Introduction To Field Theory

Course delivered in 2019 by

Professor Michael Spannowsky

Durham University

Notes taken by

Richie Dadhley

richie.s.dadhley@durham.ac.uk

Acknowledgements

These are my notes on the 2019 lecture course "Introduction to Field Theory" taught by Professor Michael Spannowsky at Durham University as part of the Particles, Strings and Cosmology Msc. For reference, the course lasted 4 weeks and was lectured over 24 hours.

I have tried to correct any typos and/or mistakes I think I have noticed over the course. I have also tried to include additional information that I think supports the taught material well, which sometimes has resulted in modifying the order the material was taught. Obviously, any mistakes made because of either of these points are entirely mine and should not reflect on the taught material in any way.

I have also used/borrowed from Professor David Tong's notes throughout the course (particularly at the start). This notes are a must for any new students to QFT and can be found via

http://www.damtp.cam.ac.uk/user/tong/qft/qft.pdf

The material is presented in the order taught. This generally coincides with Professor Tong's notes, however there are small deviations here and there. Probably the biggest difference is the fact that we present the LSZ theorem in this course, whereas that is saved for the Advanced QFT course at Cambridge. Equally, normal ordering is not introduced until lecture 7, where Wick's theorem is discussed, whereas Prof. Tong introduces it much earlier (in order to drop the infinities in expectation values).

I would like to extend a message of thanks to Professor Spannowsky for teaching this course. I would also like to thank Professor Tong for uploading his notes to the internet.

If you have any comments and/or questions please feel free to contact me via the email provided on the title page.

For a list of other notes/works I have available, visit my blog site

www.richiedadhley.com

These notes are not endorsed by Professor Spannowsky or Durham University.

Richie Dadhley

Contents

1 | Introduction & Motivation

1.1 The Battle Between Quantum Mechanics & Special Relativity

The first question we should ask ourselves is "why does quantum field theory (QFT) even exist and what is it used for?" The answer to this question comes from the fact that special relativity and quantum mechanics (QM) don't get along, and we need some method to make them consistent. Let's see why.

1.1.1 Causality

In the pre-Einstein ages, we often just had formulae for the forces between objects and just got on with using them. For example, let's say we have two particles of charges q_1 and q_2. Let's plonk them down somewhere and label their separation by d. Then the rules of electromagnetism tell us that the force between these two particles is given by

$$F_{EM} \propto \frac{q_1 q_2}{d^2}. \tag{1.1}$$

We were happy with this, until Einstein came along and told us about a little thing called *causality*. This is basically the statement that information cannot travel faster than the speed of light. In the language of special relativity that is 'information cannot travel between two spacelike separated points'. This wreaks havoc for Equation (1.1). Why? Well because there is no time dependence! Therefore if we were to move one of the charges, q_1 say, then the value of d would change, and Equation (1.1) tells us that q_2 would feel this change in force *instantly*. In other words, the information contained in the fact that we have moved q_1 is transmitted to q_2 (which could be arbitrarily far away) instantaneously; not good! We can fix this problem if instead we think of the force being mediated by some field that travels at the speed of light, and this is exactly what we do.

Remark 1.1.1. Note that when q_1 and q_2 are close by, the fact that light travels so fast means that the difference between the result obtained by Equation (1.1) and the field theory description will be negligible. This is the usual story of non-relativistic vs. relativistic theories.

Ok, that's all very well and good, but what on earth does that have to do with QM? Well in quantum mechanics we have operators which act on the states of our Hilbert space and do something to them. This is information! If we couple this idea with the fact that we want to be able to act with operators at any point in space and time[1] we see that we are in a very similar situation to the electromagnetism case.

[1] From now on I shall only say spacetime, and not 'space and time' as the latter tends to promote the idea of thinking of space and time as different things.

Recall that one of the weird and wonderful things about QM is that it is not always possible to simultaneously measure two observables. Mathematically we write this as the statement that their commutator is non-vanishing:

$$[\hat{\mathcal{O}}_A, \hat{\mathcal{O}}_B] \neq 0$$

for non-compatible operators. This essentially tells us that the action of $\mathcal{O}_A$ on our state speaks to $\mathcal{O}_B$ and alters its action and visa versa. However if we are going to preserve causality, it is important that two spacelike separated operators must not speak to each other. In terms of our commutator relations, this is[2]

$$[\hat{\mathcal{O}}_A(x), \hat{\mathcal{O}}_B(y)] = 0 \qquad \text{if } (x-y)^2 < 0, \tag{1.2}$$

for *any* two operators $\mathcal{O}_A$ and $\mathcal{O}_B$. To emphasise, the above condition holds even if $\mathcal{O}_A$ and $\mathcal{O}_B$ are non-compatible.

The collection of observables form a quantum field in spacetime and signals propagate as disturbances in this field. We interpret these disturbances as *particles*.[3]

1.1.2 Multiparticle States

Another problem that arises when trying to marry QM with special relativity is the conservation of particle number. In regular quantum mechanics, we work hard to define our Hilbert space[4] and then we simply act on it. The game is a bit different in QFT, though.

Recall that in QM we have the famous *Heisenberg uncertainty relation*, which tells us that you cannot know both the position and momentum to arbitrary accuracy, and that the lower bound is

$$\Delta L \Delta \mathbf{p} \geq \frac{\hbar}{2} \qquad \Longrightarrow \qquad \Delta \mathbf{p} \geq \frac{\hbar}{2\Delta L}. \tag{1.3}$$

This tells us that if we constrain a particle to a small region of space (say by putting it into a really small box) then the uncertainty in momentum goes through the roof.

Recall separately that special relativity tells us that energy and mass are essentially the same thing $E = mc^2$, and that in a general frame the energy of a particle is given by its mass and 3-momentum:

$$E^2 = m^2c^4 + \mathbf{p}^2c^2.$$

If we consider a very light particle, this relation can be approximated as

$$E \approx |\mathbf{p}|c \tag{1.4}$$

Now we have a problem! Imagine constraining the particle into a box of size $L = \hbar/2mc$ then Equation (1.3) tells us that the uncertainty in the momentum is

$$\Delta \mathbf{p} \geq mc,$$

[2] We are working with a field theorist signature of $(+,-,-,-)$.

[3] On a technical note, the particle interpretation in QFT is actually much more subtle. A global notion of particle requires a global notion of time. As we are working in Minkowski spacetime we of course have a global time coordinate and so this is not a problem. However for a curved spacetime things become more subtle and you get some highly non-trivial results, namely that black holes have temperature; Hawking Radiation.

[4] For *a lot* more info on this see Simon Rea and my joint notes on Dr. Frederic Schuller's Quantum Theory course.

and then Equation (1.4) tells us that the uncertainty in energy is

$$\Delta E \geq mc^2.$$

This is enough energy to create a new particle!

So we see that in QFT we need to adapt our Hilbert space so that it accounts for these multiparticle states. This construction is known as a Fock space, and mathematically can be written as

$$\mathcal{F} = \bigoplus_{n=0}^{\infty} \mathcal{H}^{\otimes n}, \tag{1.5}$$

where $\mathcal{H}$ is our 'one particle state'. To make the above seem a little less abstract, the two particle state is an element of the space $\mathcal{H}^{\otimes 2} := \mathcal{H} \otimes \mathcal{H}$, and so each n above corresponds to allowing n-particle states. The $\oplus$ symbol just means we need to 'sum over' these different combinations to account for the fact that our system can transition from a n-particle state to a $(n+1)$-particle state.[5]

Remark 1.1.2. Anyone familiar with the basic idea of QFT will know that what we never really just create a single particle in this way, but instead they always appear is so-called *pair production*. This is essentially just the statement that a particle is always produced with its antiparticle, but we'll return to this later.

QFT will allow us to study how such particle creation processes occur and interpret them in terms of interacting field. Being a quantum theory, what we will obtain is the *probability* that such a process occurs, known as *cross sections*.

1.1.3 Fields, Not Particles

The above argument should actually give us a reasonably intuitive idea to why we want to consider fields instead of particles. We have just seen that even in the simple case of Minkowski spacetime where the notion of a particle makes sense (see footnote 3), the notion of how many particles we have is very blurred. This is not something that lends well to our understanding, and so we would really like to think of things differently, and this is where the fields come in. As mentioned above, the fields are the fundamental objects and we obtain a particle interpretation by the disturbances in the field. To quote Weinberg:

"Quantum fields are the basic ingredients of the Universe, particles are just bundles of energy and momentum made out of them."

So QFT theory is, as the name suggests, a study of *fields*, not particles. This might seem like a pedantic point to make, however this switch in mentality is actually very important for a solid understanding of the material that is about to come.

[5] For more info on the symbols used, see either the QM notes mentioned in footnote 4 or any decent linear algebra textbook.

1.2 Units & Range

1.2.1 Units

Experienced physicists will know that physical units (or dimensions) are important, and must be kept track of. As you may be aware, it turns out that *all* physical units can be characterised by seven others, known as SI units. The three most commonly encountered (and the only ones used in this course) are mass, length and time. We write the dimensions of a quantity by square brackets and use the letters $[M]$, $[L]$ and $[T]$ for the three units, respectively.

The constants that appear in our equations will, of course, sometimes carry units. The three dimensionful constants in particle physics are: the reduced Planck's constant, $\hbar$; the speed of light c; and Newton's gravitational constant G_N. These have dimensions

$$[c] = \frac{[L]}{[T]}, \qquad [\hbar] = \frac{[M][L]^2}{[T]}, \quad \text{and} \quad [G_N] = \frac{[L]^3}{[M][T]^2}.$$

It is the former two that will show up a lot (as they already have above) and particle physicists decided one day they had had enough of writing them out and so decided to switch to a different set of units, known as *natural units*. These are defined by setting $\hbar = 1 = c$. From the above relation this results in

$$[L] = [T] = [M]^{-1}.$$

We can therefore just characterise everything in terms of one of these three units. It is standard to give this privilege to mass, and we obtain *mass dimensions*. These are just numbers, and correspond to the power of $[M]$ appearing. For example

$$[G_N] = [M]^{-3}[M]^{-1}[M]^2 = -2.$$

Now recall Equation (1.3). This tells us that position (which has length dimension) has the opposite dimension to 3-momentum. Then Equation (1.4) tells us that energy has the same dimension as 3-momentum. We therefore have

$$[E] = 1. \tag{1.6}$$

This result is also trivially obtained from $E = m$ (as $c = 1$). We can therefore express these dimensions in units of energy. Due to the energy scales that appear in particle physics, we do not choose to use the Joule but instead use the much more convenient electron volt, eV. These often appear with the prefixes M, G, T for 'mega', 'tera' and 'giga', respectively.

Example 1.2.1. As an example, the mass of an electron in these units is

$$m_e = 0.511 MeV.$$

1.2.2 Range

Perhaps an obvious question to now ask is "what energy scale does QFT apply to?"

1.3 Classical Particle Mechanics

In order to make the discussion of QFT easier, let's first look at classical field theory.

1.3.1 Configuration Space & Generalised Coordinates

We can describe the configuration of a system by a set of parameters. For example, the position of a particle on $\mathbb{R}^2$ is simply given by its coordinates (x, y), say. Perhaps a less obvious example is the colour of something. This is described by its RGB value, and so is parameterised by 3 numbers. We can think of these parameters as coordinates for our system. This coordinate space need not be continuous, as it would be for both the above examples, but could form a lattice. For example if we whether something is 'right way up' or 'upside down' would form two points on a configuration space.

The question is, do constraints on the system effect the parameters, and therefore effect the coordinate systems? The answer is yes. To see why lets contrast a free particle with one constrained to some surface, a circle, say. The free particle can have any position and so we need to consider the entirety of $\mathbb{R}^2$, that is it can have any (x, y) values. The constrained particle, however, can only take values such that $x^2 + y^2 = r^2$. If instead we used polar coordinates, we could completely specify the particle's position just by its angle around the circle. So this constraint has reduced the number of parameters needed, and thus altered our configuration space. A similar example for the colour of an object would be if we fixed the R value but allowed G and B to vary still.

Note all of these constructions were made relative to some 'reference point': for the particle's position we reference the values to the origin; for the colours we reference it to white, which has RGB=000.

Definition. [Generalised Coordinates] We call the minimum number of independent parameters needed to describe a system relative to some reference point the **generalised coordinates**. We often denote these with a $\{q_a\}$.

We use the name 'generalised' coordinates, because, as we have seen, we don't need the same number of generalised coordinates as we do 'real' coordinates (i.e. those used to define a coordinate system). Note that the configuration space is given by the span of the generalised coordinates.

Example 1.3.1. The generalised coordinates for the free particle are (x, y), whereas for the constrained particle we just have one, θ.

Example 1.3.2. The values (q_1, q_2, q_3) could be the generalised coordinates for either a single particle in 3-dimensions or for 3 particles in 1-dimension. This example highlights that in order to differentiate two systems, we need to know how the configuration space is constructed.

1.3.2 Principle Of Least Action & Euler-Lagrange

The configuration space tells us the configuration of a system, but in physics we also want to know about the *dynamics* of a system. That is, we want to know how the configuration varies through time. For the case of a particle's position, this corresponds to knowing its momentum. In the language of generalised coordinates, we call these momenta the *generalised velocities*. We often denote these by $\{\dot{q}_a\}$. The span of $\{q_a\} \cup \{\dot{q}_a\}$ is called the *phase space*. Points in the phase space correspond to *states* of the system, and the evolution of states are given by paths on phase space.

In classical physics, the dynamics of a particle is given by a phase space where q_a and p_i are the position and momentum of the particle. For non-accelerating particles, we know that we can simply specify the position and momentum of a particle at some point and from that extract out its evolution. In other words, each point in phase space only has one path going through it, and so by specifying the state of a system we know its evolution.

The question is "how do we find this evolution?" The answer comes from the *principle of least action*. These tell us the equations of motion for a system (i.e. how it evolves through time) and so will give us exactly what we want.

The procedure is as follows: we obtain the *Lagrangian* L for our system in terms of q and $\dot{q}$, integrate it from initial to final time to obtain the *action*, then we extremise this action under the constraint that the start and end points remain fixed, and obtain the shortest path between those points. The calculation is as follows:

$$\begin{aligned}
\delta S &= \int_{t_i}^{t_f} dt \delta L\big(q_a(t), \dot{q}_a(t)\big) \\
&= \int_{t_i}^{t_f} dt \left(\frac{\partial L}{\partial q_a} \delta q_a + \frac{\partial L}{\partial \dot{q}_a} \delta \dot{q}_a \right) \\
&= \int_{t_i}^{t_f} dt \left[\frac{\partial L}{\partial q_a} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a} \right) \right] \delta q_a + \left[\frac{\partial L}{\partial \dot{q}_a} \delta q_a \right]_{t_i}^{t_f} \\
&= \int_{t_i}^{t_f} dt \left[\frac{\partial L}{\partial q_a} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a} \right) \right] \delta q_a,
\end{aligned}$$

where we have used

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a} \delta q_a \right) = \frac{\partial L}{\partial \dot{q}_a} \delta \dot{q}_a + \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a} \right) \delta q_a,$$

integrated by parts, and used the fact that $\delta q_a = 0$ at t_i and t_f. Now setting the variation to zero, and using the fact that δq_a was arbitrary (up the constraints) we can conclude

$$\frac{\partial L}{\partial q_a} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a} \right) = 0, \tag{1.7}$$

known as the *Euler-Lagrange equations*. These are exactly the equations of motion of the system we set out to obtain.

Practically speaking, the Lagrangian is given by the kinetic energy, T, minus the potential energy, U. We shall now give a few examples to show that Equation (1.7) does indeed give us the expected equations of motion.

Example 1.3.3. Consider the motion of single particle in 1-dimension with a potential $U(x)$. The Lagrangian is

$$L(x, \dot{x}) = \frac{1}{2} m \dot{x}^2 - U(x),$$

and so out Euler-Lagrange equations are:

$$m\ddot{x} = -U'(x),$$

which is exactly Newton's law of motion.

1.4 Lagrangian Field Theory

The above example was to do with a single *particle*, but what we're trying to study here are *field* theories. We therefore need to adapt the above procedure to this case.

The first change we need to make is to the generalised coordinates. For the particle case, a was an index which labelled the different $q(t)$s. However in field theory, we must also introduce a label for the position spatial of the field, and so our generalised coordinates are replaced by

$$q_a(t) \longrightarrow \phi(x) := \phi_a(\vec{x}, t),$$

where $\vec{x}$ is a treated as a label. Note that for the field, we are dealing with a system with an infinite number of degrees of freedom: one for each value of $\vec{x}$.

We can also show by considering a chain of connected springs (Do this later from his notes)[6] that we must also replace the Lagrangian with a Lagrangian *density*, $\mathcal{L}$, related by

$$L(t) = \int d^3\vec{x}\, \mathcal{L}\big(\phi_a(x), \partial_\mu \phi_a(x)\big). \tag{1.8}$$

Our Euler-Lagrange equations then become

$$\frac{\partial \mathcal{L}}{\partial \phi_a} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \right) = 0 \tag{1.9}$$

The symmetries of the physics are inserted at the Lagrangian level, by which we mean if we want a theory that is spatial-rotationally invariant then we require

$$R(\varphi_x, \varphi_y, \varphi_z)\mathcal{L} = \mathcal{L},$$

where $R(\varphi_x, \varphi_y, \varphi_z)$ is a general spatial rotation.

In this course we will only consider Lagrangians that obey the following conditions

(i) Lorentz invariance; i.e. $\mathcal{L}(\Lambda x) = \mathcal{L}(x)$ where Λ is a general Lorentz transformation,

(ii) Local Lagrangians; i.e. all of our fields will be evaluated at the same event in spacetime,

(iii) Only contain single derivatives of the fields; i.e. we don't want anything of the form $\partial_\mu \partial_\nu \phi$,

(iv) Renormalisable Lagrangians; this is something that will not really be covered explained in this course, but basically the coefficients in our Lagrangian must have positive dimension. This essentially amounts to not having more than d powers of the field, where d is the dimension of the spacetime.

These four conditions will actually massively reduce the number of possible Lagrangians we can consider. We now give a brief justification for each:

(i) We are considering relativistic particles and so we really want our theories to be Lorentz invariant,

[6] I plan to come back and fill this in later. Just right now don't think it's worth the time when I have so many other courses to be focusing on.

(ii) This massively helps with avoiding non-causal theories,

(iii) There are two reasons: one is that we would need to alter our Euler-Lagrange equations to account for higher derivatives, and also adapt our phase space[7]; secondly because we are fundamentally interested in 4-dimensional spacetime, a Lagrangian that included a $\partial_\mu\partial_\nu\phi$ term would be non-renormalisable. This comes from the fact that the action has vanishing mass dimension[8] and so, in order to cancel the $[d^dx] = -d$[9] contribution from the integral we require $[\mathcal{L}] = d$. Now if we want to study kinematics, we need a kinetic energy term in the Lagrangian, this takes the form

$$\mathcal{L} = (\partial_\mu\phi)(\partial^\mu\phi) + ...,$$

but the derivatives have dimension $[\partial_\mu] = 1$[10] and so we get $4 = 2 + 2[\phi]$, or $[\phi] = 1$. If we were then to include a second order derivative term, Lorentz invariance would require it appears as

$$\mathcal{L} = a(\partial_\mu\partial_\nu\phi)(\partial^\mu\partial^\nu\phi),$$

but this would then require $[a] = -2$, which gives a non-renormalisable theory.

(iv) We don't need to understand this here, but basically it causes a huge pain in perturbation theory.

Notation. As we have just seen in the explanation of (iii) above, Lorentz invariance requires that any derivative term comes with the corresponding 'raised' derivative term. It is therefore very useful to introduce a common notion to simply the expressions. We define

$$(\partial_\mu\phi)^2 := \partial_\mu\phi\partial^\mu\phi = \eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi.$$

This notation can appear a bit confusion because it appears to break Einstein summation convention (the left-hand side seems to have a μ index but on the right-hand side it is summed over). For this reason some people (and I am one of these people) prefer to define the notation

$$(\partial\phi)^2 := \partial_\mu\phi\partial^\mu\phi. \tag{1.10}$$

This is the notation that will be used in these notes from now on, unless some potential confusion might arise, in which case we shall write the full expression (i.e. like the right-hand side above). Similarly we shall use

$$\partial^2 := \partial_\mu\partial^\mu. \tag{1.11}$$

Example 1.4.1. One of the simplest Lagrangians that meets the above conditions is

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 \\ &= \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}\nabla^2\phi - \frac{1}{2}m^2\phi^2,\end{aligned} \tag{1.12}$$

[7] Basically you get equations of motion that contain second order derivatives, and so it is not enough to just specify q and $\dot{q}$ on an initial time surface, but we must also specify $\ddot{q}$.

[8] Basically because $S \sim \hbar$.

[9] Note that integration increases the dimension and so has units of length.

[10] They take away length dimension.

where $\phi = \phi(x)$ is a real scalar field. The Euler-Lagrange equations for this are simply

$$(\partial^2 - m^2)\phi = 0.$$

This is known as the *Klein-Gordan equation*. It is very important to note here that this is the *classical* Klein-Gordan equation, not the quantum version. As we will see later it has an quantum analogue (which looks exactly the same) which gives the equations of motion of spineless particles, i.e. quantum scalars.

Example 1.4.2. The Klein-Gordan equation above takes the form of a wave equation (i.e. it is second order in both temporal and spatial derivatives). This is nice,but we know in QM that the wave equation is replaced by the Schrödinger equation, which is linear in time derivatives and quadratic is spatial derivatives. We can get a *classical* analogue to this by considering a complex scalar field ψ. The Lagrangian is

$$\mathcal{L} = \frac{i}{2}(\psi^*\dot{\psi} - \dot{\psi}^*\psi) - \nabla\psi^* \cdot \nabla\psi - m\psi^*\psi.$$

You treat ψ and ψ^* as independant fields. For the ψ^* field we have,

$$\frac{\partial\mathcal{L}}{\partial\psi^*} = \frac{i}{2}\dot{\psi} - m\psi, \qquad \frac{\partial\mathcal{L}}{\partial\dot{\psi}^*} = -\frac{i}{2}\psi, \quad \text{and} \quad \frac{\partial\mathcal{L}}{\partial\nabla\psi^*} = -\nabla\psi,$$

and so the equation of motion for this field is

$$i\frac{\partial\psi}{\partial t} = -\nabla^2\psi + m\psi.$$

The ψ field gives a similar equation of motion.

This *looks* a lot like the Schrödinger equation, but as was emphasised above this is a *classical* theory. It does not have any of properties of the Schrödinger equation, e.g. it doesn't have a probabilistic interpretation, as is the key to quantum theory.

2 | Symmetries & Hamiltonian Field Theory

2.1 Lorentz Invariance

As we explained last lecture, we want to consider theories that are Lorentz invariant. Will therefore be testing for this often and so we want to clarify the notion and conventions here.

As in (i) last lecture, we denote a general Lorentz transformation via the symbol Λ. We think of Lorentz transformations as acting on the spacetime coordinates $\{x^\mu\}$[1] and so it is useful to express them as matrices. We therefore have

$$x^\mu \longrightarrow \widetilde{x}^\mu = \Lambda^\mu{}_\nu x^\nu$$

as our transformation. We also have the condition

$$\Lambda^\mu{}_\tau \eta^{\tau\sigma} \Lambda^\nu{}_\sigma = \eta^{\mu\nu}. \tag{2.1}$$

Lorentz transformations are essentially spatial rotations and boosts, and we can work out the entries of the matrix $\Lambda^\mu{}_\nu$ by consider the action on the $\{x^\mu\}$. We give an example of each below.

Example 2.1.1. Let's set $(x^0, x^1, x^2, x^3) = (t, x, y, z)$. First let's consider a rotation around the z axis. This does nothing to the t axis and obviously doesn't effect z. We therefore just set the SO(2) rotation in the xy-plane by angle θ. In other words we have

$$\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \longrightarrow \begin{pmatrix} t \\ x\cos\theta - y\sin\theta \\ y\sin\theta + x\cos\theta \\ z \end{pmatrix},$$

and so we can conclude[2]

$$(\Lambda^\mu{}_\nu) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

[1] In fact this interpretation is a bit misleading, but that's not important here. For an explanation of what I mean, see my notes on Dr. Schuller's Winter School on Gravity and Light, Section 13.3.

[2] In these notes I shall try and be consistent and put brackets around indexed objects to indicate that we mean the matrix.

Example 2.1.2. Following a similar idea to the previous example, we can show that for a boost along the x axis we have

$$(\Lambda^{\mu}{}_{\nu}) = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $\gamma = 1/\sqrt{1-v^2}$.

2.1.1 Active vs. Passive Transformations

There is a subtlety in taking transformations that can be wonderfully confusing the first time you hear it, and this is the difference between an active and passive transformation. This short section just aims to clear up the differences so that the notation that follows in these notes is not confusing.

Let's say we want shift the temperature field in a room to the right. That is the temperature your left-hand neighbour is currently feeling you will feel after the transformation. There are essentially two ways to achieve this: firstly we could actually move the air particles themselves and so physically move the temperature to the right; secondly we could leave the air where it is and instead we, the people,[3] could move to the left. In both cases you will end up feeling the temperature your left-hand neighbour previously felt, the question is "what's the difference?"

The answer is that the former is an *active* transformation whereas the latter is a *passive* one. Put in a more mathematically meaningful way, an active transformation is one where the thing itself (in this case the temperature field) is moved, whereas a passive transformation is one where we simply shift the underlying reference frame/coordinate system (in this case, us).

Those passionate about the concepts of relativity should now throw their arms up in protest. Why? Well because a passive transformation clearly depends on the choice of coordinate system and so should not be a physical thing. In contrast, the active transformation makes perfect sense without reference to any coordinate system. Put in terms of the example above, the passive transformation only makes sense if we are in the room and move ourselves and explain the shift, whereas the active transformation makes perfect sense even if we're not in the room at all.

It is for this reason that we actually consider active transformations in field theory. That is we actually consider shifting the *fields* in our problems and not the coordinate system. As we have just seen, when we *measure* something (i.e. we personally measure the change in temperature), we can think of the active transformation as a passive one going in the other direction (i.e. we move left so that 'the temperature moves right'). This is why in what is to follow we shall write the transformation on fields as follows

$$\phi(x) \longrightarrow \widetilde{\phi}(x) = \phi\big(\Lambda^{-1}x\big). \tag{2.2}$$

The above equation is actually a particular case for the action of the *representation* of the Lorentz transformation on fields. In this case we are acting on a scalar field and so no $D[\Lambda]$

[3] Not to be confused with this.

factors appear. For a general field we have

$$\phi^{\mu\nu\cdots}(x) \longrightarrow D[\Lambda]^{\mu}_{\sigma}\big(D[\Lambda]^{\nu}_{\tau}\ldots\big)\phi^{\sigma\tau\cdots}\big(\Lambda^{-1}x\big).$$

2.2 Noether's Theorem

Theorem 2.2.1 (Noether's Theorem)**.** *Every continuous symmetry of the Lagrangian gives rise to a conserved current, which we label $j^{\mu}(x)$, such that the equations of motion imply*

$$\partial_{\mu}j^{\mu} = 0. \tag{2.3}$$

As we will see, symmetries and their conserved currents are incredibly important in field theory, and so Noether's theorem is an invaluable tool to field theorists.

Proof. In order to prove Noether's theorem it is very helpful to work infinitesimally. As we are considering continuous symmetries, we can always do this.

Now the first thing we note is that the equations of motion, Equation (1.9), were obtained by varying the action and dropping surface terms. We therefore see that these will be completely unaffected for any transformation that takes the form

$$\mathcal{L} \longrightarrow \mathcal{L} + \epsilon\partial_{\mu}F^{\mu}, \tag{2.4}$$

where F^{μ} is some function of the fields and ϵ is some small parameter.

