

Quantum Field Theory Lecture Notes

Silvia Nagy

September 8, 2026

Contents

1	Introduction	2
1.1	Why quantum field theory?	2
1.2	Conventions	6
2	Special relativity and Lorentz invariance	6
2.1	Review of Lorentz invariance	6
2.2	Group theory of Lorentz group	9
2.3	Translations and the Poincaré group	11
2.4	Realizing transformations on fields	12
2.4.1	Scalar fields	12
2.4.2	Vector fields	14
3	Lagrangian methods and classical field theory	14
3.1	Lagrangians for classical mechanics	14
3.2	Lagrangian methods for classical field theory	15
3.2.1	Action of a real scalar field	16
3.3	Noether's theorem	17
3.3.1	Example: a complex scalar field	18
3.3.2	Proof of Noether's theorem	18
3.3.3	Alternative way to find the current	20
3.4	Hamiltonian formalism in classical mechanics	20
4	Quantum field theory in canonical formalism	22
4.1	Quantizing the simple harmonic oscillator	23

4.2	Quantizing the free complex scalar field	25
4.2.1	Operator-valued Fourier expansions	26
4.2.2	Algebra of creation/annihilation operators	27
4.2.3	Hamiltonian and the energy of the vacuum	29
4.2.4	Single particle states	30
4.2.5	The Fock space	32
4.3	Propagators and causality	33
4.3.1	Lorentz-invariant integration measures	33
4.3.2	Commutators and the light-cone	34
4.3.3	Particle propagation	35
4.4	Feynman propagator	36
5	Interacting quantum field theories	40
5.1	The interaction picture	41
5.2	Wick's theorem	44
5.3	Feynman diagrams and Feynman rules	47
5.4	Wrapping up	52

These are the lecture notes for Quantum Field Theory course for the 2026 STFC HEP Summer School in Durham. They are based on an Advanced Quantum Theory course I taught at Durham, and the lecture notes borrow a lot from previous courses by Nabil Iqbal and Charlotte Sleight. Please send any errors to silvia.nagy@durham.ac.uk. Some useful books are:

- Peskin & Schroeder, "An Introduction to Quantum Field Theory" – this is in many ways the standard reference in the field.
- Anthony Zee, "QFT in a Nutshell" – a delightful book, a bit thin on the actual calculations but great for inspiration (and bed-time reading, and learning about new applications of QFT).
- Mark Srednicki, "Quantum Field Theory" – nicely complementary to P&S. A version of the book is available for free from the author's website: <http://web.physics.ucsb.edu/~mark/qft.html>

These lectures probably cover more than we will be able to go over during the week! Feel free to use them as a resource for future study, along with the books above, and many other reviews and books available on the subject.

1 Introduction

1.1 Why quantum field theory?

This is a course on **quantum field theory**. Quantum field theory is an amazing and wonderful subject. It is the mathematical framework behind our most precise understanding of the natural world, and it is really

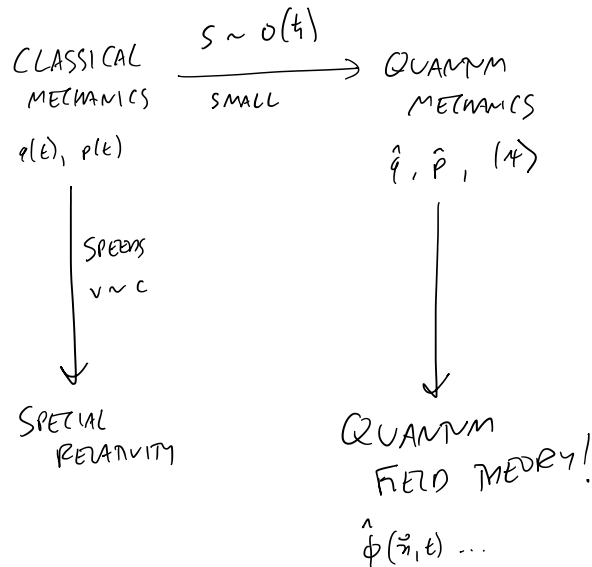


Figure 1.1: A square showing different kinematic regimes of nature

the language in which essentially all modern research in quantum physics is done. I am happy to say that by the end of this year you will not only understand the principles behind it, but you will also be able to use it to *calculate* observable things.

I should also say at the outset that quantum field theory can be hard. It is a bit of a conceptual leap, and the calculations that are involved can be difficult. There are also many different ways to say the same thing, as you will see. This is part of the fun, but as always the things that are the most fun invariably involve a bit of a challenge.

So before the inevitable blizzard of indices and infinite-dimensional integrals that will follow, let's first understand: **why** do we need quantum field theory?

Let's start by drawing a map of physical theories, as in Figure 1.1.

The goal is to understand things that are both small, *and* that move fast. Some examples are:

1. **...light.** You may have learned in an earlier quantum mechanics class that there are “light particles” called photons. (If you haven't learned this, don't worry; it's hard to explain it properly without quantum field theory). You might expect that light particles are “small”, and they move at the speed of light. Thus a careful quantum mechanical treatment of light will invariably require quantum field theory.
2. **Particle colliders**, e.g. the LHC at CERN – these are devices that make individual particles go very, *very* fast. It then smashes them together. To figure out what happens in this process requires quantum field theory. (This is maybe a good time to stress that in reality this “smashing” operation is a kind of microscope – we do this so that we can see what is happening at small scales. We will learn more about this later).
3. **Very hot things.** If you heat something up, the particles in it move around more quickly. If you heat it up *enough*, you might worry that they move around near the speed of light. You then need QFT to understand this. This sort of thing does not happen very often, but it did happen in the early

universe – the “Big Bang” was basically an explosion of arbitrarily high temperature, and to describe what happens afterwards requires the formalism of quantum field theory.

So the point is that quantum field theory is basically what you get if you combine quantum mechanics and special relativity. It may surprise you that this needs a whole course; why can’t we just make quantum mechanics relativistically invariant and move on with our life? (In other words, why isn’t this subject called “relativistic quantum mechanics” instead of “quantum field theory”?).

It turns out a great deal of new physics enters. Let’s heuristically try and understand why. This is really for inspiration, later on we will do this all precisely. Consider a free particle with mass m . In non-relativistic classical mechanics, the energy for the particle would be

$$E_{\text{non-rel}} = \frac{\vec{p}^2}{2m} \quad (1.1)$$

where $\vec{p}$ is the 3-vector momentum of the particle. This is not the right equation for a *relativistic* particle, however. The correct equation there turns out to be:

$$E_{\text{rel}} = \sqrt{\vec{p}^2 c^2 + m^2 c^4} \approx mc^2 + \frac{\vec{p}^2}{2m} + \mathcal{O}(p^4) \quad (1.2)$$

You may recall this equation from an earlier class on relativity; if not, we will recap quite a lot of it soon so don’t worry.

In the last equality I performed a Taylor expansion in powers of p^2 . Note that if we set $p = 0$ we get the world’s most famous equation $E = mc^2$: this says is that in relativity a single particle has an *inextricable amount of energy* mc^2 called the rest mass¹ (the thing we normally call “kinetic energy” is basically a correction on top of that). The reason why this formula is so profound is that it tells you that particles always have energy, even when they are not moving² – however, reversing the logic you might imagine that if we supply enough energy, we might be able to create a particle! In particular if we supply energy $E = 2mc^2$ then we can create a particle/anti-particle pair.

So far I have only discussed relativity; now let us include quantum mechanics. In quantum mechanics, we know that energy *fluctuates*. More precisely, what this means is that we will often consider a state that is not an energy eigenstate, which means that it will have some spread of energies ΔE . If we ever end up with

$$\Delta E \approx 2mc^2 \quad (1.3)$$

then it suggests that the *fluctuations* of energy will be enough to create a particle/anti-particle pair. In other words, there exist states where the particle number will change. But this is not allowed in traditional quantum mechanics – you always write down the Schrodinger equation for 1 particle, or 2 particles, or some fixed number N , which you fix to start with. Relativity tells us this is not enough. We need a new formalism entirely, one where the state space allows the particle number to change. That is the formalism of **quantum field theory**.

Let us push on this a bit more and figure out when we expect this new formalism to be experimentally relevant. Imagine putting the particle in a box of size L . Recall the Heisenberg uncertainty relation

$$\Delta q \Delta p \geq \frac{\hbar}{2} \quad (1.4)$$

What does this mean for our particle in the box? We see that because the position is known to within an accuracy of at least L , there is an uncertainty in the momentum of at least

$$\Delta p \geq \frac{\hbar}{2L} \quad (1.5)$$

¹Just a note that in some sources m_0 will be used instead of m in reference to the rest mass.

²Of course the liberation of this energy is how nuclear power works, and the conversion factor of c^2 is one of the reasons it is so powerful.

Now I want to use this to figure out the typical uncertainty in ΔE .

$$E_{\text{rel}} = \sqrt{\vec{p}^2 c^2 + m^2 c^4} \approx |\vec{p}|c \quad (1.6)$$

where the last equality is true in the limit $p \rightarrow \infty$. If we use the last expression to figure out ΔE then we have

$$\Delta E \sim c \Delta p \sim \frac{c \hbar}{2L} \quad (1.7)$$

(Note that self-consistency means that this is actually the correct expression only as we take L very small, so let's do that from now on).

Now from above we know that something interesting will happen when we have $\Delta E \sim 2mc^2$, which means that

$$\Delta E \sim 2mc^2 \quad \rightarrow \quad \boxed{L \sim \frac{\hbar}{mc}} \quad (1.8)$$

This is a very interesting formula. Notice that it combines $\hbar$ (for quantum) and c (for relativity) and it says that if you try to confine a particle to a box of size smaller than size $\frac{\hbar}{mc}$ then the quantum fluctuations of energy will threaten to create particles from the vacuum. For a given particle this number is called the (reduced) *Compton* wavelength:

$$L_{\text{Compton}} \equiv \frac{\hbar}{mc} \quad (1.9)$$

For an electron this wavelength is 10^{-12} m or so.

Now I haven't proven to you that relativistic quantum mechanics doesn't work; it can be an instructive exercise to just ignore what I told you above and *try* and build a relativistic version of quantum mechanics and check that things go wrong. You get confusing things like negative probabilities, negative energies, etc. I may discuss some of these things as we go on; it requires a bit of machinery of Lorentz-invariance.

But first let me tell you where we are going. In traditional classical mechanics the degrees of freedom are the numbers $q_a(t)$, $p_a(t)$. Here a runs over the degrees of freedom; for example for a particle in three dimensions we would have $a \in \{1, 2, 3\}$.

In quantum mechanics we learned how to *canonically quantize* these. The process of canonical quantization is (for us) an algorithm: we stick in a classical theory, we do some manipulations, and then we end up with a quantum theory, where we obtain *operators* $\hat{q}_a, \hat{p}_a$ that obey a Heisenberg uncertainty relation like (1.4).

Now let's think about classical *field* theory. In classical field theory, the basic degree of freedom – i.e. the analogue of $q(t)$ – is a *field* $\phi(\vec{x}, t)$. In this example, I am talking about a real scalar field, which is a function from space and time to the real numbers:

$$\phi : \mathbb{R}^{3,1} \rightarrow \mathbb{R} \quad (1.10)$$

we can have other kinds of fields; for example a complex scalar field, or a vector field such as $\vec{E}(x, t)$, the electric field from elementary electrodynamics. Note that this object $\phi(\vec{x}, t)$ takes on a different value at *each point in space*, so it is a lot of degrees of freedom – infinitely many, in fact.

In classical field theory $\phi(\vec{x}, t)$ will have a conjugate momentum which I will call $\pi(\vec{x}, t)$. We will discuss the equations that these two things obey. Now to get quantum field theory from this, we again canonically quantize this theory, and we end up with a quantum system with *operators* $\hat{\phi}(\vec{x})$ and $\hat{\pi}(\vec{x})$. This process – i.e. quantizing fields rather than simple degrees of freedom like $q_a(t)$ – is usually called *second quantization*. In my opinion (...not shared by everyone) calling this “second” quantization is somewhat confusing terminology and is largely for historical reasons.

Note that we have infinitely many operators – one quantum operator at *each point in space*, and thinking about how to organize this information is one of the main goals of this course.

Now let me tell you what will happen once we understand the physics above. We will understand the following facts about the universe:

1. There are different, yet *totally indistinguishable* copies of elementary particles like the electron.
2. There is a relationship between the *statistics* of particles (i.e. the behavior of their exchange) and their *spin* (i.e. what happens if you rotate them).
3. Anti-particles exist.
4. Particles can be created and destroyed. (We already talked about this).
5. Finally, and perhaps most profoundly: the things that we call *forces* can be imagined as being caused by the *exchange of particles*.

1.2 Conventions

We will use the following conventions: from now on we will mostly set $\hbar = c = 1$. Our metric signature will be $(-, +, +, +)$; note that many quantum field theory textbooks use a different convention, so stay vigilant for sign differences.

2 Special relativity and Lorentz invariance

We will now review Lorentz invariance. However, before doing that, let's discuss something much simpler: we will warm up very quickly with *rotational invariance*.

2.1 Review of Lorentz invariance

Einstein defined special relativity by the following two postulates:

1. **The principle of relativity: the laws of nature are the same in all inertial frames.**
2. **The speed of light is the same in all inertial frames.**

I haven't defined an inertial frame yet. Informally, two inertial frames are related in the following way: if I am at rest in an inertial frame, and you are moving past me at a constant speed, then you are also in an inertial frame. I'll give a more formal definition slightly below.

Now let us put some mathematics on these ideas. We have space and time, which I will denote as $\mathbb{R}^{1,3}$; the "1" here separates the time from the space.

We label a point in spacetime by $x^\mu = (t, x^1, x^2, x^3) = (x^0, x^1, x^2, x^3)$ (I will sometimes call the time component x^0). Note that I have put the μ index "up"; this is called the contravariant 4-vector³ and this notation carries some information and is important in what follows.

Now consider the following transformation to a new coordinate system x'^μ :

$$x'^\mu = \begin{pmatrix} t' \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma \left(t - \frac{v}{c^2} x^1 \right) \\ \gamma \left(x^1 - vt \right) \\ x^2 \\ x^3 \end{pmatrix} \quad \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2.1)$$

³I will never use that name again and just call it "up".

This is called a **Lorentz boost**. It mixes together space and time, and you should think about it as a fancy version of the rotations we discussed previously.

Let's think about what it means for a second. Consider a particle sitting at rest in the origin of unprimed coordinate system, which means that it extends along the following line:

$$\begin{pmatrix} t \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}_{\text{particle}} = \begin{pmatrix} \tau \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (2.2)$$

Here τ is a parameter which runs along the worldline of the particle. What is happening in the primed coordinate system? There we have

$$\begin{pmatrix} t' \\ x'^1 \end{pmatrix}_{\text{particle}} = \begin{pmatrix} \gamma\tau \\ -v\gamma\tau \end{pmatrix} \quad (2.3)$$

(I am ignoring $x^{2,3}$ as they don't do anything). In other words, in the primed coordinate system the particle is *moving*, it does not have a constant x'^1 . Thus the primed coordinates correspond to a frame that is moving with respect to the original frame. We will see what they are for shortly.

Note that if we take the limit $c \rightarrow \infty$ – this is the *non-relativistic limit* – this transformation becomes;

$$x'^\mu = \begin{pmatrix} t' \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} t \\ x^1 - vt \\ x^2 \\ x^3 \end{pmatrix} \quad (2.4)$$

which is called a “Galilean boost”, familiar from Newtonian mechanics. From now on I am going to set $c = 1$, and it will never appear again in any of our formulas. **Exercise:** how then do we recover the non-relativistic limit?

Now, the point is the following: recall that in the section on rotational invariance, we realized that some things (lengths of vectors, dot products, etc.) were invariant under rotations, and some things weren't. We now want to build the same formalism for Lorentz boosts: in other words, what is invariant under Lorentz boosts?

I will first define the spacetime Minkowski metric $\eta_{\mu\nu}$ with “down” indices, which is defined to be

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.5)$$

The point of the spacetime metric is to allow us to “lower” indices in the following way:

$$x_\mu \equiv (x_0, x_1, x_2, x_3) = \sum_\nu \eta_{\mu\nu} x^\nu = (-x^0, x^1, x^2, x^3) \quad (2.6)$$

This is of course just matrix multiplication with the η matrix. This “lower” vector is called the covariant vector. From now on, we will use the Einstein summation convention and not write down the sums explicitly; whenever you see a repeated index you should imagine a sum, i.e.

$$x_\mu = \sum_\nu \eta_{\mu\nu} x^\nu = \eta_{\mu\nu} x^\nu \quad (2.7)$$

Now, let me define for you another object: $\eta^{\mu\nu}$ with only up indices. This is defined as the matrix inverse of the metric tensor with only down indices. Component-wise you can just take the inverse and find

$$\eta^{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.8)$$

(i.e. it's numerically the same). But because it is the inverse, we can now invert (2.7) to find:

$$x^\mu = \eta^{\mu\nu} x_\nu \quad (2.9)$$

So $\eta^{\mu\nu}$ *raises* indices. By the way, note that the statement that $\eta^{\mu\nu}$ is the inverse of $\eta_{\mu\nu}$ can be written as

$$\eta^{\mu\nu} \eta_{\nu\rho} = \delta_\rho^\mu \quad (2.10)$$

where δ_ρ^μ is the identity matrix:

$$\delta_\rho^\mu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.11)$$

Now the point of all of this fancy machinery is to allow us to construct an inner product between two vectors x^μ and y^μ as follows: the dot product is defined to be

$$x \cdot y = x^\mu y^\nu \eta_{\mu\nu} = x^T \eta y = (x^0, \vec{x}) \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y^0 \\ \vec{y} \end{pmatrix} = -x^0 y^0 + x^1 y^1 + x^2 y^2 + x^3 y^3 \quad (2.12)$$

This is very similar to the familiar dot product that you know and love from ordinary $\mathbb{R}^3$, except for the extra minus sign in front of the time component of the vector. This minus sign can be thought of as the ultimate origin of the (...observationally quite obvious...) fact that time is different from space. You should now convince yourselves that all of the following ways of computing this dot product give you the same answer:

$$x \cdot y = x^\mu y_\mu = x_\mu y^\mu = x_\mu y_\nu \eta^{\mu\nu} \quad (2.13)$$

In other words, you can always contract an upper and a lower index in this way. This is the nice thing about this notation; it basically makes it impossible to mess up the dot product, *unless* you write something like this:

$$x^\mu y^\mu \quad (2.14)$$

with both indices up. You should not do this (and indeed even writing this leaves me with a terrible feeling of discomfort).

Now let us ask the following question, which should now seem quite natural given our discussion of two dimensional rotational invariance: what is the most general transformation of the coordinates that leaves this dot product between two vectors invariant? In other words, what is the most general 4×4 matrix $\Lambda = \Lambda^\mu_\nu$ such that I can write:

$$x' \cdot y' = (x')^\mu (y')_\mu = x^\mu y_\mu = x \cdot y \quad (x')^\mu = \Lambda^\mu_\nu x^\nu \quad y'^\mu = \Lambda^\mu_\nu y^\nu \quad (2.15)$$

We see that the equation to be satisfied is:

$$x^T \Lambda^T \eta \Lambda y = x^T \eta y \quad \rightarrow \quad \boxed{\Lambda^T \eta \Lambda = \eta} \quad (2.16)$$

or, in indices, the same equation becomes:

$$\boxed{\Lambda^\mu_\nu \eta_{\mu\rho} \Lambda^\rho_\sigma = \eta_{\nu\sigma}} \quad (2.17)$$

A matrix satisfying this property is an element of a group called $O(1, 3)$, or the **Lorentz group**. You should imagine that this matrix “preserves the Minkowski metric”.

It is helpful to convince ourselves that the boost introduced back in (2.1) does satisfy this relation. Suppressing the $x^{2,3}$ directions, from there we have

$$\Lambda^\mu_\nu = \gamma \begin{pmatrix} 1 & -v \\ -v & 1 \end{pmatrix} \quad (2.18)$$

and now just explicitly calculate that

$$\gamma^2 \begin{pmatrix} 1 & -v \\ -v & 1 \end{pmatrix}^T \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -v \\ -v & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad (2.19)$$

as claimed.

