epsilon}$. Its β-function at one loop is thus

$$\beta = -2\epsilon g_4 + \frac{3g_4^2}{16\pi^2}. \tag{12.89}$$

We can therefore find a fixed point by canceling these two terms against each other, leading to

$$g_4^* = \frac{32\pi^2\epsilon}{3}. \tag{12.90}$$

If we take $\epsilon \to \frac{1}{2}$ this gives a rather large coupling in $d = 3$, but we can boldly press ahead and see what we find for the anomalous dimension γ_{ϕ^2}.

For small ϵ we can compute γ_{ϕ^2} by studying the renormalization of the composite operator ϕ^2. So far we have not discussed how to compute correlation functions of composite operators, but the basic idea is simple: start with the pieces of the operator at different points, and then bring them together ignoring any diagrams with propagators connecting the pieces of the operator. In particular for ϕ^2 we subtract a factor of $G_F(0)$, removing the obvious divergence proportional to the identity operator as we bring the two ϕ's together. This renormalization of composite operators is called **normal ordering**, and it must be done even in free field theory to define a sensible composite operator. We can also understand normal ordering in the operator approach, where the divergence arises from the term

$$\Phi^2(x) \supset \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{1}{2\sqrt{\omega_{\vec{p}}\omega_{\vec{p}'}}} e^{i(p-p')x} a_{\vec{p}} a^\dagger_{\vec{p}'}, \tag{12.91}$$

which has the divergent vacuum expectation value

$$\langle \Phi^2(x) \rangle = \int \frac{d^{d-1}p}{(2\pi)^{d-1}} \frac{1}{2\omega_{\vec{p}}} = G_F(0). \tag{12.92}$$

[87] There are other more modern (and more rigorous) approaches to doing this calculation, but for the most part they require substantial numerical work while the ϵ-expansion gives quick analytic results that already work pretty well.

What normal ordering does in free field theory is re-order all products of a's and $a^\dagger$'s so that the $a^\dagger$'s are to the left of the a's, ensuring a vanishing vacuum expectation value. It is convenient to instead compute correlation functions of the rescaled operator $\frac{1}{2}\phi^2$, as Feynman diagrams for these have symmetry factors that work in the way we are familiar with (the $1/2$ is similar to the $1/4!$ we put in front of ϕ^4, and cancels the two ways that incoming propagators can be attached to the operator). The leading diagrams arising from an insertion of $\frac{1}{2}[\phi^2]_0$ are shown in figure 34. Evaluating these diagrams we see that at this order the only effect is to multiply the Fourier transform

$$\frac{1}{2}[\phi^2(p)]_0 = \frac{1}{2}\int dx e^{-ip\cdot x}[\phi^2(x)]_0 \tag{12.93}$$

by a factor

$$N_{\phi^2} = \left[1 - \frac{\lambda}{2}\mathcal{I}(p) + O(\lambda^2)\right], \tag{12.94}$$

where $\mathcal{I}(p)$ is our old friend

$$\mathcal{I}(q) = \int \frac{d^d\ell}{(2\pi)^d}\frac{1}{\ell^2+m^2}\frac{1}{(\ell+q)^2+m^2} = \frac{1}{16\pi^2}\left(\frac{1}{\epsilon} - \gamma + \log(4\pi) - \int_0^1 dx \log\left(m^2 + x(1-x)q^2\right)\right). \tag{12.95}$$

Here we are interested in the massless case, so rewriting things in terms of g_4 we have

$$N_{\phi^2} = \left[1 - \frac{g_4}{32\pi^2}\left(\frac{1}{\epsilon} - \gamma + \log(4\pi) + 2 + \log\frac{\Lambda^2}{p^2}\right) + O(g_4^2)\right]. \tag{12.96}$$

Absorbing the finite one-loop contributions into a rescaling of the cutoff via

$$\log\Lambda'^2 = \frac{1}{\epsilon} - \gamma + \log(4\pi) + 2 + \log\Lambda^2, \tag{12.97}$$

we can write this as

$$\begin{aligned} N_{\phi^2} &= \left[1 - \frac{g_4}{32\pi^2}\log\left(\frac{\Lambda'^2}{p^2}\right) + O(g_4^2)\right] \\ &= \left(\frac{\Lambda'^2}{p^2}\right)^{-\frac{g_4}{32\pi^2}}\left[1 + O(g_4^2)\right] \end{aligned} \tag{12.98}$$

Therefore we can remove the cutoff dependence by defining the renormalized operator

$$\phi^2 = \Lambda'^{\frac{g_4}{16\pi^2}}[\phi^2]_0. \tag{12.99}$$

which means that the anomalous dimension of ϕ^2 is

$$\gamma_{\phi^2} = \frac{g_4}{16\pi^2}. \tag{12.100}$$

At the fixed point (12.90) we therefore have

$$\gamma_{\phi^2} = \frac{2\epsilon}{3}, \tag{12.101}$$

so from (12.86) the critical exponent ν is given by

$$\nu = \frac{1}{2} + \frac{\epsilon}{6} + O(\epsilon^2). \tag{12.102}$$

Boldly setting $\epsilon = 1/2$ to get to $d = 3$ we therefore have the prediction

$$\nu \approx \frac{7}{12} \approx .583. \tag{12.103}$$

This prediction is fairly close to the experimental value $.625 \pm .006$, which is nice. We can do better working to higher order in ϵ, the state of the art calculation based on this method gives $.6290 \pm .0025$,[88] which is in quite impressive agreement with the experimental value! In the next semester we will meet many other experimental successes of quantum field theory.

[88] These values are from Zinn-Justin's book "Quantum field theory and critical phenomena".