Ok, so let's consider some infinitesimal transformation of the field

$$\phi_a(x) \longrightarrow \phi_a(x) + \epsilon\delta\phi_a(x).$$

This gives a change in the action, of the general form[4]

$$\begin{aligned}
\epsilon\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi_a}\epsilon\delta\phi_a + \frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi_a)}\partial_{\mu}(\epsilon\phi_a) \\
&= \left[\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_{\mu}\left(\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi_a)}\right)\right]\epsilon\delta\phi_a + \epsilon\partial_{\mu}\left(\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi_a)}\delta\phi_a\right) \\
&= \epsilon\partial_{\mu}\left(\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi_a)}\delta\phi_a\right),
\end{aligned}$$

where we have used Equation (1.9) to set the square bracket term to zero. Now if this transformation is a symmetry of the system then we also have

$$\epsilon\delta\mathcal{L} = \partial_{\mu}F^{\mu},$$

and so taking these two results away from each other gives us

$$j^{\mu} = \frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi_a)}\delta\phi_a - F^{\mu}(\phi_a) \qquad \Longrightarrow \qquad \partial_{\mu}j^{\mu} = 0. \tag{2.5}$$

So this is our conserved current. We get one of these for each continuous $\delta\phi$ transformation and so we have proved Noether's theorem. □

[4] We use the fact that ϵ is a constant to take it outside the derivative.

Before doing some examples, first there is an important comment to make. As we have been careful to say above, Noether's theorem tells us that we get a conserved *current*. This is a stronger result then saying we get a conserved *charge*. Indeed a conserved current gives rise to a conserved charge, and we get this charge as follows:

$$\partial_\mu j^\mu = 0 \qquad \Longrightarrow \qquad \frac{\partial j^0}{\partial t} = -\nabla \cdot \vec{j}.$$

Integrating both sides of this over all of space, we have

$$\begin{aligned} \frac{\partial}{\partial t}\int_{\mathbb{R}^3} d^3\vec{x}\, j^0 &= -\int_{\mathbb{R}^3} d^3\vec{x}\, \nabla\cdot\vec{j} \\ &= -\oint d\vec{s}\cdot\vec{j} \\ &= 0, \end{aligned}$$

where the last line follows from the fact that we assume $\vec{j}$ falls off sufficiently fast as $|\vec{x}| \to \infty$. We can therefore define our conserved charge by

$$Q := \int_{\mathbb{R}^3} d^3\vec{x}\, j^0. \tag{2.6}$$

This is a *globally* conserved charge.

The reason that the existence of a conserved current is stronger than the existence of a globally conserved charge is that the current tells us that charge is conserved *locally*. This is easily seen by repeating the above procedure but now just integrating over some finite volume, V. We obtain

$$\frac{dQ_V}{dt} = -\oint_{\partial V} \vec{j}\cdot d\vec{S},$$

where ∂V is the boundary of V. This equation tells us the the charge leaving a volume V in time t (left-hand side) is equal to the current flowing through the boundary of the volume (right-hand side).

This is a much more powerful statement than simply that charge is conserved globally. For further clarity, if charge is conserved locally everywhere then we know that it is conserved globally (simply take your local region to be infinitely big). However if we only know that charge is conserved globally it could be possible that a charge disappears at some point $x \in \mathbb{R}^3$ and miraculously reappears at some other point $y \in \mathbb{R}^3$ far away from x. In other words the charge 'teleported' through space.

2.2.1 The Energy-Momentum Tensor

An important example of Noether's theorem comes from considering translations in spacetime. Recall that in particle mechanics, translations in space gave rise to momentum conservation, whereas translations in time gave rise to energy conservation. We are now dealing with a relativistic theory and so want to avoid this splitting of spacetime as much as possible, and so we want to find a 4-dimensional generalisation of the above conservation laws. This is exactly the *energy-momentum tensor*. We denote it by a T.

Remark 2.2.2. Some people also refer to the energy-momentum tensor as the stress-energy tensor or the stress-momentum tensor, or even the stress-energy-momentum tensor. I will try to use energy-momentum tensor everywhere here, but I might sometimes just call it the stress tensor. Apologies for this in advance.

Ok, let's derive this lovely chappy.

Consider a continuous translation of the spacetime coordinates. As above, we shall work infinitesimally, and so we have

$$x^\mu \longrightarrow x^\mu - \epsilon^\mu,$$

where we use a minus sign so that the transformation of the field is positive, see the discussion of active vs. passive transformations above. We can Taylor expand the field to obtain the transformation

$$\phi_a(x) \longrightarrow \phi_a(x) + \epsilon^\mu \partial_\mu \phi_a(x),$$

and similarly the Lagrangian transforms as

$$\mathcal{L}(x) \longrightarrow \mathcal{L}(x) + \epsilon^\mu \partial_\mu \mathcal{L}(x).$$

Comparing this to Equation (2.4), we see that $F^\mu = \epsilon^\mu \mathcal{L}$. Our conserved current is therefore

$$\begin{aligned} j^\mu &= \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \epsilon^\nu \partial_\nu \phi_a - \epsilon^\mu \mathcal{L} \\ &= \epsilon^\nu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \partial_\nu \phi_a - \delta^\mu_\nu \mathcal{L} \right), \end{aligned}$$

and so we have a conserved current for each ϵ^ν, i.e. for the translations in each direction. These are the components of our energy-momentum tensor.

$$T^\mu{}_\nu := (j^\mu)_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \partial_\nu \phi_a - \delta^\mu_\nu \mathcal{L}. \tag{2.7}$$

Being a conserved current it satisfies

$$\partial_\mu T^\mu{}_\nu = 0 \qquad \forall \nu = 0, 1, ..., d. \tag{2.8}$$

In spacetime (i.e. $d = 4$) we have four currents and so we have 4 charges. The ν index above tells us which direction we translated in, and so in correspondance with the comment made at the start of this section we want $\nu = 0$ to correspond to energy conservation and $\nu = 1, 2, 3$ to correspond to momentum conservation. We therefore define[5]

$$E := \int d^3\vec{x}\, T^{00}, \quad \text{and} \quad P^i := \int d^3\,\vec{x} T^{0i} \tag{2.9}$$

to be the total energy and the (i-th component of the) total momentum of the field configuration, respectively.

Remark 2.2.3. Note in deriving Equation (2.7), we didn't say anything about the actual form of our Lagrangian, and so this result holds for generic $\mathcal{L}$.

[5]Note we have raised the ν index here. As we are working in Minkowksi spacetime this really isn't a big deal, however as these are definitions you should be careful in more general spacetimes as some non-trivial factors will appear.

2.2.2 Angular Momentum & Boost Symmetries

So we have done spacetime translations, but recall that we have also restricted ourselves to Lagrangians that are Lorentz invariant. Given that Lorentz transformations are continuous, Noether's theorem tells us that we should have some conserved currents to go along with them. Keeping in line with the logic applied to the translations, what kind of conservation do we expect? Well in particle mechanics spatial rotations give rise to conservation of angular momentum, so we expect some generalisation of this. As we are considering the whole set of Lorentz transformations, we will also derive the symmetries corresponding to boosts, whatever these may be.

As always, we want to work infinitesimally, and its a fact that infinitesimal Lorentz transformations can be written as

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu}, \tag{2.10}$$

for some infintesimal $\omega^{\mu}{}_{\nu}$, and where Equation (2.1) tells us

$$\omega^{\mu\nu} = -\omega^{\nu\mu}. \tag{2.11}$$

From Equation (2.2), we have, after Taylor expanding

$$\phi(x) \longrightarrow \phi(x^{\mu}) - \omega^{\mu}{}_{\nu} x^{\nu} \partial_{\mu}\phi(x) \qquad \Longrightarrow \qquad \delta\phi = -\omega^{\mu}{}_{\nu} x^{\nu} \partial_{\mu}\phi.$$

Similarly we obtain

$$\delta\mathcal{L} = -\omega^{\mu}{}_{\nu} x^{\nu} \partial_{\mu}\mathcal{L}.$$

Now we use a clever trick. Firstly we note that $\omega^{\mu}{}_{\nu}$ is a constant so we can take it inside the derivative. Next we note that

$$\partial_{\mu} x^{\nu} = \delta^{\nu}_{\mu},$$

along with $\omega^{\mu}{}_{\mu} = 0$ by Equation (2.11). We can therefore also take the x^{ν} inside the derivative and obtain

$$\delta\mathcal{L} = -\partial_{\mu}\big(\omega^{\mu}{}_{\nu} x^{\nu} \mathcal{L}\big).$$

This is now of the form Equation (2.4), with $F^{\mu} = \omega^{\mu}{}_{\nu} x^{\nu} \mathcal{L}$. So our conserved currents are

$$\begin{aligned} j^{\mu} &= -\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi)} \omega^{\sigma}{}_{\nu} x^{\nu} \partial_{\sigma}\phi + \omega^{\mu}{}_{\nu} x^{\nu} \mathcal{L} \\ &= -\omega^{\sigma}{}_{\nu} \left[\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi)} \partial_{\sigma}\phi + \delta^{\mu}{}_{\sigma} \mathcal{L} \right] x^{\nu} \\ &= -\omega^{\sigma}{}_{\nu} T^{\mu}{}_{\sigma} x^{\nu}. \end{aligned}$$

As above we can split this into the individual currents, one for each $\omega^{\sigma}{}_{\nu}$, and obtain

$$(\mathcal{J}^{\mu})^{\rho\sigma} = x^{\rho} T^{\mu\sigma} - x^{\sigma} T^{\mu\rho}, \qquad \text{with} \qquad \partial_{\mu}(\mathcal{J}^{\mu})^{\rho\sigma} = 0 \qquad \forall \rho, \sigma = 0, .., 3. \tag{2.12}$$

Each one of these has a corresponding conserved charge. We get the particle result, i.e. angular momentum, when $\rho, \sigma = 1, 2, 3$:

$$Q^{ij} = \int d^3\vec{x} \left(x^i T^{0j} - x^j T^{0i} \right).$$

The question is "what charge do the boosts give us?" Well these correspond to $\rho = 0, \sigma = 1, 2, 3$, which give

$$Q^{0i} = \int d^3\vec{x}\,(x^0 T^{0i} - x^i T^{00}).$$

What is this? Well we know that its temporal derivative vanishes, so using $x^0 = t$, we have

$$\begin{aligned}\frac{dQ^{0i}}{dt} &= \int d^3\vec{x}\,T^{0i} + \int d^3\vec{x}\,t\frac{\partial T^{0i}}{\partial t} - \frac{d}{dt}\int d^3\vec{x}\,x^i T^{00} \\ 0 &= P^i + \frac{dP^i}{dt} - \frac{d}{dt}\int d^3\vec{x}\,x^i T^{00},\end{aligned}$$

where we've used Equation (2.9). But we've already seen that P^i is a constant and so we simply get

$$\frac{d}{dt}\int d^3\vec{x}\,x^i T^{00} = \text{constant},$$

which is the statement that the center of energy of the field travels with constant velocity.

Remark 2.2.4. Note that Equation (2.11) agrees with the number of Lorentz transformations. That is, it is an antisymmetric, 4×4 matrix and so has $4 \times 3/2 = 6$ independant entries. So it has 6 basis elements, 3 of these correspond to spatial rotations and the other 3 are our boosts.

2.2.3 Internal Symmetries

In the above calculations we considered transformations that did something to the spacetime itself, that is the x^μs changed. The question is "is this the only kind of symmetry we can have?" The answer is no, but in order to see it we need to have at least two fields in our Lagrangian. When this is the case we can consider transformations in the plane spanned by these two fields. If our transformations give rise to a symmetry, then we call them *internal symmetries*. Let's look at some examples.

Example 2.2.5. Consider a theory with 2 real scalar fields ϕ_1 and ϕ_2. We can package these together into a two column representation simply as

$$\vec{\phi} := \begin{pmatrix}\phi_1 \\ \phi_2\end{pmatrix}.$$

We can consider these spanning some 2-dimensional plane as indicated diagrammatically below. Let's then consider the Lagrangian

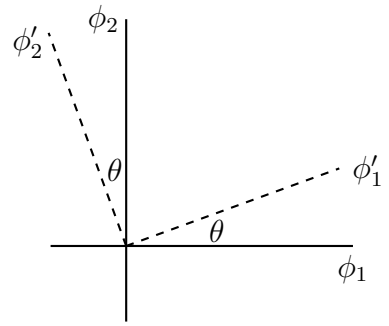
$$\begin{aligned}\mathcal{L} &= \frac{1}{2}(\partial_\mu\vec{\phi})\cdot(\partial^\mu\phi) - \frac{1}{2}m^2\vec{\phi}\cdot\vec{\phi} \\ &= \frac{1}{2}\left[(\partial\phi_1)^2 + (\partial\phi_2)^2\right] - \frac{1}{2}m^2\left[\phi_1^2 + \phi_2^2\right]\end{aligned}$$

where on the first line we have written out the multiple of the derivatives explicitly to make the dot-product clear. This Lagrangian is invariant under rotations in the $\phi_1 - \phi_2$ plane, i.e. under

$$\begin{aligned}\phi_1 &\longrightarrow \phi_1' = \phi_1\cos\theta + \phi_2\sin\theta \\ \phi_2 &\longrightarrow \phi_2' = -\phi_1\sin\theta + \phi_2\cos\theta.\end{aligned}$$

This is a continuous transformation and so we can work infinitesimally and find the conserved current. We have

$$\begin{aligned} \phi_1' &\approx \phi_1 + \theta\phi_2 \qquad \Longrightarrow \qquad \delta\phi_1 = \theta\phi_2 \\ \phi_1' &\approx -\theta\phi_1 + \phi_2 \qquad \Longrightarrow \qquad \delta\phi_2 = -\theta\phi_1. \end{aligned}$$



The Lagrangian doesn't change, so our conserved current is simply

$$\begin{aligned} j^\mu &= \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_1)}\delta\phi_1 + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_2)}\delta\phi_2 \\ &= (\partial^\mu\phi_1)\phi_2 - (\partial^\mu\phi_2)\phi_1. \end{aligned}$$

We can check that this obeys $\partial_\mu j^\mu = 0$. First we need the Euler-Lagrange equations for our action. We have

$$0 = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_1)}\right) - \frac{\partial\mathcal{L}}{\partial\phi_1} = (\partial^2 + m^2)\phi_1$$

and similarly for ϕ_2. Now take the derivative of our current:

$$\begin{aligned} \partial_\mu j^\mu &= (\partial^2\phi_1)\phi_2 + \partial^\mu\phi_1\partial_\mu\phi_2 - (\partial^2\phi_2)\phi_1 - \partial^\mu\phi_2\partial_\mu\phi_1 \\ &= m^2\phi_1\phi_2 - m^2\phi_2\phi_1 \\ &= 0, \end{aligned}$$

where we used the Euler-Lagrange equations to get to the third line, and then used the fact that we have real scalars so $\phi_1\phi_2 = \phi_2\phi_1$.

This symmetry is known as a *global SO(2) symmetry*. The name makes sense: its global (i.e. the whole plane is rotated) and its a rotation in 2-dimensions. It is actually a specific case of the more general global SO(N) symmetry, which has

$$\frac{N(N-1)}{2}$$

conserved currents for N fields.

Example 2.2.6. We can reformulate the previous example by considering the complex scalar field

$$\psi = \frac{1}{\sqrt{2}}\big(\phi_1 + i\phi_2\big).$$

The above Lagrangian then becomes[6]

$$\mathcal{L} = \partial_\mu \psi^* \partial^\mu \psi - m^2 \psi^* \psi. \tag{2.13}$$

We are now rotating in the complex plane so instead of considering a SO(2) rotation we consider the complex U(1) rotation

$$\psi \longrightarrow e^{i\alpha}\psi.$$

We already know that this is a symmetry because its exactly the same as the previous example, but what does the current look like in this complexified case? Well simple calculation will give

$$j^\mu = i\big[(\partial^\mu \psi^*)\psi - \psi^*(\partial^\mu \psi)\big] \tag{2.14}$$

Exercise

Derive the above conserved current. *Hint: Note that* $\psi^* \longrightarrow e^{-i\alpha}\psi^*$ *and then work infinitesimally.*

Remark 2.2.7. Maybe put a comment here in line with the non-abelian comment made in Tong, page 18.

Internal symmetries will prove to be vital for the study of particle physics as a QFT. We will see that the charges arising from such symmetries correspond to things such as electric charge and particle number.

2.3 Hamiltonian Field Theory

Above we have constructed classical field theory in terms of Lagrangians. It is true that we can extend this Lagrangian approach to QFT by using so-called *path integrals*, and there are advantages to doing that (most notably that the Lorentz invariance is manifest throughout), however in this course we shall use the Hamiltonian approach instead. This uses so-called *canonical quantisation* to promote the fields to operators.

In particle mechanics, we define the Hamiltonian to be

$$H(p,q) = \sum_a p_a \dot{q}_a - L(q,\dot{q}),$$

where

$$p_a := \frac{\partial L}{\partial \dot{q}_a}$$

is known as the *conjugate momentum*. The idea is to eliminate $\dot{q}$ wherever we can and replace it with p.

In field theory we do a very similar thing, but now we have to use the Lagrangian density, $\mathcal{L}$, and we define the *conjugate momentum density* and *Hamiltonian density*

[6]Note that we don't have a factor of $1/2$ in this expression as was the case for the real scalar field. This is just because of the $1/\sqrt{2}$ factor above.

$$\pi(x) := \frac{\partial \mathcal{L}}{\partial \dot{\phi}(x)} \tag{2.15}$$

$$\mathcal{H} = \pi(x)\dot{\phi}(x) - \mathcal{L}(x). \tag{2.16}$$

Again we favour replacing any $\dot{\phi}(x)$ dependence with $\pi(x)$ once we know the relation. As with Equation (1.8), we define the Hamiltonian as the spatial integral over the Hamiltonian density,

$$H(t) = \int d^3\vec{x}\, \mathcal{H}(x). \tag{2.17}$$

Note in this line we appear to have broken Lorentz symmetry as we have picked a preferred time to define our Hamiltonian. However, as long as we are careful not to do anything silly, we should be alright because we started with a Lorentz invariant theory. This is what we mean by the Lorentz symmetry not being manifest: it is not obvious just by looking at the equations that we have Lorentz symmetry, in contrast to the Lagrangian formalism where spacetime indices were always summed over and so it was clear.

Example 2.3.1. As an example let's consider the real scalar field mentioned above with Lagrangian[7]

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}(\partial\phi_a)^2 - \frac{1}{2}m^2\phi_a^2 \\ &= \frac{1}{2}\dot{\phi}_a^2 - \frac{1}{2}(\nabla\phi_a)^2 - \frac{1}{2}m^2\phi_a^2.\end{aligned}$$

The conjugate momentum is therefore

$$\pi_a(x) = \dot{\phi}_a(x),$$

and the Hamiltonian density is

$$\begin{aligned}\mathcal{H} &= \dot{\phi}_a^2 - \mathcal{L} \\ &= \frac{1}{2}\dot{\phi}_a^2 + \frac{1}{2}(\nabla\phi_a)^2 + \frac{1}{2}m^2\phi_a^2 \\ &= \frac{1}{2}\pi_a^2 + \frac{1}{2}(\nabla\phi_a)^2 + \frac{1}{2}m^2\phi_a^2,\end{aligned}$$

where in the last line we have done our procedure of swapping out the $\dot{\phi}$s for πs using the relation above.[8] The integral over these terms give three contributions to the Hamiltonian, they are

$$H = \int d^3\vec{x}\left(\underbrace{\frac{1}{2}\pi_a^2}_{\text{kinetic}} + \underbrace{\frac{1}{2}(\nabla\phi_a)^2}_{\text{shear}} + \underbrace{\frac{1}{2}\phi_a^2}_{\text{mass}}\right).$$

[7] It is incredibly common in the field theory world to forget to say 'density', as I have done here (and probably above too). The symbols should tell you what we mean: densities are normally given in fancy curly font.

[8] For clarity, you don't just swap $\dot{\phi} \longrightarrow \pi$, but you use the relation for π in terms of $\dot{\phi}$ to eliminate the $\dot{\phi}$s. In this particular case, $\pi = \dot{\phi}$, and it does correspond to just swapping them.

3 | Canonical Quantisation & Free Klein Gordan Field

3.1 Canonical Quantisation

Recall the idea behind quantisation in QM is to promote the generalised coordinates in the classical theory and 'promote' them to operators acting on the Hilbert space of the quantum theory. This recipe is called *canonical quantisation*. The Poisson bracket[1] relation between the generalised coordinates magically transformed into a commutation relation between the operators. That is[2]

$$\begin{aligned} \{q_a, a_b\} = \{p^a, p^b\} = 0 \qquad &\longrightarrow \qquad [\hat{q}_a, \hat{q}_b] = [\hat{p}^a, \hat{p}^b] = 0 \\ \{q_a, p^b\} = \delta_a^b \qquad &\longrightarrow \qquad [\hat{q}_a, \hat{p}^b] = i\delta_a^b \end{aligned}$$

As we are working in Minkowski spacetime, we can lower the indices on the ps easily and obtain

$$[\hat{q}_a, \hat{q}_b] = [\hat{p}_a, \hat{p}_b] = 0 \quad \text{and} \quad [\hat{q}_a, \hat{p}_b] = i\delta_{ab}.$$

We adopt the same philosophy for fields, and promote them to *operator valued functions*. However we have a small problem: for the particle mechanics case we just had a finite number of generalised coordinates and so it was easy to do, whereas for the fields there's an infinite number, one for each point $\vec{x}$ in space. We treat this spatial dependence as a label and so need something analogous to the δ_{ab} above. The answer is obviously the usual delta function. We use the Schrödinger picture, where all time dependence appears in the states $|\psi\rangle$ which obey the Schrödinger equation

$$i\frac{d\,|\psi\rangle}{dt} = H\,|\psi\rangle\,,$$

and the operators themselves are time independent, and write[3]

$$[\phi_a(\vec{x}), \phi_b(\vec{y})] = [\pi_a(\vec{x}), \pi_b(\vec{y})] = 0, \quad \text{and} \quad [\phi_a(\vec{x}), \pi_b(\vec{y})] = i\delta^{(3)}(\vec{x} - \vec{y})\delta_{ab}. \tag{3.1}$$

Remark 3.1.1. Note we have really butchered our manifest Lorentz invariance here: we completely separated space and time and made the operators only functions of space! Of course it must be true that we're alright to do this and still get a Lorentz invariant theory, otherwise we wouldn't be doing it in these notes. However this choice of doing things will introduce some factors here and there (e.g. to the measures in integrals) to ensure Lorentz invariance.

[1] Note to self: maybe put something about Poisson brackets above when discussing generalised coordinates.
[2] There are factors of $\hbar$s in these equations, but in natural units they go bye-bye.
[3] You can work in the Heisenberg picture and define what are known as 'equal time' commutation relations.

3.2 Free Theories

Recall that the typical goal of QM is to find the spectrum (i.e. eigenvalues) of operators, in particular the Hamiltonian. We now want to do a similar thing for QFTs, however this turns out to be an incredibly hard thing to do, as we now have an uncountably infinite number of degrees of freedom (one for each $\vec{x}$ value)! The question is "can we somehow get around this?" The answer is "yes, but it restricts the type of theories we consider." What we do is consider theories in which each degree of freedom evolved independently to all others. This is essentially the statement that $\phi(\vec{x})$ and $\phi(\vec{y})$ don't talk to each other unless $\vec{x} = \vec{y}$. For a reasonably self explanatory reason, we call these theories *free theories*.

Remark 3.2.1. Of course free theories are boring from a physical perspective (as nothing interacts so there are essentially no forces!), and we want to study interacting theories. We will return to these later, and study them as perturbations using Feynman diagrams, but first we need to develop the mathematical tools of free theory. So hold tight, more interesting stuff is coming.

Ok, so how do we go about quantising the free fields and finding their spectrum? Well the answer is to consider our lovely friend the classical free Klein-Gordan field. Recall that the Euler-Lagrange equations gave us the Klein-Gordan equation, Equation (1.11):

$$(\partial^2 + m^2)\phi = 0.$$

As we are working in Minkowski spacetime, we have a global notion of time[4] and so we can decompose these fields in terms of their Fourier transform

$$\begin{aligned}\phi(\vec{x},t) &= \int \frac{d^3\vec{p}}{(2\pi)^3}\,\frac{1}{2}\left[\widetilde{\phi}(\vec{p},t)e^{i\vec{p}\cdot\vec{x}} + \widetilde{\phi}^*(\vec{p},t)e^{-i\vec{p}\cdot\vec{x}}\right] \\ &= \int \frac{d^3\vec{p}}{(2\pi)^3}\,\widetilde{\phi}(\vec{p},t)e^{i\vec{p}\cdot\vec{x}},\end{aligned}$$

where to get to the second line we have identified $\widetilde{\phi}^*(-\vec{p},t) = \widetilde{\phi}(\vec{p},t)$. Using this coordinate system, our Klein-Gordan equation reduces to (dropping the tilde and using the argument to indicate which function it is)

$$\left(\frac{\partial^2}{\partial t^2} + (\vec{p} + m^2)\right)\phi(\vec{p},t) = 0.$$

Then the keen-eyed person notes that *for each value* of $\vec{p}$ this corresponds to its own harmonic oscillator with frequency

$$\omega_{\vec{p}} = \sqrt{\vec{p}^2 + m^2}. \tag{3.2}$$

To stress the point, we get a harmonic oscillator each *every single* value of $\vec{p}$ independently from any other value. The general field $\phi(\vec{x},t)$ is then simply given by a linear superposition of (an infinite number of) harmonic oscillators. We therefore have, at least classically, achieved the goal above to frame the theory in such a way that each degree of freedom (each $\vec{p}$) evolves independently from the others. Now what we want to do is quantise it.

[4] Again this comment is just made because things are different in general spacetimes, see footnote 3 from the first lecture.

Remark 3.2.2. Note that in order to take the Fourier expansion above we need to assume that the field $\phi(\vec{x}, t)$ die off sufficiently quickly as $|\vec{x}| \to \infty$. In order to guarantee this, we use fields from the so-called *Schwartz space*. A precise definition (with a lot more context for why they're useful in QM) can be found in Simon Rea and my notes on Dr. Schuller's course on quantum theory, but for here we shall just say they are functions that die off at infinity as do their derivatives.

3.2.1 A Quantum Field Theorist's Best Friend: The Harmonic Oscillator

Notation. In this section, and most likely in everything to follow from now on, I am going to drop the hats on quantum operators. I might reinstate them at some points for clarity, but we shall see.

As we have been doing above, let's forget about field theory for a minute and study regular old quantum mechanics. Recall that the quantum harmonic oscillator (QHO) has the Hamiltonian (in 1-dimension)

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2, \tag{3.3}$$

where p and q are the canonical operators obeying

$$[q, p] = i. \tag{3.4}$$

We can rewrite this Hamiltonian in a nicer form by introducing the *creation* and *annihilation operators*: $a^\dagger$ and a, respectively. These are defined such that

$$q = \frac{1}{\sqrt{2\omega}}(a + a^\dagger), \quad \text{and} \quad p = -i\sqrt{\frac{\omega}{2}}(a - a^\dagger). \tag{3.5}$$

Exercise

Show that Equation (3.4) and Equation (3.5) give

$$[a, a^\dagger] = 1. \tag{3.6}$$

Then using this result show that the Hamiltonian Equation (3.3) can be rewritten as

$$H = \omega\left(a^\dagger a + \frac{1}{2}\right). \tag{3.7}$$

Finally show that

$$[H, a^\dagger] = \omega a^\dagger, \quad \text{and} \quad [H, a] = -\omega a. \tag{3.8}$$

We can now begin to look at the spectrum of the QHO. The *ground state* $|0\rangle$ is defined via

$$a\,|0\rangle = 0 \tag{3.9}$$

and so we see from Equation (3.7) that the ground state has energy

$$E_0 = \frac{\omega}{2}.$$

Now using Equation (3.8), we see that

$$\begin{aligned} H(a^\dagger |0\rangle) &= a^\dagger(H|0\rangle) + \omega(a^\dagger|0\rangle) \\ &= (E_0 + \omega)(a^\dagger|0\rangle), \end{aligned}$$

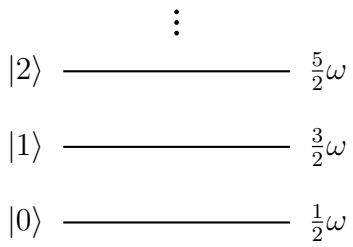
and so we see that $a^\dagger|0\rangle$ is an eigenvector of H with energy $E = E_0 + \omega$. We therefore define *excited states* as[5]

$$|n\rangle := (a^\dagger)^n |0\rangle . \tag{3.10}$$

Extending the calculation above, these states have energy

$$E_n = \left(n + \frac{1}{2}\right)\omega, \tag{3.11}$$

and so we build a spectrum for our theory as integer steps in ω, as indicated pictorially below.