2.2 Group theory of Lorentz group

That was an example. Let us now understand the full set of Λ for which this is true. First, let’s take the determinant of the equation (2.16). We then find that

$$(\det \Lambda)^2 \det \eta = \det \eta \quad \rightarrow \quad \det \Lambda = \pm 1 \quad (2.20)$$

Thus the set of Λ fall into two disjoint sets, depending on the sign of their determinant. Let’s first consider the set with positive determinant. These are the Λ ’s that are continuously connected to the identity⁴. As you might remember from GMP, it is then a useful thing to consider the Lie algebra: to remind you how this works, consider a rotation of the following form:

$$\Lambda = \exp(\mathcal{M}) \quad (2.21)$$

where $\mathcal{M}$ is a 4×4 matrix which we can write in indices as $\mathcal{M}^\mu_\nu$. We now want to understand the space of possible $\mathcal{M}$ ’s, which are the generators of the group. To understand this, let us start by assuming that $\mathcal{M}$ is small; in that case we can Taylor expand in powers of it, to find:

$$\Lambda = \mathbf{1} + \mathcal{M} + \mathcal{O}(\mathcal{M}^2) \quad (2.22)$$

Now let us plug this into (2.17). We find that

$$(\delta^\mu_\nu + \mathcal{M}^\mu_\nu) \eta_{\mu\rho} (\delta^\rho_\sigma + \mathcal{M}^\rho_\sigma) = \eta_{\nu\sigma} \quad (2.23)$$

which simplifies to

$$\eta_{\nu\sigma} + \mathcal{M}^\mu_\nu \eta_{\mu\sigma} + \eta_{\nu\rho} \mathcal{M}^\rho_\sigma = \eta_{\nu\sigma} \quad (2.24)$$

and now finally we can write this as

$$\mathcal{M}_{\sigma\nu} + \mathcal{M}_{\nu\sigma} = 0 \quad (2.25)$$

in other words, $\mathcal{M}$ (with both indices down) is antisymmetric. This turns out to be the only constraint on the $\mathcal{M}$ ’s. There are 6 linearly independent 4×4 antisymmetric matrices. Let’s call a basis for these matrices $M^{(a)}$ where (a) runs from 1 to 6. Now consider the following objects

$$(M^{\rho\sigma})^\mu_\nu = \eta^{\sigma\mu} \delta^\rho_\nu - \eta^{\rho\mu} \delta^\sigma_\nu \quad (2.26)$$

⁴Strictly speaking, one should also impose $\Lambda^0_0 \geq 1$, in order to restrict to the so-called proper autochronous Lorentz group, but this will be automatically satisfied in our construction.

Now the notation is actually doing a lot of work here, so let me unpack it very carefully.

Here μ and ν are spacetime (or matrix) indices; thus for each value of ρ and σ we have a 4×4 matrix.

$$M^{\rho\sigma} = \begin{pmatrix} a & b & c & d \\ \dots & & & \end{pmatrix} \quad (2.27)$$

However the indices ρ, σ each run from 0 to 3, and from the definition above you can see explicitly that $M^{\rho\sigma} = -M^{\sigma\rho}$. Thus the set of antisymmetric combinations of ρ and σ label the 6 possible matrices that we can have.

Now to make this seem a little less crazy let me write some examples out carefully. First let's do some M 's who have a single 0 index. Expanding out the definition (2.26), we have

$$(M^{01})^\mu{}_\nu = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (2.28)$$

Note that notation M^{01} makes it clear which two dimensions are being mixed up. It won't surprise you that we have

$$(M^{02})^\mu{}_\nu = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (2.29)$$

and so on. You can do the same for the other components (there are some interesting signs here and there).

We are now in a position to write down the most general Lorentz transformation, which is the exponential of an arbitrary sum over all of these possible matrices. This takes the form

$$\mathcal{M} = \omega_{\rho\sigma} M^{\rho\sigma} \quad \Lambda(\omega_{\rho\sigma}) = \exp(\omega_{\rho\sigma} M^{\rho\sigma}) \quad (2.30)$$

where $\omega_{\rho\sigma}$ is a antisymmetric tensor $\omega_{\rho\sigma} = -\omega_{\sigma\rho}$ that allows us to choose an arbitrary linear combination of the M 's. (Note that we have suppressed the μ and ν indices in the above in order to avoid clutter, but keep in mind that $\mathcal{M}$ and Λ are matrices !)

Now the Lorentz transformations form a *group*, called the Lorentz group. As you recall from GMP, the commutators of the generators play an important role. You can explicitly check from the definition (2.26) that the commutators of the M 's are:

$$[M^{\mu\nu}, M^{\rho\sigma}] = -\eta^{\nu\rho} M^{\mu\sigma} + \eta^{\mu\rho} M^{\nu\sigma} + \eta^{\nu\sigma} M^{\mu\rho} - \eta^{\mu\sigma} M^{\nu\rho} \quad (2.31)$$

This defines the *Lorentz algebra*.

Now let's think physically about all of this. The transformations M^{0i} correspond to boosts in the x^i direction: for example if you work out

$$\Lambda(\omega_{01}) = \exp(\omega_{01} M^{01}) = \begin{pmatrix} \cosh(\omega_{01}) & \sinh(\omega_{01}) & 0 & 0 \\ \sinh(\omega_{01}) & \cosh(\omega_{01}) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.32)$$

This is in fact exactly the same as (2.18) if we map the boost velocity v to the ω_{01} by $\sinh \omega = -\frac{v}{\sqrt{1-v^2}}$. ω_{01} is often called the *rapidity*.

Now let's consider those M 's with both spatial indices. We find:

$$(M^{12})^\mu{}_\nu = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (2.33)$$

Exponentiating this we get

$$\Lambda(\omega_{12}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\omega_{12}) & -\sin(\omega_{12}) & 0 \\ 0 & \sin(\omega_{12}) & \cos(\omega_{12}) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.34)$$

which is clearly a rotation that mixes together the x^1 and the x^2 directions and so is a rotation about the x^3 axis. We see that the group of spatial rotations $SO(3)$ is a subgroup of $SO(1,3)$.

Details on how to do these matrix exponentials will be given in the homework.

Finally let's discuss those transformations which do *not* have $\det\Lambda = 1$; from earlier, this means that they will have $\det\Lambda = -1$, and some examples are **time reversal**, which acts as $T : (x^0, x^1, x^2, x^3) \rightarrow (-x^0, x^1, x^2, x^3)$ – so in matrix notation it is

$$\Lambda_T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.35)$$

Another important one is **parity** P , which changes the direction of all of the spatial directions: $P : (x^0, x^1, x^2, x^3) \rightarrow (x^0, -x^1, -x^2, -x^3)$. We then have

$$\Lambda_P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (2.36)$$

These clearly both have determinant -1 . One can get the whole Lorentz group $O(1,3)$ by composing these transformations with the infinitesimal ones studied above. **Exercise:** what is the form of parity if we had only two spatial (and one time) dimension? Why is it different?

2.3 Translations and the Poincaré group

We have just understood Lorentz transformations, and I argued that these are the relevant generalization of the idea of “rotation” to a relativistic spacetime. We should also now think about space-time translations, i.e. transformations of the form:

$$x^\mu \rightarrow x'^\mu = x^\mu - a^\mu \quad (2.37)$$

where a^μ is a constant 4-vector. These are symmetries of (sufficiently small) regions of empty space in our universe. Regarding a^μ : it's 0-th component corresponds to a translation of the origin of *time* (saying that the laws of physics are the same now and in the future) and its spatial components correspond to a translation of the origin of *space* (saying that the laws of physics are the same in this classroom and in the next).

The combination of these together with the Lorentz transformations results in the following 10-parameter symmetry group, called the **Poincaré group**:

$$x^\mu \rightarrow x'^\mu = \Lambda^\mu{}_\nu x^\nu - a^\mu \quad (2.38)$$

It has ten generators, the six $M^{\rho\sigma}$ and four translations P_μ . In the next section I will write down more explicitly how P_μ acts on fields defined on spacetime.

Finally, it's worth noting that the translations don't commute with the Lorentz transformations. The full structure of commutators is:

$$\begin{aligned} [M^{\mu\nu}, M^{\rho\sigma}] &= -\eta^{\nu\rho} M^{\mu\sigma} + \eta^{\mu\rho} M^{\nu\sigma} + \eta^{\nu\sigma} M^{\mu\rho} - \eta^{\mu\sigma} M^{\nu\rho} \\ [P^\mu, M^{\nu\rho}] &= -\eta^{\mu\nu} P^\rho + \eta^{\mu\rho} P^\nu \\ [P^\mu, P^\nu] &= 0 \end{aligned} \tag{2.39}$$

This is called the **Poincaré algebra**. I have not quite proved the second of these relations for you yet. We will see how to do so shortly.

2.4 Realizing transformations on fields

We have seen how coordinate transformations act on the coordinates x , i.e. in an index-free notation as:

$$x' = \Lambda x - a \tag{2.40}$$

Now let us come to the basic ingredient in this class: consider a **scalar field** $\phi(x)$, which is a map $\phi : \mathbb{R}^{1,3} \rightarrow \mathbb{R}$. $\phi(x)$ will also transform under coordinate transformations, so we can say that $\phi(x)$ transforms in a **representation of the Poincaré group**. We will now understand what this means.

2.4.1 Scalar fields

The basic rule is the following: when you go from an unprimed to a primed coordinate system, everything changes, including the scalar field $\phi(x)$. Let us call the new scalar field $\phi'(x')$: the primed field at the primed point is determined by the unprimed field at the unprimed point, i.e.

$$\boxed{\phi'(x') = \phi(x)} \tag{2.41}$$

Let us work out explicitly how this works for translations, i.e. consider the transformation

$$x'^\mu = x^\mu - a^\mu \tag{2.42}$$

So we have that

$$\phi'(x') = \phi'(x - a) = \phi(x) \tag{2.43}$$

where the last equality came from (2.41). We now have

$$\phi'(x) = \phi(x + a) \tag{2.44}$$

Let's start with an infinitesimal translation; expanding the right hand side in powers of a we have

$$\phi'(x) = \phi(x) + a^\mu \frac{\partial}{\partial x^\mu} \phi(x) + \mathcal{O}(a^2) \tag{2.45}$$

Thus we see that under an infinitesimal translation, the transformation of the scalar field is

$$\phi'(x) = \phi(x) + \delta\phi(x) \quad \boxed{\delta\phi(x) = a^\mu \partial_\mu \phi(x)} \tag{2.46}$$

Now let us consider finite transformations. For any transformation continuously connected to the identity, we can obtain the finite transformation by exponentiating some algebra element, i.e. there exists a P_μ such that

$$\phi'(x) = \phi(x + a) = \exp(a^\mu P_\mu) \phi(x) \tag{2.47}$$

In this case, by expanding both sides in powers of a we see that

$$P^\mu = \frac{\partial}{\partial x^\mu} \quad (2.48)$$

i.e. the generator of translations on the scalar field is by *derivatives*. This should seem philosophically reasonable.

Let us now come to the slightly more complicated problem of Lorentz transformations. This is similar except that the transformation of x is given by (2.30), i.e.

$$x'^\mu = \Lambda^\mu_\nu x^\nu = \exp(\omega_{\rho\sigma} M^{\rho\sigma})^\mu_\nu x^\nu \quad (2.49)$$

Let us now expand this out for infinitesimal ω ; we have then:

$$x'^\mu = x^\mu + \delta x^\mu = x^\mu + \omega_{\rho\sigma} (M^{\rho\sigma})^\mu_\nu x^\nu + \mathcal{O}(\omega^2) \quad (2.50)$$

Now we play the same game as before: we have

$$\phi'(x') = \phi(x) \quad (2.51)$$

Formally speaking, we would like to find a set of generators $L_{\rho\sigma}$ so that the following expression is true:

$$\phi'(x) = \exp(\omega_{\rho\sigma} L^{\rho\sigma}) \phi(x) = \phi(x) + \omega_{\rho\sigma} L^{\rho\sigma} \phi(x) + \mathcal{O}(\omega^2) \quad (2.52)$$

It's perhaps a good time to understand the difference between $L_{\rho\sigma}$ and $M_{\rho\sigma}$: here $M_{\rho\sigma}$ is a set of 4×4 matrices which are linear operators which realize the Lorentz algebra on the *coordinates* x^μ . However $L_{\rho\sigma}$ is going to be a set of *differential* operators, which are linear operators who realize the Lorentz algebra on the *field* $\phi(x)$. In the language of GMP, the field $\phi(x)$ and the coordinate x^μ form different representations of the Lorentz group.

Our task now is to determine the form of $L_{\rho\sigma}$.

To do this, we can proceed algorithmically just as before. We have

$$\phi'(x + \delta x) = \phi(x) \quad \rightarrow \quad \phi'(x) = \phi(x - \delta x) \quad (2.53)$$

Now we work for small δx to find

$$\phi'(x) = \phi(x) - \delta x^\mu \partial_\mu \phi(x) = \phi(x) - (\omega_{\rho\sigma} (M^{\rho\sigma})^\mu_\nu x^\nu) \partial_\mu \phi(x) + \mathcal{O}(\omega^2) \quad (2.54)$$

From here we can basically read off the form of the $L_{\rho\sigma}$. Plugging in the form of $M^{\rho\sigma} = (M^{\rho\sigma})^\mu_\nu = \eta^{\sigma\mu} \delta^\rho_\nu - \eta^{\rho\mu} \delta^\sigma_\nu$ from (2.26) and comparing with (2.52) we find

$$L^{\rho\sigma} = x^\sigma \frac{\partial}{\partial x_\rho} - x^\rho \frac{\partial}{\partial x_\sigma} \quad (2.55)$$

This is the desired expression. Note what it says: if you perform a Lorentz transformation, the field changes – the amount that it changes depends on how far you are from the origin of the transformation, which makes complete sense.

Now the L 's and the P 's form a representation of the Poincaré group, and thus satisfy the following algebra:

$$[L^{\mu\nu}, L^{\rho\sigma}] = -\eta^{\nu\rho} L^{\mu\sigma} + \eta^{\mu\rho} L^{\nu\sigma} + \eta^{\nu\sigma} L^{\mu\rho} - \eta^{\mu\sigma} L^{\nu\rho} \quad (2.56)$$

$$[P^\mu, L^{\nu\rho}] = -\eta^{\mu\nu} P^\rho + \eta^{\mu\rho} P^\nu \quad (2.57)$$

$$[P^\mu, P^\nu] = 0, \quad (2.58)$$

which you are welcome to check explicitly.

2.4.2 Vector fields

There are other kinds of fields than scalar fields. Let us think for a second about a *vector* field, e.g. $A^\mu(x)$. This will be important when we discuss gauge theories later on in the course. This is different from a scalar, because it has an extra index which must also be transformed. In other words, the transformation law is

$$A'^\mu(x') = \Lambda^\mu_\nu A^\nu(x) \quad (2.59)$$

We can now go through the same analysis as before to determine the form of the Lorentz generator. I'll leave you to go through this as an exercise: it becomes

$$L^{\rho\sigma} = M^{\rho\sigma} + x^\sigma \frac{\partial}{\partial x^\rho} - x^\rho \frac{\partial}{\partial x^\sigma} \quad (2.60)$$

The extra M here acts on the space of $A_\mu(x)$ at a single point and rotates it just as though it was a vector; the other parts act on the coordinates x .

Finally there is another type of field which will be important, called a *spinor field*, which again you have studied in GMP. I will discuss the Lorentz transformation properties of these in due course, but for now I feel we have done enough group theory and I would like to get to some dynamics.

3 Lagrangian methods and classical field theory

We now move towards field theory. First, Lagrangian methods for classical mechanics.

3.1 Lagrangians for classical mechanics

In classical mechanics imagine we have a system with N coordinates, where $a \in \{1, \dots, N\}$. Then we can organize our understanding with a **Lagrangian** $L(q_1, \dots, q_N, \dot{q}_1, \dots, \dot{q}_N)$, where $\dot{q}_a = \frac{dq_a}{dt}$. A typical form of the Lagrangian is

$$L(q_1, \dots, q_N, \dot{q}_1, \dots, \dot{q}_N) = \sum_{a=1}^N \frac{1}{2} m \dot{q}_a^2 - V(q_1, \dots, q_N) \quad (3.1)$$

i.e. of the form “kinetic energy” minus “potential energy”. The point of the Lagrangian is that we use it to determine the *action*:

$$S[q_a] = \int_{t_i}^{t_f} dt L(q_a, \dot{q}_a) \quad (3.2)$$

The action is a map from the space of particle trajectories to $\mathbb{R}$, and the **principle of least action** tells us that the solution to the classical equation of motion is the one that extremizes the action, i.e. if we consider a variation of the path $q_a(t) \rightarrow q_a(t) + \delta q_a(t)$, then the variation of the action should be stationary. In other words we demand:

$$0 = \delta S[q_a] = \int_{t_i}^{t_f} dt \delta L(q_a, \dot{q}_a) \quad (3.3)$$

$$= \int_{t_i}^{t_f} dt \sum_{a=1}^N \left(\frac{\partial L}{\partial q_a} \delta q_a + \frac{\partial L}{\partial \dot{q}_a} \delta \dot{q}_a \right) \quad (3.4)$$

$$= \int_{t_i}^{t_f} dt \sum_{a=1}^N \left[\frac{\partial L}{\partial q_a} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_a} \right) \right] \delta q_a(t) \quad (3.5)$$

where in the second equality we have integrated by parts and neglected a boundary term (formally, we demand that $\delta q_a(t_i) = \delta q_a(t_f) = 0$).

Demanding that this holds for the space of all variations $\delta q_a(t)$ we find that the quantity multiplying $\delta q_a(t)$ must vanish, which gives us the **Euler-Lagrange equations of motion**, which is the following set of ODEs.

$$\frac{\partial L}{\partial q_a} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_a} \right) = 0 \quad (3.6)$$

This should be familiar. Note that we have obtained N equations of motion from a single scalar quantity L ; this is a computationally very useful thing and is one of the reasons why we formulate classical mechanics in terms of Lagrangians.

3.2 Lagrangian methods for classical field theory

We now want to write down a Lagrangian for a classical scalar field $\phi(x) = \phi(t, \vec{x})$. Let me begin by making an analogy with normal classical mechanics; there we have $q_a(t)$ – now consider having a run over the discrete N sites of a cubic lattice, and imagine making N bigger and bigger until the lattice sites approximate a continuum $\mathbb{R}^3$. Then we see that $q_a(t) \rightarrow \phi(t, \vec{x})$, and $\sum_{a=1}^N \rightarrow \int d^3x$. Basically the spatial coordinate $\vec{x}$ plays the role of the index a labelling the degrees of freedom.

We are thus led to consider a Lagrangian which takes the following form:

$$L = \int d^3x \mathcal{L}(\phi(x), \partial_\mu \phi(x)) \quad (3.7)$$

Here $\mathcal{L}$ is called the **Lagrangian density**. Let us insert this into the action to find:

$$S[\phi] = \int dt d^3x \mathcal{L}(\phi(x), \partial_\mu \phi(x)) \quad (3.8)$$

This is the general form of the field theories that we will consider. Let me note a few nice things about it:

1. Note that in the integral we have $d^4x = dx^0 dx^1 dx^2 dx^3$; in other words space and time look like they're treated on a similar footing, which is crucial for relativistic invariance. We will explicitly check relativistic invariance soon.
2. This has **locality** built in: fields only couple to each other at the same space-time point, i.e. given this form you will never see a term in the action that looks like $\phi(t, \vec{x})\phi(t, \vec{y})$. This appears to be a basic requirement for a field theory to be well-behaved.

Now let us derive the analogue of the Euler-Lagrange equations. We again demand that the action is stationary under a variation $\phi(x) \rightarrow \phi(x) + \delta\phi(x)$, so we have

$$0 = \delta_\phi S = \int d^4x \delta \mathcal{L} = \int d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi} \delta \phi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \partial_\mu \delta \phi \right] \quad (3.9)$$

$$= \int d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) \right] \delta \phi(x) \quad (3.10)$$

Now this gives us the classical Euler-Lagrange equations for a field theory, i.e.

$$\boxed{\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) = 0} \quad (3.11)$$

Note that everything here was a function of space-time, so the action formalism gives us a way to obtain a PDE from a single scalar density $\mathcal{L}(\phi, \partial_\mu \phi)$.

Some important notation: when a field configuration satisfies the Euler-Lagrange equations, we say that it is **on-shell**. Whenever we are looking at an on-shell field configuration the action is stationary under small variations about this field configuration.

**** (non-examinable)** What would happen to (3.11) if we allowed the Lagrangian density to depend on double derivatives of the fields, i.e. $\mathcal{L} = \mathcal{L}(\phi(x), \partial_\mu \phi(x), \partial_\mu \partial_\nu \phi(x))$?**

3.2.1 Action of a real scalar field

Let us now describe some conditions that we want for our action S (equivalently, our Lagrangian density $\mathcal{L}$):

1. The action S should be invariant under Poincaré transformations. Note we have

$$S = \int d^4x \mathcal{L}(\phi, \partial_\mu \phi) \quad (3.12)$$

Let's quickly check how the measure transforms under Poincaré transformations: under $x' = x - a$ we have $d^4x' = d^4x$. Under $x' = \Lambda x$ we have $d^4x' = |\det \Lambda| d^4x = d^4x$. So we see that the integral doesn't bring in any new features, and invariance of the action really only requires that the Lagrangian density $\mathcal{L}$ be invariant under Poincaré.

2. The action S should result in non-trivial dynamics, i.e. we should get an interesting PDE out of (3.11). In practice, this means that $\mathcal{L}$ must depend on $\partial_\mu \phi$.