As the notation suggests (and as can easily be seen from Equation (3.5)) the creation and annihilation operators are Hermitian conjugates to each other. This translates into bra-ket notation in terms of left and right actions:

$$\big(a|\psi\rangle\big)^\dagger = \langle\psi| a^\dagger. \tag{3.12}$$

3.2.2 Spectrum Of Quantum Klein-Gordan Field

Ok so we now know how to get the spectrum of a single QHO. We have also seen that the Klein-Gordan field is essentially an infinite sum (i.e. an integral) of QHOs, so now all we need to do to get the spectrum of the Hamiltonian is integrate over an infinite number of creation and annihilation operators, labelled by $\vec{p}$ — $a^\dagger_{\vec{p}}$ and $a_{\vec{p}}$. To do this we express ϕ and π as

$$\begin{aligned} \phi(\vec{x}) &= \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\Big[a_{\vec{p}}\, e^{i\vec{p}\cdot\vec{x}} + a^\dagger_{\vec{p}} e^{-i\vec{p}\cdot\vec{x}}\Big] \\ \pi(\vec{x}) &= \int \frac{d^3\vec{p}}{(2\pi)^3}(-i)\sqrt{\frac{\omega_{\vec{p}}}{2}}\Big[a_{\vec{p}}\, e^{i\vec{p}\cdot\vec{x}} - a^\dagger_{\vec{p}} e^{-i\vec{p}\cdot\vec{x}}\Big], \end{aligned} \tag{3.13}$$

where the form of these expressions makes sense when comparing to Equation (3.5).

Remark 3.2.3. Note that it is only at this point that we are now considering the *quantum* Klein-Gordan field. This is in contrast why I was careful to emphasise before that we were considering the classical Klein-Gordan field. The equations of motion are still the same, but now they are quantum expressions.

[5] We're ignoring normalisation here, i.e. $\langle n|n\rangle \neq 1$. This is not important for our present discussion.

This looks nice, but what about the commutators?

Claim 3.2.4. Equation (3.1) hold, if and only if,

$$[a_{\vec{p}}, a_{\vec{q}}] = \left[a^\dagger_{\vec{p}}, a^\dagger_{\vec{q}}\right] = 0, \quad \text{and} \quad \left[a_{\vec{p}}, a^\dagger_{\vec{q}}\right] = (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q}). \tag{3.14}$$

This claim is not complicated to prove, but a bit of a pain to write out and so I, lovingly, decided to set them as an exercise.

Exercise

Prove Claim 3.2.4. Note it is an if and only if statement so you need to show it both ways, i.e. the ϕ/π commutators imply the $a/a^\dagger$ ones and visa versa. *Hint: If you get very stuck, Prof. Tong sketches one on page 24 of his notes.*

Remark 3.2.5. Equation (3.14) seem strange: the left-hand side is a commutator between operators whereas the right-hand side is a delta distribution? The way we wrap our heads around this is to remember that any physical results in QM appear in inner products $\langle\psi| A |\psi\rangle$, which can be written as integrals and so the delta function makes sense.

Using Equation (3.12), and the fact that our states are orthogonal, we see that it is only the terms that contain $a_{\vec{p}} a^\dagger_{\vec{p}}$ or $a^\dagger_{\vec{p}} a_{\vec{p}}$ that contribute to expectation values. We will see an example of this below when finding the momentum of the first excited state.

This idea, and similar ones where we get $\mathbb{C}$-numbers on the right-hand side, comes up again and again in QFT. What we have to keep telling ourselves is that "remember we're sandwiching this between states!" and then go from there.

Ok, so we have already derived that the Hamiltonian for this system is

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

We can substitute Equation (3.13) in and obtain

$$\begin{aligned} H &= \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{\omega_{\vec{p}}}{2}\left[a_{\vec{p}} a^\dagger_{\vec{p}} + a^\dagger_{\vec{p}} a_{\vec{p}}\right] \\ &= \int \frac{d^3\vec{p}}{(2\pi)^3}\,\omega_{\vec{p}}\left[a^\dagger_{\vec{p}} a_{\vec{p}} + \frac{1}{2}(2\pi)^3\delta^{(3)}(0)\right], \end{aligned} \tag{3.15}$$

where to get to the last line we have used Equation (3.14).

Exercise

Another lovely exercise for you: Obtain the first line of Equation (3.15). *Hint: Again if you get really stuck, Prof. Tong has done this in his notes, also on page 24.*

Remark 3.2.6. All jokes aside about the above exercises, it is actually a really beneficial exercise to do and will test your understanding of a reasonable amount of the information leading up to here. So honestly at least give them a go.

3.2.3 A Quantum Field Theorist's Least Favourite Friend: Infinities

We can extend Equation (3.9) to the field theory case to define the ground state as

$$a_{\vec{p}}\,|0\rangle = 0 \qquad \forall \vec{p}. \tag{3.16}$$

We can then use Equation (3.15) to find the ground state energy:

$$H\,|0\rangle = \left[\int d^3\vec{p}\,\frac{1}{2}\omega_{\vec{p}}\delta^{(3)}(0)\right]|0\rangle\,,$$

which is... infinite?! Uh oh this does not seem good at all, but what did we do wrong? The answer is actually nothing. QFT is filled with infinities and, as Prof. Tong explains "each tells us something important, usually that we're doing something wrong, or asking the wrong question". What we need to do when infinities arise is ask where they come from and what that implies. So let's investigate this one.

Firstly the bad news... it's actually two infinities! The first one comes from the fact that we're considering the entirety of $\mathbb{R}^3$. So instead let's put our theory inside a box of size L. We adopt the usual idea where we prescribe periodic boundary conditions to the sides, and so get the *flat torus*.[6] Finally we just let $L \to \infty$ to get our result back. We can then use

$$(2\pi)^3\delta^{(3)}(0) = \lim_{L\to\infty}\int_{-L/2}^{L/2} d^3\vec{x}\,e^{i\vec{x}\cdot\vec{p}}\Big|_{\vec{p}=0} = \lim_{L\to\infty}\int_{-L/2}^{L/2} d^3\vec{x} = V,$$

where V is the volume of our theory. Ok so the $(2\pi)^3\delta(0)$ infinity is just because we're considering the total energy E_0 instead of the *energy density*

$$\varepsilon_0 := \frac{E_0}{V} = \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{2}\omega_{\vec{p}}.$$

This is still infinite though! Why? well because we take the positive root in Equation (3.2) and so we're summing an infinite number of positive numbers! This seems like a more tricky beast to tame, however then a light bulb goes off in our heads and we realise that this is just the ground state energy, and so, from the extension of Equation (3.11), it will appear in *all* the energies. Then remembering that all we can measure physically is energy *differences* we realise that we are free to just 'take this away' from every energy in our system (as the difference wont be effected), so this is what we shall do.

You might not be very comfortable with 'subtracting infinity', and if that is the case, allow me to provide a calming remark: we are mere mortals pretending to be Gods. In less poetic (and perhaps offensive) words: we have assumed that the theory we have written down is valid up to arbitrarily large momentum. This corresponds to arbitrarily large energies, or arbitrarily low length scales. Nature is highly unlikely to agree with us and so plays the trump card of "you need to cut-off your theory at high momenta!" If we do this, the ground state energy density will become finite and then we can comfortably take it away from all energies and proceed as if nothing ever happened. Infinities of this kind are called *ultra-violet divergences*.

[6] If you're confused why I say flat torus, either google it or think about what happens geometrically when you make these periodic boundary conditions.

3.2.4 Finally, The Spectrum

Now that we have tamed our infinities, we can proceed to finding the spectrum of the Klein-Gordan field. Continuing with the extension of the QHO, we have

$$[H, a_{\vec{p}}]\,|0\rangle = -\omega_{\vec{p}}\, a_{\vec{p}}\,|0\rangle\,, \quad \text{and} \quad [H, a^\dagger_{\vec{p}}] = \omega_{\vec{p}}\, a^\dagger_{\vec{p}}|0\rangle\,, \tag{3.17}$$

which encourages us to define

$$|\vec{p}\rangle := a^\dagger_{\vec{p}}\,|0\rangle\,.$$

But now we have different creation/annihilation operators and so can excite the ground state in different ways. We therefore use the notation

$$|\vec{p}, \vec{q}, ...\rangle := \left(a^\dagger_{\vec{p}}\, a^\dagger_{\vec{q}} ...\right)|0\rangle\,.$$

Then finally using Equation (3.17) gives us

$$H\,|\vec{p}, \vec{q}, ...\rangle = (\omega_{\vec{p}} + \omega_{\vec{q}} + ...)\,|0\rangle\,.$$

This exhausts our spectrum.[7]

Remark 3.2.7. Just as with the states $|n\rangle$ defined in Equation (3.10), our states $|\vec{p}\rangle$ are not normalised. They are orthogonal though. We shall return to the normalisation later.

3.2.5 Interpreting The Eigenstates: Particles

So we have our spectrum, now we want to interpret what they mean physically. Well let's focus on the first excited state

$$|\vec{p}\rangle = a^\dagger_{\vec{p}}|0\rangle\,.$$

We have already seen that this has energy

$$E_{\vec{p}} = \omega_{\vec{p}} = \sqrt{\vec{p}^2 + m^2},$$

and have already remarked that this is the energy of a relativistic particle. But could this just be some coincidence? Well let's look at the *physical* momentum given by Equation (2.9). Straight forward calculation using the Lagrangian for our system gives

$$T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - \eta^{\mu\nu}\mathcal{L},$$

and so the momentum is

$$P^i = \int d^3\vec{x}\,\dot{\phi}(x)\partial^i\phi(x),$$

which as a operator we can write[8]

$$\vec{P} = -\int d^3\vec{x}\,\pi(\vec{x})\nabla\phi(\vec{x})$$

[7] And perhaps it has exhausted you getting to this point.

[8] Recall that we're using the field theorist's convention of $(+,-,-,-)$.

If we then use Equation (3.13), we get

$$\begin{aligned}\vec{P} &= \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{\vec{p}}{2}\left[a^\dagger_{\vec{p}}a_{\vec{p}} + a_{\vec{p}}a^\dagger_{\vec{p}} + a_{\vec{p}}a_{\vec{-p}} + a^\dagger_{\vec{p}}a^\dagger_{\vec{-p}}\right] \\ &= \int \frac{d^3\vec{p}}{(2\pi)^3}\vec{p}a^\dagger_{\vec{p}}a_{\vec{p}},\end{aligned}$$

where to get to the last line we have used the commutation relation between $a/a^\dagger$ and then dropped the $\delta(0)$ term as we did for the energy, and then 'dropped' the aa and $a^\dagger a^\dagger$ terms using the argument of Remark 3.2.5. So we see that the first excited state has momentum

$$\vec{P}\,|\vec{p}\rangle = \vec{p}\,|\vec{p}\rangle\,,$$

which is exactly what we expect for a particle. We therefore *interpret* the state $|\vec{p}\rangle$ to be a particle with 3-momentum $\vec{p}$. It is important to note that this is simply an interpretation, what we're really dealing with is an excitation of a field.

Exercise

Using the definition

$$J^i = \epsilon^{ijk}\int d^3\vec{x}\,(\mathcal{J}^0)^{jk},$$

where ϵ^{ijk} is the Levi-Civita tensor density and $\mathcal{J}$ is the classical angular momentum defined in Equation (2.12), show that the state $|\vec{p}=0\rangle$ has

$$J^i\,|\vec{p}=0\rangle = 0.$$

The result of the above exercise tells us that the our interpreted particle has no spin (i.e. no internal angular momentum). That is, our quantisation of the Klein-Gordan field gives rise to a spin-0 particle.

3.2.6 Multiparticle States

Above we just considered the first excited state, but we have already seen that we can apply the creation operators multiple times, i.e. we have

$$|\vec{p}_1,...,\vec{p}_n\rangle = a^\dagger_{\vec{p}_1}...a^\dagger_{\vec{p}_n}\,|0\rangle\,.$$

As we have used n creation operators, we refer to these states are n-particle states and interpret it as n particles with 3-momenta $(\vec{p}_1,...,\vec{p}_n)$. As the creation operators commute with each other we have

$$|\vec{p},\vec{q}\rangle = |\vec{q},\vec{p}\rangle\,,$$

and so the particles are symmetric under interchange. We therefore see that these particles are *bosons*. This also agrees with the fact that they are spin-0, as per the above exercise.

Now recall that in the first lecture we showed when we unite special relativity with QM we have to account for particle number changing. We said that we do this by changing our

Hilbert space to be a Fock space, Equation (1.5). We can clarify a bit what we meant there now. If $\mathcal{H}$ is the Hilbert space of our 1-particle states (i.e. $|\vec{p}\rangle \in \mathcal{H}$), then we have

$$|\vec{p}_1, ..., \vec{p}_n\rangle \in \mathcal{H}^{\otimes n}.$$

We then take the direct sum over all these different states so that our total Hilbert space accounts for the particle creation process. That is, let's say we start off with a state $|\psi\rangle \in \mathcal{H}^{\otimes n}$ and then something happens to it and causes the production of a new particle. The new state would then be an element of $\mathcal{H}^{\otimes(n+1)}$. If we don't take the Fock space construction as above, this process would mean that we leave our Hilbert space, which is a big no-no.

3.2.7 Interpretation Of $\phi(\vec{x})\,|0\rangle$

The next thing we can ask is how does the action of $\phi(\vec{x})$ fit into our particle interpretation? Well let's calculated it and see. We have

$$\begin{aligned}\phi(\vec{x})\,|0\rangle &= \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\left[a_{\vec{p}}\,e^{i\vec{p}\cdot\vec{x}} + a^\dagger_{\vec{p}}\,e^{-i\vec{p}\cdot\vec{x}}\right]|0\rangle \\ &= \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}e^{-i\vec{p}\cdot\vec{x}}\,|\vec{p}\rangle\,.\end{aligned}$$

Now compare this to the result from QM[9]

$$|x\rangle = \int dp e^{-ipx}\,|p\rangle\,.$$

We can therefore interpret $\phi(\vec{x})\,|0\rangle$ as a particle at position $\vec{x}$.

3.2.8 Number Operator

We have just talked about constructing a whole Fock space to take into account the fact that particle number can change. Well in our free theory this construction is actually not important because nothing is interacting and so there's no chance for more particles to be produced. We can show this explicitly by defining the *number operator*

$$N := \int \frac{d^3p}{(2\pi)^3}a^\dagger_{\vec{p}}a_{\vec{p}}, \tag{3.18}$$

which satisfies

$$N\,|\vec{p}_1, ..., \vec{p}_n\rangle = n\,|\vec{p}_1, ..., \vec{p}_n\rangle\,.$$

The proof that particle number is conserved in a free theory is the content of the next exercise.

Exercise

Using the above definition and Equation (3.15) (with the delta term dropped) show that

$$[H, N] = 0. \tag{3.19}$$

[9] This result comes from inserting a complete set of states, $\mathbb{1} = \int dp\,|p\rangle\,\langle p|$. See any decent QM book for details.

This tells us that once we are in a particular $\mathcal{H}^{\otimes n}$ in the Fock space, we will never leave it in a free theory. This makes them a bit boring to study. Fortunately later we will allow for interactions in our theory and see that the above result no longer holds in general. The sad news is free theories are the only ones we can solve exactly, and so in order to study interacting theories we are going to have to develop the mathematics of Feynman diagrams. These are essentially life saving tools to calculate things in perturbation expansions of interacting QFTs.

4 | Restoring Lorentz Invariance & Causality

4.1 Operator Valued Distributions

Recall that last lecture we gave an interpretation of our results as particles. There is a problem with this interpretation: they are momentum eigenstates and so, by Heisenberg uncertainty principle, are not localised in position space at all. This problem essentially stems from the fact that in QM the position and momentum eigenstates are not good elements of the Hilbert space as the normalise to delta-functions. The QFT equivalent is that the operators $\phi(\vec{x})$ and $a_{\vec{p}}$ are not good operators on the Fock space as the resulting states are not normalisable. Explicitly we have

$$\langle 0|\, a_{\vec{p}} a^\dagger_{\vec{p}} \,|0\rangle = \langle \vec{p}|\vec{p}\rangle = (2\pi)^3\delta^{(3)}(0), \quad \text{and} \quad \langle 0|\, \phi(\vec{x})\phi(\vec{x}) \,|0\rangle = \langle \vec{x}|\vec{x}\rangle = \delta(0).$$

These are known as *operator valued distributions*, and the way we fix this problem is by smearing them over space. For example the wavepacket

$$|\varphi\rangle = \int \frac{d^3\vec{p}}{(2\pi)^3} e^{-i\vec{p}\cdot\vec{x}} \varphi(\vec{p}) \,|\vec{p}\rangle \,,$$

is partially localised in both momentum and position space. A typical state is an element of the Schwarz space and is of the form

$$\varphi(\vec{p}) = e^{-\vec{p}^2/2m^2}.$$

4.2 Relativistic Normalisation

This problem has come from the fact that we haven't been careful to maintain Lorentz invariance. This is what we were warning about in Remark 3.1.1. So what's broken our Lorentz invariance? The answer is that

$$\langle \vec{p}|\vec{p}\rangle = (2\pi)^3\delta^{(3)}(0),$$

which is a 3-delta function, i.e. we have lost a coordinate! Now it's possible that we're lucky and this object is still Lorentz invariant, however this is not the case. We see this easily by taking a Lorentz transformation

$$p^\mu \longrightarrow \Lambda^\mu{}_\nu p^\nu.$$

If our expectation value is going to be unchanged, we would need the transformation on the states to be unitary, i.e.

$$|\vec{p}\rangle \longrightarrow U(\Lambda)\,|\vec{p}\rangle\,,$$

but there is absolutely no reason why this would be the case.

So what do we do? We need to normalise our one-particle states so that we somehow get a Lorentz invariant inner product. We do this using two tricks.

4.2.1 Trick 1

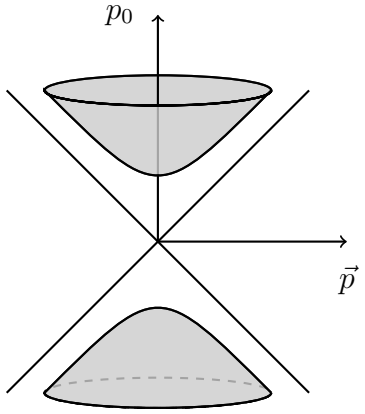
The first thing we note is that the integral $\int d^4p$ is manifestly Lorentz invariant, provided the integrand is also Lorentz invariant. Then we note that the relativistic dispersion relation

$$p_\mu p^\mu = m^2$$

is also manifestly Lorentz invariant. This can be written as

$$p_0^2 = \vec{p}^2 + m^2 = \omega_{\vec{p}}^2,$$

which has two branches, $p_0 = \pm\omega_{\vec{p}}^2$, as depicted below.



The choice of branch is Lorentz invariant, so w.l.o.g. we take $p_0 > 0$. We we know the integral

$$\int \frac{d^4p}{(2\pi)^4} 2\pi\delta(p^2 - m^2)\Theta(p_0)$$

where Θ is the Heaviside function, is Lorentz invariant. We can rewrite this in polar coorindates as

$$\int \frac{d^3\vec{p}}{(2\pi)^3}\int_0^\infty dp_0\,\delta(p_0^2 - \vec{p}^2 - m^2) = \int \frac{d^3\vec{p}}{(2\pi)^3}\int_0^\infty \frac{dp_0}{2p_0}\delta(p^0 - \omega_{\vec{p}}) = \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{2\omega_{\vec{p}}}$$

where the middle term comes from

$$\delta\big(f(x) - f(x_0)\big) = \frac{1}{|f'(x_0)|}\delta(x - x_0).$$

The thing we started with was Lorentz invariant and so we know that that the measure

$$\int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{2\omega_{\vec{p}}} \tag{4.1}$$

is a Lorentz invariant measure.

4.2.2 Second Trick

Now we use

$$\int \frac{d^3\vec{p}}{(2\pi)^3} \langle 0| a_{\vec{q}} a^\dagger_{\vec{p}} |0\rangle = \int \frac{d^3\vec{p}}{(2\pi)^3} (2\pi)^3 \delta^{(3)}(\vec{p}-\vec{q}) = 1,$$

which is clearly Lorentz invariant (it's just a number). We want to write this in terms of our Lorentz invariant measure, Equation (4.1). That is:

$$\int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{2\omega_{\vec{p}}} \langle 0| \left(\sqrt{2\omega_{\vec{q}}}\, a_{\vec{q}}\right)\left(\sqrt{2\omega_{\vec{p}}}\, a^\dagger_{\vec{p}}\right) |0\rangle ,$$

and conclude that if we normalise our one-particle states as

$$|\vec{p}\rangle = \sqrt{2\omega_{\vec{p}}}\, a^\dagger_{\vec{p}} |0\rangle = \sqrt{2E_{\vec{p}}}\, a^\dagger_{\vec{p}} |0\rangle \tag{4.2}$$

where on the second equality we have just used E, for energy, as this is standard notation, then our inner products become Lorentz invariant.

4.3 Heisenberg Picture

As we remarked above, we slaughtered our Lorentz invariance through our quantisation procedure. Even besides the above mix of the measure, we could have easily seen (or at least been cautious) by the fact that we are working in the Schrödinger picture, so our operators $\phi(\vec{x})$ only depend on space, and not time. Equally our one states evolve in time, with the evolution given by the Schrödinger equation, which implies

$$|\psi(t)\rangle_S = e^{-iH(t-t_0)} |\psi(t_0)\rangle_S .$$

So we would be closer to having manifest Lorentz invariance if we stripped the time dependence from the states and put it in the operators. This is exactly what the Heisenberg picture does. The relation between states is

$$|\psi\rangle_H = e^{iH(t-t_0)} |\psi(t)\rangle_S \qquad \Longrightarrow \qquad \frac{d}{dt} |\psi\rangle_H = 0. \tag{4.3}$$

The operators are related by

$$\mathcal{O}_H = e^{iHt} \mathcal{O}_S e^{-iHt} \tag{4.4}$$

which tells us

$$\frac{d}{dt}\mathcal{O}_H = i[H, \mathcal{O}_H]. \tag{4.5}$$

So in our field theory our Schrödinfer picture operators $\phi(\vec{x})/\pi(\vec{x})$ become the Heisenberg operators

$$\phi(x) = e^{iHt}\phi(\vec{x})e^{-iHt}, \quad \text{and} \quad \pi(x) = e^{iHt}\pi(\vec{x})e^{-iHt}, \tag{4.6}$$

where, as is standard in field theory, we have dropped the H/S subscripts and used the arguments to tell us which picture we're in (i.e. we've used $\phi(x) = \phi(\vec{x}, t)$). We adapt the commutation relations Equation (3.1) to be *equal time* commutations:

$$[\phi_a(\vec{x}, t), \phi_b(\vec{y}, t)] = [\pi_a(\vec{x}, t), \pi_b(\vec{y}, t)] = 0, \quad \text{and} \quad [\phi_a(\vec{x}, t), \pi_b(\vec{y}, t)] = i\delta^{(3)}(\vec{x}-\vec{y})\delta_{ab}. \tag{4.7}$$

We can now study to see if the Heisenberg equations, Equation (4.5), for $\phi(x)$ and $\pi(x)$ are equivalent to the Klein-Gordan equation. Using

$$H = \int d^3\vec{x}\,\frac{1}{2}\big(\pi^2 + (\nabla\phi)^2 + m^2\phi^2\big),$$

and the commutations above

$$\begin{aligned}\frac{d}{dt}\phi &= \frac{i}{2}\Big[\int d^3\vec{y}\,\pi^2(y), \phi(x)\Big] \\ &= \frac{i}{2}\int d^3y\Big(\pi(y)[\pi(y),\phi(x)] + [\pi(y),\phi(x)]\pi(y)\Big) \\ &= i\int d^3\vec{y}\,\pi(y)(-i)\delta^{(3)}(\vec{y}-\vec{x}) \\ &= \pi(x).\end{aligned}$$

That is,

$$\dot{\phi}(x) = \pi(x). \tag{4.8}$$

Exercise

Using Equation (4.5) show that

$$\dot{\pi}(x) = (\nabla^2 - m^2)\phi(x). \tag{4.9}$$

Combining Equations (4.8) and (4.9), we have

$$(\nabla^2 - m^2)\phi(x) = \dot{\pi} = \ddot{\phi},$$

or more familiarly,

$$(\partial^2 + m^2)\phi = 0,$$

the Klein-Gordan equation! We refer to this as the *correspondence principle*.

Exercise

Show that

$$e^{iHt}a_{\vec{p}}\,e^{-iHt} = e^{-iE_{\vec{p}}}a_{\vec{p}}, \quad \text{and} \quad e^{iHt}a^\dagger_{\vec{p}}\,e^{-iHt} = e^{iE_{\vec{p}}}a^\dagger_{\vec{p}}.$$

Hint: Use

$$e^{iHt}a^\dagger_{\vec{p}}\,e^{-iHt} = \sum_{n=0}^{\infty}\frac{(iHt)^n}{n!}a^\dagger_{\vec{p}}\,e^{-iHt}, \quad \text{and} \quad [H, a^\dagger_{\vec{p}}] = E_{\vec{p}}\,a^\dagger_{\vec{p}}.$$

We can find the mode expansion of $\phi(x)$ in the Heisenberg picture by starting with the Schrödinger expression

$$\phi(\vec{x}) = \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\Big(e^{i\vec{p}\cdot\vec{x}}a_{\vec{p}} + e^{-i\vec{p}\cdot\vec{x}}a^\dagger_{\vec{p}}\Big)$$

and then use Equation (4.4) and the result of the previous exercise to obtain

$$\begin{aligned}\phi(x) &= \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left(e^{-iE_{\vec{p}}t+i\vec{p}\cdot\vec{x}}a_{\vec{p}} + e^{iE_{\vec{p}}t-i\vec{p}\cdot\vec{x}}a^\dagger_{\vec{p}}\right)\\ &= \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left(a_{\vec{p}}\,e^{-ipx} + a^\dagger_{\vec{p}}\,e^{ipx}\right)\Big|_{p_0=E_{\vec{p}}}.\end{aligned} \tag{4.10}$$

You can then quickly arrive at the expression for $\pi(x)$ by using Equation (4.8), giving

$$\pi(x) = \int \frac{d^3\vec{p}}{(2\pi)^3} i\sqrt{\frac{E_{\vec{p}}}{2}}\left(a^\dagger_{\vec{p}}\,e^{ipx} - a_{\vec{p}}\,e^{-ipx}\right)\Big|_{p_0=E_{\vec{p}}}. \tag{4.11}$$

These look a lot like Equation (3.13), but now in terms of 4-vectors. This seems more manifestly Lorentz invariant!

So we see that $\phi(x)$ is a superposition of plane waves $e^{\pm ip_0 t}$, but the coefficients create and destroy particles. We can therefore think of the field as a 'hammer' which bashes the vacuum and shakes quanta out of it. That is the operator acting on the vacuum creates a particle:

$$\phi(\vec{x},0)\,|0\rangle = \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{2E_{\vec{p}}}e^{-i\vec{p}\cdot\vec{x}}\,|\vec{p}\rangle\,, \qquad \Longrightarrow \qquad \langle\vec{p}|\,\phi(\vec{x},0)\,|0\rangle = e^{-i\vec{p}\cdot\vec{x}}.$$

If we act on an n particle state instead, because $\phi(x)$ contains both creation and annihilation operators, it has a amplitude to create both $(n+1)$ and $(n-1)$ particle states.

4.4 Causality

As we have mentioned above, in order for our theory to be causal, we want spacelike separated operators to commute, that is

$$[\mathcal{O}_1(x), \mathcal{O}_2(y)] = 0 \qquad \forall (x-y)^2 < 0.$$

This statement is obviously made in the Heisenberg picture, as the operators are functions of spacetime, not just space. So we need to study the commutator between two fields $[\phi(x), \phi(y)]$, but where now we don't impose equal time, i.e. $x^0 \neq y^0$ in general. We define

$$\Delta(x-y) := [\phi(x), \phi(y)] \tag{4.12}$$

Plugging in the definition Equation (4.10), we have

$$\begin{aligned}\Delta(x-y) &= \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{d^3\vec{q}}{(2\pi)^3}\frac{1}{2\sqrt{E_{\vec{p}}E_{\vec{q}}}}\left[a_{\vec{p}}\,e^{ipx} + a^\dagger_{\vec{p}}\,e^{-ipx}\,, a_{\vec{q}}\,e^{iqy} + a^\dagger_{\vec{q}}\,e^{-iqy}\right]\\ &= \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{d^3\vec{q}}{(2\pi)^3}\frac{1}{2\sqrt{E_{\vec{p}}E_{\vec{q}}}}\left(e^{-ipx+ipy} + e^{ipx-ipy}\right)(2\pi)^3\delta^{(3)}(\vec{p}-\vec{q})\\ &= \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{2E_{\vec{p}}}\left(e^{-ip(x-y)} - e^{ip(x+y)}\right)\end{aligned}$$

The first thing we note is that this expression is Lorentz, as we have the correct measure, Equation (4.1), and the exponential contain 4-vectors. We can therefore choose a specific reference frame in order to evaluate its timelike/spacelike properties.

(i) First look at the spacelike separations, i.e. $(x-y)^2 > 0$. Let's work in a reference frame where $(x-y) = (0, \vec{a})$, then we have

$$\Delta(0,\vec{a}) = \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{E_{\vec{p}}}\Big(e^{i\vec{p}\cdot\vec{a}} - e^{-i\vec{p}\cdot\vec{a}}\Big).$$

Now note that $E_{\vec{p}} = \sqrt{\vec{p}^2 + m^2}$ and so $E_{-\vec{p}} = E_{\vec{p}}$, and also that the measure is invariant under $\vec{p} \longrightarrow -\vec{p}$, as you get a minus sign but the integration limits flip. So we can freely change the sign in the second exponential, and then the two terms cancel, i.e.

$$\Delta(0,\vec{a}) = 0.$$

So our theory is causal!

(ii) Let's also check that it doesn't vanish for timelike separations (which would be an obvious problem). Again can pick a frame, so simply choose $(x-y) = (t, 0)$, then we have

$$\Delta(t,0) = \int \frac{d^3\vec{p}}{(2\pi)^3}\Big(e^{-iE_{\vec{p}}\,t} - e^{iE_{\vec{p}}\,t}\Big) \sim \int d^3\vec{p}\,\sinh\big(\sqrt{\vec{p}^2+m^2}\,t\big),$$

which does not vanish (its a Bessel function).

Remark 4.4.1. Note that the objects on the right-hand side of Equation (4.12) are operators. However we've just shown that it comes out to be a $\mathbb{C}$-number. Essentially what we have done is sandwich it in between two vacuum states, as per Remark 3.2.5. In other words we really should have written

$$\langle 0|\,[\phi(x),\phi(y)]\,|0\rangle\,,$$

but its common field theory notation to drop the bra and ket. We will continue to do this next lecture, but this is just another remark to go with Remark 3.2.5 to remind us that everything needs to be sandwiched.

5 | Complex Scalar Field & Propagators

5.1 Complex Scalar Field

Recall the *classical* Lagrangian for the complex scalar field, Equation (2.13),

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

We treat ψ and ψ^* as independent fields and get the equations of motion

$$\big(\partial^2 + m^2\big)\psi = 0, \quad \text{and} \quad \big(\partial^2 + m^2\big)\psi^* = 0.$$

When we quantise this theory we again treat each field as an independent degree of freedom. In quantising we replace the complex conjugate with the Hermitian conjugate. We want to obtain an expansion equivalent to that of Equation (3.13), but we have to note something: the field ψ is *not* real and so the corresponding quantum operator will *not* be Hermitian. We therefore need to have different labels for our creation and annihilation operators in the expansion for ψ. We then get the expansion for $\psi^\dagger$ by taking the Hermitian conjugate. The result is

$$\begin{aligned} \psi &= \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_{\vec{p}}}} \Big(b_{\vec{p}}\, e^{i\vec{p}\cdot\vec{x}} + c^\dagger_{\vec{p}}\, e^{-i\vec{p}\cdot\vec{x}} \Big) \\ \psi^\dagger &= \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_{\vec{p}}}} \Big(c_{\vec{p}}\, e^{i\vec{p}\cdot\vec{x}} + b^\dagger_{\vec{p}}\, e^{-i\vec{p}\cdot\vec{x}} \Big) \end{aligned} \tag{5.1}$$

Remark 5.1.1. Note that the dagger comes with the negative exponential. This is because we want the particles we create to have positive frequency, which corresponds to $e^{-i\vec{p}\cdot\vec{x}}$.

Classically we had the result $\pi = \partial\mathcal{L}/\partial\dot{\psi}$, Equation (2.15). We turn this into a quantum relation, and using $\partial\mathcal{L}/\partial\dot{\psi} = \dot{\psi}^*$, we obtain the operator expressions

$$\begin{aligned} \pi &= \int \frac{d^3\vec{p}}{(2\pi)^3} i\sqrt{\frac{E_{\vec{p}}}{2}} \Big(- c_{\vec{p}} e^{i\vec{p}\cdot\vec{x}} + b^\dagger_{\vec{p}} e^{-i\vec{p}\cdot\vec{x}} \Big) \\ \pi^\dagger &= \int \frac{d^3\vec{p}}{(2\pi)^3} (-i)\sqrt{\frac{E_{\vec{p}}}{2}} \Big(b_{\vec{p}} e^{i\vec{p}\cdot\vec{x}} - c^\dagger_{\vec{p}} e^{-i\vec{p}\cdot\vec{x}} \Big) \end{aligned} \tag{5.2}$$

Remark 5.1.2. I have chosen to write the above equations so that the creation operator always appears to the far most right. This means that the bs and cs move around, obviously. There are two reasons I am writing them like this: it's how Prof. Spannowsky writes them, and it then they take the same form as Equation (3.13). Obviously it doesn't matter which order I write them in, I just make this remark because Prof. Tong writes them so that the b term is always to the left (see pages 33/34 of his notes).

We want to keep the same form for the *equal time* commutation relations as for the real scalar field, i.e.

$$[\psi(\vec{x}), \pi(\vec{y})] = i\delta^{(3)}(\vec{x}-\vec{y}), \qquad [\psi(\vec{x}), \pi^\dagger(\vec{y})] = 0, \tag{5.3}$$

and similarly with Hermitian conjugates everywhere,[1] and the vanishing ones

$$[\psi(\vec{x}), \psi(\vec{y})] = 0 = [\pi(\vec{x}), \pi(\vec{y})].$$

Exercise

Show that the above commutation relations hold if, and only if, we also have

$$[b_{\vec{p}}\, b^\dagger_{\vec{p}}] = (2\pi)^3\delta^{(3)}(\vec{p}-\vec{q}), \quad \text{and} \quad [c_{\vec{p}}\, c^\dagger_{\vec{p}}] = (2\pi)^3\delta^{(3)}(\vec{p}-\vec{q}), \tag{5.4}$$

and all others vanishing, i.e.

$$[b_{\vec{p}}\, b_{\vec{p}}] = 0 = [b_{\vec{p}}\, c_{\vec{p}}]$$

etc.

Exercise

Derive the following expression for the Hamiltonian,

$$H = \frac{1}{2}\int \frac{d^3\vec{p}}{(2\pi)^3} E_{\vec{p}} \left(b^\dagger_{\vec{p}} b_{\vec{p}} + c^\dagger_{\vec{p}} c_{\vec{p}}\right). \tag{5.5}$$

Hint: Derive the Hamiltonian density and use Equations (5.1) and (5.2). You will need to drop an infinity term as we did for the real scalar field.

In our particle interpretation, Equation (5.5) tells us that the complex theory contains two particles of the same energy, and therefore mass. We can show that they also both have vanishing spin. The states of our Hilbert space are made up of products of these two particle types, we use a semi-colon in the bra/ket to indicate the different types. For example

$$|\vec{p}_1, \vec{p}_2; \vec{q}_1, \vec{q}_2\rangle := b^\dagger_{\vec{p}_1}\, b^\dagger_{\vec{p}_2} c^\dagger_{\vec{q}_1}\, c^\dagger_{\vec{q}_2}\, |0\rangle\,.$$

We refer to the particles created by c as *anti-particles* and the particles created by b as *particles*. Note that for the real scalar field we have $b = c$ and so the 'particle is its own antiparticle'.

[1] Note you can simply use $[A, B]^\dagger = -[A^\dagger, B^\dagger]$, along with the is on the right-hand, to get these results.

The classical complex field conserved current, Equation (2.14), can be upgraded to the quantum theory and from that we can obtain the conserved charge

$$Q = \int \frac{d^3\vec{p}}{(2\pi)^3}\left(c^\dagger_{\vec{p}}c_{\vec{p}} - b^\dagger_{\vec{p}}b_{\vec{p}}\right) = N_c - N_b, \tag{5.6}$$

which says that the number of antiparticles minus the number of particles is conserved *locally*. Of course in the free theory we have that N_c and N_b are separatly conserved, so this statement is nothing new. However, as we keep teasing, in the interacting theory the number of each type of particle will no longer be a conserved charge, but Equation (5.6) will still hold. The locality of this result will translate into the fact that particles and antiparticles are always created in pairs, and they can annihilate to produce something different (e.g. a photon).

Remark 5.1.3. As we just said, these are spin-0 particles and so do not correspond to Fermions (i.e. electrons etc), but they give us a taste of what's to come. We can refer to this complex scalar field as a 'poor man's Fermion'.

5.2 Propagators

At the end of the last lecture we checked to see if our theory was causal by checking that the operators commuted for spacelike separation. We could have worded this question differently, and perhaps more 'particle physicsy'. We could have asked "what is the probability for a particle to propagate from point y to point x?" If x and y are spacelike separated, we would want the answer to be a big fact zero, otherwise the particles would be travelling faster than light. The propagation is given by the 2-point function $\langle 0|\, A(x)A(y)\,|0\rangle$ where A is our field in the Heisenberg picture. So we need to consider this calculation theory by theory.

5.2.1 Real Scalar Field

First let's consider the real scalar field. Our propagator is

$$\begin{aligned}
D(x-y) &= \langle 0|\,\phi(x)\phi(y)\,|0\rangle \\
&= \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{d^3\vec{q}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\frac{1}{\sqrt{2E_{\vec{q}}}}\langle 0|\left(a_{\vec{p}}\,e^{-ipx} + a^\dagger_{\vec{p}}\,e^{ipx}\right)\left(a_{\vec{q}}\,e^{-iqy} + a^\dagger_{\vec{q}}\,e^{iqy}\right)|0\rangle \\
&= \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{d^3\vec{q}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\frac{1}{\sqrt{2E_{\vec{q}}}}\langle 0|\,a_{\vec{p}}\,a^\dagger_{\vec{q}}|0\rangle\, e^{-ipx+iqy} \\
&= \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{2E_{\vec{p}}}e^{-ip(x-y)},
\end{aligned}$$

where we have used

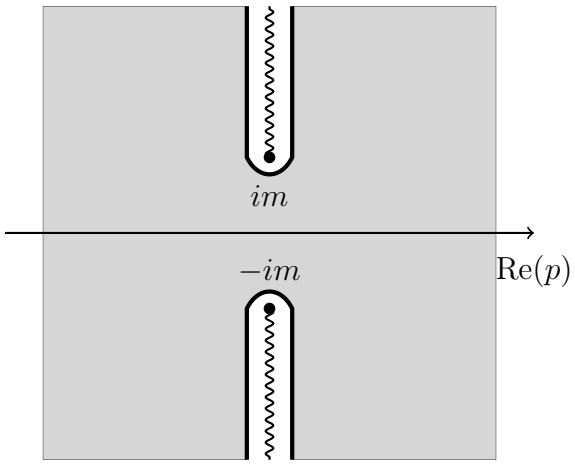
$$\langle 0|\,a^\dagger_{\vec{p}} = 0 = a_{\vec{q}}\,|0\rangle\,, \quad \text{and} \quad \langle 0|\,a_{\vec{p}}\,a^\dagger_{\vec{q}}|0\rangle = (2\pi)^3\delta^{(3)}(\vec{p}-\vec{q}).$$

Again this result is Lorentz invariant and so we can pick any frame we like. We want to check spacelike separations $(x-y)^2 < 0$, so choose a frame with $(x-y) = (0,\vec{r})$, and work in polar

coordinates, $d^3\vec{p} = p^2 \sin^2\theta d\theta d\varphi dp$. This gives us

$$\begin{aligned} D(x-y) &= \frac{2\pi}{(2\pi)^3}\int_0^\infty dp \frac{p^2}{2E_p}\frac{e^{ipr}-e^{-ipr}}{ipr} \\ &= -\frac{i}{2(2\pi)^2 r}\int_{-\infty}^{\infty} dp \frac{pe^{ipr}}{\sqrt{p^2+m^2}} \end{aligned}$$

This is an integral in the complex plane with two poles at $p = \pm im$. We therefore need to take two branch cuts as indicated in the figure below. We do the integral over the shaded region and extend out to infinity.



We then define $\rho = -ip$ and obtain

$$D(x-y) = \frac{1}{4\pi^2 r}\int_m^\infty d\rho \frac{\rho e^{-\rho r}}{\sqrt{\rho^2-m^2}},$$

which in the limit $r \to \infty$ (which is *very* spacelike separated) gives

$$D(x-y) \sim e^{-m|\vec{x}-\vec{y}|},$$

where we've put $r = |\vec{x}-\vec{y}|$ back in.

Ah this does not appear good... we've shown that the amplitude for a particle to propagate on a spacelike curve decays exponentially, which is small but non-zero. What was the difference to what we did at the end of last lecture? Well last lecture we had the (sandwiched) *commutator*, which in terms of the propagators is

$$[\phi(x), \phi(y)] = D(x-y) - D(y-x).$$

But the above formula has $|\vec{x}-\vec{y}|$, so it doesn't change when we switch $x \longleftrightarrow y$. This is why the result vanished last lecture. So what does this correspond to in our particle propagation terms? Well because we're considering a spacelike path, their is no Lorentz invariant way to order events, so we have to consider both the propagation $x \to y$ and $y \to x$, but these have the same amplitude, and so cancel.

5.2.2 Complex Scalar Field

What about for the scalar field? Well first consider the exercise

Exercise

Define

$$\begin{aligned} D_b(x-y) &:= \langle 0|\, \psi(x)\psi^\dagger(y)\, |0\rangle\,, \\ D_c(x-y) &:= \langle 0|\, \psi^\dagger(x)\psi(y)\, |0\rangle\,. \end{aligned} \tag{5.7}$$

Then using the relevant definitions and commutators show these both give

$$\int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{2E_{\vec{p}}}e^{-ip(x-y)}$$

Hint: The subscripts b/c should become justified relatively quickly.

For the complex scalar case we therefore have

$$\left[\psi(x), \psi^\dagger(y)\right] = D_b(x-y) - D_c(y-c),$$

and a similar calculation to the real scalar gives that this vanishes for spacelike separations. So we give the same particle propagation tale, but now with an interesting difference. As we see in Equation (5.7), $D_b(x,y)$ corresponds to an *antiparticle* propagating from $y \to x$ whereas $D_c(y-x)$ corresponds to a *particle* propagating from $x \to y$. So now the cancellation in amplitude comes from a particle going one way while an antiparticle going the other way! So causality requires that every particle have a corresponding antiparticle with the same mass but opposite quantum numbers. This isn't actually a new interpretation, it's just that for the real scalar field the particle is its own antiparticle so we didn't notice.

5.2.3 Feynman Propagator

As we shall see, for interacting field theories perhaps the most important object is the so-called *Feynman propagator*:

$$\Delta_F(x-y) = \langle 0|\, \mathcal{T}\phi(x)\phi(y)\, |0\rangle = \begin{cases} D(x-y) & x^0 > y^0 \\ D(y-x) & y^0 > x^0 \end{cases}, \tag{5.8}$$

where the cases on the right hand side defines $\mathcal{T}$, the *time ordering* operator:

$$\begin{aligned} \mathcal{T}\big(\mathcal{O}_1(t_1)\mathcal{O}_2(t_2)\big) &= \mathcal{O}_1(t_1), \mathcal{O}_2(t_2)\Theta(t_1-t_2) + \mathcal{O}_2(t_2), \mathcal{O}_1(t_1)\Theta(t_2-t_1) \\ &= \begin{cases} \mathcal{O}_1(t_1), \mathcal{O}_2(t_2) & t_1 > t_2 \\ \mathcal{O}_1(t_2), \mathcal{O}_2(t_1) & t_2 > t_1 \end{cases}. \end{aligned} \tag{5.9}$$

Physically what this tells us is that first the particle is created and then destroyed. It will be incredibly useful for us to write this expression as a 4-integral, and this comes in the form of *Feynman's trick*.

Claim 5.2.1. We can write the Feynman propagator as

$$\Delta_F(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\epsilon} e^{-ip(x-y)}, \tag{5.10}$$

where $\epsilon > 0$ but infinitesimal.

Proof. We prove this claim by taking a complex contour integral and using Cauchy's theorem. First let's rewrite the right-hand side of Equation (5.10) slightly:

$$\begin{aligned}
\frac{1}{p^2 - m^2 + i\epsilon} &= \frac{1}{p_0^2 - \vec{p}^2 - m^2 + i\epsilon} \\
&= \frac{1}{p_0^2 - E_{\vec{p}}^2 + i\epsilon} \\
&= \frac{1}{p_0^2 - (E_{\vec{p}} - i\epsilon)^2} \\
&= \frac{1}{p_0 - (E_{\vec{p}} - i\epsilon)} \frac{1}{p_0 + (E_{\vec{p}} - i\epsilon)}
\end{aligned}$$

where to get to the penultimate line we used the fact that ϵ is infinitesimal, and relabelled $\epsilon = 2E_{\vec{p}}\epsilon$. Explicitly, we have

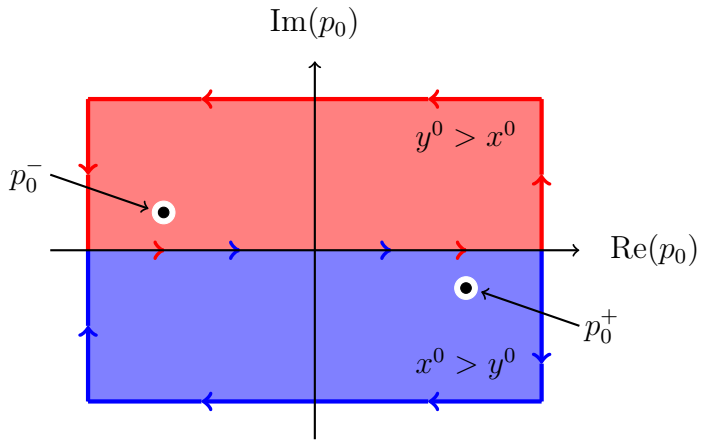
$$(E_{\vec{p}} - i\epsilon)^2 = E_{\vec{p}}^2 - 2iE_{\vec{p}}\,\epsilon - \epsilon^2$$

then we drop the ϵ^2 term and, because $E_{\vec{p}} > 0$, we can redefine $\epsilon \to 2E_{\vec{p}}\epsilon$ which is still positive and infinitesimal.

We now have an integral with poles

$$p_0^{\pm} = \pm(E_{\vec{p}} - i\epsilon)$$

as indicated in the following diagram.



The shaded regions are meant to indicate how we close our contours for the relevant $(x^0 - y^0)$ sign. The arrows are meant to indicate which way round we close each contour (so $x^0 > y^0$ is clockwise and $y^0 > x^0$ is anticlockwise). We pick up the relative poles as indicated.

Let's consider the case $x^0 > y^0$, then we pick up p_0^+ pole, and using Cauchy's theorem,

$$\oint dz \frac{f(z)}{(z-z_0)} = (2\pi i) f(z_0),$$

we get the residue $-1/2E_{\vec{p}}$, where the minus sign comes from that fact that we're doing a clockwise integral,[2] and obtain

$$\int \frac{d^4}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\epsilon} e^{-ip(x-y)} = \int \frac{d^3p}{(2\pi)^3} \frac{2\pi i}{2\pi} \frac{i}{2E_{\vec{p}}} e^{-iE_{\vec{p}}(x^0-y^0)+i\vec{p}\cdot(\vec{x}-\vec{y})},$$

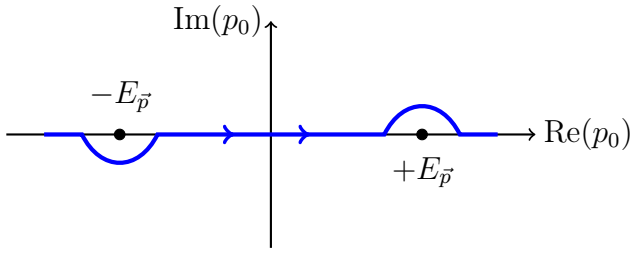
but the right-hand side (after cancelling) is just $D(x-y)$. □

Exercise

Finish the proof above. That is show

$$\Delta_F(x-y) = D(y-x) \qquad \text{if} \qquad y^0 > x^0.$$

Remark 5.2.2. If you are not 100% comfortable with the idea of putting the $i\epsilon$ in the denominator, you can obtain the same result but now the poles like on the real axis. The contour you need to take to get the Feynman propagator is drawn below.



Exercise

Show that the diagram in the above remark will lead to the same result for the Feynman propagator. *Hint: If you get stuck, this is how Prof. Tong does it. But I advise you try it first using the calculation done in the proof above for guidance.*

The Feynman propagator is, in fact, also a Green's function[3] for the Klein-Gordan equation. We see this easily using Equation (5.10) (with $\epsilon = 0$, as we don't need to use a contour integral here):

$$\begin{aligned}
(\partial^2 + m^2)\Delta_F(x-y) &= \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2} (\partial^2 + m^2) e^{-ip(x-y)} \\
&= \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2} \big(-p^2 + m^2 \big) e^{-ip(x-y)} \\
&= -i \int \frac{d^4p}{(2\pi)^4} e^{-ip(x-y)} \\
&= -i\delta^{(4)}(x-y).
\end{aligned}$$

[2] If this doesn't make sense, basically in Cauchy's theorem you use the anticlockwise contour integral.
[3] Basically a function that gives a delta function when acted on by a differential operator.

This is a really nice result as it allows us to 'invert' the Klein-Gordan equation, i.e. turn it from a differential equation to a integral one.[4]

5.3 Building Interacting Theories

As we have seen, in free theories our equations of motion are linear in the fields, e.g. the Klein-Gordan equation

$$(\partial^2 + m^2)\phi = 0.$$

We solved these theories by a Fourier analysis and saw that all the different modes decoupled. That is we never had a $a_{\vec{p}}a_{\vec{q}}$ term in our Hamiltonian etc. This lead to us showing that the number operator in such theories in conserved, Equation (3.19). This is why we called them free theories.

Obviously the in the real world things do interact, and we have hinted a few times that this will lead to the number operator not being conserved anymore. We need some way, then, to construct interacting theories and the rest of this lecture is dedicated to exactly that.

We get interactions in our theory when the equations of motion are not linear in the fields. Recalling that the equations of motion come from varying the Lagrangian (i.e. they're the Euler-Lagrange equations). So if we want interaction terms in our equations of motion, we're going to have to add interaction terms into the Lagrangian:

$$\mathcal{L} = \mathcal{L}_{\text{free}} + \mathcal{L}_{\text{int}}.$$

This will also lead to interaction terms in the Hamiltonian, which we want.

However, in QFT we cannot just arbitrarily change our Lagrangian. That is, we need to keep our 'suitable Lagrangian' conditions we introduced previously. We shall state them again here, but now in terms of the additional $\mathcal{L}_{\text{int}}$:

(i) We only want local interactions, in order to preserve causality. That is something of the form

$$\int d^3y \phi(x)\phi^2(y)$$

is *not* allowed.

(ii) $\mathcal{L}_{\text{int}}$ is a Lorentz scalar.

(iii) $\mathcal{L}_{\text{int}}$ respects the internal symmetries. For example, if our free theory was U(1) invariant (i.e. $\psi \to e^{i\alpha}\psi$) then we cannot have something like $\psi\psi$ in $\mathcal{L}_{\text{int}}$.

(iv) Renormalisability. We explained this before, but for a reminder, in 4-dimensions, we would have to truncate the following Lagrangian at the ϕ^4 term, as $[\lambda_{i>4}] < 0$ which is non-renormalisable.

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda_3}{3!}\phi^3 - \frac{\lambda_4}{4!}\phi^4 - \frac{\lambda_5}{5!}\phi^5 - ...$$

Note here we can see $\mathcal{L}_{\text{free}}$ as the real scalar field.

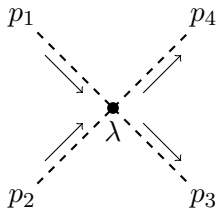
[4] For more info on why this is the case, just Google what a Green's function is.

5.3.1 Examples

Example 5.3.1. As an example consider the popular ϕ^4 theory. This has Lagrangian

$$\mathcal{L} = \underbrace{\frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2}_{\mathcal{L}_{\text{free}}} - \underbrace{\frac{\lambda}{4!}\phi^4}_{\mathcal{L}_{\text{int}}}. \tag{5.11}$$

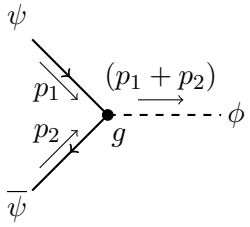
In terms of Feynman diagrams (to come shortly) the interaction Lagrangian will correspond to something of the form



Example 5.3.2. In the above theory we only had one field ϕ, but we can have (and we will definitely need) interactions between different fields. As an example, we have the *scalar Yukawa* theory, with Lagrangian

$$\mathcal{L} = \underbrace{\frac{1}{2}(\partial\phi)^2 - m^2\phi^2}_{\text{Klein-Gordan}} + \underbrace{(\partial_\mu\psi^*)(\partial^\mu\psi) - M^2\psi^*\psi}_{\text{Complex Scalar}} - \underbrace{g\psi^*\psi\phi}_{\text{interaction}}. \tag{5.12}$$

This is the interaction between a Klein-Gordan field of mass m and a complex scalar field of mass M.



Exercise

Find the equations of motion for Equations (5.11) and (5.12).

As the result of the previous exercise shows, our equations of motion are non-linear and so we can no longer solve them by Fourier decomposition. This is exactly what we wanted, and will lead to different modes coupling together. As we will see, we solve these theories in a perturbative way; we assume the couplings λ/g are small and expand around the free theory. We are saved from hideously long expression by the famous *Feynman diagrams*. These things really are a field theorist's best friend.[5]

[5]Sorry Harmonic Oscillators, you've been pipped at the post.

6 | Interaction Picture & Dyson's Formula

So far we have used the Schrödinger picture and the Heisenberg picture. Recall that the difference between these two pictures was where the time dependence sat: for the Schrödinger picture the states are time dependant, whereas in the Heisenberg picture it is the operators that are time dependant. There is a third picture, the *interaction picture*, which is a sort of hybrid of the other two. As the name suggests, it is useful when studying interacting systems. What we do is consider small perturbations from some well-understood Hamiltonian, H_0. That is, we write the Hamiltonian as

$$H = H_0 + H_{\text{int}}.$$

We treat H_{int} like a Schrödinger Hamiltonian, and H_0 like a Heisenberg one. That is, the time dependence of the states is governed by H_{int} while the time dependence of the operators is governed by H_0. As we did for the other pictures, we denote the states/operators in the interaction picture with a subscript I, and they are given as follows:

$$|\psi(t)\rangle_I = e^{iH_0 t}\,|\psi(t)\rangle_S\,, \quad \text{and} \quad \mathcal{O}_I(t) = e^{iH_0 t}\mathcal{O}_S e^{-iH_0 t}. \tag{6.1}$$

From this we ca work out how the states $|\psi\rangle_I$ depend on time. We simply use the Schrödinger equation

$$i\frac{d}{dt}\,|\psi\rangle_S = H_S\,|\psi\rangle_s$$

along with $H_S = (H_0 + H_{\text{int}})_S$. We get

$$\begin{aligned} i\frac{d}{dt}\Big(e^{-iH_0 t}\,|\psi\rangle_I\Big) &= \big(H_0 + H_{\text{int}}\big)_S e^{-iH_0 t}\,|\psi\rangle_I \\ H_0\Big(e^{-iH_0 t}\,|\psi\rangle_I\Big) + e^{-iH_0 t}\left(i\frac{d}{dt}\,|\psi\rangle_I\right) &= \big(H_0 + H_{\text{int}}\big)_S e^{-iH_0 t}\,|\psi\rangle_I\,, \end{aligned}$$

so we get

$$i\frac{d}{dt}\,|\psi\rangle_I = (H_{\text{int}})_I\,|\psi\rangle_I\,, \qquad \text{with} \qquad (H_{\text{int}})_I := e^{iH_o t}(H_{\text{int}})_S e^{-iH_0 t}. \tag{6.2}$$

Notation. From now on we shall simply write $H_I(t)$ instead of $(H_{\text{int}})_I(t)$ for obvious reasons. However we should be careful not to confuse $H_I(t)$ with H_{int} itself: the former is the latter in the interaction picture, while the latter is a small perturbation to our Hamiltonian. Note in particular that the former is *always* a function of time whereas the latter is time independent in the Schrödinger picture.