The simplest lagrangian density that satisfies these two properties is the following **action of a real scalar field**⁵:

$$\boxed{\mathcal{L}(\phi, \partial_\mu \phi) = -\frac{1}{2} \partial_\mu \phi \partial^\mu \phi - V(\phi)} \quad (3.13)$$

Here $V(\phi)$ is a function of one variable – it is called the **potential of the scalar field**. We will spend a great deal of time thinking about implications of this action. Let's expand it out a little bit:

$$\mathcal{L} = -\frac{1}{2} \eta^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi) \quad (3.14)$$

$$= \frac{1}{2} (\partial_t \phi)^2 - \frac{1}{2} \vec{\nabla} \phi \cdot \vec{\nabla} \phi - V(\phi) \quad (3.15)$$

This is relativistically invariant because we have contracted all of the indices using the Minkowski metric tensor; this introduced an interesting minus sign between the derivatives. Note that by analogy with the case for a particle, we can consider the first term a kind of “kinetic energy” and the second two terms a sort of “potential energy”.

The simplest choice for $V(\phi)$ is

$$V(\phi) = \frac{1}{2} m^2 \phi^2 \quad (3.16)$$

What is the meaning of the parameter m ? At this point I have to tell you what this theory describes. With this choice of potential it turns out that the *quantum* theory described by this action describes a system of *non-interacting particles with mass m* . I have not justified this radical statement at all yet, so you should be very skeptical of this, but I'm going to show you this is true.

⁵** (non-examinable) Try guessing the form of the Lagrangian for a field A_μ (perhaps you have seen this already), and a field $T_{\mu_1 \dots \mu_n}^{\nu_1 \dots \nu_n}$ such that we have Poincaré invariance.**

Let us now work out the Euler-Lagrange equations for this system. I will do this somewhat carefully – we start with

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) \quad (3.17)$$

Now let's work out both of the terms. For the first we simply have

$$\frac{\partial \mathcal{L}}{\partial \phi} = -\frac{\partial V}{\partial \phi} = -m^2 \phi \quad (3.18)$$

For the second we need to work out

$$\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} = \frac{\partial}{\partial (\partial_\mu \phi)} \left(-\frac{1}{2} \eta^{\rho\sigma} \partial_\rho \phi \partial_\sigma \phi \right) \quad (3.19)$$

In the above I have taken the term from (3.14) and renamed some indices to avoid confusion. Now in constructing this derivative let us start by making sure that we understand the following expression:

$$\frac{\partial}{\partial (\partial_\sigma \phi)} (\partial_\rho \phi) = \delta_\rho^\sigma \quad (3.20)$$

The fact that the left-hand side vanishes if $\sigma \neq \rho$ and is 1 if $\sigma = \rho$ actually means that the right-hand side must be equal to the Kronecker delta, which means the index placement must be as shown. The fact that the ρ index is down should not be too surprising, but the fact that σ index is up might be. You can verify that the Lorentz transformation properties of this expression do in fact make sense.

Using this expression we find

$$\frac{\partial}{\partial (\partial_\mu \phi)} \left(-\frac{1}{2} \eta^{\rho\sigma} \partial_\rho \phi \partial_\sigma \phi \right) = -\frac{1}{2} \eta^{\rho\sigma} (\delta_\rho^\mu \partial_\sigma \phi + \partial_\rho \phi \delta_\sigma^\mu) = -\eta^{\mu\sigma} \partial_\sigma \phi = -\partial^\mu \phi \quad (3.21)$$

Putting both of these pieces in we find:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) = -m^2 \phi - \partial_\mu (-\partial^\mu \phi) \quad (3.22)$$

or

$$\boxed{\partial^\mu \partial_\mu \phi - m^2 \phi = 0} \quad (3.23)$$

This is the **Klein-Gordon equation**. It is a linear PDE that describes a fully relativistically invariant wave equation.

Let us expand out the structure $\partial^\mu \partial_\mu = \eta^{\mu\nu} \partial_\mu \partial_\nu = -\partial_t^2 + \vec{\nabla} \cdot \vec{\nabla}$ – this structure should be familiar to you from studying the wave equation in other courses. Note the speed of wave propagation – which is given by the ratio of the spatial and time derivatives – is 1. This is because of my choice of units; if use more human units it would have been the speed of light c .

**** (Non-examinable) Derive the Euler-Lagrange equations for the Lagrangians you wrote in footnote 5. ****

3.3 Noether's theorem

We will now discuss **Noether's theorem**, probably the deepest and most important result in mathematical physics. It is basically the following statement: **a continuous symmetry results in a conserved current**. Let me first explain what I mean by a continuous symmetry.

3.3.1 Example: a complex scalar field

I will start with the simplest example, also as an excuse to consider a slightly different field theory. Let us imagine not a real scalar field but a *complex* scalar field Φ . The action for the complex scalar field is

$$S[\Phi, \Phi^*] = \int d^4x \left(-(\partial_\mu \Phi^*)(\partial^\mu \Phi) - m^2 \Phi^* \Phi \right) \quad (3.24)$$

Some notational points: this field has both a real and imaginary part, and it is customary to encode this in the field Φ and its conjugate Φ^* . Clearly the action is invariant under the following $U(1)$ symmetry, a phase rotation of Φ :

$$\Phi \rightarrow e^{i\alpha} \Phi \quad \Phi^* \rightarrow e^{-i\alpha} \Phi^* \quad (3.25)$$

Let's just verify this very carefully:

$$S[\Phi e^{i\alpha}, \Phi^* e^{-i\alpha}] = \int d^4x \left(-(\partial_\mu \Phi^* e^{-i\alpha})(\partial^\mu \Phi e^{i\alpha}) - m^2 \Phi^* e^{-i\alpha} \Phi e^{i\alpha} \right) \quad (3.26)$$

$$= \int d^4x \left(-(\partial_\mu \Phi^*)(\partial^\mu \Phi) - m^2 \Phi^* \Phi \right) = S[\Phi, \Phi^*] \quad (3.27)$$

Here α is a constant. This is called a **global continuous symmetry**; it is **global** because it acts everywhere in space and time in the same way since α is a constant. There are also symmetries which are *not* global symmetries: you will learn about those in due course in the part of the course on gauge theory.

It is **continuous** because it is a symmetry for all values of α , $e^{i\alpha} \in U(1)$. Importantly, we can consider its infinitesimal version:

$$\Phi \rightarrow (1 + i\alpha)\Phi + \mathcal{O}(\alpha^2) \quad \Phi^* \rightarrow (1 - i\alpha)\Phi^* + \mathcal{O}(\alpha^2) \quad (3.28)$$

We can also have *discrete* symmetries, which are not continuous. In fact the action above has one:

$$\Phi \rightarrow \Phi^* \quad \Phi^* \rightarrow \Phi \quad (3.29)$$

This symmetry is often called **charge conjugation**, or C for short. (It will turn out that in the quantum theory this symmetry switches particles and antiparticles; again for now you just have to take my word for it). Note it does not have an infinitesimal version, and there isn't an immediate version of Noether's theorem for it.

3.3.2 Proof of Noether's theorem

We will now prove Noether's theorem. Consider a field theory described by a set of scalar fields ϕ_a , with an action $S[\phi_a]$. Now let us imagine that the theory is invariant under a continuous symmetry. This means that there exists an infinitesimal variation of the fields

$$\phi_a(x) \rightarrow \phi_a(x) + \epsilon \Delta \phi_a(x) \quad (3.30)$$

under which the action is invariant, i.e.

$$S[\phi_a + \epsilon \Delta \phi_a] - S[\phi_a] = 0 \quad \rightarrow \quad \delta_\epsilon S[\phi_a] = 0 \quad (3.31)$$

Here ϵ is an infinitesimal symmetry parameter; it is analogous to the α that I used above in the $U(1)$ example.

Now, let us perform a trick: let's consider a variation where we allow $\epsilon(x)$ to be an arbitrary function of space and time, i.e.

$$\phi_a(x) \rightarrow \phi_a(x) + \epsilon(x) \Delta \phi_a(x) \quad (3.32)$$

The action now is *not* invariant – why would it be? However it *must* vanish when $\epsilon(x)$ is a constant (in which case it reduces to the case above), so the variation takes the form

$$\delta_\epsilon S[\phi_a] = \int d^4x j^\mu \partial_\mu \epsilon(x) \quad (3.33)$$

(Here I have secretly also assumed that the field theory is local). Here j^μ is some function of the fields which depends on the theory in question. Now let's integrate by parts and neglect boundary terms to find

$$\delta_\epsilon S[\phi_a] = - \int d^4x (\partial_\mu j^\mu) \epsilon(x) \quad (3.34)$$

This relation holds for *any* choice of field configuration ϕ^a . But now let us consider field configurations ϕ_a that satisfy the Euler-Lagrange equations of motion, i.e. they are *on-shell*. In that case the variation of the action is zero for *any* variation of the fields, including (3.32). So *on-shell* we find

$$0 = \delta_\epsilon S[\phi_a] = - \int d^4x (\partial_\mu j^\mu) \epsilon(x) \quad (3.35)$$

This holds for any choice of $\epsilon(x)$; thus whenever the fields satisfy the equations of motion we have

$$\boxed{\partial_\mu j^\mu = 0} \quad (3.36)$$

i.e. there exists a divergenceless vector field j^μ , which is usually called a “conserved current”. Thus we have proven Noether's theorem. We note that this approach also gives us an algorithm to construct the conserved current.

Now let us discuss for a second what it *means* to have a divergenceless current. Basically it means that there exists a conserved *charge*. Let's explain this carefully: the equation (3.36) reads:

$$\partial_t j^0(t, x^i) + \partial_i j^i(t, x^i) = 0 \quad (3.37)$$

Now let us integrate this equation over all *space*:

$$\partial_t \int d^3x j^0 + \int d^3x \partial_i j^i(t, x^i) = 0 \quad (3.38)$$

The last term is a total derivative, which means we can neglect it all fields die off quickly enough at infinity. In that case we find the following expression

$$\frac{d}{dt} Q(t) = 0 \quad Q \equiv \int d^3x j^0(t, x^i) \quad (3.39)$$

In other words, there is a **conserved charge** Q , which is independent of time.

Let us now work this out explicitly in the case of the complex scalar field. We consider the infinitesimal variation from (3.28)

$$\Delta\Phi(x) = i\alpha\Phi(x) \quad \Delta\Phi^* = -i\alpha\Phi(x)^* \quad (3.40)$$

Now as suggested above we make the symmetry parameter α depend on space and consider the infinitesimal variation:

$$\Delta\Phi(x) = i\alpha(x)\Phi(x) \quad \Delta\Phi^* = -i\alpha(x)\Phi(x)^* \quad (3.41)$$

And now let's consider the infinitesimal variation. We treat Φ and Φ^* as independent variables to find from (3.24):

$$\delta_\alpha S[\Phi, \Phi^*] = \int d^4x (-(\partial_\mu \Delta\Phi^*)(\partial^\mu \Phi) - (\partial_\mu \Phi^*)(\partial^\mu \Delta\Phi) - m^2 \Delta\Phi^* \Phi - m^2 \Phi^* \Delta\Phi) \quad (3.42)$$

$$= \int d^4x (-(\partial_\mu (-i\alpha\Phi^*))(\partial^\mu \Phi) - (\partial_\mu \Phi^*)(\partial^\mu (i\alpha\Phi)) - m^2 (-i\alpha\Phi^*)\Phi - m^2 \Phi^* (i\alpha\Phi)) \quad (3.43)$$

$$= - \int d^4x \partial_\mu \alpha [i(\partial^\mu \Phi^*)\Phi - i\Phi^* \partial^\mu \Phi] \quad (3.44)$$

Indeed, as promised, this looks exactly the same as (3.33), which lets us immediately read off the explicit expression of the conserved current:

$$j^\mu = -i [(\partial^\mu \Phi^*)\Phi - \Phi^*(\partial^\mu \Phi)] \quad (3.45)$$

We can also now construct the conserved charge Q :

$$Q = \int d^3x j^0 = i \int d^3x [(\partial_t \Phi^*)\Phi - \Phi^*(\partial_t \Phi)] \quad (3.46)$$

It can be instructive to show that the current is conserved using the ordinary form of the Euler-Lagrange equations for the complex scalar field.

3.3.3 Alternative way to find the current

Recall that the action is written in terms of the Lagrangian density $\mathcal{L}(\phi_a, \partial_\mu \phi_a)$. In many situations we find that it is not the integrated action which is invariant under the symmetry, but also the Lagrangian itself. In that case there's a shortcut to finding the current.

Consider the symmetry variation of the fields (note there is an implicit sum over a here):

$$\delta S = \int d^4x \delta \mathcal{L} = \int d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi^a} \epsilon \Delta \phi_a + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \partial_\mu (\epsilon \Delta \phi^a) \right) = 0 \quad (3.47)$$

If the Lagrangian itself is invariant under the symmetry, then for constant ϵ we have:

$$\delta \mathcal{L} = \epsilon \left(\frac{\partial \mathcal{L}}{\partial \phi^a} \Delta \phi_a + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \partial_\mu (\Delta \phi^a) \right) = 0 \quad (3.48)$$

Now recall that the Euler-Lagrange equations of motion tell us that on-shell we have:

$$\frac{\partial \mathcal{L}}{\partial \phi^a} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \right) = 0 \quad (3.49)$$

Let's use this formula to eliminate $\frac{\partial \mathcal{L}}{\partial \phi^a}$. We see from (3.48) that

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \right) \Delta \phi_a + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \partial_\mu (\Delta \phi^a) = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \Delta \phi_a \right) = 0 \quad (3.50)$$

Thus we have identified a divergenceless current

$$j^\mu = \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \Delta \phi_a \right) \quad (3.51)$$

This is often a quick way to find the symmetry current. It's instructive to check that for the complex scalar field this gives the same answer as (3.45). I stress that the formula above only works if the Lagrangian itself is invariant, which is often the case but is strictly a stronger condition than the integrated action being invariant. There is a way to improve this equation to suit the general case (see e.g. p18 of Peskin & Schroeder) but under those circumstances I usually just use the algorithm of the previous part, which always works.

3.4 Hamiltonian formalism in classical mechanics

Finally, let us briefly remind ourselves how the Hamiltonian works in classical mechanics. Given a classical mechanical degree of freedom $q_a(t)$, we can construct its *canonical momentum*

$$p_a \equiv \frac{\partial L}{\partial \dot{q}_a} \quad (3.52)$$

We then construct the Hamiltonian by the following algorithm:

$$H(q_a, p_a) \equiv \sum_a p_a \dot{q}_a - L(q_a, \dot{q}_a) \quad (3.53)$$

Now let's do the same thing for field theory. Our basic degree of freedom is the field $\phi(t, \vec{x})$. We now construct its canonical momentum $\pi(t, \vec{x})$ from the Lagrangian *density* via:

$$\pi(t, \vec{x}) \equiv \frac{\partial \mathcal{L}}{\partial(\dot{\phi}(t, \vec{x}))} \quad (3.54)$$

Let's work this out explicitly with the example of the real scalar field

$$\mathcal{L} = -\frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - m^2\phi^2 = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}\vec{\nabla}\phi \cdot \vec{\nabla}\phi - \frac{1}{2}m^2\phi^2 \quad (3.55)$$

So we find

$$\pi(t, \vec{x}) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \dot{\phi} \quad (3.56)$$

And we can now find the Hamiltonian density $\mathcal{H}$ by the same sort of formula, where as usual we eliminate $\dot{\phi}$ in terms of π .

$$\mathcal{H} = \pi(t, \vec{x})\dot{\phi}(t, \vec{x}) - \mathcal{L} \quad (3.57)$$

$$= \pi^2 - \left(\frac{1}{2}\pi^2 - \frac{1}{2}\vec{\nabla}\phi \cdot \vec{\nabla}\phi - \frac{1}{2}m^2\phi^2 \right) \quad (3.58)$$

$$= \frac{1}{2}\pi^2 + \frac{1}{2}\vec{\nabla}\phi \cdot \vec{\nabla}\phi + \frac{1}{2}m^2\phi^2 \quad (3.59)$$

Note that the Hamiltonian density is defined pointwise in space. The full Hamiltonian H is the integral of the Hamiltonian density over all of space

$$H = \int d^3x \mathcal{H} = \int d^3x \left(\frac{1}{2}\pi^2 + \frac{1}{2}\vec{\nabla}\phi \cdot \vec{\nabla}\phi + \frac{1}{2}m^2\phi^2 \right) \quad (3.60)$$

As usual in classical mechanics, this can be interpreted as the *energy* of the system. It is an integral of a positive-definite quantity, so the energy is positive, which makes sense. One can also show that this energy is conserved (it is in fact the Noether charge associated with time translational symmetry ⁶).

Note that even though the underlying system was Lorentz invariant, we brutally destroyed this manifest Lorentz invariance by doing various non-covariant things using a particular choice of time to define $\cdot \equiv \frac{\partial}{\partial t}$. This is always an issue with Hamiltonian methods. The final answers will actually always turn out to be Lorentz invariant though intermediate steps generally will not be.

^{6**} See exercise class 1: Derive (3.60) using Noether's theorem for a time translation symmetry. Note that you will either need to use the more general form for the conserved current, given on p18 of Peskin & Schroeder, or use the trick of making the global parameter local (this is the method we used in the exercise class). **

4 Quantum field theory in canonical formalism

We are now ready to start thinking about quantum theories. We will begin in what is probably the more standard and familiar fashion, using a Hamiltonian; this is called the **canonical formalism** (to be distinguished from the **path integral formalism**).

Again we remind ourselves how this works in quantum mechanics. We start with a classical system with coordinates q_a and conjugate momenta p_a . There is also a Hamiltonian $H(q_a, p_a)$, constructed (perhaps) from the Lagrangian $L(q_a, \dot{q}_a)$ as described previously. These degrees of freedom have **Poisson brackets** which are

$$\{q_a, q_b\} = 0 \quad \{p_a, p_b\} = 0 \quad \{q_a, p_b\} = \delta_{ab} \quad (4.1)$$

Now there is an algorithm (usually called “quantization”) that takes this structure and makes it into a quantum theory. It works by promoting the classical variables into Hermitian operators $\hat{q}_a, \hat{p}_a$ which now act on some appropriate Hilbert space. The structure of the Poisson brackets is now realized in commutators. There are two ways to think about quantum mechanics – the so-called *Schrodinger picture* where the states $|\psi\rangle$ evolve in time and the operators are time-independent, in which case the algebra of operators is

$$[\hat{q}_a, \hat{q}_b] = 0 \quad [\hat{p}_a, \hat{p}_b] = 0 \quad [\hat{q}_a, \hat{p}_b] = i\hbar\delta_{ab} \quad (4.2)$$

One can also think about the Heisenberg picture: in this case the operators evolve in time: we have $\hat{q}_a(t), \hat{p}_b(t)$, and it should be understood that the commutator algebra above holds only at **equal time**, i.e:

$$[\hat{q}_a(t), \hat{q}_b(t)] = 0 \quad [\hat{p}_a(t), \hat{p}_b(t)] = 0 \quad [\hat{q}_a(t), \hat{p}_b(t)] = i\hbar\delta_{ab} \quad (4.3)$$

(If we considered commutators at unequal times we would find different and generically more complicated commutation relations which depend explicitly on the time delay).

From now on I will set $\hbar \rightarrow 1$ again.

We now want to perform the same process for field theory, i.e. we want to *quantize* the classical theory of the free scalar field that we introduced previously. Recall that for the real scalar field we have the classical variables $\phi(t, x)$ and their conjugate momenta $\pi(t, x)$. We now want to construct quantum operators $\hat{\phi}(t, \vec{x})$ and $\hat{\pi}(t, \vec{x})$ which obey the following **canonical commutation relations**⁷:

$$[\hat{\phi}(t, \vec{x}), \hat{\phi}(t, \vec{y})] = 0 \quad [\hat{\pi}(t, \vec{x}), \hat{\pi}(t, \vec{y})] = 0 \quad [\hat{\phi}(t, \vec{x}), \hat{\pi}(t, \vec{y})] = i\delta^{(3)}(\vec{x} - \vec{y}) \quad (4.4)$$

Here the last commutator is the interesting one; note that the delta function in space is the continuum field theoretical analogue of the δ_{ab} in (4.3). Our task now is to build the quantum theory that satisfies this: understand the quantum version of the Hamiltonian, understand the state space, figure out its eigenvalues, etc.

This is in general a difficult task; we will do it for the free theory (i.e. that where the potential $V(\phi)$ is quadratic), where the problem turns out to be tractable. To understand why this is, let’s remind ourselves of the Euler-Lagrange equation for the free real scalar field, which we derived a few lectures ago:

$$(\partial_\mu \partial^\mu - m^2) \phi(t, \vec{x}) = \left(-\partial_t^2 + \vec{\nabla}^2 - m^2\right) \phi(t, \vec{x}) = 0 \quad (4.5)$$

This is a linear partial differential equation in a system with translational invariance. Whenever you have translational invariance, it makes sense to go to momentum space, i.e. to write down the following expansion:

$$\phi(t, \vec{x}) = \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}} \tilde{\phi}(t, \vec{k}) \quad (4.6)$$

⁷** (Non-examinable): what is the field theory version of (4.1)? What is the general definition of the Poisson bracket $\{f, g\}$ in field theory (think about the coordinate dependence)? **

Note in passing that this can be inverted as follows:

$$\tilde{\phi}(t, \vec{k}) = \int d^3x e^{-i\vec{k}\cdot\vec{x}} \phi(t, x) \quad (4.7)$$

by using the usual Fourier transformation formulas. Now consider acting on this with the Euler-Lagrange operator; each ∇^2 brings down a $-\vec{k}^2$ and we find:

$$\left(\partial_t^2 + \vec{k}^2 + m^2\right) \tilde{\phi}(t, \vec{k}) = 0 \quad (4.8)$$

(Strictly speaking we find the integral of this equation multiplied by $e^{i\vec{k}\cdot\vec{x}}$, but that operator is invertible). This is exactly the equation for a simple harmonic oscillator with frequency $\omega_{\vec{k}} = \sqrt{\vec{k}^2 + m^2}$:

$$\left(\partial_t^2 + \omega_{\vec{k}}^2\right) \tilde{\phi}(t, \vec{k}) = 0 \quad (4.9)$$

Thus, at least classically, **each Fourier mode of the free scalar field is a harmonic oscillator**. We do know how to solve the quantum theory of the harmonic oscillator: so let's review that and then return to the field-theoretical problem.