Exercise

Convince yourself that $[H_I(t_1), H_I(t_2)] \neq 0$ in general.

Ok this looks great but it's useless unless we can actually solve it, so what do we do? The answer is we derive what is known as *Dyson's formula*.

6.1 Dyson's Formula

We start with the ansatz[1]

$$|\psi(t)\rangle_I = U(t,t_0)\,|\psi(t_0)\rangle_I\,,$$

where $U(t,t_0)$ is a *unitary time evolution operator* satisfying

$$U(t_1,t_2)U(t_2,t_3) = U(t_1,t_3), \quad \text{and} \quad U(t,t) = \mathbb{1}.$$

Using Equation (6.2), we can easily show

$$i\frac{d}{dt}U(t,t_0) = H_I(t)U(t,t_0). \tag{6.3}$$

Now if we were just considering wavefunctions here instead of operators then the solution to this differential equation would just be

$$U(t,t_0) = \exp\left(-i\int_{t_0}^{t} H_I(t')dt'\right). \tag{6.4}$$

However we are considering non-commuting operators, and so we have to be careful about ordering. To be more explicit, consider the Taylor expansion of the above formula, we get

$$U(t,t_0) = \mathbb{1} - i\int_{t_0}^{t} H_I(t')dt' + \frac{(-i)^2}{2}\left(\int_{t_0}^{t} H_I(t')dt'\right)^2 + ...,$$

and we want the derivative to leave us with

$$i\frac{d}{dt}U(t,t_0) = H_I(t)\left[\mathbb{1} - i\int_{t_0}^{t} H_I(t')dt' + \frac{(-i)^2}{2}\left(\int_{t_0}^{t} H_I(t')dt'\right)^2 + ...\right].$$

However if we take the derivative, we get terms like

$$-\frac{1}{2}\left(\int_{t_0}^{t} H_I(t')dt'\right)H_I(t) - \frac{1}{2}H_I(t)\left(\int_{t_0}^{t} H_I(t')dt'\right).$$

The second term is the kind of thing we want but the first term has the $H_I(t)$ on the wrong side! We cannot simply 'pull it through' as the operators are non-commuting. So we need to alter Equation (6.4) somehow to account for this. The obvious[2] thing to try is taking a time ordering.

[1] As with all good physics derivations do, our guess will turn out to miraculously be correct...

[2] Or perhaps only obvious when you know its the answer...

Claim 6.1.1. The solution of Equation (6.3) is given by *Dyson's formula*

$$U(t,t_0) = \mathcal{T}\exp\left(-i\int_{t_0}^{t} H_I(t')dt'\right), \tag{6.5}$$

where $\mathcal{T}$ is the time ordering operator defined above, Equation (5.9).

Remark 6.1.2. Before presenting the proof, let's make a quick remark. Recall that if the result of an integral is finite, then we can use[3]

$$\left(\int_a^b f(x)dx\right)^2 = \int_a^b dx\int_a^b dy f(x)f(y).$$

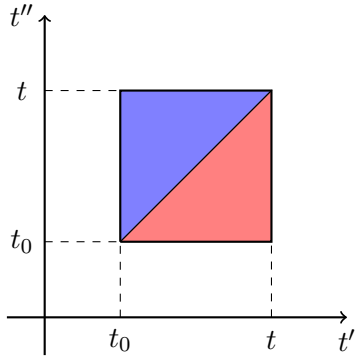
Now in QFT we consider bounded operators, and so the integral over the Hamiltonian is finite, so we can turn our exponential expansion above into a sum of higher order integrals, rather then powers of integrals. This is obviously a *massive* help. In particular, our expansion becomes

$$U(t,t_0) = \mathbb{1} - i\int_{t_0}^{t} dt' H_I(t') + \frac{(-i)^2}{2}\left[\int_{t_0}^{t} dt'\int_{t'}^{t} dt'' H_I(t'')H_I(t') + \int_{t_0}^{t} dt'\int_{t_0}^{t'} dt'' H_I(t')H_I(t'')\right] + \ldots$$

Note the integration limits on the quadratic terms. The first one is the condition that $t'' > t'$ as the lower limit for t'' is t', while the second one is the condition that $t' > t''$ as the upper limit for t'' is t'. These two terms are in fact equal as

$$\begin{aligned}\int_{t_0}^{t} dt'\int_{t'}^{t} dt'' H_I(t'')H_I(t') &= \int_{t_0}^{t} dt''\int_{t_0}^{t''} dt' H_I(t'')H_I(t') \\ &= \int_{t_0}^{t} dt'\int_{t_0}^{t'} dt'' H_I(t')H_I(t''),\end{aligned}$$

where the left had side says t'' has lower limit t' while the first expression on the right hand side says t' has supper limit t''. These are clearly the same statement. The last expression is obtained by simple change of variables $t' \longleftrightarrow t''$. We can also see this diagrammatically as the two following triangle areas: the blue area corresponds to the left hand side above (i.e. $t'' > t'$), whereas the red area is the final expression (i.e. $t' > t''$).



[3]If you haven't seen this before, it's worth convincing yourself why this is true.

So our expansion becomes

$$U(t,t_0) = \mathbb{1} - i\int_{t_0}^{t} dt' H_I(t') + (-i)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' H_I(t')H_I(t'') + ...$$

Proof. Given the final result from the above remark, the proof of Dyson's formula is almost trivial. We simply note that our left most integral is *always* the one with upper limit t, and so when we act with the time derivative we are always just getting the latest time result (as the expression is time ordered). We can therefore pull it out of the time ordering and get the result. In terms of maths, that is

$$\begin{aligned} i\frac{d}{dt}\left[\mathcal{T}\exp\left(-i\int_{t_0}^{t} H_I(t')dt'\right)\right] &= \mathcal{T}\left[H_I(t)\exp\left(-i\int_{t_0}^{t} H_I(t')dt'\right)\right] \\ &= H_I(t)\left[\mathcal{T}\exp\left(-i\int_{t_0}^{t} H_I(t')dt'\right)\right], \end{aligned}$$

which is the result we want. □

Dyson's formula might look a bit daunting, but when we remember that in the interaction picture we're considering small perturbations, we can truncate our expansion and it becomes a lot nicer. Now, recalling that we said in order to study interacting filed theories we were going to take a perturbation around a free theory, we see why Dyson's formula is so useful. It will allow us to use something called *Wick's theorem* and then define the all-so-useful Feynman diagrams. But first let's talk about scattering.

6.2 Scattering

Definition. We define the *scattering matrix*, or *S-matrix* for short, to be the amplitude to go from some initial state $|i\rangle$ to a final state $|f\rangle$:

$$S_{fi} := \langle f|\, S\, |i\rangle := \lim_{t_\pm\to\infty} \langle f(t_+)|\, U(t_+,t_-)\, |i(t_-)\rangle_I\,, \tag{6.6}$$

where the states are in the interaction picture

When we study scattering in QFT we make one big assumption:

> The initial and final states, which we collectively call *asymptotic states*, are eigenstates of the free Hamiltonian. .

The basic idea is we want to say that the initially the fields/particles are so far apart that they do not interact with each other at all, and so do not feel in the interaction Hamiltonian. As Prof. Tong points out,[4] at first this seems like a reasonable thing to do, but then we realise it's actually not as solid an assumption as we might think. We shall illustrate one of the problems here.

[4] Pages 54-55.

Consider the case of a real Klein-Gordan field, ϕ, in some potential $V(x)$, the Lagrangian is

$$\mathcal{L} = \mathcal{L}_{KG} - V(x)\phi,$$

then the equations of motion become

$$(\partial^2 + m^2)\phi = -V(x).$$

So our asymptotic states condition tells us that we need $V(x) \to 0$ as $|x| \to \infty$. This still seems somewhat reasonable. However now let's consider the interaction term to be a ϕ^4 term. Recall the Lagrangian is

$$\mathcal{L} = \mathcal{L}_{KG} - \frac{\lambda}{4!}\phi^4,$$

and you should have shown in the exercise that the equation of motion is

$$(\partial^2 + m^2)\phi = -\frac{\lambda}{3!}\phi^3.$$

This poses a more significant problem, as we can't 'turn off' our interaction at $|t| \to \infty$; the field is defined everywhere! Despite this we continue on with our assumption that the asymptotic states are eigenstates of the free theory. We therefore rewrite Equation (6.7) as

$$S_{fi} := \langle f | S | i \rangle := \lim_{t_\pm \to \infty} \langle f(t_+) | U(t_+, t_-) | i(t_-) \rangle , \tag{6.7}$$

where the states are now eigenstates for the free theory.

We can actually define a different matrix, called the *transition matrix*, by

$$T_{fi} := \langle f | (S - \mathbb{1}) | i \rangle ,$$

which removes the $\mathbb{1}$ term in the Taylor expansion of $U(t, t_0)$.

6.2.1 Particle Decay

Ok let's actually derive our first interaction result. Consider scalar Yakawa theory, and recall that the Lagrangian is

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - m^2\phi^2 + \partial_\mu\psi^*\partial^\mu\psi - M\psi^*\psi - g\psi^*\psi\phi.$$

This has[5]

$$H_{\text{int}} = g \int d^3x \, \psi_x^\dagger \psi_x \phi_x.$$

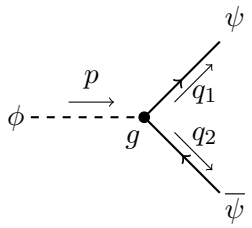
We're going to 'flip' the diagram we gave for this at the end of last lecture, and consider the decay process

$$\phi(p) \to \overline{\psi}(q_2)\psi(q_1),$$

where ϕ is the particle created by $a^\dagger$, ψ the one created by $b^\dagger$ and $\overline{\psi}$ the *anti*particle created by $c^\dagger$. The diagram[6] looks like

[5] This notation is meant to mean everything in the bracket is a function of x. Just want to save writing all the brackets and clutterig notation.

[6] We will understand how to draw these soon.



Our asymptotic states are given by

$$|i\rangle = \sqrt{2E_{\vec{p}}}\, a^\dagger_{\vec{p}}\, |0\rangle =: |\phi(p)\rangle$$
$$|i\rangle = \sqrt{4E_{\vec{q}_1} E_{\vec{q}_2}}\, b^\dagger_{\vec{q}_1}\, c^\dagger_{\vec{q}_2}\, |0\rangle =: \left|\psi(q_1)\overline{\psi}(q_2)\right\rangle$$

So what is our scattering amplitude? Well we are just considering the first order expansion, i.e. one power of H_I (higher powers would correspond to terms like $\overline{\psi}(p_1)\psi(p_2) \to \phi(q) \to \overline{\psi}(p_3)\psi(p_4)$ etc), so we have

$$U(-\infty, \infty) = \mathbb{1} - ig \int d^4x \psi^\dagger_x \psi_x \phi_x + \mathcal{O}(g^2),$$

and so the transition matrix to leading order in g is

$$T_{fi} = ig \langle f| \int d^4x\, \psi^\dagger_x \psi_x \phi_x\, |i\rangle\,.$$

We now need to substitute in $\psi^\dagger_x$, ψ_x and ϕ_x. First let's just do ϕ_x. We use the definition

$$\phi(x) = \int \frac{d^3\vec{k}}{(2\pi)^3} \frac{1}{\sqrt{2E_{\vec{k}}}} \left(a_{\vec{k}}\, e^{-ikx} + a^\dagger_{\vec{k}}\, e^{ikx} \right),$$

and note that the $a^\dagger_{\vec{k}}$ term will act on $|i\rangle$ and produce a two particle state. That is (up to factors of $\sqrt{2E}$)

$$a^\dagger_{\vec{k}}\, |i\rangle = a^\dagger_{\vec{k}}\, a^\dagger_{\vec{p}} |0\rangle = |\phi(p)\phi(k)\rangle\,.$$

Then noting that the $\psi/\psi^\dagger$ terms are not going to be able to 'undo' this, we will get a zero projection onto the final state $|f\rangle$. So we just need to consider the $a_{\vec{k}}$ term. So using the usual trick of commuting $a_{\vec{k}}$ with the $a^\dagger_{\vec{p}}$ and picking up a $(2\pi)^3\delta^{(3)}(\vec{k} - \vec{p})$, we get

$$T_{fi} = -ig \langle f| \int d^4x\, \psi^\dagger_x \psi_x e^{-ipx}\, |0\rangle\,.$$

Exercise

Using the definitions for ψ and $\psi^\dagger$ show that

$$\langle f|\, \psi^\dagger_x \psi_x \text{ “} = \text{” } e^{i(q_1+q_2)x}\, \langle 0|\,.$$

Hint: Again you will argue away terms that will vanish when taking the inner product. This is why I have put "=".

Using the result of the above exercise we get

$$T_{fi} = -ig \int d^4x\, e^{i(q_1+q_2-p)x} = -ig(2\pi)^4 \delta^{(4)}(p - q_1 - q_2).$$

This result is basically just telling us that 4-momentum must be conserved at the vertex, a very nice physical result. What kind of restrictions does this put on the process? Well the result is Lorentz invariant so we can pick the decaying ϕs rest frame to evaluate it. In this frame we have $p = (m, 0, 0, 0)$, so the delta function splits into

$$\delta^{(4)}(p - q_1 - q_2) = \delta(m - q_1^0 - q_2^0)\delta^{(3)}(\vec{q}_1 + \vec{q}_2),$$

and so we see the decay can only happen if $m \geq 2M$.

Obviously we expect momentum conservation to be something that comes up a lot in our interaction calculations, and indeed it turns out that in general

$$T_{fi} = i(2\pi)^4 M_{fi}\, \delta^{(4)}(p_i - p_f), \tag{6.8}$$

where p_i and p_f are the total 4-momentum initial and final state, respectively. So essentially our goal is to find the M_{fi}s, and these are given the, somewhat disappointing, name *matrix elements*. We often also group the i with it and call iM_{fi} the matrix elements.

6.2.2 ϕ^4 Scattering

Ok so we've computed our first scattering matrix, but it corresponded to a decay, so now let's find the first *scattering* scattering matrix. That is we want something like

$$F(p_1)G(p_2) \to F(p_3)G(p_4),$$

where F and G are some fields. We could consider higher order terms in the expansion for the scalar Yakawa theory as mentioned above, however as we will see these are incredibly hard to compute and will require the machinery of Wick contractions. So we want something that to first order in H_I will give us 4 fields (two in, two out). The easiest example is of course ϕ^4 theory, which has

$$H_{\text{int}} = \frac{\lambda}{4!} \int d^3x\, \phi_x \phi_x \phi_x \phi_x,$$

and corresponds simply to the scattering process

$$\phi(p_1)\phi(p_2) \to \phi(p_3)\phi(p_4).$$

The calculation of the transition matrix is left as an exercise.

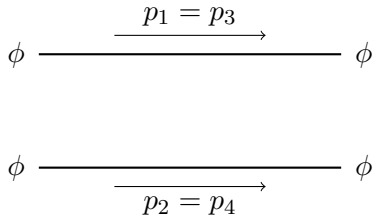
Exercise

Show that the above interaction Hamiltonian leads to

$$T_{fi} = -i\lambda\delta^{(4)}(p_1 + p_2 - p_3 - p_4).$$

Note that the factor of 4! has gone in the result to the above exercise. This is actually the reason it's included in the original expression, so that the matrix elements don't have any additional factors. Its an example of a what we call *symmetry factors*, which will become more clear after studying Wick contractions.

Remark 6.2.1. Note that for the scalar Yakawa decay we could have actually used the S-matrix instead of the transition matrix. This is because the two only differ by the inner product between the initial and final state, which for the scalar Yakawa theory vanishes. However for the ϕ^4 scattering, this inner product is non-vanishing and doesn't actually correspond to a scattering. Diagrammatically it looks like,



It is therefore important to actually distinguish between S_{fi} and T_{fi}, however it is often the case that we write S_{fi} and really mean T_{fi}. In these notes I shall try be careful here, but its quite possible that I will forget and make a mistake, so keep on your toes.

7 | Wick's Theorem

Last lecture we showed how we could use Dyson's formula to find the S-matrix for a simple decay and simple scattering. We made a comment about how would could have got a scattering in the scalar Yakawa theory, but that it would involve a quadratic term from Dyson's formula, and so would be very hard to do in the manner above. However, obviously we want to be able to calculate such terms (otherwise we never pass first order in perturbation theory, and we loose *a lot* of information), so how do we do it? The answer is we be smart about how we write things.

7.1 Normal Ordering

Ok we just said we "be smart", but what does that actually entail? We shall use the scalar Klein-Gordan field to explain. Recall that we can write this field as the expansion

$$\phi(x) = \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\Big(a_{\vec{p}}\, e^{-ipx} + a^\dagger_{\vec{p}} e^{ipx}\Big).$$

Well let's define

$$\begin{aligned}\phi^+(x) &= \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}} a_{\vec{p}}\, e^{-ipx} \\ \phi^-(x) &= \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}} a^\dagger_{\vec{p}}\, e^{ipx},\end{aligned} \tag{7.1}$$

so that we simply have

$$\phi(x) = \phi^+(x) + \phi^-(x).$$

Remark 7.1.1. It might seem strange that we define $\phi^+(x)$ to be the one that has the negative exponential. The reason for this is that e^{-ipx} corresponds to a mode with positive frequency, and so we use $\phi^+(x)$. Also note that the dagger does *not* come with $\phi^+(x)$, as you might think. This is actually important because it is fact that $\phi^\pm(x)$ correspond to annihilation/creation operators, respectively, that we will use.

Now, as we just mentioned in the remark above, $\phi^+(x)$ comes with an annihilation operator, so if it were to appear to the far right in a vacuum expectation value, the result would be zero. The same thing is true for $\phi^-(x)$ being to the far left. That is

$$\phi^+(x)\,|0\rangle = 0 = \langle 0|\,\phi^-(x).$$

Similarly, we can see that

$$\phi^+(x)\,|\vec{p}\rangle = e^{-ipx}\,|0\rangle\,, \quad \text{and} \quad \langle\vec{p}|\,\phi^-(x) = e^{ipx}\,\langle 0|\,.$$

We call the procedure of putting all the annihilation operators to the right *normal ordering*.

Definition. [Normal Ordering] We denote the *normal ordered* string of operators $\mathcal{O}_1(x_1)...\mathcal{O}_n(x_n)$ as

$$:\mathcal{O}_1(x_1)...\mathcal{O}_n(x_n):\,, \tag{7.2}$$

which tells us to put all the annihilation operators to the right.

Remark 7.1.2. Normal ordering is 'linear' in the sense that

$$:AB+CD: \;=:AB:+:CD:\,.$$

Remark 7.1.3. Note we could have introduced this (as Prof. Tong does) when talking about the energy of the ground state being infinite. That is we could write the Hamiltonian for the single free scalar field as

$$:H: \;= \int \frac{d^3\vec{p}}{(2\pi)^3} E_{\vec{p}} a^\dagger_{\vec{p}} a_{\vec{p}},$$

so that

$$:H:\;|0\rangle = 0.$$

Notation. The colon notation used here for normal ordering is very common, however some authors, including Peskin and Schroeder, denote normal ordering by an N. So Equation (7.2) would be written $N(\mathcal{O}_1(x_1)...\mathcal{O}_n(x_n))$.

7.2 Wick's Theorem

Ok normal ordering seems great, but Dyson's formula comes in the form of time ordered products, not normal ordered ones. So the question becomes "can we convert one into the other?" Well let's look at an example and see.

7.2.1 Recovering The Propagator

Let's use Equation (7.1) to try investigate the link. Let's start with the simple case of the product of two fields. Let's define

$$\phi^\pm(x_i) = \phi^\pm_i,$$

and let's consider $x_1^0 > x_2^0$. Then the time ordered product is[1]

$$\begin{aligned}\mathcal{T}(\phi_1\phi_2) &:= \phi_1\phi_2 \\ &= (\phi_1^+\phi_1^-)(\phi_2^+\phi_2^-) \\ &= \phi_1^+\phi_2^+ + \phi_1^+\phi_2^- + \phi_1^-\phi_2^+ + \phi_1^-\phi_2^- \\ &= \phi_1^+\phi_2^+ + \phi_2^-\phi_1^+ + [\phi_1^+, \phi_2^-] + \phi_1^-\phi_2^+ + \phi_1^-\phi_2^- \\ &=: \phi_1\phi_2: + [\phi_1^+, \phi_2^-] \\ &=: \phi_1\phi_2: + D(x_1 - x_2),\end{aligned}$$

where we have used the commutator to make every term have a ϕ^+ to the far right and the definition of the propagator $D(x-y)$.

Exercise

Following a similar method to the above, show that if $x_2^0 > x_1^0$, then

$$\mathcal{T}(\phi_1\phi_2) =: \phi_1\phi_2: + D(x_2 - x_1).$$

So together with the result of the above exercise, we conclude that

$$\mathcal{T}(\phi_1\phi_2) =: \phi_1\phi_2: + \begin{cases} D(x_1 - x_2) & \text{if } x_1^0 > x_2^0 \\ D(x_2 - x_1) & \text{if } x_2^0 > x_1^0 \end{cases}.$$

Hang on, this looks familiar... the cases term is just the Feynman propagator, Equation (5.8). So we get

$$\mathcal{T}(\phi_1\phi_2) =: \phi_1\phi_2: + \Delta_F(x_1 - x_2). \tag{7.3}$$

Remark 7.2.1. Note that once again we have an expression that mixes operators and $\mathbb{C}$-numbers: both $\mathcal{T}(\phi_1\phi_2)$ and $:\phi_1\phi_2:$ are operators, but the Feynman propagator is a $\mathbb{C}$-number. We have made this comment several times throughout this course, but make it here again to keep us on our toes.

7.2.2 Wick Contractions & Wick's Theorem

So we obtained a really nice expression relating the time ordering and normal ordering for two real scalar fields, and their difference was the Feynman propagator. We now want to try and generalise this result to a general relation between time ordering and normal ordering. In order to do that, it is convenient to introduce a notation for *Wick contractions*.

[1] Be careful not to confuse the colon in :=, which means "is defined to be equal to" to the colons for normal ordering. I considered using another type font for the colon, but decided against it.

Definition. [Wick Contraction] We define the *Wick contraction*, or just *contraction*, of a pair of fields, ϕ_i and ϕ_j, in a string of n-operators, $\phi_1...\phi_i...\phi_j...\phi_n$, to mean "replace those operators with the Feynman propagator, and leave everything else alone." We use the following notation.

$$\phi_1...\overbrace{\phi_i...\phi_j}...\phi_n = \Delta_F(x_i - x_j)\phi_1...\phi_n, \tag{7.4}$$

where the ... on the right hand side does *not* include ϕ_i and ϕ_j.[a]

[a] If you like, in standard "this element is missing" notation, we could write the right hand side as $\Delta_F(x_i - x_j)\phi_1...\hat{\phi}_i..., \hat{\phi}_j...\phi_n$, but I thought this would be confusing, as hats are used for operators in QFT.

Example 7.2.2. So we can write Equation (7.3) simply as

$$\mathcal{T}(\phi_1\phi_2) =: \phi_1\phi_2 : + \overbrace{\phi_1\phi_2}.$$

Claim 7.2.3. A similar argument/proof can be made for complex fields, where we have

$$\overbrace{\psi_1\psi_2^\dagger} = \Delta_F(x_1 - x_2), \quad \text{and} \quad \overbrace{\psi_1\psi_2} = 0 = \overbrace{\psi_1^\dagger\psi_2^\dagger},$$

so that

$$\mathcal{T}(\psi_1\psi_2^\dagger) =: \psi_a\psi_2^\dagger : + \overbrace{\psi_1\psi_2^\dagger}.$$

So we see that for the case of two real scalar fields, and claimed it for two complex scalar field, the difference between the time ordered and normal ordered products is a Wick contraction, but what about if we have more than two fields? Well this is where Wick's theorem comes in.

Theorem 7.2.4 (Wick's Theorem)**.** *For any collection of fields $\phi_1, ..., \phi_n$ the following holds*

$$\mathcal{T}(\phi_1...\phi_n) =: \phi_1...\phi_n : + : \textit{all possible contractions} : . \tag{7.5}$$

Proof. We do not prove Wick's theorem here, but an explanation of the proof (its an inductive proof) is given in Prof. Tong's notes, page 58. □

It is often very useful to break a calculation using Wick's theorem into categories, given by the number of contractions done. Every calculation will have the 0-contraction term, which is just the fully normal ordered term, and then you consider the terms in increasing number of contractions.[2]

Example 7.2.5. Let's consider 4 real Klein-Gordan fields. Then we can have either 0-contractions, 1-contraction or 2-contractions. The respective terms are:

[2] Well it depends what you want to do. In string theory it is often the term which is fully contracted that contains most of the useful information. For more details, see my notes on Dr. Shiraz Minwalla's string theory course.

- 0-contractions

$$: \phi_1\phi_2\phi_3\phi_4 :,$$

- 1-contraction:

$$: \overbrace{\phi_1\phi_2}\phi_3\phi_4 + \phi_1\overbrace{\phi_2\phi_3}\phi_4 + \phi_1\overbrace{\phi_2\phi_3\phi_4} + \phi_1\phi_2\overbrace{\phi_3\phi_4} + \phi_1\overbrace{\phi_2\phi_3\phi_4} + \phi_1\phi_2\overbrace{\phi_3\phi_4} :$$

- 2-contractions

$$: \overbrace{\phi_1\phi_2}\overbrace{\phi_3\phi_4} + \overbrace{\phi_1\overbrace{\phi_2\phi_3}\phi_4} + \overbrace{\phi_1\overbrace{\phi_2\phi_3\phi_4}} :$$

This result implies that

$$\langle 0|\,\mathcal{T}(\phi_1\phi_2\phi_3\phi_4)\,|0\rangle = \Delta_F(x_1-x_2)\Delta_F(x_3-x_4)+\Delta_F(x_1-x_3)\Delta_F(x_2-x_4)+\Delta_F(x_1-x_4)\Delta_F(x_2-x_3),$$

which is a Green's function for the Klein-Gordan operator (recall that the Feynman integral is a Green's function).

7.2.3 Wick's Theorem For States

Wick's theorem is defined for time ordered products of operators, but we can adapt it to allow us to contract with states. We do this by defining

$$\overbrace{\phi(x)\,|\vec{p}\rangle} = e^{-ipx}, \quad \text{and} \quad \overbrace{\langle\vec{p}|\,\phi(x)} = e^{ipx}, \tag{7.6}$$

and then Wick's theorem is simply adapted to include contractions between fields and states. This allows us to calculate the S-matrix completely using Wick's theorem.

7.3 Scalar Yakawa Scattering

We can now return to the problem of trying to find the $\mathcal{O}(g^2)$ term in scalar Yakawa theory. Recall the Lagrangian is,[3]

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 + \frac{1}{2}\partial_\mu\psi^\dagger\partial^\mu\psi - M^2\psi^\dagger\psi - g\phi\psi^\dagger\psi.$$

Similarly to Equation (7.1) we define

$$\phi_b^+(x) = \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}b_{\vec{p}}\,e^{-ipx}, \quad \text{and} \quad \phi_b^-(x) = \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}b_{\vec{p}}^\dagger\,e^{ipx}$$
$$\phi_c^+(x) = \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}c_{\vec{p}}\,e^{-ipx} \quad \text{and} \quad \phi_c^-(x) = \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}c_{\vec{p}}^\dagger\,e^{ipx}$$

so that

$$\psi(x) = \phi_b^+(x) + \phi_c^-(x), \quad \text{and} \quad \psi^\dagger(x) = \phi_c^+(x) + \phi_b^-(x),$$

[3] Using $\psi^\dagger$ now instead of ψ^* because we will only consider the quantum theory from now on.

where we note that each term contains a b and a c, in correspondance with Equation (5.1). Then, we recalling the definitions

$$|\psi(p)\rangle := \sqrt{2E_{\vec{p}}}\, b^\dagger_{\vec{p}}|0\rangle\,, \quad \text{and} \quad |\overline{\psi}(p)\rangle := \sqrt{2E_{\vec{p}}}\, c^\dagger_{\vec{p}}|0\rangle\,,$$

our Wick contractions are

$$\overbrace{\psi(x)\,|\psi(p)\rangle} = \overbrace{\psi^\dagger(x)\,|\overline{\psi}(p)\rangle} = e^{-ipx}, \quad \text{and} \quad \langle\overbrace{\psi(p)|\,\psi^\dagger(x)} = \langle\overbrace{\overline{\psi}(p)|\,\psi(x)} = e^{ipx},$$

and all others vanishing.

So what we're looking at is $\psi\psi \to \phi \to \psi\psi$, so our initial and final states are

$$|i\rangle = |\psi(p_1)\psi(p_2)\rangle\,, \quad \text{and} \quad \langle f| = \langle\overline{\psi}(p_3)\overline{\psi}(p_4)|\,.$$

The transition matrix is then

$$\begin{aligned} T_{fi} = &-ig\,\langle\overline{\psi}(p_3)\overline{\psi}(p_4)|\int d^4x\,\psi^\dagger_x\psi_x\phi_x\,|\psi(p_1)\psi(p_2)\rangle \\ &+\frac{(-ig)^2}{2!}\,\langle\overline{\psi}(p_3)\overline{\psi}(p_4)|\,\mathcal{T}\Big(\int d^4x\,d^4y\,\psi^\dagger_x\psi_x\phi_x\psi^\dagger_y\psi_y\phi_y\Big)\,|\psi(p_1)\psi(p_2)\rangle + \mathcal{O}(g^3). \end{aligned}$$

We can use Wick's theorem to evaluate this. First we note that the first term vanishes. Why? Well because we have nothing to contract the ϕ_x with, so we have to leave it as a normal ordered term, which then acts on the vacuum giving a vanishing result. We will see in terms of Feynman diagrams soon why such a term can't exist. Now what about the second term? Well we get three different types of contractions which are conceptually different they are

(i) Fully Connected,

(ii) Connected, and

(iii) Disconnected.

We shall consider the fully connected diagrams in this lecture and derive the corresponding Feynman rules, and then talk about connected and disconnected diagrams next lecture.[4]

7.3.1 Fully Connected

There are two unique types of fully connected contractions for the term above, they are

$$-\frac{g^2}{2!}\int d^4x\,d^4y\;\langle\overbrace{\overline{\psi}(p_3)\underbrace{\overline{\psi}(p_4)|\,:\,\psi^\dagger_x}\psi^\dagger_y\overbrace{\phi_x}\overbrace{\phi_y}\overbrace{\psi_x}\underbrace{\psi_y}:\,|\psi(p_1)}\psi(p_2)\rangle + (x \longleftrightarrow y),$$

and

$$-\frac{g^2}{2!}\int d^4x\,d^4y\;\langle\overbrace{\overline{\psi}(p_3)\underbrace{\overline{\psi}(p_4)|\,:\,\psi^\dagger_x\psi^\dagger_y}\overbrace{\phi_x}\overbrace{\phi_y}\overbrace{\psi_x}\underbrace{\psi_y}:\,|\psi(p_1)}\psi(p_2)\rangle + (x \longleftrightarrow y).$$

The $(x \longleftrightarrow y)$ terms are symmetries factors that will remove the $1/2$ factors, just like ones we mentioned before that got rid of the $1/4!$ in ϕ^4 theory. We will see soon in terms of Feynman diagrams why they are symmetry factors.

[4] Just because that's what happened on the course.

Evaluating the contractions using the rules we have above, these two terms become, respectively

$$\int d^4x\, d^4y \left[e^{i(p_3-p_1)x}\, (-ig)\Delta_F^{\phi}(x-y)(-ig)\, e^{i(p_4-p_2)y} \right], \tag{7.7}$$

and

$$\int d^4x\, d^4y \left[e^{i(p_4-p_1)x}\, (-ig)\Delta_F^{\phi}(x-y)(-ig)\, e^{i(p_3-p_2)y} \right]. \tag{7.8}$$

Exercise

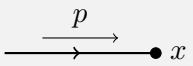
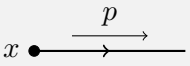
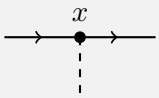
Check that the Wick contractions do indeed result in the above two expressions.

7.3.2 Feynman Rules — Position Space

We are finally at a place to draw a Feynman diagram and derive the corresponding Feynman rules. Really its a flip of a coin which you do first (draw the diagram or define the rules), here I've decided to present the rules and then use them to draw the diagram as I think this is perhaps more instructive.

The Feynman rules are a set of instructions on how to draw the different terms in the transition matrix, so that you can manipulate the expressions diagrammatically, and then at the end convert them back to equations should you need to.

The *position space* Feynman rules for scalar Yakawa scattering $\psi\psi \to \phi \to \psi\psi$ are:

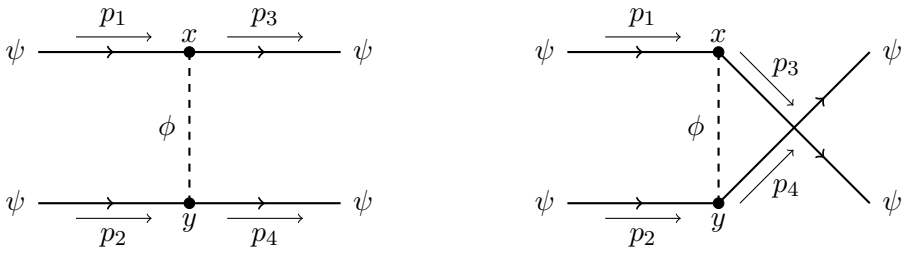
Type	Diagram	Maths Expression
Incoming Particle		$e^{-ipx} = \overbrace{\psi_x \left\vert \vec{p} \right\rangle}$
Outgoing Particle		$e^{ipx} = \overbrace{\left\langle \vec{p} \right\vert \psi_x}$
Propagator		$\Delta_F^{\phi}(x-y) = \overbrace{\phi_x \phi_y}$
Vertex		$ig \int d^4x$

Let's clarify the notation/terminology of the rules.

- We take time to run from left to right. So the initial state is to the far left and final state is to the far right. Note, though, that some authors choose to have time run vertically. Peskin and Schroeder use this latter convention.
- We use solid lines to indicate complex fields (or nucleons) and dotted lines to indicate real fields (or mesons).
- We use a filled in circle to indicate a vertex, and we label it by its integration variable.

- The arrow above the lines indicate the value of the 4-momentum and the direction it flows.
- The arrows on the nucleon line tells us about the flow of the charge (e.g. Fermion number), if the line points in the same direction as the momentum arrow we have a particle, and if they point oppositely, then we have an antiparticle. This is why we don't draw an arrow on the dotted line; a real scalar field is its own antiparticle.
- The arrows for a charge must always be conserved at a vertex. By this we mean one must always go in and one must always leave.
- The incoming and outgoing particles are known as *external line* because they only connect to one vertex, whereas the propagator is called an *internal line* because it is connected to two vertices.
- We usually denote the 4-momentum on the external lines with p_i, and in the position space picture we don't need to label the momentum of the internal lines.
- Sometimes we also include the symbols of the particles next to the lines. This is particularly the case when we have multiple different species particle of the same type (e.g. electrons and muons) in the same diagram.

Using these rules, we can construct the Feynman diagrams for Equations (7.7) and (7.8). They are, respectively:



Remark 7.3.1. The symmetry factor $x \longleftrightarrow y$ in terms of the Feynman diagrams is seen by the fact that if we relabel $x \longleftrightarrow y$ then the diagram looks exactly the same apart from the relabelling. So there is a symmetry in the diagram, hence the name.

Remark 7.3.2. Note that we can see from the Feynman diagrams that the order g term wont contribute, as we said before. That's just because there is no way to start with two ψ lines and end up with two ψ lines using only one vertex.[5] This illustrates one of the incredible powers of Feynman diagrams: it allows us to see almost instantly which terms do not contribute to the transition matrix.

Exercise

Convince yourself that these diagrams are correct. *Hint: This exercise should make it clear why we collected the terms in Equations* (7.7) *and* (7.8) *ad we did.*

[5] If you don't see this, try and do it and you'll quickly see what I mean.

7.3.3 Feynman Rules — Momentum Space

As it says in the box above, the above rules are in position space (our integrals are over x and y), but the Feynman propagator is given by Equation (5.10),

$$\Delta_F(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{ie^{-ip(x-y)}}{p^2 - m^2 + i\epsilon},$$

and so we can replace the position integrals with momentum ones. This gives us the *momentum space* Feynman rules. Let's do this explicitly for Equation (7.7). Plugging the Feynman propagator in, and using k as the momentum variable in correspondance with the rules above, we have

$$\int d^4x\, d^4y \left[e^{i(p_3-p_1)x} (-ig) \int \frac{d^4k}{(2\pi)^4} \frac{ie^{-ik(x-y)}}{k^2 - m^2 + i\epsilon} (-ig)\, e^{i(p_4-p_2)y} \right],$$

combining the x and y exponentials and then integrating over x and y gives us two delta functions, that is we get

$$\int \frac{d^4k}{(2\pi)^4} (2\pi)^4 \delta^{(4)}(p_1 - p_3 - k)(-ig) \frac{i}{k^2 - m^2 + i\epsilon} (-ig)(2\pi)^4 \delta^{(4)}(p_2 - p_4 + k),$$

and then integrating over k leaves us with[6]

$$(2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4) \left[(-ig) \frac{i}{(p_1 - p_3)^2 - m^2 + i\epsilon} (-ig) \right],$$

which, comparing to Equation (6.8), tells us

$$iM = (-ig) \frac{i}{(p_1 - p_3)^2 - m^2 + i\epsilon} (-ig).$$

Exercise

Show that Equation (7.8) leads to

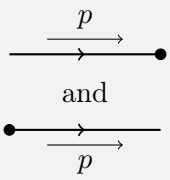
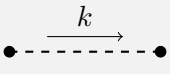
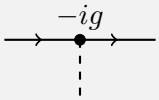
$$(2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4) \left[(-ig) \frac{i}{(p_1 - p_4)^2 - m^2 + i\epsilon} (-ig) \right],$$

but note the difference of the intermediate steps. *Hint: Be careful about delta functions.*

So how do the Feynman rules change? Well we note now that x and y don't appear in the final expression so we don't need to label the nodes any more. Then we note in the second step, the two delta functions simply correspond to conservation of 4-momentum at the vertices, i.e. $\delta^{(4)}(p_1 - p_3 - k)$ is the momentum flowing into the x vertex above. Then to get the final answer to integrated over the internal momentum k, which gave use the final delta function.

[6] Note we could also use $(p_2 - p_4)^2$ in the denominator.

The *momentum space* Feynman rules for scalar Yakawa scattering $\psi\psi \to \phi \to \psi\psi$ are:

Type	Diagram	Maths Expression
External lines		1
Propagator		$\int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i\epsilon}$
Vertex		$-(ig)$

So we have the rough[7] following procedure to obtain the maths expression from a momentum-space Feynman diagram:

- Take the position space Feynman diagram, and given every internal line a momentum labelled by k_i.
- Next to each vertex write down the coupling strength (in the case of scalar Yakawa this is $-ig$).
- Impose momentum conservation at each vertex with a factor of the coupling strength, i.e. write $(-ig)(2\pi)^4\delta^{(4)}(\sum_i k_i)$, where the sum is taken over the momentum flowing *into* the vertex. That is, if the momentum flows out, put a minus sign in the sum, as we did in $\delta^{(4)}(p_1 - p_3 - k)$ and $\delta^{(4)}(p_2 - p_4 + k)$.
- Include a factor of
$$\int \frac{d^4k}{(2\pi)^4} \frac{i}{k_i^2 - m^2 + i\epsilon}$$
for all the internal ϕ lines, and the same for internal ψ lines but now with $m \to M$.

Remark 7.3.3. Some authors do not include the integral factor $\int \frac{d^4k}{(2\pi)^4}$ in the propagator term expression, but instead add another step at the end that says "finally integrate over all internal momentum". Of course this amounts to exactly the same thing, but this remark is just included to avoid confusion when comparing to other literature.

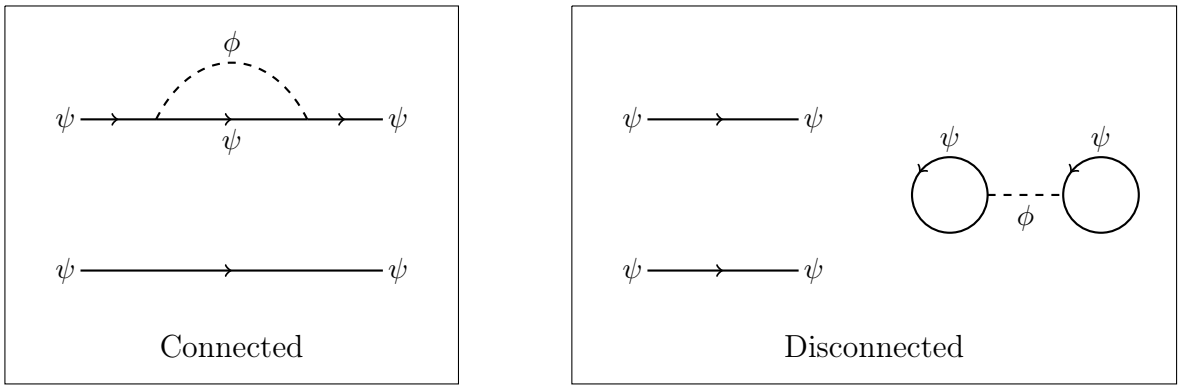
Exercise

Draw the scalar Yakawa scattering Feynman diagrams above in momentum space.

[7] We will need to be more clear for the case of spin-1/2 particles later.

8 | LSZ Theorem, Loop & Amputated Diagrams

Last lecture we discussed fully connected diagrams and derived the Feynman rules for $\psi\psi \to \psi\psi$ scattering. We also said that there were two other types of conceptually different diagrams, namely *connected* and *disconnected* diagrams. As diagrams they look like:



Intuitively, we would say these don't really correspond to scattering as the initial states don't interact with each other. However we need some way to actually prove this, and this is where the so-called LSZ theorem[1] comes in. First we need some definitions and a proposition.

Definition. [n-Point Green's Function] We define the *n-point Green's function* to be the time ordered vacuum expectation value of n Heisenberg picture[2] field operator. That is

$$G_n(x_1, ..., x_n) := \langle 0 | \mathcal{T}\big[\phi(x_1)...\phi(x_n)\big] | 0 \rangle . \tag{8.1}$$

Definition. [Wavefunction Normalisation] We define the so-called *wavefunction normalisation* to be

$$Z := | \langle p_i | \phi(x) | 0 \rangle |^2,$$

so that

$$\langle p_1, ..., p_n | \phi(t = \pm\infty, \vec{x}) | q_1, ..., q_m \rangle = \lim_{t \to \pm\infty} Z^{-1/2} \langle p_1, ..., p_n | \phi(x) | q_1, ..., q_m \rangle . \tag{8.2}$$

[1] Named after Harry Lehmann, Kurt Symanzik and Wolfhart Zimmermann.

Remark 8.0.1. What Equation (8.2) does is it allows us to turn an asymptotic field $\phi(t = \pm\infty, \vec{x})$ into a field at a general time under a limit. The factor of $\sqrt{Z}$ is included just to normalise the result, hence the name.

Definition. [Left-Right Derivative Action][3] Let $\phi_1(x)$ and $\phi_2(x)$ be to fields on the space-time. Then we define the *left-right derivative action* as

$$\phi_1(x)\overset{\leftrightarrow}{\partial}_\mu \phi_2(x) := \phi_1(x)\partial_\mu\phi_2(x) - \big(\partial_\mu\phi_1(x)\big)\phi_2(x). \tag{8.3}$$

Proposition 8.0.2. *For the real scalar field, Equation (4.10), we have*

$$\begin{aligned} a_{\vec{p}} &= \frac{i}{\sqrt{2E_{\vec{p}}}} \int d^3\vec{x}\, e^{ip\cdot x} \overset{\leftrightarrow}{\partial}_0 \phi(x) \\ a^\dagger_{\vec{p}} &= -\frac{i}{\sqrt{2E_{\vec{p}}}} \int d^3\vec{x}\, e^{-ip\cdot x} \overset{\leftrightarrow}{\partial}_0 \phi(x) \end{aligned} \tag{8.4}$$

Remark 8.0.3. Note that because we only integrate over space but $\phi(x)$ is a function of space-time, our creation and annihilation operators are functions of time! This will prove crucial in the proof of the LSZ theorem below.

Exercise

Prove Equation (8.4). *Hint: You will want to use the constraint $p_0 = E_{\vec{p}}$ present in Equation (4.10) along with $E_{\vec{p}} = \sqrt{\vec{p}^2 + m^2} = E_{-\vec{p}}$ at the end.*

We choose here to first state the result of the LSZ theorem and then derive it, as this way the heavy manipulations that follow have some guiding light.

Theorem 8.0.4 (LSZ Reduction Formula)**.** *We can express the S-matrix[4] between an initial m-particle state and a n-particle final particle state in terms of the $(n+m)$-point Green's function as follows:*

$$\begin{aligned} T_{fi} =& \big(iZ^{-1/2}\big)^{n+m} \int dy_1 ... dy_n \int dx_1 ... dx_m\, e^{i(p_j\cdot y_j - q_i\cdot x_i)} \\ &\times \left(\prod_{i=1}^{m}(\partial^2_{x_i} + m^2)\right)\left(\prod_{j=1}^{n}(\partial_{y_j} + m^2)\right) G_{n+m}(y_1, ..., y_n, x_1, ..., x_m) \\ &+ \textit{all non-fully connected stuff}, \end{aligned} \tag{8.5}$$

where we assume an implicit sum in the exponential.

Before proving this theorem it is worth stating what it allows us to do. Basically we note that the right-hand side of Equation (8.5) splits into a contribution that contains only information about the fully connected diagrams (we will prove this), and a part that contains only the non-fully connected stuff. We can therefore simply restrict our analysis to processes

[4] Again we are really interested in the Transition matrix. I have tried to use Ts here, but I might have missed some, and written S by accident. Apologies for that, please feel free to point these out to me.

that are only fully connected in a well defined way. That is, a priori there is no clear way to see that the result of the full S-matrix calculation wont 'mix' fully connected and non-fully connected stuff, but the LSZ theorem tells us that it indeed does. This statement does *not* mean that the connected diagrams do not contribute to the full S-matrix value, but simply that we are allowed to restrict ourselves to considering only fully-connected diagrams.

Proof. Unfortunately for those who don't like long proofs, this one comes from sheer brute force,[5] so prepare yourself. Firstly we need to clarify the notation that follows: by Dysons formula, the S-matrix is given by a bunch of time evolution operators, and we also take our states to asymptotic (i.e. at $t \to \pm\infty$). Therefore the creation operators that we extract from our initial states *do not* simply act on the final state and annihilate them, as they are separated in time. For this reason we shall use subscripts 'in' and 'out' on the creation/annihilation operators to avoid confusion. Similarly we will denote the initial/final states using the following notation

$$\langle p_1 ... p_n; out| \,, \quad \text{and} \quad |q_1 ... q_m; in\rangle \,.$$

We use this notation as it is the one used by Prof. Weigand in his notes.

We therefore have

$$\begin{aligned} T_{fi} &= \langle p_1 ... p_n; out | q_1 ... q_m; in \rangle \\ &= \sqrt{2E_{\vec{q}_1}} \, \langle p_1 ... p_n; out| \, a^\dagger_{\vec{q}_1} \, |q_2 ... q_m; in\rangle \\ &= \frac{1}{i} \int d^3\vec{x}_1 \, \langle p_1 ... p_n; out| \, e^{-iq_1\cdot x_1} \overset{\leftrightarrow}{\partial}_0 \phi(t=-\infty, \vec{x}) \, |q_2 ... q_m; in\rangle \\ &= \frac{1}{i} \lim_{t\to-\infty} Z^{-1/2} \int d^3\vec{x}_1 \, \langle p_1 ... p_n; out| \, e^{-iq_1\cdot x_1} \overset{\leftrightarrow}{\partial}_0 \phi(x) \, |q_2 ... q_m; in\rangle \\ &= \frac{1}{i\sqrt{Z}} \lim_{t\to-\infty} \int d^3\vec{x}_1 e^{-iq_1\cdot x_1} \overset{\leftrightarrow}{\partial}_0 \, \langle p_1 ... p_n; out| \, \phi(x) \, |q_2 ... q_m; in\rangle \end{aligned}$$

where we have used the definition $|q\rangle = \sqrt{2E_{\vec{q}}} a^\dagger_{\vec{q}} |0\rangle$ and Equations (8.2) and (8.4). Note on the third line we have $\phi(t=-\infty, \vec{x})$, this is because we want to consider an initial state, which we take to be at $t = -\infty$. Indeed for this exact reason, we should really denote the creation operator on the second line as something like

$$(a_{\text{in}})^\dagger_{\vec{q}_1} \,,$$

so that we know it only acts on in/initial states.

What we now want to do is get the $\phi(x)$ to act on a final state, in which case we need to change the limit from $t \to -\infty$ to $t \to +\infty$. We achieve this by using the simple result

$$\lim_{t\to-\infty} f(t) = \lim_{t\to+\infty} f(t) - \lim_{T\to\infty} \int_{-T}^{T} \frac{d}{dt} f(t),$$

so our S-matrix splits into two terms:

$$\begin{aligned} T_{fi} =& \frac{1}{i\sqrt{Z}} \lim_{t\to+\infty} \int d^3\vec{x}_1 e^{-iq_1\cdot x_1} \overset{\leftrightarrow}{\partial}_0 \, \langle p_1 ... p_n; out| \, \phi(x) \, |q_2 ... q_m; in\rangle \\ &+ \frac{1}{i\sqrt{Z}} \int d^4x_1 \, \partial_0 \Big(e^{-iq_1\cdot x_1} \overset{\leftrightarrow}{\partial}_0 \, \langle p_1 ... p_n; out| \, \phi(x) \, |q_2 ... q_m; in\rangle \Big), \end{aligned}$$

[5] I'm following the one given by Prof. Weigand, starting on page 49.

where we notice that the second term contains a 4-integral. Let's call the first A and the second B, and consider them in turn. In A we can use Equation (8.2) backwards to get

$$A = \frac{1}{i\sqrt{Z}} \int d^3\vec{x}_1 e^{-iq_1\cdot x_1} \overset{\leftrightarrow}{\partial}_0 \langle p_1 ... p_n; out| \, \phi(t = +\infty, \vec{x}) \, |q_2 ... q_m; in\rangle \, ,$$

and then we can use Equation (8.4) backwards to give

$$A = \sqrt{2E_{\vec{q}_1}} \, \langle p_1 ... p_n; out| \, (a_{\text{out}})^\dagger_{\vec{q}_1} \, |q_2 ... q_m; in\rangle \, ,$$

where, in agreement with the comment made above, we have included a subscript 'out' to tell us this acts on the final states. Now the action of a creation operator to the left gives a delta function, and so we get

$$A = 2E_{\vec{q}_1} (2\pi)^3 \sum_{k=1}^{n} \delta^{(3)}(\vec{p}_k - \vec{q}_1) \, \langle p_1 ... \hat{p}_k ... p_n; out | q_2 ... q_m; in\rangle \, ,$$

where we have used the standard maths notation where a hatted entry in a string of terms is missing (i.e. $\hat{p}_k$ is missing from the final states). This corresponds to a connected, *but not fully connected*, diagram as we have the k-th final particle just corresponding to the 1st initial particle (it is just a straight line in a diagram).

Now what about B? Well we use Equation (8.3) to get

$$\begin{aligned} B &= \frac{1}{i\sqrt{Z}} \int d^4x_1 \, \partial_0 \Big(e^{-iq_1\cdot x_1} \partial_0 \langle ... \rangle - \partial_0 \big(e^{-iq_1\cdot x_1} \big) \langle ... \rangle \Big) \\ &= \frac{1}{i\sqrt{Z}} \int d^4x_1 \, \Big(e^{-iq_1\cdot x_1} \partial_0^2 \langle ... \rangle - \partial_0^2 \big(e^{-iq_1\cdot x_1} \big) \langle ... \rangle \Big) \end{aligned}$$

Next note that

$$\partial_0^2 e^{-ip\cdot x} = -p_0^2 e^{-ip\cdot x} = (p^2 + (\vec{p})^2) e^{-ip\cdot x} = (m^2 - \nabla^2) e^{-ip\cdot x},$$

which we can apply to the second term on the second line above. We then integrate by parts twice[6] to move the ∇^2 to act on $\langle ... \rangle$ and then use

$$\partial^2 := \partial_0^2 - \nabla^2$$

to give us

$$B = \frac{1}{i\sqrt{Z}} \int d^4x_1 e^{-iq_1\cdot x_1} \big(\partial^2 + m^2 \big) \, \langle p_1 ... p_n; out| \, \phi(x) \, |q_2 ... q_m; in\rangle$$

This is exactly of the form of one of the product terms in Equation (8.5)! We can then apply a similar argument to all the other initial states, noting that

$$[a^\dagger_{\vec{q}_i}, a^\dagger_{\vec{q}_j}] = 0 \qquad \forall i, j \in \{1, ..., m\},$$

to obtain the full product from 1 to m in Equation (8.5).

[6] Note, as always, we assume boundary terms vanish at spatial infinity.