I note that if the potential had not been quadratic in ϕ this wouldn't have worked and we would immediately have been stuck. Indeed solving such "interacting" quantum field theories is a huge and hard problem in physics.

4.1 Quantizing the simple harmonic oscillator

First, we study the *classical* harmonic oscillator. The simple harmonic oscillator with frequency ω in one dimension has a single coordinate $q(t)$ and the following Lagrangian:

$$L(q, \dot{q}) = \frac{1}{2} \dot{q}^2 - \frac{1}{2} \omega^2 q^2 \quad (4.10)$$

We can easily verify that the Euler-Lagrange equations are

$$\ddot{q}(t) = -\omega^2 q(t) \quad (4.11)$$

The general solution is

$$q(t) = \frac{1}{\sqrt{2\omega}} (a e^{-i\omega t} + b e^{+i\omega t}) \quad (4.12)$$

where a and b are integration constants and the factor of $\sqrt{2\omega}$ is for later convenience.

Now: let us recall that $q(t)$ is a real number, so $q(t) = q^*(t)$. Thus we need $b = a^*$:

$$q(t) = \frac{1}{\sqrt{2\omega}} (a e^{-i\omega t} + a^* e^{+i\omega t}) \quad (4.13)$$

Next, let's go to the Hamiltonian formalism. From our familiar expressions for the Hamiltonian we have

$$p = \frac{\partial L}{\partial \dot{q}} = \dot{q} \quad H = p\dot{q} - L = \frac{1}{2} (p^2 + \omega^2 q^2) \quad (4.14)$$

Now, let's quantize! We promote everything to operators acting on a Hilbert space, so $p \rightarrow \hat{p}$ and $q \rightarrow \hat{q}$. We will work in the Heisenberg picture, in which the operators depend on time. The Hamiltonian becomes

$$\hat{H} = \frac{1}{2} (\hat{p}^2 + \omega^2 \hat{q}^2) \quad (4.15)$$

Now what is the operator analogue of the expansion (4.13)? We write

$$\hat{q}(t) = \frac{1}{\sqrt{2\omega}} (\hat{a}e^{-i\omega t} + \hat{a}^\dagger e^{+i\omega t}) \quad (4.16)$$

where now $\hat{a}$ and $\hat{a}^\dagger$ are quantum operators that we take to be time-independent. In the context of the harmonic oscillator they are called **ladder operators**. In that case we can compute $\hat{p}$ as:

$$\hat{p}(t) = \dot{\hat{q}}(t) = -i\sqrt{\frac{\omega}{2}} (\hat{a}e^{-i\omega t} - \hat{a}^\dagger e^{+i\omega t}) \quad (4.17)$$

Now the fundamental canonical commutation relations that we need are the following:

$$[\hat{q}(t), \hat{p}(t)] = i \quad [\hat{q}(t), \hat{q}(t)] = [\hat{p}(t), \hat{p}(t)] = 0 \quad (4.18)$$

What does this imply for the commutation relations of a and $a^\dagger$? We can solve the two equations above for $\hat{a}$ and its conjugate:

$$\hat{a} = \left(\sqrt{\frac{\omega}{2}} \hat{q}(t) + \frac{i}{\sqrt{2\omega}} \hat{p}(t) \right) e^{+i\omega t} \quad \hat{a}^\dagger = \left(\sqrt{\frac{\omega}{2}} \hat{q}(t) - \frac{i}{\sqrt{2\omega}} \hat{p}(t) \right) e^{-i\omega t} \quad (4.19)$$

Working out the commutators by using the equal-time commutation relations for $\hat{p}$ and $\hat{q}$ we find them to be:

$$[\hat{a}, \hat{a}^\dagger] = 1 \quad [\hat{a}, \hat{a}] = [\hat{a}^\dagger, \hat{a}^\dagger] = 0 \quad (4.20)$$

Next, let's construct the Hamiltonian. Expressing $\hat{p}$ and $\hat{q}$ in terms of the a 's, we get:

$$\hat{H} = \frac{\omega}{2} (\hat{a}^\dagger \hat{a} + \hat{a} \hat{a}^\dagger) = \omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) \quad (4.21)$$

where in the last equality we used the known commutator to swap the order around of $\hat{a}$ and $\hat{a}^\dagger$.

Now we come to the key point of the ladder operators: they diagonalize the Hamiltonian! To see this, let's compute the commutator:

$$[\hat{H}, \hat{a}^\dagger] = \left[\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right), \hat{a}^\dagger \right] = \omega \hat{a}^\dagger [\hat{a}, \hat{a}^\dagger] = \omega \hat{a}^\dagger \quad (4.22)$$

Similarly, we can compute

$$[\hat{H}, \hat{a}] = -\omega \hat{a} \quad (4.23)$$

This means that these two operators raise and lower the energy! To understand this, consider an energy eigenstate $|E\rangle$ that has energy E , i.e. it satisfies the following equation:

$$\hat{H}|E\rangle = E|E\rangle \quad (4.24)$$

(I stress that in the above expression E is a number, not an operator; this is an eigenvalue equation). Now let's consider the state $\hat{a}^\dagger|E\rangle$. What happens if we act on it with the Hamiltonian?

$$\hat{H}\hat{a}^\dagger|E\rangle = \left(\hat{a}^\dagger \hat{H} + [\hat{H}, \hat{a}^\dagger] \right) |E\rangle = \left(\hat{a}^\dagger \hat{H} + \omega \hat{a}^\dagger \right) |E\rangle = (E + \omega) \hat{a}^\dagger |E\rangle \quad (4.25)$$

In other words the state $\hat{a}^\dagger|E\rangle$ is also an eigenstate of the Hamiltonian, with energy $E + \omega$: $\hat{a}^\dagger$ has raised the energy by the fixed amount ω . Similarly, we can consider the state $\hat{a}|E\rangle$, i.e.

$$\hat{H}\hat{a}|E\rangle = \left(\hat{a} \hat{H} + [\hat{H}, \hat{a}] \right) |E\rangle = (E - \omega) \hat{a} |E\rangle \quad (4.26)$$

and we see that $\hat{a}$ has lowered the energy by an amount ω .

These two operators now allow us to construct the full set of energy eigenstates. We define a state $|0\rangle$ called the **ground state** by the condition that $\hat{a}|0\rangle = 0$. Note that the energy of the ground state is:

$$\hat{H}|0\rangle = \omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) |0\rangle = \frac{\omega}{2} |0\rangle \quad (4.27)$$

Now we can construct a tower of excited states by acting repeatedly with the raising operators:

$$|n\rangle = (\hat{a}^\dagger)^n |0\rangle \quad \hat{H}|n\rangle = \omega \left(n + \frac{1}{2} \right) |n\rangle \quad (4.28)$$

This solves the problem of the single simple harmonic oscillator. Now on to quantum field theory, which is really many (artfully arranged) simple harmonic oscillators.

4.2 Quantizing the free complex scalar field

Now we do something very similar for the free complex scalar field, which has the following classical action:

$$S[\phi, \phi^*] = \int d^4x \left(-(\partial_\mu \phi^*)(\partial^\mu \phi) - m^2 \phi^* \phi \right) \quad (4.29)$$

First let's solve the classical system. Recall from earlier that the equation of motion in position and in momentum space is:

$$\left(-\partial_t^2 + \vec{\nabla}^2 - m^2 \right) \phi(t, \vec{x}) = 0, \quad \left(\partial_t^2 + \omega_{\vec{k}}^2 \right) \tilde{\phi}(t, \vec{k}) = 0 \quad \omega_{\vec{k}} = \sqrt{\vec{k}^2 + m^2} \quad (4.30)$$

As we described earlier, the momentum space equation of motion is just that of a harmonic oscillator, so following what we did above, we can now write down the most general solution to the momentum space equation of motion:

$$\tilde{\phi}(\vec{k}, t) = \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(a_{\vec{k}} e^{-i\omega_{\vec{k}} t} + b_{-\vec{k}}^* e^{i\omega_{\vec{k}} t} \right) \quad (4.31)$$

This is directly analogous to (4.12); here $a_{\vec{k}}, b_{\vec{k}}$ are momentum-dependent Fourier coefficients. Now we go to position space by:

$$\phi(\vec{x}, t) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(a_{\vec{k}} e^{-i\omega_{\vec{k}} t} + b_{-\vec{k}}^* e^{i\omega_{\vec{k}} t} \right) e^{+i\vec{k} \cdot \vec{x}} \quad (4.32)$$

Now it is conventional to make the following change of variables in the second term: $\vec{k} \rightarrow -\vec{k}$. In that case this flips all the k 's on that term and we find the following slightly nicer expression:

$$\phi(\vec{x}, t) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(a_{\vec{k}} e^{-i\omega_{\vec{k}} t + i\vec{k} \cdot \vec{x}} + b_{\vec{k}}^* e^{i\omega_{\vec{k}} t - i\vec{k} \cdot \vec{x}} \right) \quad (4.33)$$

This is nicer as both expressions in the exponent are now of a Lorentz-invariant form.

Note that for the harmonic oscillator above q was real, and so there was a relation between a and b . Similarly, if we were doing the real scalar field here we would find such a relation ($a = b$ in this case), but for the complex scalar field they are independent.

Now we are getting ready to quantize! First, let's work out the Hamiltonian. If we expand out the Lagrangian we find

$$\mathcal{L} = \dot{\phi} \dot{\phi}^* - (\vec{\nabla} \phi \cdot \vec{\nabla} \phi^*) - m^2 \phi^* \phi \quad (4.34)$$

Now we need to determine the canonical momenta. Recall that ϕ and ϕ^* are independent variables: so the momentum conjugate to ϕ and ϕ^* respectively are:

$$\pi_\phi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \dot{\phi}^* \quad \pi_{\phi^*} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}^*} = \dot{\phi} \quad (4.35)$$

For notational convenience from now on I will refer to these as

$$\pi = \pi_\phi \quad \pi^* = \pi_{\phi^*} \quad (4.36)$$

and now the Hamiltonian is

$$\mathcal{H} = \pi \dot{\phi} + \pi^* \dot{\phi}^* - \mathcal{L} = \dot{\phi} \dot{\phi}^* + (\vec{\nabla} \phi \cdot \vec{\nabla} \phi^*) + m^2 \phi^* \phi \quad (4.37)$$

4.2.1 Operator-valued Fourier expansions

It is now time to quantize! The idea is the fields $\phi(t, \vec{x})$ and $\pi(t, \vec{x})$ acquire hats and become operators in a Hilbert space!

$$\phi(t, \vec{x}) \rightarrow \hat{\phi}(t, \vec{x}) \quad (4.38)$$

Similarly for their conjugates (which we now write with a dagger to denote the fact that it is now a Hermitian conjugate of an operator, and not just the complex conjugate of a complex number):

$$\phi^*(t, \vec{x}) \rightarrow \hat{\phi}^\dagger(t, \vec{x}) \quad (4.39)$$

So, in the Heisenberg picture, we have the following expansion of the operator:

$$\hat{\phi}(\vec{x}, t) = \int \frac{d^3 k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(\hat{a}_{\vec{k}} e^{-i\omega_{\vec{k}} t + i\vec{k} \cdot \vec{x}} + \hat{b}_{\vec{k}}^\dagger e^{i\omega_{\vec{k}} t - i\vec{k} \cdot \vec{x}} \right) \quad (4.40)$$

where the a 's and b 's are now operators. Similarly, we have:

$$\hat{\phi}^\dagger(\vec{x}, t) = \int \frac{d^3 k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(\hat{a}_{\vec{k}}^\dagger e^{+i\omega_{\vec{k}} t - i\vec{k} \cdot \vec{x}} + \hat{b}_{\vec{k}} e^{-i\omega_{\vec{k}} t + i\vec{k} \cdot \vec{x}} \right) \quad (4.41)$$

This is the Hermitian conjugate of the above expression.

Let us also work out the Fourier transform of the canonical momenta. We do this the same way as for the harmonic oscillator: we begin with the classical definition:

$$\hat{\pi}(\vec{x}, t) = \dot{\hat{\phi}}^\dagger(\vec{x}, t) = i \int \frac{d^3 k}{(2\pi)^3} \sqrt{\frac{\omega_{\vec{k}}}{2}} \left(\hat{a}_{\vec{k}}^\dagger e^{i\omega_{\vec{k}} t - i\vec{k} \cdot \vec{x}} - \hat{b}_{\vec{k}} e^{-i\omega_{\vec{k}} t + i\vec{k} \cdot \vec{x}} \right) \quad (4.42)$$

and finally, same for the conjugate:

$$\hat{\pi}^\dagger(\vec{x}, t) = \dot{\hat{\phi}}(\vec{x}, t) = i \int \frac{d^3 k}{(2\pi)^3} \sqrt{\frac{\omega_{\vec{k}}}{2}} \left(-\hat{a}_{\vec{k}} e^{-i\omega_{\vec{k}} t + i\vec{k} \cdot \vec{x}} + \hat{b}_{\vec{k}}^\dagger e^{+i\omega_{\vec{k}} t - i\vec{k} \cdot \vec{x}} \right) \quad (4.43)$$

Now the point of all of these expansions is that we see that everything is determined by the operators $\hat{a}_{\vec{k}}$ and $\hat{b}_{\vec{k}}$. It turns out that these operators can again be used to diagonalize the Hamiltonian of the free quantum field theory and determine its spectrum: they will turn out to have a simple physical interpretation: **they create and annihilate particles!**

To understand this, let's first recall the canonical commutation relations between the scalar field and its canonical conjugate:

$$\boxed{[\hat{\phi}(\vec{x}, t), \hat{\pi}(\vec{y}, t)] = [\hat{\phi}^\dagger(\vec{x}, t), \hat{\pi}^\dagger(\vec{y}, t)] = i\delta^{(3)}(\vec{x} - \vec{y})} \quad (4.44)$$

All of the rest of the (equal-time) commutation relations are zero:

$$[\hat{\phi}, \hat{\phi}] = [\hat{\phi}, \hat{\pi}^\dagger] = \dots = 0 \quad (4.45)$$

I would now like to ask the question: what does this commutation relation imply for the commutators of the a 's and b 's? It turns out that it implies the following relation:

$$[\hat{a}_{\vec{k}}, \hat{a}_{\vec{k}'}^\dagger] = (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}') \quad (4.46)$$

$$[\hat{b}_{\vec{k}}, \hat{b}_{\vec{k}'}^\dagger] = (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}') \quad (4.47)$$

with all other commutators zero.

Before proving this, let's think for a second about what this means: it means that we can imagine each Fourier mode as being two independent harmonic oscillators: One coming from a and one coming from b . The delta function means that as far as this algebra is involved the Fourier modes don't talk to each other.

4.2.2 Algebra of creation/annihilation operators

Now let's prove this. This is the sort of calculation that comes up often in field theory, involving a lot of manipulation of integrals. Again as this is the first one, we will do it in quite a lot of detail.

We will first invert (4.40) to (4.43) for a, b as a function of ϕ, π . Before getting into it let's note the familiar Fourier identity:

$$\int d^3x e^{i\vec{p} \cdot \vec{x}} = (2\pi)^3 \delta^{(3)}(\vec{p}) \quad (4.48)$$

Now we multiply (4.40) with $e^{i\vec{k}' \cdot \vec{x}}$ and integrate over all $\vec{x}$:

$$\int d^3x e^{i\vec{k}' \cdot \vec{x}} \hat{\phi}(\vec{x}, t) = \int d^3x \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(\hat{a}_{\vec{k}} e^{-i\omega_{\vec{k}}t + i(\vec{k} + \vec{k}') \cdot \vec{x}} + \hat{b}_{\vec{k}}^\dagger e^{i\omega_{\vec{k}}t - i(\vec{k} - \vec{k}') \cdot \vec{x}} \right) \quad (4.49)$$

$$= \int d^3k \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(\hat{a}_{\vec{k}} e^{-i\omega_{\vec{k}}t} \delta^{(3)}(\vec{k} + \vec{k}') + \hat{b}_{\vec{k}}^\dagger e^{i\omega_{\vec{k}}t} \delta^{(3)}(\vec{k} - \vec{k}') \right) \quad (4.50)$$

$$= \frac{1}{\sqrt{2\omega_{\vec{k}'}}} \left(\hat{a}_{-\vec{k}'} e^{-i\omega_{\vec{k}'}t} + \hat{b}_{\vec{k}'}^\dagger e^{i\omega_{\vec{k}'}t} \right) \quad (4.51)$$

where it can be helpful to remember that $\omega_{-\vec{k}} = \omega_{\vec{k}}$. We have expressed a sum of a and $b^\dagger$ in terms of ϕ . To extract a or b by itself we need another linearly independent combination of them. We can find this by doing some very similar manipulations on $\hat{\pi}^\dagger(\vec{x}, t)$ in (4.43) to obtain

$$\int d^3x e^{i\vec{k} \cdot \vec{x}} \hat{\pi}^\dagger(\vec{x}, t) = i\sqrt{\frac{\omega_{\vec{k}}}{2}} \left(-\hat{a}_{-\vec{k}} e^{-i\omega_{\vec{k}}t} + \hat{b}_{\vec{k}}^\dagger e^{i\omega_{\vec{k}}t} \right) \quad (4.52)$$

Now we can multiply the first equation with $\sqrt{\omega_{\vec{k}}}$ and the second by $\frac{i}{\sqrt{\omega_{\vec{k}}}}$ to extract $a_{\vec{k}}$ by itself:

$$\hat{a}_{\vec{k}} = \int d^3x e^{-i\vec{k} \cdot \vec{x} + i\omega_{\vec{k}}t} \left(\hat{\phi}(\vec{x}, t) \sqrt{\frac{\omega_{\vec{k}}}{2}} + i\hat{\pi}^\dagger(\vec{x}, t) \frac{1}{\sqrt{2\omega_{\vec{k}}}} \right) \quad (4.53)$$

(note I took $\vec{k} \rightarrow -\vec{k}$ in order to get $a_{\vec{k}}$ rather than $a_{-\vec{k}}$). Similarly we can extract $\hat{b}_{\vec{k}}^\dagger$ by taking the opposite linear combination:

$$\hat{b}_{\vec{k}}^\dagger = \int d^3x e^{i\vec{k} \cdot \vec{x} - i\omega_{\vec{k}}t} \left(\hat{\phi}(\vec{x}, t) \sqrt{\frac{\omega_{\vec{k}}}{2}} - i\hat{\pi}^\dagger(\vec{x}, t) \frac{1}{\sqrt{2\omega_{\vec{k}}}} \right) \quad (4.54)$$

And as expected we can obtain the corresponding expressions for the daggered operators by taking the Hermitian conjugate, i.e.

$$\hat{a}_{\vec{k}}^\dagger = \int d^3x e^{+i\vec{k}\cdot\vec{x} - i\omega_{\vec{k}}t} \left(\hat{\phi}^\dagger(\vec{x}, t) \sqrt{\frac{\omega_{\vec{k}}}{2}} - i\hat{\pi}(\vec{x}, t) \frac{1}{\sqrt{2\omega_{\vec{k}}}} \right) \quad (4.55)$$

$$\hat{b}_{\vec{k}} = \int d^3x e^{-i\vec{k}\cdot\vec{x} + i\omega_{\vec{k}}t} \left(\hat{\phi}^\dagger(\vec{x}, t) \sqrt{\frac{\omega_{\vec{k}}}{2}} + i\hat{\pi}(\vec{x}, t) \frac{1}{\sqrt{2\omega_{\vec{k}}}} \right) \quad (4.56)$$

Okay, now let us consider the equal-time commutator of $\hat{a}_{\vec{k}}$ with $\hat{a}_{\vec{k}'}^\dagger$. We have:

$$\begin{aligned} [\hat{a}_{\vec{k}}, \hat{a}_{\vec{k}'}^\dagger] &= \frac{1}{2} \int d^3x d^3x' e^{it(\omega_{\vec{k}} - \omega_{\vec{k}'}) - i\vec{k}\cdot\vec{x} + i\vec{k}'\cdot\vec{x}'} \times \\ &\left(\sqrt{\omega_{\vec{k}}\omega_{\vec{k}'}} [\hat{\phi}(\vec{x}, t), \hat{\phi}^\dagger(\vec{x}', t)] - i\sqrt{\frac{\omega_{\vec{k}}}{\omega_{\vec{k}'}}} [\hat{\phi}(\vec{x}, t), \hat{\pi}(\vec{x}', t)] + i\sqrt{\frac{\omega_{\vec{k}'}}{\omega_{\vec{k}}}} [\hat{\pi}^\dagger(\vec{x}, t), \hat{\phi}^\dagger(\vec{x}', t)] + \frac{1}{\sqrt{\omega_{\vec{k}}\omega_{\vec{k}'}}} [\hat{\pi}^\dagger(\vec{x}, t), \hat{\pi}(\vec{x}', t)] \right) \end{aligned} \quad (4.57)$$

The commutators of ϕ with $\phi^\dagger$ and π with $\pi^\dagger$ automatically vanish. The other two give two delta functions, and we find:

$$[\hat{a}_{\vec{k}}, \hat{a}_{\vec{k}'}^\dagger] = \frac{1}{2} \int d^3x d^3x' e^{it(\omega_{\vec{k}} - \omega_{\vec{k}'}) - i\vec{k}\cdot\vec{x} + i\vec{k}'\cdot\vec{x}'} \left(\sqrt{\frac{\omega_{\vec{k}}}{\omega_{\vec{k}'}}} \delta^{(3)}(\vec{x} - \vec{x}') + \sqrt{\frac{\omega_{\vec{k}'}}{\omega_{\vec{k}}}} \delta^{(3)}(\vec{x} - \vec{x}') \right) \quad (4.58)$$

Now we can use the delta function to do the $\vec{x}'$ integral, setting $\vec{x}'$ to $\vec{x}$ everywhere:

$$[\hat{a}_{\vec{k}}, \hat{a}_{\vec{k}'}^\dagger] = \frac{1}{2} \int d^3x e^{it(\omega_{\vec{k}} - \omega_{\vec{k}'}) - i\vec{k}\cdot\vec{x} + i\vec{k}'\cdot\vec{x}} \left(\sqrt{\frac{\omega_{\vec{k}}}{\omega_{\vec{k}'}}} + \sqrt{\frac{\omega_{\vec{k}'}}{\omega_{\vec{k}}}} \right) \quad (4.59)$$

and now the remaining $\vec{x}$ integral gives us a delta function, i.e.

$$[\hat{a}_{\vec{k}}, \hat{a}_{\vec{k}'}^\dagger] = \frac{1}{2} (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}') e^{it(\omega_{\vec{k}} - \omega_{\vec{k}'})} \left(\sqrt{\frac{\omega_{\vec{k}}}{\omega_{\vec{k}'}}} + \sqrt{\frac{\omega_{\vec{k}'}}{\omega_{\vec{k}}}} \right) \quad (4.60)$$

Finally, because everything is multiplying a delta function which only has support when $\vec{k} = \vec{k}'$, we can freely set them to be equal to each other. The two square-rooted expressions each become 1 and together cancel the 2; thus we find eventually the desired result:

$$\boxed{[\hat{a}_{\vec{k}}, \hat{a}_{\vec{k}'}^\dagger] = (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}')} \quad (4.61)$$

I will leave it to you to check the only other nonzero commutator is

$$\boxed{[\hat{b}_{\vec{k}}, \hat{b}_{\vec{k}'}^\dagger] = (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}')} \quad (4.62)$$

with all else vanishing.