We now need to do the same thing for the final states. However now we have

$$\langle p| = \sqrt{2E_{\vec{p}}}\langle 0|\,(a_{\text{out}})_{\vec{p}},$$

and from

$$[a_{\vec{p}_i}, a^\dagger_{\vec{q}_j}] \neq 0$$

we cannot simply commute the annihilation operators past $\phi(x)$ to get them to act on the initial state, that is

$$\begin{aligned}\langle p_2...p_n; out|\,\phi(y_1^0=\infty,\vec{y}_1)\phi(x_1^0=-\infty,\vec{x}_1)\,|q_1...q_m; in\rangle \\ \neq \langle p_2...p_n; out|\,\phi(x_1^0=-\infty,\vec{x}_1)\phi(y_1^0=\infty,\vec{y}_1)\,|q_1...q_m; in\rangle\,.\end{aligned}$$

This is when we notice that Equation (8.5) contains a Green's function, which contains time-ordering and so we realise we are saved! More explicitly, we will get something of the form

$$\begin{aligned}&\lim_{x_1^0\to-\infty}\lim_{y_1^0\to+\infty}\int d^3\vec{x}_1\int d^3\vec{y}_1 e^{-iq_1\cdot x_1}\overset{\leftrightarrow}{\partial}_{x_1^0} e^{ip_1\cdot y_1}\overset{\leftrightarrow}{\partial}_{y_1^0}\,\langle p_2...p_n; out|\,\phi(y_1)\phi(x_1)\,|q_2...q_m; in\rangle \\ &=\lim_{x_1^0\to-\infty}\int d^3\vec{x}_1 e^{-iq_1\cdot x_1}\overset{\leftrightarrow}{\partial}_{x_1^0}\lim_{y_1^0\to+\infty}\int d^3\vec{y}_1 e^{ip_1\cdot y_1}\overset{\leftrightarrow}{\partial}_{y_1^0}\,\langle p_2...p_n; out|\,\phi(y_1)\phi(x_1)\,|q_2...q_m; in\rangle\end{aligned}$$

where we notice that the x_1^0 limit is taken after the y_1^0 one, so during the latter we can take x_1^0 to be some fixed finite value, so we can just consider the y_1 part of the expression. Now consider

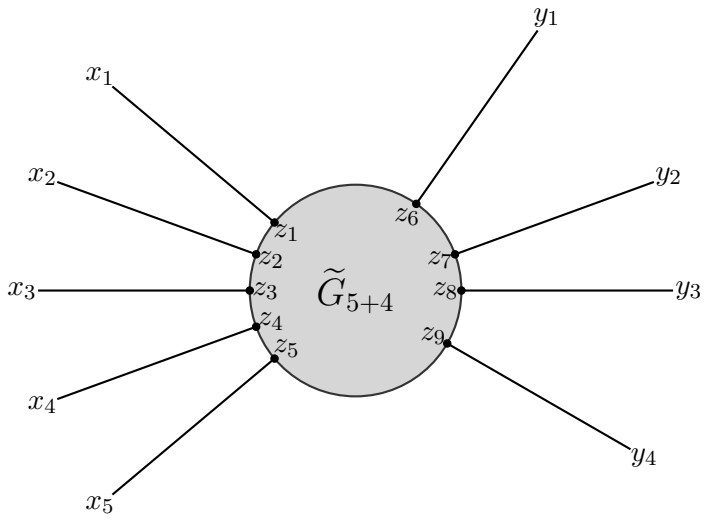
$$\begin{aligned}&\lim_{T\to\infty}\int_{-T}^{T} dy_1^0\partial_{y_1^0}\Big(\int d^3\vec{y}_1 e^{ip_1\cdot y_1}\overset{\leftrightarrow}{\partial}_{y_1^0}\langle p_2...p_n; out|\,\mathcal{T}\big[\phi(y_1)\phi(x_1)\big]\,|q_2...q_m; in\rangle\Big) \\ &\qquad=\lim_{y_1^0\to+\infty}\int d^3\vec{y}_1 e^{ip_1\cdot y_1}\overset{\leftrightarrow}{\partial}_{y_1^0}\langle p_2...p_n; out|\,\phi(y_1)\phi(x_1)\,|q_2...q_m; in\rangle \\ &\qquad\quad-\lim_{y_1^0\to-\infty}\int d^3\vec{y}_1 e^{ip_1\cdot y_1}\overset{\leftrightarrow}{\partial}_{y_1^0}\langle p_2...p_n; out|\,\phi(x_1)\phi(y_1)\,|q_2...q_m; in\rangle\,.\end{aligned}$$

The first term on the right hand side is what we want, (compare it to the proof above), while the second term on the right hand side will again give a connected, but not fully connected, diagram as we get a $a_{\vec{p}_1}$ term acting on the initial state. So we finally see that we can replace the expression we had above for the Green's function (the time-ordered expression) at the cost of picking up a connected, but not fully connected, term. Doing this for all the final states will then give us all the terms in the product from $j=1$ to $j=n$ needed in Equation (8.5). Note also that all the other factors needed for Equation (8.5) also appear, i.e. all the exponentials and the $iZ^{-1/2}$ factors. □

Ok so we have that the fully connected part of the scattering process is given by a $(n+m)$-point Green's function. However the LSZ formula is quite a large formula and so we want to simplify it somehow. We do this by introducing the *truncated Green's function* as follows

$$\begin{aligned}G_{n+m}(y_1,...,y_n,x_1,...,x_m) = \int d^4z_1...d^4dz_{m+n}\Delta_F(x_1-z_1)...\Delta_F(x_m-z_m) \\ \times\,\Delta(y_1-z_{m+1})...\Delta_F(y_n-z_{n+m})\widetilde{G}_{n+m}(z_1,...,z_{n+m}),\end{aligned} \tag{8.6}$$

where Δ_F is the Feynman propagator and $\widetilde{G}_{n+m}$ is our truncated $(n+m)$-point Green's functions. Essentially what we're doing is saying is propagate our initial states to the points $\{z_1, ..., z_m\}$ and similarly the final states with $\{z_{m+1}, ..., z_{m+n}\}$, and characterise all the interaction stuff in the middle using some Green's function, $\widetilde{G}$, as we've tried to indicate below diagrammatically.



Why would we introduce the truncated Green's function? Well we recall that the Feynman propagator is a Green's function of the Klein-Gordan operator:

$$(\partial^2 + m^2)\Delta_F(x - z) = 0,$$

so in the LSZ formula all of the Klein-Gordan operators will just give us delta functions, which we can integrate over, leaving us with

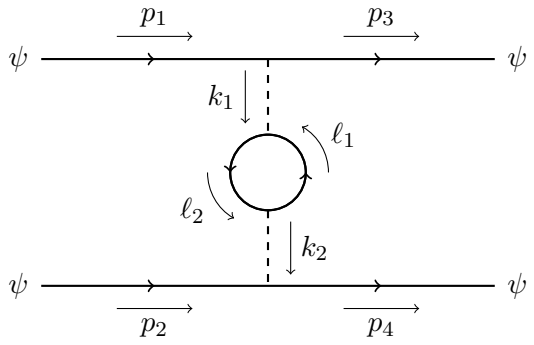
$$\begin{aligned} T_{fi}^{F.C.} &= \left(iZ^{-1/2}\right)^{n+m} \int dy_1 ... dy_n \int dx_1 ... dx_m e^{i(p_j \cdot y_j - q_i \cdot x_i)} \widetilde{G}_{n+m}(x_1 ..., x_m, y_1, ..., y_n) \\ &= \left(iZ^{-1/2}\right)^{n+m} \widetilde{G}_{n+m}(q_1, ..., q_m, p_1, ..., p_n), \end{aligned} \tag{8.7}$$

where the second line follows from the delta functions we get from the exponentials, and where the superscript F.C. means "fully connected". This result tells us that we can compute the entire fully connected S-matrix by just considering the momentum space truncated Green's function! This is a massively convenient result.

Remark 8.0.5. It is worth stressing again that the LSZ theorem does *not* tell us that the connected, but not fully connected, diagrams vanish but that we can ignore them if we want to just consider a scattering process where everything interacts. In fact the former diagrams will contribute to renormalisation and so are very important.

8.1 Loop Diagrams

So far we have only considered what are known as *tree level diagrams*. This essentially just means that all of our vertices are connected to our external lines. However this is clearly not the only type of diagram we can have and we can have vertices internally. For example in the scalar Yakawa theory the following is a valid diagram to order g^2:



where we haven't labelled the internal particles and the $(-ig)$ factors to avoid cluttering the diagram. This is called a *loop diagram* for obvious reasons. Loop diagrams are a bit of a pain because we don't have enough delta functions to remove all of our integrals. That is, recall that in the momentum space Feynman rules we include a factor of[7]

$$\int \frac{d^4k_i}{(2\pi)^2} \frac{i}{k_i^2 - m_{\phi/\psi}^2 + i\epsilon}$$

for each internal line and a delta function of the momentum flowing in at each vertex. For example, in the above diagram the k_1 and k_2 momentum are obtained from the delta functions with the external momenta, i.e. $k_1 = p_1 - p_3$ and $k_2 = p_4 - p_2$. However we cannot find the values of ℓ_1 and ℓ_2, but at best can obtain a relation between them. This is just because ℓ_1 and ℓ_2 will both appear in two delta functions, namely

$$\delta^{(4)}(k_1 + \ell_1 - \ell_2), \quad \text{and} \quad \delta^{(4)}(\ell_2 - k_2 - \ell_1),$$

and so when we integrate over one of the two ℓs we kill both delta functions and we are left with the integral over the other ℓ and no delta function to remove it! To be more explicit, if we did the integral over ℓ_2, say, we would get

$$\ell_2 = k_1 + \ell_1, \quad \text{and} \quad \ell_2 = k_2 + \ell_1,$$

so if we equate the two terms we just get

$$k_1 = k_2 \qquad \Longrightarrow \qquad p_1 + p_2 = p_3 + p_4,$$

where the implication comes from the momentum conservation on the ks. This just tells us the 4-momentum initially and finally are equivalent, and does not allow us to conclude what ℓ_1 is.

Remark 8.1.1. As we said when we introduced the momentum space Feynman rules, it is common to alter the rules slightly to so just include the factor

$$\frac{i}{k^2 - m^2 + i\epsilon}$$

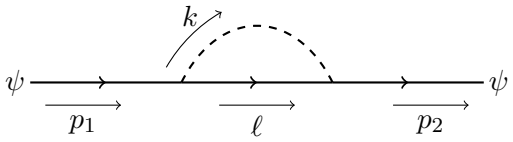
for each propagator and then impose momentum conservation at each vertex. If we use these rules, then when we have loops we insert the additional rule of "and then integrate over all undetermined momenta", i.e. ℓ_1 in the above example.

[7] Note that before we only defined this for ϕ propagators, but here we say we get the same factor for ψ ones but we simply replace the mass with the ψ mass. From now on we shall use $m_\phi = m$ and $m_\psi = M$.

We also get loop diagrams in connected, but not fully connected, diagrams, as in the next exercise.

Exercise

Show that this diagram



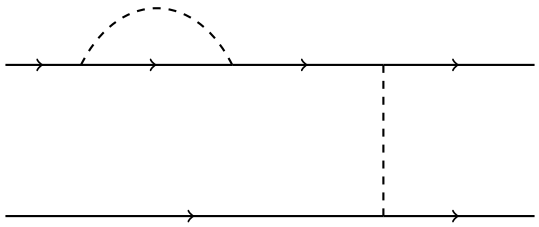
corresponds to

$$(2\pi)^4\delta^{(4)}(p_2-p_1)\left[\int\frac{d^4\ell}{(2\pi)^4}(-ig)^2\left(\frac{i}{(p_1-\ell)^2-m^2+i\epsilon}\right)\left(\frac{i}{\ell^2-M^2+i\epsilon}\right)\right]$$

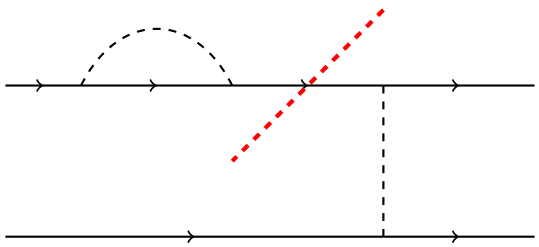
Remark 8.1.2. Note that the result of the previous exercise diverges logarithmically in ℓ. This is easily seen because we essentially have $1/\ell^4$ and we do $d^4\ell$. This kind of divergence is known as a *UV divergence*.

8.2 Amputated Diagrams

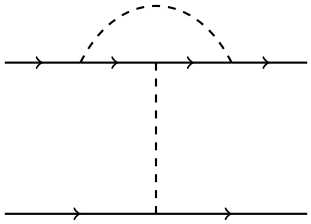
Note in all the Feynman diagrams we have just drawn, we always draw the ingoing and outgoing particles as 'pure' lines, by which I mean nothing is attached to them. An example of something we haven't drawn is the following:



The reason we have never considered these is because what we've been drawing are so-called *amputated diagrams*. You get amputated diagrams by saying "can I cut through the diagram and remove something by only cutting through one line?" If the answer is yes, do it. So for the above example we would cut along the dotted red line below:



which gives us the scattering diagram we drew before. It is important that you only cut through 1 line, i.e. we can not amputate the following



The reason we amputate diagrams is because the LSZ theorem tells us to consider only amputated diagrams. This is seen from the truncated Greens function expression: we consider just propagating our initial and final states in/out without any interactions and then only consider the interactions that happen between all the particles in the internal parts, see the truncated $\widetilde{G}$ picture above.

8.3 Full Feynman Rules For Scalar Yakawa

For completeness/have it all in one place, the table below gives the full set of (momentum space) Feynman rules for Scalar Yakawa theory, the only real change to the table given before is that we can now have antiparticle states $\overline{\psi}$ and we can have internal $\psi/\overline{\psi}$ states and external ϕ states.

Here are some notes to go along with the table:

- As before, multiply all the factors from each part of the diagram together, and include a factor of $(-ig)(2\pi)^4\delta^{(4)}\big(\sum_i k_i\big)$ where k_i is positive if it flows into the vertex and negative if it flows out. Then integrate over all the delta functions, which imposes momentum conservation at every vertex.

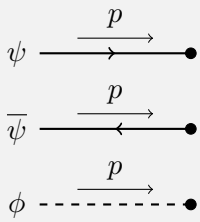
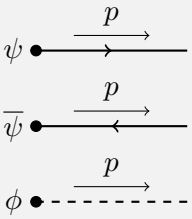
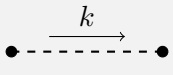
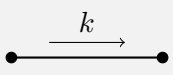
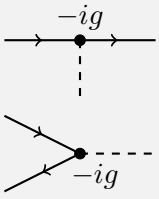
- As mentioned above, if you have a loop in the diagram, you will not have enough delta functions to remove all the integrals, and so you will be left with an integral over the *loop momenta*.

- Note that for the antiparticle $\overline{\psi}$ the charge flow and the momentum flow point in *opposite* directions. This is how we distinguish particle from antiparticle in Feynman diagrams.

- Note that for the internal ψ propagator we don't label the flow of charge (i.e. the arrow on the line itself). This is because there is no way for us to test to see if it is a 'particle going forward in time'[8] or an 'antiparticle going backwards in time'. Obviously this is just lingo to mean there is no way we can test to see which way the arrow points. In either case we get the same result, and so we can just leave it as is. There is no harm in putting the arrow, of course, just if you do you then have to make sure all the other arrows internally agree, i.e. you must always have one arrow in and one arrow out at every vertex.

[8] Recall time flows left to right in our diagrams.

- As mentioned at the end of the last bullet point, we always need one charge arrow in and one charge arrow out of every vertex. In the $\psi\psi \to \psi\psi$ process this meant we could not annihilated the two initial ψs (i.e. they didn't meet at a vertex) as we would then have two charge arrows into the vertex and zero out. However the $\overline{\psi}$s charge arrow points oppositely to ψs, and so now they *can* annihilate. We draw this diagramatically at the bottom of the table.

Type	Diagram	Maths Expression
Incoming External Lines		1
Outgoing External Lines		1
ϕ Propagator		$\int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i\epsilon}$
ψ Propagator		$\int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - M^2 + i\epsilon}$
Vertices		$-(ig)$

9 | ϕ^4 Theory & Cross Sections

9.1 ϕ^4 Theory

So we have looked at scalar Yakawa theory in some detail and derived the full set of Feynman rules for this theory. We now want to do the other spin-0 theory we have been studying during this course, ϕ^4 theory. Recall that here we have

$$H_I = \int d^3\vec{x} \frac{\lambda}{4!}\phi^4(x),$$

so we can use Dyson's formula and Wick's theorem to find the the S-matrix, or more correctly T-matrix, as mentioned before. Let's consider the scattering

$$\phi(p_1)\phi(p_2) \to \phi(p_3)\phi(p_4).$$

Remark 9.1.1. With the LSZ theorem in mind, we shall only consider fully connected diagrams, which corresponds to only considering terms where everything is contracted. This should be reasonably clear, but if it is not, go back to the definition of what Wick contractions give us and convince yourself this is true.

9.1.1 First Order

To first order in λ we have

$$\begin{aligned} T_{fi} = \langle f| S - \mathbb{1} |i\rangle &= \left\langle \vec{(}p_4)\vec{(}p_3)\right| \mathcal{T}\bigg(-\frac{i\lambda}{4!}\int d^4x \phi_x\phi_x\phi_x\phi_x \bigg) |\vec{p}_1\vec{p}_2\rangle \\ &= -\frac{i\lambda}{4!}\int d^4x \, \langle \vec{p}_4\vec{p}_3| \, \phi_x\phi_x\phi_x\phi_x \, |\vec{p}_1\vec{p}_2\rangle \end{aligned}$$

There is only one topologically distinct[1] contraction given by

$$T_{fi} = -\frac{i\lambda}{4!}\int d^4x \, \langle \vec{p}_4\vec{p}_3| \, \overbrace{\phi_x\phi_x}\overbrace{\phi_x\phi_x} \, |\vec{p}_1\vec{p}_2\rangle + \text{permutations},$$

where by permutations we mean simply things like contract that $\vec{p}_4$ with any of the other 3 ϕ_xs etc.

[1] I.e. the diagrams are indistinguishable. For those interested, I think the topology condition of 'distinction' used here is the notion of homotopy, but I could be wrong.

Exercise

Convince yourself that there are exactly 4! different permutations for the above contraction. *Remark: As mentioned before, this result is exactly why we include the factor of* $1/4!$ *in the Lagrangian in the first place, i.e. it is a symmetry factor.*

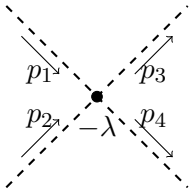
Exercise

Show that the Wick contractions above lead to

$$iM = -i\lambda.$$

Hint: Use the result from the previous exercise along with Equation (7.6).

The Feynman rules rules are the same as before but now we have the vertex factor $-\lambda$. The above scatting corresponds to the diagram

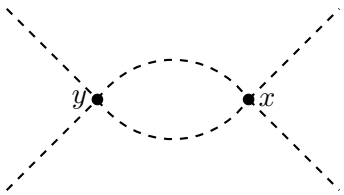


9.1.2 Second Order & Symmetry Factors

At order λ^2 we get the following contractions

$$T_{fi} = \frac{1}{2!}\left(-\frac{i\lambda}{4!}\right)^2 \int d^4x\, d^4y\, \langle \vec{p}_4\vec{p}_3|\, \phi_x\phi_x\phi_x\phi_x\phi_y\phi_y\phi_y\phi_y\, |\vec{p}_1\vec{p}_2\rangle + \text{permutations}$$

We also get the terms with $x \longleftrightarrow y$. Recall that the contraction between a ϕ_x and a ϕ_y means that we have a ϕ propagating between x and y, so in position space the diagram looks like[2]



How do we find the number of permutations for this diagram? Well let's just consider the diagram as above, i.e. don't consider switching $x \longleftrightarrow y$, and then at the end we can just double the answer at the end. So how do we find this? The answer is to go one Wick contraction at a time. Let's first considering the contractions with the final states: these contract with ϕ_xs, of which there are 4, so the first contraction (say with $\vec{p}_4$) has 4 choices,

[2]We should label the momentum on the external states, but to save me doing some extra Tikz I've left them out.

and then the second contraction ($\vec{p}_3$ in this case, has 3 left options. To be a bit more explicit, we have for the first contraction

$$\overbrace{\langle\vec{p}_4\vec{p}_3|\,\phi_x\phi_x\phi_x\phi_x} + \overbrace{\langle\vec{p}_4\vec{p}_3|\,\phi_x\phi_x\phi_x\phi_x} + \overbrace{\langle\vec{p}_4\vec{p}_3|\,\phi_x\phi_x\phi_x\phi_x} + \overbrace{\langle\vec{p}_4\vec{p}_3|\,\phi_x\phi_x\phi_x\phi_x},$$

and then for the second contraction each term above has three options, e.g. for the last term we have

$$\overbrace{\langle\vec{p}_4\vec{p}_3|\,\phi_x\phi_x\phi_x\phi_x} + \overbrace{\langle\vec{p}_4\vec{p}_3|\,\phi_x\phi_x\phi_x\phi_x} + \overbrace{\langle\vec{p}_4\vec{p}_3|\,\phi_x\phi_x\phi_x\phi_x}.$$

So from this part we get $4 \cdot 3$ perturbations. Similarly the contractions of the initial states with the ϕ_ys give another $4 \cdot 3$ perturbations. Then we have just left with the contractions between the two remaining ϕ_xs and ϕ_ys, of which there are only two choices[3]

$$\overbrace{\phi_x\phi_x\phi_y\phi_y} + \overbrace{\phi_x\phi_x\phi_y\phi_y}.$$

So in total we have

$$(4 \cdot 3) \cdot (4 \cdot 3) \cdot 2 \cdot 2 = 4!4!,$$

where the second 2 factor is the $x \longleftrightarrow y$ factor. So the numerical prefactor of the Wick contraction is

$$\frac{1}{2}\frac{4!4!}{4!4!} = \frac{1}{2},$$

so we get a *symmetry factor* of 1/2. Symmetry factors arise when elements of the diagrams can be indistinguishably interchanged.

9.2 Decay Rates & Cross Sections

So far we have found the amplitude for an interaction to occur, but in quantum theory the amplitudes themselves aren't measurable. What we measure is the *probability* for a process to occur, and that's what we want to talk about now. However there appears to be a problem: the probability is given by taking the absolute value squared of the amplitudes,

$$P = \frac{|\,\langle f|\,S - \mathbb{1}\,|i\rangle\,|^2}{\langle f|f\rangle\,\langle i|i\rangle}, \tag{9.1}$$

but we've already seen that

$$\langle f|\,S - \mathbb{1}\,|i\rangle \sim \delta^{(4)}(p_I - p_F),$$

where I/F stands for initial/final. So the probability is given by squaring a delta function... what does this even mean? The answer is related to what we discussed earlier in the course when we had delta functions popping up, its to do with the fact that our states are defined everywhere and are normalised to delta functions.[4] This basically means they are not square-integrable states. The claim is that if we put the theory in a box of finite volume V we will get square-integrable states, and this result will be smooth in the limit $V \to \infty$. This takes care of the spatial part of the delta function, but we need to also take care of the temporal

[3] Note we cannot contract ϕ_x with ϕ_x and similarly for ϕ_ys.

[4] Note this also tells us that the denominator of Equation (9.1) is also two delta functions!

delta function and so we consider the process happening in some finite time T, and then again take the limit $T \to \infty$. This will give us some formula for the probability, at which point we have to realise that we cannot measure the final state momentum perfectly, but that they will be smeared over some region in momentum space. To account for this, we must therefore integrate over this region. We call this part of the formula the *Lorentz invariant phase space measure*, or LIPS for short, and it is given by

$$d\Pi = (2\pi)^4 \delta^{(4)}(p_F - p_I) \prod_{\text{final states}} \frac{d^3\vec{p}_i}{(2\pi)^3} \frac{1}{2E_{\vec{p}_i}}. \tag{9.2}$$

We get the result for the differential probability per unit time to be

$$\frac{dP}{T} = |iM_{fi}|^2 V d\Pi \prod_{\text{initial states}} \frac{1}{2VE_{\vec{p}_i}} \tag{9.3}$$

We refer to the product over initial states as the *flux factor*.

We do not provide a derivation of the proof here as little insight is gained from the steps and we just spent almost an entire lecture deriving the LSZ theorem. However a proof using the "put it in a box" argument can be found in Prof. Michael Luke's notes.[5]

9.2.1 Decay Rates

Ok so let's look at a decay process. Here we only have one initial particle, or energy E, and so we get

$$d\Gamma := \left.\frac{dP}{T}\right|_{1 \text{ particle}} = \frac{1}{2E}|iM_{fi}|^2 d\Pi,$$

where we note that the volume factors in Equation (9.3) have cancelled out. This is very good because we want to take the limit $V \to \infty$ without hitting any problems. We call $d\Gamma$ the *differential decay rate probability*.[6] If we go to the particle's rest frame we can replace $E = m$ (where m is the mass of our particle), giving

$$d\Gamma = \frac{1}{2m}|iM_{fi}|^2 d\Pi$$

To get the full decay probability, we need to consider the fact that our particle could decay into a multitude of different states with different numbers (and/or types) or particles. This gives us

$$\Gamma = \frac{1}{2m} \sum_{\text{final states}} \int |iM_{fi}|^2 d\Pi, \tag{9.4}$$

known as the *decay width* of the particle. The decay width tells us the probability for the particle to decay in any way shape or form, and so it is clearly related to the lifetime of the particle by

$$\tau = \frac{1}{\Gamma} \tag{9.5}$$

[5] Note to self: maybe come back later and do the proof. Just don't think its worth the time typing it out now.

[6] Or maybe differential decay probability per unit time? Basically "rate" here means take it per unit time.

Remark 9.2.1. We call Γ the decay *width* because of the uncertainty principle: there is an uncertainty in the energy of the particle, and so there is an uncertainty in the lifetime, and so we do not get an exact value for the decay probability but a spread, or width.

Remark 9.2.2. As we stated it, the sum in Equation (9.4) only accounts for final states that are distinguishable (i.e. the number of particles and/or the type of particles change). However we just saw in the previous section that certain processes come with symmetry factors, and so we should slightly alter our expression to include these. That is we send

$$\Gamma \to S \cdot \Gamma,$$

where S is the symmetry factor.

9.2.2 Cross Sections

Now let's consider the system with two initial particles. Obviously in this case we need them to collide. The way we do this experimentally is to send two beams of particles flying at each other and get them to collide. Obviously not every particle collides, and so we need some measure of the fraction that do. It is obvious that the cross sectional area of the beams will effect the number of measurements that occur, and so we define the *flux*, F, to be the number of incoming particles per unit area per unit time. We can then characterise the number of scatterings per unit time via

$$dN = F d\sigma,$$

where $d\sigma$ is called the *differential cross section*, which is the probability for a scattering to happen in the solid angle (θ, ϕ). It is given by

$$d\sigma := \frac{\text{Differential Probability}}{\text{Unit Time} \times \text{Unit Area}} = \frac{1}{4E_1E_2V}\frac{1}{F}|iM_{fi}|^2 d\Pi, \tag{9.6}$$

where E_1/E_2 are, of course, the energies of our two initial state particles.

At first this looks a bit problematic because we still have a factor of V on the right-hand side. However, we still need to insert the flux. So what is it? Well imagine that the beam is moving with 3-velocity $\vec{v}$ perpendicular to some plane of area A. If the density of the particles in the beam if ρ, then the total number of particles passing through the plane in time t is simply

$$N = A|\vec{v}|\rho t, \qquad \implies \qquad F = \frac{N}{At} = |\vec{v}|\rho.$$

In the procedure above when we put the theory in a box, we did it in a normalisation where there was exactly 1 particle in each volume V and so $\rho = 1/V$. So if we go to the centre of mass frame of the collision, we have

$$F = \frac{|\vec{v}_1 - \vec{v}_2|}{V},$$

where $\vec{v}_1/\vec{v}_2$ are the 3-velocities of our beams. Plugging this into Equation (9.6), we get

$$d\sigma = \frac{1}{4E_1E_2}\frac{1}{|\vec{v}_1 - \vec{v}_2|}|iM_{fi}|^2 d\Pi. \tag{9.7}$$

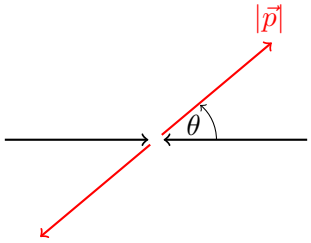
Of course we get the full cross section by the same idea as for the decay rate: integrate of the LIPS, sum over all the different final states and then multiply by the symmetry factor.

9.2.3 Two Particle Final States

As we have said, the above formulas hold for general final states, however there is a particularly nice final state, namely that with two particles in it. Why is this a nice final state? Well we can see this from our LIPS measure, Equation (9.2), which becomes

$$d\Pi_2 = (2\pi)^4\delta^{(4)}(p_1 + p_2 - p_I)\frac{d^3\vec{p}_1}{(2\pi)^3 2E_{\vec{p}_1}}\frac{d^3\vec{p}_2}{(2\pi)^3 2E_{\vec{p}_2}}.$$

The two $d^3\vec{p}_i$s give us 6 degrees of freedom in the final state (i.e. the 3 components of each final state particle), however the delta function kills 4 of these degrees of freedom leaving us with just 2. These two degrees of freedom can be made very simple by going to the centre of mass frame for the collision. Here we know that the final state particles must have equal and opposite momentum, and so we can categorise the whole phase space by simply specifying the modulus of the momentum and the angle at which these particles fly off relative to the collision axis. In the diagram below, the red arrows are meant to be the final state particles.



Doing this explictly, we have

$$\begin{aligned}
\Pi_2 &= \frac{1}{(2\pi)^2}\int \frac{d^3\vec{p}_1}{2E_{\vec{p}_1}}\frac{d^3\vec{p}_2}{2E_{\vec{p}_2}}\delta^{(4)}(p_1 + p_2 - p_I) \\
&= \frac{1}{(2\pi)^2}\int \frac{d^3\vec{p}_1}{2E_{\vec{p}_1}}\frac{d^3\vec{p}_2}{2E_{\vec{p}_2}}\delta(E_1 + E_2 - E_T)\delta^{(3)}(\vec{p}_1 + \vec{p}_2) \\
&= \frac{1}{(2\pi)^2}\int \frac{d^3\vec{p}_1}{4E_{\vec{p}_1}^2}\delta(E_1 + E_2 - E_T) \\
&= \int \frac{d\Omega}{(2\pi)^2}\int \frac{d|\vec{p}_1|}{4E_{\vec{p}_1}E_{\vec{p}_2}}|\vec{p}_1|^2\delta\big(E_1 + E_2 - E_T\big)
\end{aligned}$$

where E_T is the total energy of the system. For clarity, we have gone to the centre of mass frame to split the delta function, done the $d^3\vec{p}_2$ integral and then gone to polar coordinates.