4.2.3 Hamiltonian and the energy of the vacuum

We now want to write the Hamiltonian of our quantum theory in terms of these creation and annihilation operators. From (4.37) we have that the classical Hamiltonian is:

$$H = \int d^3x \left(\pi\pi^* + \vec{\nabla}\phi \cdot \vec{\nabla}\phi^* + m^2\phi^*\phi \right) \quad (4.63)$$

We now want to figure out the quantum Hamiltonian: and we do this by simply making each of these classical variables into operators:

$$\hat{H} = \int d^3x \left(\hat{\pi}\hat{\pi}^\dagger + \vec{\nabla}\hat{\phi} \cdot \vec{\nabla}\hat{\phi}^\dagger + m^2\hat{\phi}^\dagger\hat{\phi} \right) \quad (4.64)$$

Now writing this in terms of the creation and annihilation operators is one of those “straightforward but tedious” exercises that you will do for your homework. After the dust settles you find the following answer

$$\hat{H} = \int \frac{d^3k}{(2\pi)^3} \omega_{\vec{k}} \left(\hat{a}_{\vec{k}}^\dagger \hat{a}_{\vec{k}} + \hat{b}_{\vec{k}} \hat{b}_{\vec{k}}^\dagger \right) \quad (4.65)$$

which I now rewrite using the commutator of b and $b^\dagger$ to be:

$$\hat{H} = \int \frac{d^3k}{(2\pi)^3} \omega_{\vec{k}} \left(\hat{a}_{\vec{k}}^\dagger \hat{a}_{\vec{k}} + \hat{b}_{\vec{k}}^\dagger \hat{b}_{\vec{k}} + (2\pi)^3 \delta^{(3)}(0) \right) \quad (4.66)$$

We see that we are instructed to evaluate the delta function at zero, where it’s infinite. This is really rather confusing (and is maybe the first of many infinities in this subject).

For now, let us build the Hilbert space – we begin by constructing the **vacuum state** $|0\rangle$, in precisely the same way as we did for the simple harmonic oscillator in – the vacuum is defined to be the state who satisfies:

$$a_{\vec{k}}|0\rangle = 0 \quad b_{\vec{k}}|0\rangle = 0 \quad \text{for all } \vec{k} \quad (4.67)$$

In other words, it is the state that is **empty of particles**. This is the state that is meant to be empty space.

Now let’s answer the first question that we can do with this formalism: what is the energy of empty space, i.e. what happens if we act with $\hat{H}$ on this state? We have

$$\hat{H}|0\rangle = \int \frac{d^3k}{(2\pi)^3} \omega_{\vec{k}} \left(\hat{a}_{\vec{k}}^\dagger \hat{a}_{\vec{k}} + \hat{b}_{\vec{k}}^\dagger \hat{b}_{\vec{k}} + (2\pi)^3 \delta^{(3)}(0) \right) |0\rangle = \int \frac{d^3k}{(2\pi)^3} \omega_{\vec{k}} (2\pi)^3 \delta^{(3)}(0) |0\rangle \quad (4.68)$$

In other words, the vacuum is indeed an energy eigenstate with an eigenvalue E_0 but this eigenvalue is

$$E_0 = \int \frac{d^3k}{(2\pi)^3} \omega_{\vec{k}} (2\pi)^3 \delta^{(3)}(0) \quad (4.69)$$

and thus the energy of the vacuum appears to be infinite.

This may appear perplexing. Let’s disentangle a few infinities. First, the $\delta^{(3)}(0)$ — this actually turns out to be related to the infinite volume of space. More precisely, consider doing the following regulated integral:

$$(2\pi)^3 \delta^{(3)}(p) = \lim_{V \rightarrow \infty} \int_V d^3x e^{ipx} \quad (4.70)$$

This means that in a finite box of size V , $\delta^{(3)}(0)$ should be interpreted as

$$(2\pi)^3 \delta^{(3)}(0) = \lim_{V \rightarrow \infty} \int_V d^3x = \lim_{V \rightarrow \infty} V \quad (4.71)$$

So one source of infinity comes from the fact that space is infinite; let's divide out by this to compute the vacuum energy *density*:

$$\epsilon_0 = \frac{E_0}{V} = \int \frac{d^3k}{(2\pi)^3} \omega_{\vec{k}} \left(\frac{1}{2} + \frac{1}{2} \right) \quad (4.72)$$

Recall the zero point energy (4.27) for the single harmonic oscillator. This is exactly the same, except that it receives a contribution from each *momentum mode* for both the a 's and the b 's and thus is extremely infinite. There is no obvious way to get rid of this – what should we do?

One thing we can do is the following: we can simply **ignore it**: after all, there is a sense in which “only energy differences are measurable”. This is thus a good time to introduce the idea of **normal ordering**. Given an operator $\mathcal{O}$ which is made up out of a product of field operators $\hat{\phi}(\vec{x}_1)\hat{\phi}(\vec{x}_2)\cdots$ etc. we define its normal ordered version $:\mathcal{O}:$ to be as follows: it is the regular product except that we expand it out and place all of the annihilation operators (i.e. a, b) to the *right* of the creation ones (i.e. $a^\dagger, b^\dagger$).

Thus the normal ordered version of the Hamiltonian above would be:

$$:\hat{H}:= \int \frac{d^3k}{(2\pi)^3} \omega_{\vec{k}} \left(\hat{a}_{\vec{k}}^\dagger \hat{a}_{\vec{k}} + \hat{b}_{\vec{k}}^\dagger \hat{b}_{\vec{k}} \right) \quad (4.73)$$

and thus the normal-ordered version of the Hamiltonian indeed annihilates the ground state $|0\rangle$:

$$:\hat{H}:|0\rangle = 0 \quad (4.74)$$

and we have lost this pesky vacuum energy.

This may feel ad-hoc – and I suppose there is a sense where it is. It may make you feel better to realize that such shifts in the Hamiltonian can also arise in the mapping from classical to quantum Hamiltonians: for example, classically the Hamiltonian $\frac{1}{2}(\omega q - ip)(\omega q + ip)$ is exactly the same as $\frac{1}{2}(\omega^2 q^2 + p^2)$, but quantum mechanically the *first* one is $\omega a^\dagger a$ and the second one is $\omega \left(a^\dagger a + \frac{1}{2} \right)$. So the issue that we are stressing about arises from a sort of ambiguity in passing from the classical to the quantum theory.

Due to lack of time I won't discuss this further, but there is a lot of beautiful physics in this vacuum energy – for example, Google *Casimir effect* to understand an experimentally measurable manifestation of this.

4.2.4 Single particle states

Let's move on past the vacuum to consider **single-particle states**. First, we will compute the commutator:

$$[:\hat{H}:, \hat{a}_{\vec{k}}^\dagger] = \left[\int \frac{d^3k'}{(2\pi)^3} \omega_{\vec{k}'} \left(\hat{a}_{\vec{k}'}^\dagger \hat{a}_{\vec{k}'} \right), \hat{a}_{\vec{k}}^\dagger \right] \quad (4.75)$$

$$= \int \frac{d^3k'}{(2\pi)^3} \omega_{\vec{k}'} \hat{a}_{\vec{k}'}^\dagger [\hat{a}_{\vec{k}'}, \hat{a}_{\vec{k}}^\dagger] \quad (4.76)$$

$$= \int \frac{d^3k'}{(2\pi)^3} (2\pi)^3 \omega_{\vec{k}'} \hat{a}_{\vec{k}'}^\dagger \delta^{(3)}(\vec{k}' - \vec{k}) \quad (4.77)$$

$$= \omega_{\vec{k}} \hat{a}_{\vec{k}}^\dagger \quad (4.78)$$

which is familiar from the SHO discussion. Similarly, we have

$$[:\hat{H}:, \hat{a}_{\vec{k}}] = -\omega_{\vec{k}} \hat{a}_{\vec{k}} \quad (4.79)$$

and exactly the same expressions for the b 's.

$$[:\hat{H}:, \hat{b}_{\vec{k}}^\dagger] = \omega_{\vec{k}} \hat{b}_{\vec{k}}^\dagger \quad [:\hat{H}:, \hat{b}_{\vec{k}}] = -\omega_{\vec{k}} \hat{b}_{\vec{k}} \quad (4.80)$$

Now, let's consider an excited state $a_{\vec{k}}^\dagger|0\rangle$. We compute its energy to be:

$$:\hat{H} : a_{\vec{k}}^\dagger|0\rangle = \left([:\hat{H} :, a_{\vec{k}}^\dagger] + a_{\vec{k}}^\dagger : \hat{H} : \right) |0\rangle = \left(\omega_{\vec{k}} a_{\vec{k}}^\dagger + a_{\vec{k}}^\dagger : \hat{H} : \right) |0\rangle = \omega_{\vec{k}} a_{\vec{k}}^\dagger|0\rangle \quad (4.81)$$

In other words, the energy of this state is

$$\omega_{\vec{k}} = \sqrt{\vec{k}^2 + m^2} \quad (4.82)$$

which is precisely the energy of a single relativistic particle with momentum $\vec{k}$. **We have created a particle!**

Similarly, we can show by an identical calculation that the energy of the state $b_{\vec{k}}^\dagger|0\rangle$ is also $\omega_{\vec{k}}$.

It is now good to make precise our interpretation of what the a 's do versus the b 's. Recall from a few lectures ago that the complex scalar field has a $U(1)$ symmetry which is $\phi \rightarrow e^{i\alpha}\phi$. We constructed a charge Q from the Noether procedure for that symmetry. Classically that charge was (where here I have adjusted the overall sign for convenience):

$$Q = -i \int d^3x [(\partial_t \phi^*)\phi - \phi^*(\partial_t \phi)] \quad (4.83)$$

which we can now promote to the quantum theory as:

$$\hat{Q} = -i \int d^3x [\hat{\pi}\hat{\phi} - \hat{\phi}^\dagger\hat{\pi}^\dagger] \quad (4.84)$$

Plugging in the Fourier expansions of the fields we ultimately find – after doing a computation which we will see in the Problem Class – that in normal ordered form this is

$$:\hat{Q} := \int \frac{d^3k}{(2\pi)^3} \left(\hat{a}_{\vec{k}}^\dagger \hat{a}_{\vec{k}} - \hat{b}_{\vec{k}}^\dagger \hat{b}_{\vec{k}} \right) \quad (4.85)$$

This is the $U(1)$ *charge*. Note the crucial sign difference as compared to the Hamiltonian. The “b”-type particles have opposite charge to the “a”-type particles.

By Noether's theorem, this charge is classically time-independent. In the quantum theory, this means that

$$[\hat{Q}, \hat{H}] = 0 \quad (4.86)$$

which means that it is possible to diagonalize the Hamiltonian and the charge simultaneously. Indeed if we act with $\hat{Q}$ on the two states above we find:

$$:\hat{Q} : |0\rangle = 0 \quad : \hat{Q} : a_{\vec{k}}^\dagger|0\rangle = (+1)\hat{a}_{\vec{k}}^\dagger|0\rangle \quad : \hat{Q} : b_{\vec{k}}^\dagger|0\rangle = (-1)\hat{b}_{\vec{k}}^\dagger|0\rangle \quad (4.87)$$

In other words, these are all charge eigenstates. The a -type state has positive charge $+1$, but the b -type state has *negative* charge -1 . Thus a creates particles, and b creates **anti-particles**.

Finally, if you recall, Problems Class 1 we constructed the (classical) operator that measures momentum P^i , by finding the Noether charge for translational invariance. We can do this for the quantum theory as well (this will be part of Problem Sheet 3) and show that we have

$$\hat{P}^i a_{\vec{k}}^\dagger|0\rangle = k^i a_{\vec{k}}^\dagger|0\rangle \quad (4.88)$$

Let me summarize what we have learned:

1. The vacuum $|0\rangle$ is empty; provided we measure energy and charge by the normal-ordered versions, it has no energy and no charge.
2. The state $a_{\vec{k}}^\dagger|0\rangle$ is a single-particle state. It has energy $\omega_{\vec{k}}$, mass m , momentum k^i , charge $+1$
3. The state $b_{\vec{k}}^\dagger|0\rangle$ is a single-*anti* particle state. It has energy $\omega_{\vec{k}}$, mass m , momentum k^i , charge -1

Note that these particles are momentum eigenstates, which means – via the regular rules of quantum mechanics – that they are completely delocalized in space.

4.2.5 The Fock space

We have just discussed a few very special excited states. Let's now consider the full space of multi-particle states, which we can create by acting multiple times with the creation and annihilation operators, e.g. imagine the following state with N_a particles and N_b antiparticles:

$$\hat{a}_{\vec{p}_1}^\dagger \hat{a}_{\vec{p}_2}^\dagger \cdots \hat{a}_{\vec{p}_{N_a}}^\dagger \hat{b}_{\vec{q}_1}^\dagger \hat{b}_{\vec{q}_2}^\dagger \cdots \hat{b}_{\vec{q}_{N_b}}^\dagger |0\rangle = |\{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_{N_a}\}; \{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_{N_b}\}\rangle \quad (4.89)$$

In this context we can see that the combination $\int \frac{d^3k}{(2\pi)^3} \hat{a}_k^\dagger \hat{a}_k$ appearing in the expression for the Hamiltonian and charge simply counts the number of particles. To compute the charge we simply add this up over all momenta (and subtract the corresponding quantity for anti-particles). To compute the energy we add this up – weighted by a factor $\omega_{\vec{k}}$ – for all particles.

Indeed this is an energy and charge eigenstate with the following eigenvalues:

$$Q = N_a - N_b \quad H = \sum_{i=1}^{N_a} \omega_{\vec{p}_i} + \sum_{j=1}^{N_b} \omega_{\vec{q}_j} \quad (4.90)$$

As promised, we have constructed a space of states that allows particle number to change. The state space created by all possible ways of acting with creation operators on the vacuum is called the **Fock space**. We see that roughly speaking a basis for the Fock space is labeled by a set of integers $n_{\vec{k}_i}$, two for each momentum.

Finally, let's consider the two-particle state, labeled by two momenta:

$$|\vec{p}_1, \vec{p}_2\rangle = \hat{a}_{\vec{p}_1}^\dagger \hat{a}_{\vec{p}_2}^\dagger |0\rangle \quad (4.91)$$

What happens if we interchange the two particles? Because the two $a^\dagger$'s commute, we have:

$$|\vec{p}_2, \vec{p}_1\rangle = \hat{a}_{\vec{p}_2}^\dagger \hat{a}_{\vec{p}_1}^\dagger |0\rangle = \hat{a}_{\vec{p}_1}^\dagger \hat{a}_{\vec{p}_2}^\dagger |0\rangle = |\vec{p}_1, \vec{p}_2\rangle \quad (4.92)$$

In other words, because the $a^\dagger$'s commute with one another, the state is *exactly* unchanged if we interchange the two particles – there is no meaning to which one we write “first”. Note that there is no way whatsoever to distinguish the two particles; this is quite different from classical physics. We say that the particles are **bosons**, after Bengali physicist Satyendranath Bose⁸.

Finally, we have discussed creating particles that have definite momentum. We may also be interested in creating particles that have a definite *position*. As usual in quantum mechanics, we do this by adding up momentum eigenstates in the correct fashion: in our case, this is already done with our original field $\phi(\vec{x}, t)$:

$$\hat{\phi}(\vec{x}, t) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(\hat{a}_{\vec{k}} e^{-i\omega_{\vec{k}}t + i\vec{k}\cdot\vec{x}} + \hat{b}_{\vec{k}}^\dagger e^{i\omega_{\vec{k}}t - i\vec{k}\cdot\vec{x}} \right) \quad (4.93)$$

Imagine acting with this field on the vacuum $|0\rangle$, to obtain the state:

$$\hat{\phi}(\vec{x}, t)|0\rangle = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \hat{b}_{\vec{k}}^\dagger e^{i\omega_{\vec{k}}t - i\vec{k}\cdot\vec{x}} |0\rangle \quad (4.94)$$

⁸Bose grew up in Kolkata in India, later moving to Dhaka University, which is now in Bangladesh. He studied the problem of black-body radiation in quantum systems and found that the quantum statistics that we describe by the formula above resulted in the correct prefactor for Planck's law. He sent his results to Einstein, who realized their importance. He translated them to German and had them published under Bose's name in *Zeitschrift für Physik*; this is why in quantum statistical mechanics the corresponding finite-temperature distribution is referred to as Bose-Einstein statistics. It results in many striking phenomena (you may have heard of Bose-Einstein condensation?), and is one of exactly two classes of quantum statistics realized in nature – the others are **fermions**, which anti-commute with each other.

We see that the $\hat{a}$ will annihilate itself, but the term in $b_{\vec{k}}^\dagger$ will create a superposition of anti-particles, summing over all momenta with wavefunction $e^{-i\vec{k}\cdot\vec{x}}$. Thus $\phi(t, \vec{x})$ creates an **anti-particle** at position $\vec{x}, t$.

Similarly, if we act with $\phi^\dagger(\vec{x}, t)$:

$$\hat{\phi}^\dagger(\vec{x}, t) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(\hat{a}_{\vec{k}}^\dagger e^{+i\omega_{\vec{k}}t - i\vec{k}\cdot\vec{x}} + \hat{b}_{\vec{k}} e^{-i\omega_{\vec{k}}t + i\vec{k}\cdot\vec{x}} \right) \quad (4.95)$$

Now the $\hat{b}$ will annihilate itself against the vacuum, and we end up with a sum over $a_{\vec{k}}^\dagger$'s:

$$\hat{\phi}^\dagger(\vec{x}, t)|0\rangle = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \hat{a}_{\vec{k}}^\dagger e^{+i\omega_{\vec{k}}t - i\vec{k}\cdot\vec{x}} |0\rangle \quad (4.96)$$

This operator creates a **particle** at position $\vec{x}, t$.

4.3 Propagators and causality

We are now ready to ask some questions about how particles **propagate** from one point to another. As it turns out, there are multiple notions of propagator. Before getting into this, however, I want to discuss a brief mathematical interlude: how do we define a Lorentz-invariant measure on integrating over 3-momenta?

From now on a slight notational change: I am going to stop putting the hats on the operators, as we basically *only* have quantum operators from now on. It should be clear from context whether something is an operator or a number.

4.3.1 Lorentz-invariant integration measures

As it turns out, the integral $\int d^3k$ is not Lorentz invariant; if I imagine that the 3-momenta $\vec{k}$ are part of a 4-vector k_μ then the integration measure above transforms by factors of γ under a boost. Let us improve on this by constructing an invariant measure. To do this, consider the following manifestly Lorentz invariant integral:

$$\int d^4k \delta(-k_\mu k^\mu - m^2) \quad (4.97)$$

Now this is an integral over 4-momenta with a constraint that the magnitude of the 4-momenta be equal to (timelike) m^2 . Note that this defines a 3-dimensional hyperboloid in momentum space.

Let us do the integral over k_0 first:

$$\int d^3k dk_0 \delta(k_0^2 - \omega_{\vec{k}}^2) \quad (4.98)$$

We can solve the delta function by picking one or the other of:

$$k_0 = \pm \sqrt{\vec{k}^2 + m^2} = \pm \omega_{\vec{k}} \quad (4.99)$$

Consider taking the positive sign for simplicity. Now using the basic rules for delta functions we find

$$\int d^3k dk_0 \delta(k_0 - \omega_{\vec{k}}) \left| \frac{dk_0^2}{dk_0} \right|^{-1} = \int \frac{d^3k}{2\omega_{\vec{k}}} \quad (4.100)$$

Thus the point is that the measure

$$\int \frac{d^3k}{2\omega_{\vec{k}}} \quad (4.101)$$

is a Lorentz-invariant measure. It is the integral over the hyperboloid in momentum space discussed earlier, and it will often appear in calculations.

4.3.2 Commutators and the light-cone

Let us now note an uncomfortable fact – though I made a big deal about Lorentz invariance in the first part of this course, when we passed to the quantum theory we made heavy use of the Hamiltonian formalism, and we quantized by picking a rest frame. This was actually quite important, as the basic starting point was the commutation relation

$$[\phi(\vec{x}, t), \pi(\vec{y}, t)] = i\delta^{(3)}(\vec{x} - \vec{y}), \quad (4.102)$$

which took place at *equal times* in that frame.

It is not at all clear that the resulting quantum theory is really Lorentz invariant; perhaps our quantization procedure destroyed this.⁹

Let us consider a diagnostic. First, let me remind you of the idea of the light-cone: if an event happens at x , then the set of events that it can influence happen on spacetime points y that are timelike separated from it, i.e. where $(x - y)^2 < 0$. This forms the interior of a cone in $\mathbb{R}^{1,3}$.

If our theory is Lorentz-invariant, then no information should propagate faster than light. This means that any two measurements in the quantum theory that happen *outside the light-cone* should not affect each other: in other words, any operator $\mathcal{O}_1$ should commute with another $\mathcal{O}_2$ provided it is *spacelike* separated from $\mathcal{O}_2$, i.e.

$$[\mathcal{O}_1(x), \mathcal{O}_2(y)] = 0 \quad \text{if } (x - y)^2 > 0 \quad (4.103)$$

Note that the two operators above are defined at different spacetime points, and not at equal times. We are now in a position to explicitly check this for the complex scalar field. Let's compute from the Fourier expansion (4.40) and (4.41):

$$[\phi(x), \phi^\dagger(y)] = \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_{\vec{k}}} \left(e^{ik \cdot (x-y)} - e^{-ik \cdot (x-y)} \right) \quad (4.104)$$

where I have introduced the 4-momentum k -vector, whose components are

$$k_\mu = (-\omega_{\vec{k}}, \vec{k}) \quad (4.105)$$

Here the first term comes from the commutator of a and $a^\dagger$, and the second from b and $b^\dagger$. This is not obviously zero. Note that it's interesting that it is *just* a number, not an operator like the objects on the right-hand side were¹⁰. Note also that it is made up of Lorentz-invariant objects, i.e. the Lorentz-invariant integration measure (4.101) and the 4-vector inner product $k \cdot (x - y)$. Let us now consider two cases: in the first, let's imagine that the separation $(x - y)^\mu$ is a *timelike* vector. In that case we can use Lorentz-invariance to rotate it to point purely in the time direction, i.e. there exists a Lorentz boost Λ so that

$$\Lambda^\mu_\nu (x - y)^\nu = (x'^0 - y'^0, 0, 0, 0)^\mu \quad (4.106)$$

(If the existence of this Lorentz transformation is not obvious, it is instructive to construct the boost that does this yourself!). We may now use Lorentz-invariance to evaluate the right-hand side in this coordinate system to find

$$\int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_{\vec{k}}} \left(e^{-i\omega_{\vec{k}}(x^0 - y^0)'} - e^{i\omega_{\vec{k}}(x^0 - y^0)'} \right) \quad (4.107)$$

which is some oscillatory function of $(x^0 - y^0)'$ that is not zero.

⁹This is not actually as academic of a worry as it seems: it is quite possible for a quantum theory to not be invariant under the symmetries of its classical counterpart: this is called a quantum mechanical *anomaly*, and there is really a sense in which most of the mass of the universe – that of protons and neutrons – arises from such an effect.

¹⁰This particular fact is actually only true for free quantum field theories.

Now, let's turn to the interesting case where $(x - y)^\mu$ is *spacelike*; in that case we can transform it to point *purely in the spatial direction*, i.e. there exists a (different) Lorentz boost Λ so that

$$\Lambda^\mu{}_\nu (x - y)^\nu = (0, \vec{x}' - \vec{y}')^\mu \quad (4.108)$$

We now perform the computation in this coordinate system to find:

$$\int \frac{d^3k}{(2\pi)^3} \frac{1}{2\sqrt{\vec{k}^2 + m^2}} \left(e^{+i\vec{k} \cdot (\vec{x} - \vec{y})'} - e^{-i\vec{k} \cdot (\vec{x} - \vec{y})'} \right) \quad (4.109)$$

Now note that we can change the integration variable $\vec{k} \rightarrow -\vec{k}$ in the last term to see that

$$\int \frac{d^3k}{(2\pi)^3} \frac{1}{2\sqrt{\vec{k}^2 + m^2}} \left(e^{+i\vec{k} \cdot (\vec{x} - \vec{y})'} - e^{+i\vec{k} \cdot (\vec{x} - \vec{y})'} \right) = 0 \quad (4.110)$$

This holds for *any* spacelike separation where $(x - y)^2 > 0$ ¹¹. Thus we conclude that

$$[\phi(x), \phi^\dagger(y)] = 0 \quad \text{if } (x - y) \text{ spacelike} \quad (4.111)$$

as required. The same is true for any commutator of any local operators, and is actually true in any Lorentz-invariant theory, not just the free quantum field theory that we just explicitly solved: quantum field theory does *not* let one send information outside the light-cone.

4.3.3 Particle propagation

Let us now ask a slightly different question. Somewhat imprecisely: if we create an anti-particle at one space-time point $x^\mu = (x^0, \vec{x})$, what is the amplitude of finding it at a different spacetime point $y^\mu = (y^0, \vec{y})$?

As it turns out there are actually several different answers to this question. The most obvious putative answer to this is given by the following overlap:

$$\langle 0 | \phi^\dagger(y) \phi(x) | 0 \rangle \quad (4.112)$$

where from the expressions (4.94) and its Hermitian conjugate we have:

$$\phi(\vec{x}, t) | 0 \rangle = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} b_{\vec{k}}^\dagger e^{i\omega_{\vec{k}}t - i\vec{k} \cdot \vec{x}} | 0 \rangle \quad \langle 0 | \phi^\dagger(\vec{x}, t) = \langle 0 | \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} b_{\vec{k}} e^{-i\omega_{\vec{k}}t + i\vec{k} \cdot \vec{x}} \quad (4.113)$$

Putting these in we find:

$$\langle \phi^\dagger(y) \phi(x) \rangle = \int \frac{d^3k_1 d^3k_2}{(2\pi)^3 (2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}_1} 2\omega_{\vec{k}_2}}} e^{-ik_1 \cdot x + ik_2 \cdot y} \langle 0 | b_{\vec{k}_2} b_{\vec{k}_1}^\dagger | 0 \rangle, \quad (4.114)$$

where for the 4-vector k_1 and k_2 we have:

$$k_{\mu,i} = (-\omega_{\vec{k}_i}, \vec{k}_i) \quad (4.115)$$

Note that we have omitted the choice of the vacuum state in the expectation value; in general whenever you see this notation you should assume that the expectation value is taken in the vacuum, i.e. $\langle 0 | X | 0 \rangle = \langle X \rangle$. Now commuting the b past its dagger and annihilating it on the vacuum we find:

$$\langle \phi^\dagger(y) \phi(x) \rangle = \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_{\vec{k}}} e^{-ik \cdot (x - y)} \quad (4.116)$$

¹¹It's helpful to compare this to the above computation and ask what the exact difference is: there is no change of variables that takes $\omega_{\vec{k}}$ to $-\omega_{\vec{k}}$

This expression is sometimes called the **propagator** or **correlator** (as it is a correlation function of two fields). Note again that it's completely Lorentz-invariant, just like the expression for the commutator I wrote down above.

It's quite possible to evaluate it explicitly, though as usual I will not do so here. Let's just ask a more basic question: note that unlike the commutator, it does *not* vanish outside the light-cone. In fact the required integral can be done via Bessel functions¹² and we find that for large spacelike $|\vec{x} - \vec{y}| \rightarrow \infty$ it behaves as

$$\langle \phi^\dagger(y) \phi(x) \rangle \sim \exp(-m|\vec{x} - \vec{y}|) \quad (4.117)$$

What does this mean, exactly? The fact that this does not vanish means that there is *entanglement* in the vacuum of a quantum field theory. One can get some intuition for this by considering a far simpler system, the famous EPR (Einstein-Podolsky-Rosen) state for two entangled spins, which you have probably seen:

$$|\text{EPR}\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle|\uparrow\rangle + |\downarrow\rangle|\downarrow\rangle) \quad (4.118)$$

If you have this state and you compute e.g. $\langle \text{EPR} | \sigma_1^z \sigma_2^z | \text{EPR} \rangle$ you will obtain a non-zero correlation because the spins are correlated by their entanglement (i.e. if you know that spin 1 is up, then you know what spin 2 is doing). Similarly, one can interpret the correlation above as stating that in the vacuum, there is quantum entanglement between *different points in space*.

Now can we use this entanglement to send information faster than light? We have already seen above through a direct computation that the answer is no. Let's try to interpret the answer we found above in this language; note that staring at the expression for the commutators we have

$$[\phi(x), \phi^\dagger(y)] = \langle [\phi(x), \phi^\dagger(y)] \rangle = \langle \phi(x) \phi^\dagger(y) \rangle - \langle \phi^\dagger(y) \phi(x) \rangle \quad (4.119)$$

The first equality follows from the fact that the commutator is a number, and so we can evaluate it in any state; the second follows from expanding out the commutator. Now let us wrap some words around the expressions on the right: the first is the probability for a *particle* to propagate from y to x , and the second is the probability for an *anti*-particle to propagate from x to y . These exactly cancel outside the light-cone.

More poetically: the reason we cannot use the particle entanglement to send information faster than light is because pesky anti-particles appear to cancel them. This highlights an uncomfortably deep relationship between the existence of anti-particles and the causal structure of relativity.

4.4 Feynman propagator

As it turns out, the most useful propagator in quantum field theory is actually *not* the object described above, but a slightly different object called **the Feynman propagator**. To introduce it, I need to first explain the notion of **time-ordering**. A **time-ordered product** of some operators $\mathcal{O}_1(x_1), \mathcal{O}_2(x_2), \dots$ is written as

$$T(\mathcal{O}_1(x_1) \mathcal{O}_2(x_2) \cdots) \quad (4.120)$$

is defined as the correlation function that you get when you order all of the operators by *time* and then place the *earliest* ones to the *right*. So after the ordering you will find something like

$$\mathcal{O}_a(x_a) \mathcal{O}_b(x_b) \cdots \quad (4.121)$$

where $a, b, \dots$ is a permutation of the original operator labels $1, 2, \dots$ where we have

$$t_a > t_b > t_c > \cdots t_n \quad (4.122)$$

¹²Fun exercise! Perhaps depends somewhat on your definition of "fun".

So for example, the time ordered correlator of ϕ and $\phi^\dagger$ is

$$T(\phi(x')\phi^\dagger(x)) = \theta(t' - t)\phi(x')\phi^\dagger(x) + \theta(t - t')\phi^\dagger(x)\phi(x'), \quad (4.123)$$

where we have introduced the Heaviside step function $\theta(x)$, as usual defined as

$$\theta(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases} \quad (4.124)$$

We now **define** the Feynman propagator $G(x, x')$ to be

$$G(x, x') = \langle 0 | T(\phi(x')\phi^\dagger(x)) | 0 \rangle, \quad (4.125)$$

i.e. the expectation value of the time-ordered product of ϕ and $\phi^\dagger$.

It is at this point reasonable to ask **why**? I will basically dodge this question for now: it turns out that this is a quantity that is in many ways the “correct” notion of particle propagation; we will see when we start doing *interacting* quantum fields and scattering particles off of each other that this is a natural building block.

Now let’s calculate it. First the two pieces: we have already calculated $\phi^\dagger(x)\phi(x')$ in (4.116), and a very similar calculation give us the remaining piece $\phi(x')\phi^\dagger(x)$. Putting them together we find

$$G(x, x') = \theta(t' - t) \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_{\vec{k}}} e^{ik \cdot (x' - x)} + \theta(t - t') \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_{\vec{k}}} e^{-ik \cdot (x' - x)} \quad (4.126)$$

in both cases we have $k_\mu = (-\omega_{\vec{k}}, \vec{k})$ as usual. It’s worth expanding out the exponents to understand precisely what the difference is between the two branches:

$$G(x, x') = \theta(t' - t) \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_{\vec{k}}} e^{-i\omega_{\vec{k}}(t' - t) + i\vec{k} \cdot (\vec{x}' - \vec{x})} + \theta(t - t') \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_{\vec{k}}} e^{+i\omega_{\vec{k}}(t' - t) - i\vec{k} \cdot (\vec{x}' - \vec{x})} \quad (4.127)$$

So the point is that if $t' > t$, then the behavior in time is like $e^{-i\omega_{\vec{k}}(t' - t)}$, whereas if $t' < t$, then the behavior in time is the other way, i.e. $e^{+i\omega_{\vec{k}}(t' - t)}$. This is the key defining property of the Feynman propagator. (The spatial dependence also looks like it oscillates the other way on the two branches, but recall from the previous part that this is an illusion: we can always take $\vec{k} \rightarrow -\vec{k}$ in the integral, but no similar manipulation can be done for the time integral.)

Now I will show you an elegant way to write this answer. As it turns out, the Feynman propagator can be written as

$$G(x, x') = -i \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik \cdot (x' - x)}}{k_\nu k^\nu + m^2} \quad (4.128)$$

Note that here we are doing something different: we are actually integrating over all four components of the 4-vector $k^\mu = (k^0, \vec{k})$, which is not something that we’ve done before; it should thus not be obvious at all that this is similar to the expressions above. To show that it is, we will explicitly do the integral over k^0 and show that it agrees with (4.127)

It also turns out that this integral is not yet well-defined, as I will explain in a second.

We will do these integrals by contour integrals in the complex plane, which I believe everyone already knows. To remind you: if we have an *analytic* function $f(z)$ of a complex variable z – which for our purposes means that the only singularities are poles – then the anti-clockwise integral over a closed loop Γ in the complex plane is

$$\oint_{\Gamma} dz f(z) = 2\pi i \sum (\text{Residues at poles of } f(z) \text{ inside } \Gamma) \quad (4.129)$$

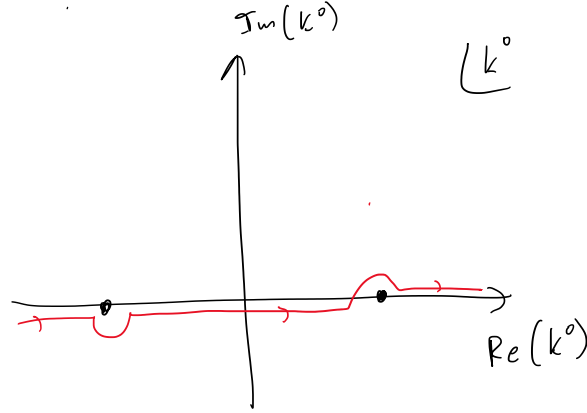


Figure 4.1: Contour for frequency integration in the complex k^0 plane for the Feynman propagator

A very simple example of this is the integral over the function $\frac{1}{z}$, which is

$$\oint_{\Gamma} dz \frac{1}{z} = 2\pi i \quad (4.130)$$

since the residue is just 1. If we do the integral clockwise instead then we pick up a different sign.

Now let's consider the integral over k^0 :

$$G(x, x') = -i \int \frac{d^3 k}{(2\pi)^3} \int_{-\infty}^{+\infty} \frac{dk^0}{(2\pi)} \frac{e^{ik^0(t'-t) - i\vec{k} \cdot (\vec{x}' - \vec{x})}}{-(k^0)^2 + \vec{k}^2 + m^2} \quad (4.131)$$

Let's hold $\vec{k}$ fixed for now and focus on the k^0 integral, which I'll call I . We see that it is done over the real line, but it actually hits two poles at $k^0 = \pm\omega_{\vec{k}}$, i.e. we have:

$$I = -i \int \frac{dk^0}{2\pi} \frac{e^{ik^0(t'-t) - i\vec{k} \cdot (\vec{x}' - \vec{x})}}{-(k^0)^2 + \vec{k}^2 + m^2} = +i \int \frac{dk^0}{2\pi} \frac{e^{ik^0(t'-t) - i\vec{k} \cdot (\vec{x}' - \vec{x})}}{(k^0 + \omega_{\vec{k}})(k^0 - \omega_{\vec{k}})} \quad (4.132)$$

These poles actually mean that the integral is ill-defined as an ordinary real integral unless we provide some more information: we need to define this as an integral where k^0 is allowed to stray into the complex l^0 plane, and we then need to explain how to go around the poles.

The Feynman propagator is **defined** by following the contour in Figure 4.1, i.e. we go *below* the pole at $-\omega_{\vec{k}}$ and *above* the pole at $+\omega_{\vec{k}}$. Let's now understand carefully how this works by computing the two branches above. To use the residue integration formula we need to consider a closed contour, which means in this case that we close the contour either above - i.e. at $k^0 \rightarrow +i\infty$, or below, at $k^0 \rightarrow -i\infty$. As usual in complex analysis, if we pick the contour so that this closure does not contribute, then the integral over the closed integral is equal to the integral which we wish to compute.

Which way we close the contour depends on the sign of $t' - t$: if $t' - t > 0$, then we should close the contour above so that $e^{ik^0(t'-t)} \sim e^{-\infty(t'-t)}$ decays exponentially. This means that we are going anti-clockwise around the pole at $k^0 = -\omega_{\vec{k}}$. The residue is everything that multiplies the pole if one sets $k^0 = -\omega_{\vec{k}}$, i.e. the from the integrand it is

$$\text{residue} = \frac{i}{2\pi} \frac{e^{-i\omega_{\vec{k}}(t'-t) - i\vec{k} \cdot (\vec{x}' - \vec{x})}}{-2\omega_{\vec{k}}} \quad (4.133)$$

so using the residue formula we find that

$$I = (2\pi i) \frac{-i}{2\pi} \frac{e^{-i\omega_{\vec{k}}(t'-t) - i\vec{k}\cdot(\vec{x}'-\vec{x})}}{2\omega_{\vec{k}}} \quad (4.134)$$

which becomes

$$I = \frac{e^{-i\omega_{\vec{k}}(t'-t) - i\vec{k}\cdot(\vec{x}'-\vec{x})}}{2\omega_{\vec{k}}} \quad t' > t \quad (4.135)$$

Now let us consider doing the same with the opposite sign for $t' - t < 0$. By the same argument, we now need to close the contour *below* so that $e^{ik^0(t'-t)} \sim e^{-\infty|t'-t|}$ decays exponentially. We then pick up the other pole at $k^0 = +\omega_{\vec{k}}$. Now the residue is again everything that multiplies that pole, if we set $k^0 = +\omega_{\vec{k}}$, i.e.

$$\text{residue} = \frac{i}{2\pi} \frac{e^{+i\omega_{\vec{k}}(t'-t) - i\vec{k}\cdot(\vec{x}'-\vec{x})}}{2\omega_{\vec{k}}} \quad (4.136)$$

Note now that we are going around the contour the *other* way, so the residue formula picks up an additional sign. Using the residue formula we find that

$$I = \frac{e^{+i\omega_{\vec{k}}(t'-t) - i\vec{k}\cdot(\vec{x}'-\vec{x})}}{2\omega_{\vec{k}}} \quad t' < t \quad (4.137)$$

The key difference here is the sign of the quantity in the exponent multiplying $t' - t$, which we see now differs depending on the sign of $t' - t$. Now putting everything together we have

$$G(x, x') = -i \int \frac{d^4 k}{(2\pi)^4} \frac{e^{-ik\cdot(x'-x)}}{k_\nu k^\nu + m^2} \quad (4.138)$$

$$= \theta(t' - t) \int \frac{d^3 k}{(2\pi)^3} \frac{1}{2\omega_{\vec{k}}} e^{-i\omega_{\vec{k}}(t'-t) + i\vec{k}\cdot(\vec{x}'-\vec{x})} + \theta(t - t') \int \frac{d^3 k}{(2\pi)^3} \frac{1}{2\omega_{\vec{k}}} e^{+i\omega_{\vec{k}}(t'-t) - i\vec{k}\cdot(\vec{x}'-\vec{x})} \quad (4.139)$$

where we have changed the sign of $\vec{k} \rightarrow -\vec{k}$ in the first term. This is exactly what we had in (4.127), which is indeed what we set out to prove.

Thus we see that the integral (4.128), together with the specification of the integral contour, give us a nice Lorentz-covariant expression for the Feynman propagator. We will make heavy use of this expression in the future. Often we capture the information in the choice of contour by writing the expression for the Feynman propagator compactly as

$$G(x, x') = -i \int \frac{d^4 k}{(2\pi)^4} \frac{e^{-ik\cdot(x'-x)}}{k^2 + m^2 - i\epsilon} \quad (4.140)$$

where $\epsilon > 0$ is a positive infinitesimal. Its role in life is to tip the poles slightly off of the real axis in precisely the way as in our contour prescription: to understand this let's solve for the location of the poles with the ϵ involved:

$$-(k^0)^2 + \vec{k}^2 + m^2 - i\epsilon = 0 \quad \rightarrow \quad k_0^2 = \omega_{\vec{k}}^2 - i\epsilon \quad (4.141)$$

which becomes

$$k_0 = \pm \omega_{\vec{k}} \sqrt{1 - \frac{i\epsilon}{\omega_{\vec{k}}^2}} \cong \pm \omega_{\vec{k}} \left(1 - \frac{i\epsilon}{2\omega_{\vec{k}}^2} \right) \quad (4.142)$$

using the fact that ϵ is infinitesimal. We see that this moves the poles off of the real axis precisely as needed (i.e. the negative and positive poles indeed move off in opposite directions). This way of writing the propagator is often called the **Feynman prescription**.

This completes the derivation of the Feynman propagator. Note that this representation is manifestly Lorentz-invariant, something which was not at all obvious from the definition in terms of time-ordering.

Teaser: I did not actually discuss this in lectures, but as we will see, this propagator has many uses. I will highlight just one of them before we move on to interacting field theories: the Feynman propagator is actually a Green's function of the Klein-Gordon equation, i.e. the PDE:

$$(\partial^2 - m^2)\phi(x) = 0 \quad (4.143)$$

To understand this, let's just directly compute

$$(\partial_x^2 - m^2)G(x, x') = -i(\partial_x^2 - m^2) \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik \cdot (x' - x)}}{k^2 + m^2 - i\epsilon} \quad (4.144)$$

$$= -i \int \frac{d^4k}{(2\pi)^4} e^{-ik \cdot (x' - x)} \frac{-k^2 - m^2}{k^2 + m^2 - i\epsilon} \quad (4.145)$$

$$= +i \int \frac{d^4k}{(2\pi)^4} e^{-ik \cdot (x' - x)} \quad (4.146)$$

$$= +i\delta^{(4)}(x' - x) \quad (4.147)$$

This means that you can use this to find classical solutions to the Klein-Gordon PDE. This also plays a more central role in the path integral formulation of QFT.

5 Interacting quantum field theories

We will now move on to quantum field theories with interactions. In this section we will work mostly with the **real** scalar field.

The free real scalar field has by now the extremely familiar action

$$S = \int d^4x \left(-\frac{1}{2}(\partial\phi)^2 - \frac{m^2}{2}\phi^2 \right) \quad (5.1)$$

We did not carefully canonically quantize it, but it is mostly a simpler version of the story from the complex scalar field. You are asked to work out some of the details in a homework problem, but basically one finds the following Fourier expansion:

$$\phi(\vec{x}, t) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(a_{\vec{k}} e^{-i\omega_{\vec{k}}t + i\vec{k} \cdot \vec{x}} + a_{\vec{k}}^\dagger e^{i\omega_{\vec{k}}t - i\vec{k} \cdot \vec{x}} \right) \quad (5.2)$$

(i.e. we have that the operator formerly called $b = a$: the anti-particle is the same as the particle), where the a 's have the familiar commutation relation

$$[a_{\vec{k}}, a_{\vec{k}'}^\dagger] = (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}') \quad (5.3)$$

and the normal ordered Hamiltonian takes the form

$$: H_0 := \int \frac{d^3k}{(2\pi)^3} \omega_{\vec{k}} a_{\vec{k}}^\dagger a_{\vec{k}} \quad (5.4)$$

We have by now completely solved this quantum theory: we know the spectrum of the Hamiltonian and we worked out the full time-dependence of the fields in the Heisenberg picture, which is given above in (5.2). The physics of this is simply of a set of non-interacting particles – the fact that they are non-interacting is clear from the energy spectrum, which simply counts the energy of each of the particles individually and adds them up.

We now like to move on to the **interacting** quantum field theory: the simplest such theory is called the $\lambda\phi^4$ theory, which has the following action:

$$S = \int d^4x \left(-\frac{1}{2}(\partial\phi)^2 - \frac{m^2}{2}\phi^2 - \frac{\lambda}{4!}\phi^4 \right) \quad (5.5)$$

The action is no longer quadratic in the field ϕ . It will turn out that the term in λ now allows the ϕ particles to interact with each other – the larger λ is, the stronger the interaction. λ is called the *coupling constant*. The first thing to note is that the simple addition of this new term makes the theory **completely unsolvable**; indeed if you can find some way to generally solve interacting quantum field theories you will definitely win a Nobel prize.

Instead, what we will do is Taylor expand the problem in powers of λ : we understand what is happening at $\lambda = 0$, and we will try to extend that understanding to a systematic expansion order by order in λ ; this is called “perturbation theory” and is one of the few tools we have to understand interacting quantum field theory.

Let us first give ourselves a goal: we are interested in computing this sort of quantity:

$$\langle 0|T(\phi(x)\phi(y))|0\rangle \quad (5.6)$$

in the interacting theory. In the free theory, we have already computed this: it is the Feynman propagator $G(x, y)$.

5.1 The interaction picture

So how will we do this? First, let’s us construct the Hamiltonian of the full theory:

$$H = H_0 + H_{int} = H_0 + \int d^3x \frac{\lambda}{4!}\phi(t_0, \vec{x})^4 \quad (5.7)$$

Now the new interaction term affects our calculation in two different ways: there is time evolution in the quantum fields $\phi(x)$, $\phi(y)$ and it also affects the choice of the vacuum state $|0\rangle$, which is no longer the same as the free vacuum.

Let us tackle the time dependence issue first: recall that if we know what the Heisenberg picture quantum field is at some time $t = t_0$, then we can figure out what it is at some later time via the usual expression for Heisenberg time evolution:

$$\phi(t, \vec{x}) = e^{iH(t-t_0)}\phi(t_0, \vec{x})e^{-iH(t-t_0)} \quad (5.8)$$

Now in general, we do not know how to exponentiate the full Hamiltonian; however we *do* know how to exponentiate the free Hamiltonian H_0 ; indeed, the time dependence arising from that Hamiltonian is exactly that which we wrote down previously in (5.2).

Let’s define a field $\phi_I(t, \vec{x})$ to be the field evolved according to only the *free* Hamiltonian, i.e.

$$\phi_I(t, \vec{x}) = e^{iH_0(t-t_0)}\phi(t_0, \vec{x})e^{-iH_0(t-t_0)} \quad (5.9)$$

This is sometimes called the **interaction picture** field. It is not equal to the “real” Heisenberg picture field $\phi(t, \vec{x})$, but if λ is small then there is a sense where it captures the “main” part of the time evolution, and we fully understand it. It is thus a useful thing to use to organize our perturbative expansion: we just need to figure out how to express everything in terms of $\phi_I(x)$.

The “real” Heisenberg picture field is related to ϕ_I by

$$\phi(t, \vec{x}) = e^{iH(t-t_0)}e^{-iH_0(t-t_0)}\phi_I(t, \vec{x})e^{+iH_0(t-t_0)}e^{-iH(t-t_0)} \quad (5.10)$$

where we are undoing the free evolution and then redoing it with the full Hamiltonian. It is now convenient to define the following **evolution operator**:

$$\boxed{U(t, t_0) \equiv e^{iH_0(t-t_0)} e^{-iH(t-t_0)}} \quad (5.11)$$

So that we have

$$\phi(t, \vec{x}) = U^\dagger(t, t_0) \phi_I(t, \vec{x}) U(t, t_0) \quad (5.12)$$

There is some temptation to combine the exponentials above: this would be a mistake, as the interaction part of H does not commute with H_0 (this is indeed the whole point; if they commuted it would be very easy to solve this problem).

We will now develop an expansion of $U(t, t_0)$ in terms of ϕ_I . To do this, let's first note that it satisfies the following differential equation:

$$i \frac{\partial}{\partial t} U(t, t_0) = i \left(\frac{\partial}{\partial t} e^{iH_0(t-t_0)} \right) e^{-iH(t-t_0)} + i e^{iH_0(t-t_0)} \left(\frac{\partial}{\partial t} e^{-iH(t-t_0)} \right) \quad (5.13)$$

$$= i \left(e^{iH_0(t-t_0)} (iH_0) e^{-iH(t-t_0)} + e^{iH_0(t-t_0)} (-iH) e^{-iH(t-t_0)} \right) \quad (5.14)$$

$$= e^{iH_0(t-t_0)} (H - H_0) e^{-iH(t-t_0)} \quad (5.15)$$

$$= e^{iH_0(t-t_0)} H_{int} e^{-iH(t-t_0)} \quad (5.16)$$

$$= e^{iH_0(t-t_0)} H_{int} e^{-iH_0(t-t_0)} e^{+iH_0(t-t_0)} e^{-iH(t-t_0)} \quad (5.17)$$

$$= \left(e^{iH_0(t-t_0)} H_{int} e^{-iH_0(t-t_0)} \right) U(t, t_0) \quad (5.18)$$

Look at the last expression – it is the interaction Hamiltonian H_{int} , evolved in time using only the free Hamiltonian H_0 as in (5.9): we might call it the interaction Hamiltonian in the interaction picture, so we will call it H_I :

$$H_I = e^{iH_0(t-t_0)} H_{int} e^{-iH_0(t-t_0)} \quad (5.19)$$

Let's understand what it is in terms of the interaction picture ϕ_I : we have

$$H_I = \int d^3x \frac{\lambda}{4!} e^{iH_0(t-t_0)} \phi(t_0, \vec{x})^4 e^{-iH_0(t-t_0)} = \int d^3x \frac{\lambda}{4!} \phi_I(t, \vec{x})^4 \quad (5.20)$$

i.e. it is simple when written in terms of the interaction picture field ϕ_I .

So to conclude, the differential equation obeyed by U is

$$i \frac{\partial}{\partial t} U(t, t_0) = H_I(t) U(t, t_0) \quad (5.21)$$

Note also the boundary condition that $U(t_0, t_0) = \mathbf{1}$. This boundary condition and differential equation **uniquely determine** $U(t, t_0)$. We will now explicitly write down a general solution for it.

The solution is called **Dyson's formula**¹³, and is really amazingly slick:

$$\boxed{U(t, t_0) = T \left[\exp \left(-i \int_{t_0}^t dt' H_I(t') \right) \right]} \quad (5.22)$$

Here the T outside the exponential defines the **time-ordered exponential**, which simply means that you should expand out the exponential and then time-order each term in the sum.

¹³Though David Tong's notes tell me that it's actually originally due to Dirac. I believe the slick presentation is due to Dyson.

Let us check that it satisfies the differential equation above: taking the derivative of (5.22) we have

$$i \frac{\partial}{\partial t} U(t, t_0) = i \frac{\partial}{\partial t} T \left[\exp \left(-i \int_{t_0}^t dt' H_I(t') \right) \right] \quad (5.23)$$

$$= T \left[H_I(t) \exp \left(-i \int_{t_0}^t dt' H_I(t') \right) \right] \quad (5.24)$$

$$= H_I(t) T \left[\exp \left(-i \int_{t_0}^t dt' H_I(t') \right) \right] \quad (5.25)$$

Note that though naively nothing in this expression commutes, we could freely take $H_I(t)$ out to the left, because the integral is only over times that are *less* than t , and thus by the magic of time-ordering $H_I(t)$ is always placed to the left of everything else and thus can be taken outside.

Finally, the boundary condition is trivial: clearly we have $U(t_0, t_0) = \mathbf{1}$. Thus, by the uniqueness of solutions to differential equations, we have indeed found that (5.22) is equal to $U(t, t_0)$ defined in (5.11).

Let me state a few properties of $U(t, t_0)$ which you can check for yourselves:

1. $U(t, t_0)$ is unitary.
2. **Composition property:** $U(t_1, t_2)U(t_2, t_3) = U(t_1, t_3)$, if $t_1 > t_2 > t_3$.
3. Taking the Hermitian conjugate evolves backwards in time: $U(t_1, t_3)[U(t_2, t_3)]^\dagger = U(t_1, t_2)$

Next, let's think about the vacuum of the interacting theory $|0\rangle$. Recall that in the free theory we had a vacuum state $|0\rangle_0$ which was defined by the condition

$$H_0|0\rangle_0 = 0 \quad (5.26)$$

This state is empty of particles. Now what can we say about the vacuum of the interacting theory $|0\rangle$?

This vacuum to be defined as follows:

$$|0\rangle = U(t_0, -\infty)|0\rangle_0 \quad (5.27)$$

This says the following: take the vacuum of the free theory in the far distant past. Evolve it forwards up to a reference time t_0 using the interaction Hamiltonian. We **take** this to be vacuum of the interacting theory.

This is somewhat heuristic and I will not actually try to justify it any further, see for example Peskin and Schroeder chapter 4 for more details. Let's just note that the norm of the true vacuum state after this maneuver will not be obviously 1. Thus we can consider computing:

$$\langle 0|0\rangle = {}_0\langle 0|U(+\infty, t_0)U(t_0, -\infty)|0\rangle_0 = {}_0\langle 0|U(+\infty, -\infty)|0\rangle_0 \quad (5.28)$$

which is called the **vacuum persistence amplitude**. It basically measures how much we mis-normalized the free vacuum state. One way to deal with this is simply to divide by it at the end.

We are now ready to write down a formal expression for the time-ordered correlation function $\langle 0|T(\phi(x)\phi(y))|0\rangle$ in the interacting theory. Let us first work this out for $x^0 > y^0 > t_0$. Recall that we convert from interaction picture field to real Heisenberg field by:

$$\phi(t, \vec{x}) = U^\dagger(t, t_0)\phi_I(t, \vec{x})U(t, t_0) \quad (5.29)$$

Placing this in and converting from free vacuum to interacting vacuum we have:

$$\langle 0|T(\phi(x)\phi(y))|0\rangle = \frac{{}_0\langle 0|U(+\infty, t_0)U^\dagger(x^0, t_0)\phi_I(x^0, \vec{x})U(x^0, t_0)U^\dagger(y^0, t_0)\phi_I(y^0, \vec{y})U(y^0, t_0)U(t_0, -\infty)|0\rangle_0}{{}_0\langle 0|U(+\infty, -\infty)|0\rangle_0} \quad (5.30)$$

$$= \frac{{}_0\langle 0|U(+\infty, x^0)\phi_I(x^0, \vec{x})U(x^0, y^0)\phi_I(y^0, \vec{y})U(y^0, -\infty)|0\rangle_0}{{}_0\langle 0|U(+\infty, -\infty)|0\rangle_0} \quad (5.31)$$

$$= \frac{{}_0\langle 0|T(\phi_I(x)\phi_I(y)U(+\infty, -\infty))|0\rangle_0}{{}_0\langle 0|U(+\infty, -\infty)|0\rangle_0} \quad (5.32)$$

It may seem that something miraculous happened in the last line – after all, there are many different U 's evolving us along different times, how did I combine them all? The key fact here is that the time-ordering will automatically ensure that every operator is placed in the right place, because it is all going from the past (in the right) to the future (in the left). You can verify that this construction also works out if $y^0 > x^0$ instead, one finds exactly the same expression.

Now we use Dyson's formula (5.22) to find the final form:

$$\boxed{\langle 0|T(\phi(x)\phi(y))|0\rangle = \frac{{}_0\langle 0|T\left[\phi_I(x)\phi_I(y)\exp\left(-i\int_{-\infty}^{\infty} dt H_I(t)\right)\right]|0\rangle_0}{{}_0\langle 0|\exp\left(-i\int_{-\infty}^{\infty} dt H_I(t)\right)|0\rangle_0}} \quad (5.33)$$

This is the key technical tool in the calculations that follow: because $H_I(t)$ is a polynomial in the field $\phi_I(x)$, basically we can expand out the exponential to any desired order in λ , and we have reduced the problem to computing expectation values of arbitrary time-ordered products of the field $\phi_I(x)$ in the vacuum.

5.2 Wick's theorem

This is a problem in bookkeeping. Let's first note that we have introduced two kinds of ordering so far: time-ordering $T(\mathcal{O}(x_1)\mathcal{O}(x_2)\dots)$ and normal-ordering $:\mathcal{O}(x_1)\mathcal{O}(x_2)\dots:$. We will now try to establish a relationship between these two kinds of ordering.

Let's start by understanding the product of two fields. Start by decomposing the field $\phi_I(x)$ as:

$$\phi_I(x) = \phi_I^+(x) + \phi_I^-(x) \quad (5.34)$$

where the two components have only annihilation and creation operators respectively, and are:

$$\phi_I^+(x) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} a_{\vec{k}} e^{ik \cdot x} \quad \phi_I^-(x) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{k}}}} a_{\vec{k}}^\dagger e^{-ik \cdot x} \quad (5.35)$$

the point of this is that

$$\phi_I^+(x)|0\rangle_0 = 0 \quad {}_0\langle 0|\phi_I^-(x) = 0 \quad (5.36)$$

Now let us consider the time-ordered product in the case where $x^0 > y^0$:

$$T(\phi_I(x)\phi_I(y)) = (\phi_I^+(x) + \phi_I^-(x))(\phi_I^+(y) + \phi_I^-(y)) \quad (5.37)$$

I want to place this in normal order: to do this we should move the ϕ_I^- to the left-hand side. Expanding out the fields, there is only one commutator we need:

$$T(\phi_I(x)\phi_I(y)) = \phi_I^+(x)\phi_I^+(y) + \phi_I^-(x)\phi_I^+(y) + \phi_I^-(y)\phi_I^+(x) + \phi_I^-(x)\phi_I^-(y) + [\phi_I^+(x), \phi_I^-(y)] \quad (5.38)$$

Now in all the terms except for the commutator, all of the terms are normal ordered: so if we have $x^0 > y^0$, we can write:

$$T(\phi_I(x)\phi_I(y)) =: \phi_I(x)\phi_I(y) : + [\phi_I^+(x), \phi_I^-(y)] \quad x^0 > y^0 \quad (5.39)$$

Now we can do the same thing in the case $y^0 > x^0$: you can verify that everything is the same, except that the final commutator is different:

$$T(\phi_I(x)\phi_I(y)) =: \phi_I(x)\phi_I(y) : + [\phi_I^+(y), \phi_I^-(x)] \quad y^0 > x^0 \quad (5.40)$$

Let us now define some useful notation: the **contraction** of two fields is denoted by a line above them, and is defined to be exactly the combination of commutators that we just obtained:

$$\overline{\phi_I(x)\phi_I(y)} = \begin{cases} [\phi_I^+(x), \phi_I^-(y)] & x^0 > y^0 \\ [\phi_I^+(y), \phi_I^-(x)] & y^0 > x^0 \end{cases} \quad (5.41)$$

Now here is the magical thing: it turns out that this combination of commutators is in fact exactly the Feynman propagator!

$$\overline{\phi_I(x)\phi_I(y)} = G(x, y) \quad (5.42)$$

The easiest way to see this is to just explicitly work out the commutators in terms of the ladder operators and compare to the expression in (4.126).

So we now have a wonderfully simple expression for the time-ordered product of two fields: it is

$$\boxed{T(\phi_I(x)\phi_I(y)) =: \phi_I(x)\phi_I(y) : + \overline{\phi_I(x)\phi_I(y)}} \quad (5.43)$$

This should make sense: consider taking the expectation value of this object in the free vacuum $|0\rangle_0$: the normal-ordered part vanishes, and we obtain the known fact that $T(\phi_I(x)\phi_I(y)) = G(x, y)$. So the point of this is to isolate the part of the correlation function that actually contributes when you take the vacuum expectation value.

We are now ready to introduce the general form of **Wick's theorem**: this tells us how to generalize the formula above to arbitrary numbers of fields, and its statement is simply.

$$\boxed{T(\phi_I(x_1)\phi_I(x_2)\cdots\phi_I(x_n)) = \sum \text{all possible contractions, with uncontracted fields normal ordered}} \quad (5.44)$$

All possible contractions is exactly what it sounds like: you look at all possible ways to draw contractions of the fields. The simplest example after two fields is four, which I will now work out. It's a bit lengthy:

$$T(\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)) =: \phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4) : \quad (5.45)$$

$$+ \overline{\phi_I(x_1)\phi_I(x_2)} : \phi_I(x_3)\phi_I(x_4) : + \overline{\phi_I(x_1)\phi_I(x_3)} : \phi_I(x_2)\phi_I(x_4) : \quad (5.46)$$

$$+ \overline{\phi_I(x_1)\phi_I(x_4)} : \phi_I(x_2)\phi_I(x_3) : + \overline{\phi_I(x_2)\phi_I(x_3)} : \phi_I(x_1)\phi_I(x_4) : \quad (5.47)$$

$$+ : \overline{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)} : + : \overline{\phi_I(x_1)\phi_I(x_2)} : \overline{\phi_I(x_3)\phi_I(x_4)} \quad (5.48)$$

$$+ \overline{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)} + \overline{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)} \quad (5.49)$$

$$+ \overline{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)} \quad (5.50)$$

The first line has zero contractions; each of the terms in the next three lines have one contraction each; and the final two lines are fully contracted. Note that every time you have a contraction, you should replace the contracted fields with the Feynman propagator, even if the fields aren't next to each other. So e.g. this term

$$\overline{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4)} = G(x_1, x_4) : \phi_I(x_2)\phi_I(x_3) : \quad (5.51)$$

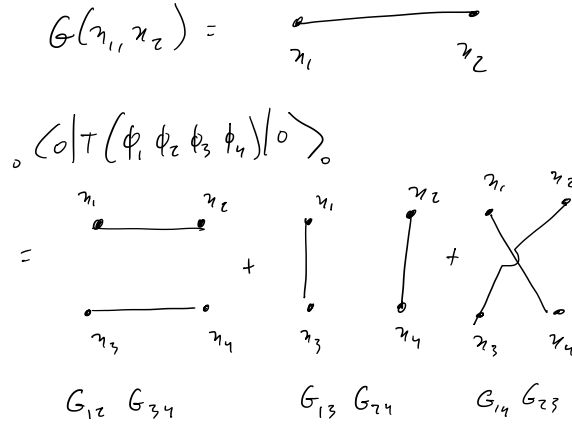


Figure 5.1: Simple Feynman diagrams showing free propagation of particles

and if you have

$$\overbrace{\phi_I(x_1)\phi_I(x_2)}\overbrace{\phi_I(x_3)\phi_I(x_4)} = G(x_1, x_2)G(x_3, x_4) \quad (5.52)$$

I have not proven Wick's theorem; it should sound vaguely plausible given how we worked out the case for two fields. It can be proved by induction, see for example Chapter 4.3 in Peskin and Schroeder.

The point of this all is of course to show you which terms are important if you calculate the expectation value; if you do this, then any term with an uncontracted field vanishes (because it is all normal ordered). So an important corollary to Wick's theorem is:

$${}_0\langle 0|T(\phi_I(x_1)\phi_I(x_2)\cdots\phi_I(x_n))|0\rangle_0 = \sum \text{terms where all fields are fully contracted} \quad (5.53)$$

For example, taking the expectation value of two fields, we have

$${}_0\langle 0|T(\phi_I(x)\phi_I(y))|0\rangle_0 = {}_0\langle 0|\left(:\phi_I(x)\phi_I(y): + \overbrace{\phi_I(x)\phi_I(y)}\right)|0\rangle_0 \quad (5.54)$$

$$= G(x, y){}_0\langle 0|0\rangle_0 = G(x, y) \quad (5.55)$$

and for four, we similarly have

$${}_0\langle 0|T(\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I(x_4))|0\rangle_0 = G(x_1, x_2)G(x_3, x_4) + G(x_1, x_3)G(x_2, x_4) + G(x_1, x_4)G(x_2, x_3) \quad (5.56)$$

Note there is now a temptation to draw **pictures** for these things. You can imagine that $G(x_1, x_2)$ represents a particle moving from x_1 to x_2 . Then the diagrammatic interpretation of the 4-point function is as shown in Figure 5.1 – i.e. one moving from x_1 to x_2 while *another* particle moves from x_3 to x_4 and so on.

This is the beginning of **Feynman diagrams**.

5.3 Feynman diagrams and Feynman rules

This sort of expansion becomes much more fun when we look at the sort of n -point functions that we get from including the actual interaction term in (5.33), which I record here for convenience:

$$\langle 0|T(\phi(x_1)\phi(x_2))|0\rangle = \frac{{}_0\langle 0|T\left[\phi_I(x_1)\phi_I(x_2)\exp\left(-i\int_{-\infty}^{\infty} dt H_I(t)\right)\right]|0\rangle_0}{{}_0\langle 0|\exp\left(-i\int_{-\infty}^{\infty} dt H_I(t)\right)|0\rangle_0} = \frac{N}{D} \quad (5.57)$$

we will call the numerator of this expression N and the denominator D . Now the idea is to expand the interaction term out in powers of λ , i.e. we have

$$\exp\left(-i\int_{-\infty}^{\infty} dt H_I(t)\right) = \exp\left(-i\int d^4x \frac{\lambda}{4!}\phi_I(x)^4\right) = 1 - i\frac{\lambda}{4!}\int d^4x \phi_I(x)^4 + \mathcal{O}(\lambda^2) \quad (5.58)$$

And similarly we can expand out both the numerator and denominator in powers of λ , i.e. formally we have

$$N = N_0 + \lambda N_1 + \lambda^2 N_2 + \dots \quad D = D_0 + \lambda D_1 + \lambda^2 D_2 + \dots \quad (5.59)$$

Let's now start with N . From above we have

$$N = {}_0\langle 0|T\left[\phi_I(x_1)\phi_I(x_2)\left(1 - i\frac{\lambda}{4!}\int d^4x \phi_I(x)^4\right) + \mathcal{O}(\lambda^2)\right]|0\rangle_0 \quad (5.60)$$

the first term in order λ^0 is

$$N_0 = {}_0\langle 0|T(\phi_I(x_1)\phi_I(x_2))|0\rangle_0 = G(x_1, x_2) \quad (5.61)$$

i.e. the result in the free theory. The effect from the interaction appears at $\mathcal{O}(\lambda^1)$:

$$\lambda N_1 = -i\frac{\lambda}{4!}\int d^4x {}_0\langle 0|T[\phi_I(x_1)\phi_I(x_2)\phi_I(x)\phi_I(x)\phi_I(x)\phi_I(x)]|0\rangle_0 \quad (5.62)$$

From Wick's theorem, we know that to determine this expectation value we have to contract these in all possible ways. Everyone should now take a minute and figure out the possible contraction patterns: how many times do we count each pattern? Draw some pictures!

Okay, it has been a minute¹⁴. There are essentially two different topological classes of diagrams. One of them looks like this:

$${}_0\langle 0|T\left[\overbrace{\phi_I(x_1)\phi_I(x_2)}\quad\overbrace{\phi_I(x)\phi_I(x)}\quad\overbrace{\phi_I(x)\phi_I(x)}\right]|0\rangle_0 = 3G(x_1, x_2)G(x, x)^2 \quad (5.63)$$

A pictorial representation of this is shown in Figure 5.2. Note this has a subtle physical interpretation: what is happening is that one particle is moving from x_1 to x_2 , and next to it a particle anti-particle pair has bubbled forth from the vacuum, come together at the intersection point x , and then re-annihilated itself. The probability of this happening is proportional to the coupling λ .

Such things are called **virtual particles**.

The factor of 3 may require some thought: the point is that x_1 and x_2 are special because they are not integrated over – there is only one way to connect the two of them. However the four $\phi_I(x)$'s are all the same. We could have chosen the first contraction in 3 ways (i.e. once you pick the first $\phi_I(x)$, you can connect it to any of the other three), and then the remaining contraction is fixed.

We will call the contribution from this pattern of contractions $\lambda N_1^{\text{disc}}$ since it consists of two disconnected pieces.

¹⁴For you.

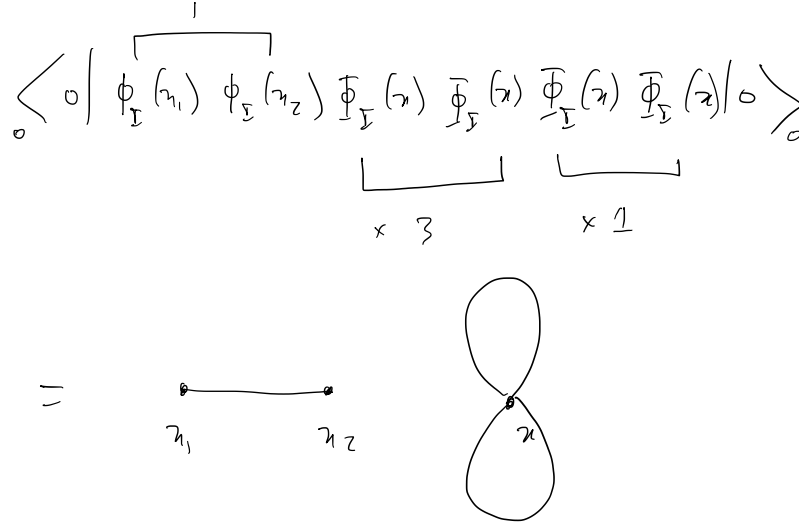


Figure 5.2: Feynman diagrams for N_1^{disc}

The other topologically distinct way to connect them is:

$${}_0\langle 0|T \left[\overbrace{\phi_I(x_1)\phi_I(x_2)\phi_I(x)\phi_I(x)}^{\text{bracket}} \overbrace{\phi_I(x)\phi_I(x)}^{\text{bracket}} \right] |0\rangle_0 = 12G(x_1, x)G(x_2, x)G(x, x) \quad (5.64)$$

We will call this pattern of contractions $\lambda N_1^{\text{conn}}$ since it consists of a single connected piece.

Here the pictorial representation in Figure 5.3 is that a particle is trying to go from x_1 to x_2 , but it spontaneously emits another particle! The particle flits around the universe before being re-absorbed again at the same space-time point x and then the “original” particle moves on to x_2 .

Now the factor of 12: there are four choices for the first connection from the external point x_1 to one of the intersection x ’s: after this there are 3 choices that remain to connect x_2 , and then only 1 choice remains for the last contraction: $12 = 4 \cdot 3 \cdot 1$.

Assembling the pieces we find that the full contribution at first order in λ is

$$\lambda N_1 = -i \frac{\lambda}{4!} \left(3 \int d^4x G(x_1, x_2) G(x, x)^2 + 12 \int d^4x G(x_1, x) G(x_2, x) G(x, x) \right) \quad (5.65)$$

Feynman diagrams are an efficient way to represent the pattern of contractions. The way we will proceed from now on is actually not in the excruciating combinatoric way that we have so far: instead we will think of a set of rules – called **Feynman rules** – that associate an analytic expression in terms of propagators with a given picture, and then we will just draw a set of pictures.

You have by now probably already understood the idea of the **position-space Feynman rules**. If I write them out explicitly for $\lambda\phi^4$ theory, they are as follows:

1. The number of external points is equal to the number of fields in the correlation function we are calculating.
2. For each line connecting two points x, y , write $G(x, y)$.

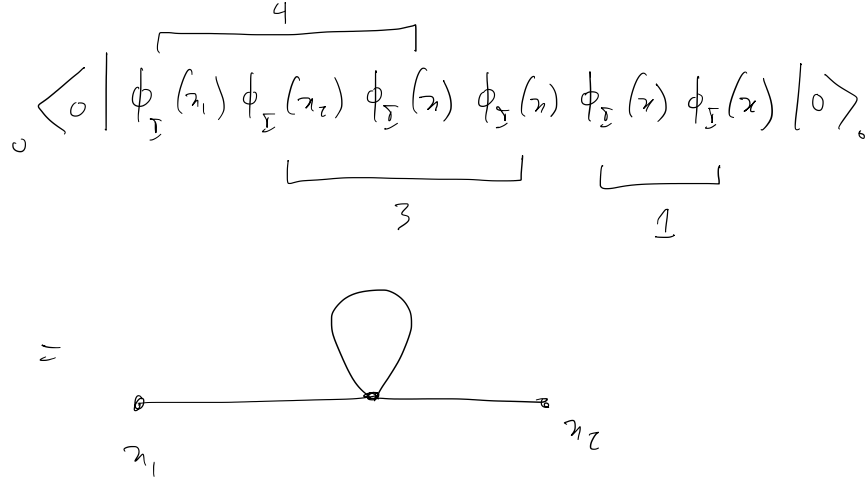


Figure 5.3: Feynman diagrams for N_1^{conn}

3. For each vertex – which necessarily has 4 legs – we write $-i\lambda \int d^4x$.
4. Divide the expression by the **symmetry factor**.

I have called these **position-space** because there is also a momentum-space version, which we will introduce in due course. (It simply involves Fourier-transforming everything).

Here the last point deserves some further explanation. The question is: how many different ways of connecting the lines represent the same analytic expression? Naively every permutation of the 4 lines coming out of each vertex represents an equivalent expression that should be counted; this factor of $4!$ exactly cancels the $4!$ in the denominator of the coupling, so there should be no combinatoric factor at all.

However it turns out that some of these $4!$ permutations are actually not independent; you can make up for them by changing some of the other things around. This will happen when the diagram has some symmetry. The degree of this symmetry is called a **symmetry factor**, and determining it is an interesting exercise in pure thought. For example, in the connected bubble diagram in Figure 5.4 you can simultaneously switch the two legs of the propagator and swap the lines coming out of the vertex, resulting in a symmetry factor of 2. In the figure-of-eight diagram in Figure 5.4 we have instead a symmetry factor of 8, as shown. This means that we have **overcounted**, and we should divide the expression by the symmetry factor.

This results in the same combinatoric factors as above:

$$\frac{3}{4!} = \frac{1}{8} \quad \frac{1}{2} = \frac{12}{4!}. \quad (5.66)$$

Assembling the pieces from the Feynman rules, we see that we arrive at the same expressions as in (5.65). We could do precisely the same thing at higher order in λ – we would just have more than one vertex. To obtain the right answer at a given order we just have to make sure we have drawn all possible diagrams.

Now let us come to the denominator D in (5.57), i.e.

$$D \equiv D_0 + \lambda D_1 + \cdots = {}_0\langle 0 | \exp \left(-i \int_{-\infty}^{\infty} dt H_I(t) \right) | 0 \rangle_0 = {}_0\langle 0 | \left(1 - i \int d^4x \frac{\lambda}{4!} \phi_I(x)^4 + \cdots \right) | 0 \rangle_0 \quad (5.67)$$

Clearly $D_0 = {}_0\langle 0 | 0 \rangle_0 = 1$. We could determine D_1 through Wick contractions. However let's use our fancy new Feynman diagrams – here we are computing the vacuum persistence amplitude with no external points,

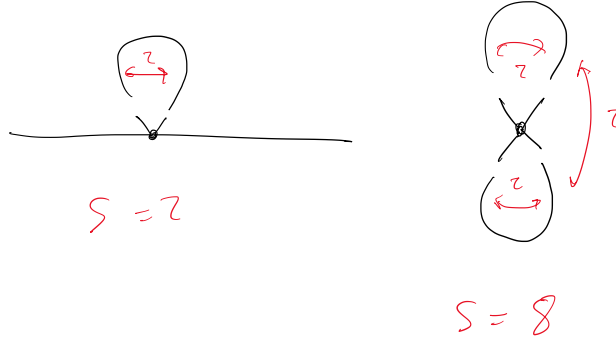


Figure 5.4: Computation of symmetry factor S for the two diagrams above.

and so we should simply draw all Feynman diagrams with only closed loops. We find at order $\mathcal{O}(\lambda)$ as in Figure 5.5:

$$\lambda D_1 = \frac{-i\lambda}{8} \int d^4x G(x, x)^2 \quad (5.68)$$

It is quite possible to compute higher order terms – you simply apply the Feynman rules above. We can do this, but for now let's work out the full expression, which involves dividing the numerator by the denominator:

$$\langle 0|T(\phi(x_1)\phi(x_2))|0\rangle = \frac{N}{D} = \frac{N_0 + \lambda N_1 + \dots}{1 + \lambda D_1 + \dots} = N_0 + \lambda(N_1 - D_1 N_0) + \mathcal{O}(\lambda^2) \quad (5.69)$$

Now let's assemble all of the pieces: we find

$$N_0 = G(x_1, x_2) \quad \lambda N_1 = \lambda N_1^{\text{disc}} + \lambda N_1^{\text{conn}} = -i\lambda \left(\frac{1}{8} \int d^4x G(x_1, x_2) G(x, x)^2 + \frac{1}{2} \int d^4x G(x_1, x) G(x_2, x) G(x, x) \right) \quad (5.70)$$

Similarly we have

$$\lambda D_1 N_0 = \frac{-i\lambda}{8} G(x_1, x_2) \int d^4x G(x, x)^2 = \lambda N_1^{\text{disc}} \quad (5.71)$$

We find in the end, up to $\mathcal{O}(\lambda)$:

$$\boxed{\langle 0|T(\phi(x_1)\phi(x_2))|0\rangle = G(x_1, x_2) - i\frac{\lambda}{2} \int d^4x G(x_1, x) G(x_2, x) G(x, x) + \dots} \quad (5.72)$$

Note that the term in $D_1 N_0$ from the denominator exactly cancels the *disconnected* diagram from the numerator! This is capturing the physical idea that the disconnected diagram in the numerator – called a **vacuum bubble** diagram, as it has no external legs – is actually capturing the physics of how the vacuum shifts due to the interaction, and has nothing to do with the propagation of the particle itself. Indeed it isn't hard to show that this pattern of cancellations actually holds to all orders in perturbation theory: the role of the denominator is simply to cancel all of the vacuum bubbles. (Remember I told you that this slightly ad-hoc prescription for adjusting the vacuum normalization would make sense?).

One thus ends up with the following extremely nice pictorial formula for the time-ordered correlation function

$$\langle 0|T(\phi(x_1)\phi(x_2))|0\rangle = \sum \text{All Feynman diagrams with 2 external points and no vacuum bubbles} \quad (5.73)$$

$$= -\frac{i\lambda}{8} \int d^4x G(x,x) G(x,x)$$

Figure 5.5: Diagram contributing to D_1

The extension to arbitrary numbers of external fields is clearly

$$\langle 0|T\phi(x_1)\phi(x_2)\cdots\phi(x_n)|0\rangle = \sum \text{All Feynman diagrams with } n \text{ external points and no vacuum bubbles} \quad (5.74)$$

This is the main idea of perturbative quantum field theory.

A brief philosophical interlude: the Feynman calculus for perturbation theory is nice for two reasons:

1. It gives a simple algorithm to work out all the terms in the perturbative expansion – just draw some pictures and apply the rules.
2. It gives a physical interpretation to each mathematical term: they represent (possibly virtual) particles running around and scattering off of each other.

You may now want to go further: for example, why don't we *evaluate* the integral over x in (5.72), to get an explicit function of x_1, x_2 ? It is quite possible to do this, but the reason I'm not doing it right now is that it is actually quite infinite. In fact, almost all integrals that involve “loops” of virtual particles are infinite – the infinity in this particular diagram arises from the fact that $G(x, x)$ represents the probability for a particle to end up exactly where it started (i.e. at point x), which is very sensitive to the short-distance structure of the theory. Dealing with this is the idea of *renormalization*, and it is actually where the real fun begins. If you are curious about this the books recommended at the start of the course are a good starting point.

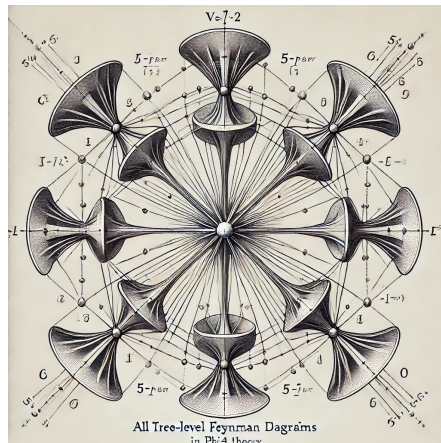


Figure 5.6: ChatGPT impression of Feynman diagrams.

5.4 Wrapping up

I want to conclude by reminding you of the list of things that I said that we would learn back on page 6, and checking whether we have indeed learned them or not.

1. *There are different, yet totally indistinguishable copies of elementary particles like the electron.*

We have learned this, at least for scalar particles: it has to do with the fact that a particle is created by an operator $a_{\vec{p}}^\dagger$ which commutes with other creation operators.

2. *There is a relationship between the statistics of particles (i.e. the behavior of their exchange) and their spin (i.e. what happens if you rotate them).*

This will require further reading, because we did not study spin yet. Sorry.

3. *Anti-particles exist.*

Yep: if we have a complex scalar field there are two different creation operators $b^\dagger$ and $a^\dagger$ with opposite charge. For the real scalar field it makes sense to think of the particle as its own antiparticle.

4. *Particles can be created and destroyed.*

Okay this one we have really studied to death. The Hilbert space of quantum field theory does not have a fixed particle number, and we saw that this has something deep to do with relativity.

5. *Finally, and perhaps most profoundly: the things that we call forces can be imagined as being caused by the exchange of particles.*

See Peskin & Schroeder for more on this important point !

The End