Now let's use the relationship

$$E_1 = \sqrt{\vec{p}_1^2 + m_1^2},$$

so that the delta function contains a $|\vec{p}_1|$, and then we use the delta-function identity

$$\delta\big(f(x)\big) = \frac{1}{|f'(x_0)|}\delta(x - x_0),$$

where $f(x_0) = 0$, which gives us a factor of

$$\left|\frac{\partial(E_1 + E_2)}{\partial|\vec{p}_1|}\right|^{-1}.$$

Now using $|\vec{p}_2| = |\vec{p}_1|$ we have

$$\frac{\partial E_1}{\partial|\vec{p}_1|} = \frac{|\vec{p}_1|}{E_1}, \quad \text{and} \quad \frac{\partial E_1}{\partial|\vec{p}_1|} = \frac{|\vec{p}_1|}{E_2},$$

and so

$$\left|\frac{\partial(E_1 + E_2)}{\partial|\vec{p}_1|}\right| = \frac{|\vec{p}_1|(E_1 + E_2)}{E_1 E_2} = \frac{|\vec{p}_1|E_T}{E_1 E_2}.$$

Putting this into our expression for the LIPS measure, we can conclude

$$d\Pi_2 = \frac{1}{16\pi^2}\frac{|\vec{p}_1|d\Omega}{E_T}$$

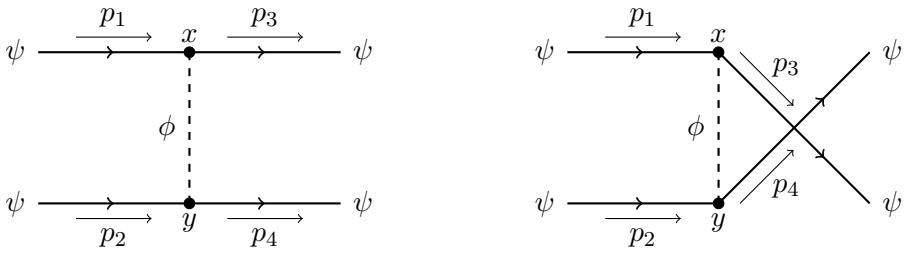
Remark 9.2.3. Note that this formula is valid for *both* the decay process and the scattering process, and we have derived it for distinguishable particles. We can get the case of identical particles easily by setting $m_1 = m_2$. Some authors, e.g. Prof. Michael Luke (where I got this proof from), say that if you take identical particles you need to include a factor of $1/2!$ to account for it. This is exactly the symmetry factor we have accounted for separately in our decay width/cross section and so we do not need to include it here.

Let's use this to find the differential cross section for $2 \to 2$ scattering, e.g. $\psi(p_1)\psi(p_2) \to \psi(p_3)\psi(p_4)$. Firstly note that we have to put $|\vec{p}_3|$ (or $|\vec{p}_4|$ in the place of $|\vec{p}_1|$ in the above formula, as it is the final state particle momentum we are talking about here.[7] We have

$$\frac{d\sigma}{d\Omega} = \frac{1}{4E_1E_2}\frac{1}{|\vec{v}_1 - \vec{v}_2|}\frac{|\vec{p}_3|}{16\pi^2 E_T}|iM_{fi}|^2. \tag{9.8}$$

9.2.4 Mandelstam Variables & Virtual Particles

So now we need to find the matrix elements. We showed before that there are two diagrams:

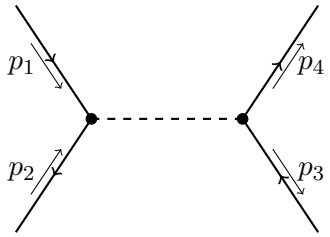


[7] I considered calling them p_3/p_4 from the beginning but that might suggest we are considering the scattering from the get go, but as the above remark says, this result also holds for the decay, so I decided against it.

and these correspond to the matrix element terms

$$(-ig)^2 \frac{i}{(p_1 - p_3)^2 - m^2 + i\epsilon}, \quad \text{and} \quad (-ig)^2 \frac{i}{(p_1 - p_4)^2 - m^2 + i\epsilon},$$

respectively. These two diagrams are referred to as t and u diagrams, respectively. As we mentioned with the Feynman rules at the end of the last lecture, in $2 \to 2$ scattering we can also have a diagram that looks like



which is known as an s diagram. Collectively these three types of diagrams describe all possible (tree level, i.e. no loops) $2 \to 2$ scattering processes, and the variables s, t, u are known as the *Mandelstam variables*. They are given by the momentum that flows through the propagator,

$$\begin{aligned} s &:= (p_1 + p_2)^2 = (p_3 + p_4)^2 \\ t &:= (p_1 - p_3)^2 = (p_2 - p_4)^2 \\ u &:= (p_1 - p_4)^2 = (p_2 - p_3)^2 \end{aligned} \tag{9.9}$$

Exercise

Show that

$$s + u + t = \sum_{i=1}^{4} m_i^2 \tag{9.10}$$

where the sum is done over all 4 external particles.

Note that s tells us the centre of mass energy squared, i.e.

$$s = E_{CM}^2. \tag{9.11}$$

We can use the Mandelstam variables to study our $\psi\psi \to \psi\psi$ scattering a bit further. Firstly note that the matrix elements given above become

$$(-ig)^2 \frac{i}{t - m^2 + i\epsilon}, \quad \text{and} \quad (-ig)^2 \frac{i}{u - m^2 + i\epsilon}. \tag{9.12}$$

Now we can go to the centre of mass frame to find two of the three variables in terms of our two free parameters $|\vec{p}_3|$ and θ. In this frame we have

$$\begin{aligned} p_1 &= (E, 0, 0, p) \\ p_2 &= (E, 0, 0, -p)) \\ p_3 &= (E, 0, p\sin\theta, p\cos\theta) \\ p_4 &= (E, 0, -p\sin\theta, -p\cos\theta). \end{aligned}$$

We can re-express these using Equation (9.11) and

$$\beta_M := \sqrt{1 - \frac{4M^2}{s}},$$

to give us

$$\begin{aligned}
p_1 &= \frac{\sqrt{s}}{2}(1, 0, 0, \beta_M) \\
p_2 &= \frac{\sqrt{s}}{2}(1, 0, 0, -\beta_M) \\
p_3 &= \frac{\sqrt{s}}{2}(1, 0, \beta_M \sin\theta, \beta_M \cos\theta) \\
p_4 &= \frac{\sqrt{s}}{2}(1, 0, -\beta_M \sin\theta, -\beta_M \cos\theta).
\end{aligned}$$

So we conclude

$$\begin{aligned}
t &= \frac{s}{2}\big(0, 0, -\beta_M \sin\theta, \beta_M(1 - \cos\theta)\big)^2 \\
&= -\frac{1}{2}s\beta_M^2(1 - \cos\theta)
\end{aligned}$$

which is *always* negative.

Exercise

Show similarly that $u < 0$ here.

So why have we bothered doing this? Well firstly we can substitute these expressions for t and u into our matrix elements and then take the complex norm squared to find our differential cross section, Equation (9.8). However it allows us to see something else quite interesting: because both t and u are negative we see the denominators of Equation (9.12) never vanish. This is the statement that the propagator is never on shell (i.e. $p^2 = m^2$ never occurs). We refer to such particles as *virtual particles*.

10 | Dirac Fields

So far all we have considered are Lorentz scalars, i.e. fields that transform as

$$\phi(x) \to \phi'(x) = \phi\big(\Lambda^{-1}x\big)$$

under Lorentz transformations. We saw (it was set as an exercise in lecture 3, the one about showing $J^i\,|\vec{p}=0\rangle = 0$) that quantised version of such fields have zero spin. Obviously we know this is not the case in the real world, and we need to study other fields that transform differently under Lorentz transformations. In particular we want to find something that once we quantise it will give us spin-1/2 particles, as these are what electrons and the like are. It is important to note that there is no a priori way to see what kind of transformation property we would need to produce spin 1/2 particles, but we would have to undergo the full calculation for a given transformation and then see what the resulting spin in. Luckily for us, however, Dirac has basically done this for us and so we do know what transformation property we require.

Remark 10.0.1. This part of the course was forced to be a bit rushed (due to lack of time) and so in order to try help expand on some of the points I am going to use Prof. Tong's notes as a heavy reference. I shall try outline most of the important points given in his notes, however I don't see the point in just copying his explanations out, so if there is anything I say that sounds confusing, I strongly recommend having a look in his notes to see if there is a more detailed explanation there.

10.1 Representations Of Lorentz Group

As we have tried to make clear in these notes, Lorentz invariance is crucial to QFT and is a vital part of the form we allow our Lagrangian to take. We have also just claimed that fields that transform in different ways under the Lorentz transformation will give rise to different types (namely different spins) of particle. In order to talk about how something transforms under the action of a group we use a *representation*, i.e.[1]

$$\phi^{ab\dots} \to \Big(D[\Lambda]^a{}_{a'} D[\Lambda]^b{}_{b'} \dots\Big)\phi^{a'b'\dots},$$

where $D[\Lambda]$ are matrices. The obvious question to ask is "what are the representations of the Lorentz group?" This is the question we shall now try answer.

[1] See my notes on Dr. Dorigoni's Group Theory course for more details.

10.1.1 The Generators

The standard way to find representations of the Lie group is to study the associated Lie algebra. We do this because the latter, being an algebra, has a basis and so are *much* easier to deal with. In order to obtain the Lie algebra, we look infinitesimally close to the identity, and we have seen, Equations (2.10) and (2.11), that for the Lorentz group this corresponds to

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu}, \qquad \text{with} \qquad \omega^{\mu\nu} = -\omega^{\nu\mu}. \tag{10.1}$$

As we have mentioned before[2] the antisymmetric nature of the $\omega^{\mu\nu}$ tells that we have $4(4-1)/2 = 6$ independent transformations, these are the 3 spatial rotations and 3 boosts. The $\omega^{\mu\nu}$ are elements elements in the Lie algebra, and as just mentioned, we want to find a basis to express these in. That is, we want to find 6 linearly independent, 4×4 antisymmetric matrices such that

$$\omega^{\mu\nu} = \Omega_a (\mathcal{M}^a)^{\mu\nu},$$

where $a \in \{1, ..., 6\}$ and Ω_a are just some numbers. We call the basis the *generators* of our Lie algebra, i.e. $(\mathcal{M}^a)^{\mu\nu}$ are the generators of (Lie algebra of) the Lorentz transformations.

In fact, for a reason that will appear clearer soon, it is beneficial to label the basis elements with 2 antisymmetric indices instead, that is:

$$\omega^{\mu\nu} = \Omega_{\rho\sigma} (\mathcal{M}^{\rho\sigma})^{\mu\nu}, \tag{10.2}$$

where the $\rho, \sigma \in \{0, ..., 3\}$ are antisymmetric, that is $\Omega_{\rho\sigma} = -\Omega_{\sigma\rho}$ and the same for $\mathcal{M}$. We note that we still only have 6 basis elements, as ρ/σ can take 4 values, but the antisymmetry gives us $4(4-1)/2 = 6$ independent choices. This is of course a necessary condition (otherwise we would be changing the dimension of our algebra!).

Remark 10.1.1. It is worth clarifying the indices here. The $\rho\sigma$ values tell us which basis element we are considering. Each $\mathcal{M}^{\rho\sigma}$ is a 4×4 matrix and the $\mu\nu$ indices tell us which element of this matrix we are considering. These two set of indices have nothing to do with each other, despite them being of very similar form.

So how do we decide on what form the basis takes? The answer is we use the one that we know gives us the answer we want and look super clever at the end for getting it correct. We use

$$(\mathcal{M}^{\rho\sigma})^{\mu\nu} = \eta^{\rho\mu}\eta^{\sigma\nu} - \eta^{\sigma\mu}\eta^{\rho\nu},$$

where on the right-hand side all the indices tell us the elements of the metric matrix, but we should keep the above remark in mind about what the indices mean in terms of our Lie algebra.

Now, looking at Equation (10.1), it is clear that we actually want to lower one of the μ/ν indices so that we can see how it acts on our fields. Using $\eta^{\mu\nu}\eta_{\nu\sigma} = \delta^{\mu}{}_{\sigma}$, this gives us

$$(\mathcal{M}^{\rho\sigma})^{\mu}{}_{\nu} = \eta^{\rho\mu}\delta^{\sigma}{}_{\nu} - \eta^{\sigma\mu}\delta^{\rho}{}_{\nu}, \tag{10.3}$$

and so we have

$$\omega^{\mu}{}_{\nu} = \frac{1}{2}\Omega_{\rho\sigma}(\mathcal{M}^{\rho\sigma})^{\mu}{}_{\nu}, \tag{10.4}$$

[2] Or at least I think it's mentioned in these notes, if not it's mentioned now.

where we have introduce a factor of 1/2 for convention.[3]

Remark 10.1.2. Note that by lowering one of the indices we have broken the antisymmetry property of $\mathcal{M}$. For examples see Prof. Tong's notes, page 82.

Exercise

Show that Equations (10.3) and (10.4) imply that

$$\omega^{\mu\nu} = \Omega^{\mu\nu}. \tag{10.5}$$

Exercise

Prove that the the above basis satisfy the Lie bracket (here just commutator) relation

$$[\mathcal{M}^{\rho\sigma}, \mathcal{M}^{\tau\rho}] = \eta^{\sigma\tau}\mathcal{M}^{\rho\rho} - \eta^{\rho\tau}\mathcal{M}^{\sigma\rho} + \eta^{\rho\rho}\mathcal{M}^{\sigma\tau} - \eta^{\sigma\rho}\mathcal{M}^{\rho\tau}. \tag{10.6}$$

Hint: Use the fact that each $\mathcal{M}^{\rho\sigma}$ is a 4×4 matrix so that

$$\left(\mathcal{M}^{\rho\sigma} \circ \mathcal{M}^{\tau\rho}\right)^{\mu}{}_{\nu} = (\mathcal{M}^{\rho\sigma})^{\mu}{}_{\chi}(\mathcal{M}^{\tau\rho})^{\chi}{}_{\nu}$$

and then use the answer to guide you. You want to end up with an expression that is a matrix labelled by $\mu\nu$, i.e. something of the form $(...)^{\mu}{}_{\nu}$.

This construction has been done infinitesimally, we can recover some finite Lorentz transformation by taking the exponential of the result, i.e.

$$\Lambda = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma}\right). \tag{10.7}$$

So we now have a way to see if something is a representation of the Lorentz transformations, namely if they satisfy Equation (10.6) then they form a representation. The question is "what satisfies this and how do we find one?" The answer we give here is that we know the answer to this question and so just show that it satisfies it. It is important to note that there is no easy way to see that the following construction satisfies Equation (10.6) a priori, but it is simply that we already know it does.

10.1.2 The Dirac Representation

We define the set of matrices $\{\gamma^0, \gamma^1, \gamma^2, \gamma^3\}$, which form a so-called *Clifford algebra*[4]

$$\{\gamma^{\mu}, \gamma^{\nu}\} = 2\eta^{\mu\nu}\mathbb{1}_{n\times n} \tag{10.8}$$

from which is follows that

[3] Note we can do this as it is essentially just a redefinition of the $\Omega_{\rho\sigma}$, which are just 6 numbers.
[4] For completeness, the curly brackets here mean the anticommutator, defined by $\{A, B\} = AB + BA$.

$$\left(\gamma^0\right)^2 = \mathbb{1}_{n\times n}, \quad \text{and} \quad \left(\gamma^i\right)^2 = -\mathbb{1}_{n\times n} \tag{10.9}$$

where $i = 1, 2, 3$. We call these the *gamma matrices*.

We have yet to specify the dimension of these matrices, i.e. what is the value of n? We can show that there is no way to satisfy both these conditions using anything less than a 4×4 matrix. There is *not* a unique set of 4×4 matrices that will satisfy these conditions, but the simplest is

$$\gamma^0 = \begin{pmatrix} \mathbb{1}_{2\times 2} & 0 \\ 0 & -\mathbb{1}_{2\times 2} \end{pmatrix}, \quad \text{and} \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \tag{10.10}$$

where $\{\sigma^i\}_{i=1,2,3}$ are the *Pauli matrices*. This is a representation of the Clifford algebra, and we refer to this particular one as the *Dirac basis*.

Remark 10.1.3. Note we can write Equation (10.10) in a more compact form using the tensor product, namely

$$\gamma^0 = \sigma^3 \otimes \mathbb{1}_{2\times 2}, \quad \text{and} \quad \gamma^j = i\sigma^2 \otimes \sigma^j.$$

This notation makes doing long manipulations with the gamma matrices easier. However in this course we will not use this notation any further.

Remark 10.1.4. The definition we have given here, Equation (10.10), differs from the one given by Prof. Tong, who uses the so-called Weyl basis. The only difference between the two is that in the *Weyl* basis we have

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1}_{2\times 2} \\ \mathbb{1}_{2\times 2} & 0 \end{pmatrix}.$$

The Dirac basis and Weyl basis are related by a simple equivalence transformation, and so they are equivalent representations of our Clifford algebra. It turns out that the Dirac basis is useful for studying massive particles, while the Weyl basis is useful for studying massless particles. It is for this reason that we also refer to the Weyl basis as the *chiral* basis (its helps to study the chirality of massless particles).

Now you would be very justified in asking what this has to do with the Lorentz group and the content of the previous subsection. The answer is it turns out we can combine these gamma matrices in such a way as to produce something which obeys Equation (10.6), and so form a representation of the Lorentz group. Ok so what is it? Well the first thing we note is that it has to contain products of 2 gamma matrices, as $\mathcal{M}^{\rho\sigma}$ has 2 indices. Perhaps the two most natural things to try (given that we want to obtain a commutator relation, Equation (10.6)) is the commutator and anticommutator. The anticommutator wont do because Equation (10.8) shows us that the commutator of these will vanish, but what about the commutator? This will turn out to work.

Exercise

Show that

$$S^{\rho\sigma} := \frac{1}{4}[\gamma^\rho, \gamma^\sigma] = \frac{1}{2}\gamma^\rho\gamma^\sigma - \frac{1}{2}\eta^{\rho\sigma}. \tag{10.11}$$

Hint: Consider the cases $\rho = \sigma$ and $\rho \neq \sigma$ separately and then use Equation (10.8) to get the $\eta^{\rho\sigma}$ term in the result.

Remark 10.1.5. As with $\mathcal{M}^{\rho\sigma}$ above, for each $\rho\sigma$ value we have a 4×4 matrix (its a product of gamma matrices), and so we need to introduce another two indices in order to label the entries of these matrices. We shall use $\alpha, \beta = 1, 2, 3, 4$ to label these.[5]

Claim 10.1.6. The matrices $S^{\rho\sigma}$ satisfy Equation (10.6) and so form a representation of the Lorentz group.

Proof. See page 84 of Prof. Tong's notes. (Or if you're feeling up for it, give it a bash yourself) □

Remark 10.1.7. Note that Equation (10.11) is satisfied at the level of the Clifford algebra. That is, we do *not* need to give a specific representation to obtain it. This is obviously a requirement if we want to use this result going forward.

10.1.3 Spinors

As we have said the $S^{\rho\sigma}$ are 4×4 matrices that form a representation of the Lorentz transformations. We have also said, being matrices they carry two indices, which we label by α and β, and so we see that the object they act on carries one index. That is

$$\psi^\alpha(x) \to S[\Lambda]^\alpha{}_\beta \psi^\beta\left(\Lambda^{-1}x\right), \tag{10.12}$$

where similarly to Equation (10.7), we have

$$S[\Lambda] = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right), \tag{10.13}$$

where we include the $[\Lambda]$ to help us distinguish this from the $S^{\rho\sigma}$s. Note that we have *the same* numbers $\Omega_{\rho\sigma}$ appearing in the exponential. This is just because we want the action of this representation to apply the same transformation to the field ψ^a as it does to x.

Remark 10.1.8. Note that both $S[\Lambda]$ and Λ are 4×4 matrices and it is entirely possible that we haven't constructed anything new. That is it could be true that $S[\Lambda]$ is equivalent (in a representation theory sense) to Λ. It turns out that we haven't done this, and a proof of this can be found in Prof. Tong's notes, pages 85 and 86.

[5] Note we do not start from 0 here. This is because we're talking about the entries of a matrix, *not* a spacetime index or a gamma index.

10.1.4 Non-Unitarity Of $S[\Lambda]$

Claim 10.1.9. The representation $S[\Lambda]$ is not completely unitary. In particular it it the boost part of the Lorentz transformations that are not unitary.

Proof. From Equation (10.13), it follows that if $S[\Lambda]$ was to be unitary then we would require

$$\left(S^{\rho\sigma}\right)^\dagger = -S^{\rho\sigma},$$

i.e. it is antihermitian, as we take $\Omega_{\rho\sigma}$ to be real numbers.[6] However it follows from Equation (10.11) that this antihermitian condition would require

$$\left(\gamma^\rho\gamma^\sigma\right)^\dagger = \gamma^\sigma\gamma^\rho,$$

for all $\rho, \sigma = 0, 1, 2, 3$. This is the statement that we require that all the gamma matrices be hermitian or antihermitian. We cannot achieve this, though, as Equation (10.9) tells us that γ^0 has real eigenvalues (i.e. hermitian) but the γ^i have imaginary eigenvalues (i.e. antihermitian). □

This result will prove very important in just a moment.

10.2 The Dirac Action & Equation

Dirac wanted to try and obtain an equation of motion that was linear in derivatives. That is he wanted something of the form

$$(A^\mu\partial_\mu - m)\psi(x) = 0,$$

where the A^μ is something included in order to ensure a Lorentz invariant result. The question is "what is A^μ?" Well given that we've just spent a bunch of time talking about them, it might seem sensible to suggest a gamma matrix, i.e. try something of the form

$$(\gamma^\mu\partial_\mu - m)\psi(x) = 0.$$

If we were to obtain this result, we would have to start from something which is bilinear in the field $\psi(x)$, as we differentiate one away in the Euler-Lagrange equations. We also note that $\psi(x)$ here is a column matrix (otherwise acting on it with the gamma matrix wouldn't make any sense),[7] so if the action meant to just be a number (which it is) we also need a transpose of $\psi(x)$ appearing to the left of ψ in our action. Then we finally recall that it is QFT we want to go on and study,[8] so its best to be safe and turn this transpose into a hermitian conjugate. Perhaps a better argument for this latter point is we want the result the action gives us to be a *real* number. We therefore make the naïve guess

$$\mathcal{L}_{\text{naïve}} = \psi^\dagger\gamma^\mu\partial_\mu\psi - m\psi^\dagger\psi.$$

[6] They are the rotation angles/boost speeds.

[7] Note for this reason we really should write $m\mathbb{1}$ for the mass term. However it is standard to be a bit sloppy with notation and just assume that people notice that the identity matrix is implicit.

[8] For emphasis, what we're doing right now is all classical. We will quantise this theory next lecture.

We need this to be Lorentz invariant (as that was one of the rules for constructing Lagrangians), so we need to check this.

We have

$$\psi(x) \to S[\Lambda]\psi\big(\Lambda^{-1}x\big) \qquad \Longrightarrow \qquad \psi^\dagger(x) \to \psi^\dagger\big(\Lambda^{-1}x\big)\big(S[\Lambda]\big)^\dagger,$$

and so the

$$\psi^\dagger(x)\psi(x) \to \psi^\dagger\big(\Lambda^{-1}x\big)\big(S[\Lambda]\big)^\dagger S[\Lambda]\psi\big(\Lambda^{-1}x\big),$$

but we just saw that $S[\Lambda]$ is not unitary, so this term is not invariant for all Lorentz transformations. Hmm... so what do we do? Well let's try and work out what went wrong and use that to help us find the correct answer. As we said in the previous subsection, we can pick a representation such that γ^0 is hermitian, $(\gamma^0)^\dagger = \gamma^0$ and the γ^is are antihermitian, $(\gamma^i)^\dagger = -\gamma^i$. From these, and $(\gamma^0)^2 = \mathbb{1}$ and the anticommutation relations, we conclude

$$\gamma^0\gamma^\mu\gamma^0 = \big(\gamma^\mu\big)^\dagger,$$

and so Equation (10.11) gives

$$\big(S^{\mu\nu}\big)^\dagger = -\gamma^0 S^{\mu\nu}\gamma^0.$$

Exercise

Use the above relation to show that

$$\big(S[\Lambda]\big)^\dagger = \gamma^0 S[\Lambda]^{-1}\gamma^0. \tag{10.14}$$

Hint: Use Equation (10.13), expanding it out and using the fact that a matrix commuted with itself.

With the result of the previous exercise in mind, we define the *Dirac adjoint*

$$\overline{\psi}(x) := \psi^\dagger(x)\gamma^0. \tag{10.15}$$

We define the Dirac adjoint, as it transforms as

$$\overline{\psi}(x) \to \psi^\dagger\big(\Lambda^{-1}x\big)\big(S[\Lambda]\big)^\dagger\gamma^0,$$

and so we have

$$\begin{aligned}
\overline{\psi}(x)\psi(x) &\to \psi^\dagger\big(\Lambda^{-1}x\big)\big(S[\Lambda]\big)^\dagger\gamma^0 S[\Lambda]\psi\big(\Lambda^{-1}\big) \\
&= \psi^\dagger\big(\Lambda^{-1}x\big)\gamma^0\gamma^0\big(S[\Lambda]\big)^\dagger\gamma^0 S[\Lambda]\psi\big(\Lambda^{-1}\big) \\
&= \psi^\dagger\big(\Lambda^{-1}x\big)\gamma^0\psi\big(\Lambda^{-1}\big) \\
&= \overline{\psi}\big(\Lambda^{-1}x\big)\psi\big(\Lambda^{-1}x\big),
\end{aligned}$$

where on the second line we have inserted $\mathbb{1} = \gamma^0\gamma^0$ and then used Equation (10.14) on the third line. So we see this bilinear is Lorentz invariant. This is exactly the kind of thing we need for our Lagrangian, but we also need to check the term with a γ^μ in it.

Exercise

Show that γ^μ satisfies

$$\big(S[\Lambda]\big)^{-1}\gamma^\mu S[\Lambda] = \Lambda^\mu{}_\nu \gamma^\mu. \tag{10.16}$$

Remark: This was set as an exercise in the course notes, so I don't want to type the answer here. However if you get really stuck, Prof. Tong's book might help...

Corollary 10.2.1. *Equation (10.16) implies that $\overline{\psi}(x)\gamma^\mu\psi(x)$ transforms as a vector under Lorentz transformations.*

Proof. Just consider the transformation (suppressing x arguments for notational brevity):

$$\begin{aligned}\overline{\psi}\gamma^\mu\psi &\to \overline{\psi}\big(S[\Lambda]\big)^{-1}\gamma^\mu S[\Lambda]\psi \\ &= \Lambda^\mu{}_\nu \overline{\psi}\gamma^\nu\psi,\end{aligned}$$

where we have used the fact that $\Lambda^\mu{}_\nu$ for given values of μ, ν is just a number so can be pushed through $\overline{\psi}$. This is exactly the transformation property of a vector under Lorentz transformations. □

This Corollary tells us that we can treat the index on the gamma matrix in the bilinear term above as if it was a Lorentz spacetime index. This therefore tells us that the term $\overline{\psi}\gamma^\mu\partial_\mu\psi$ in the Lagrangian will be Lorentz invariant. To summarise, the Dirac Lagrangian is

$$\mathcal{L} = \overline{\psi}(x)\big(i\gamma^\mu\partial_\mu - m\big)\psi(x), \tag{10.17}$$

where the factor of i is included for a reason that will be explained shortly. Varying this equation w.r.t. $\overline{\psi}(x)$ gives us the *Dirac equation*

$$\big(i\gamma^\mu\partial_\mu - m\big)\psi(x) = 0. \tag{10.18}$$

But what about if we vary Equation (10.17) w.r.t. $\psi(x)$? Well this gives

$$i\partial_\mu\overline{\psi}(x)\gamma^\mu + m\overline{\psi}(x) = 0.$$

Exercise

Take the hermitian conjugate of this expression and show it gives the Dirac equation again.

Remark 10.2.2. Ok that mysterious i, why is it there? Well the action for the Dirac field is simply

$$S = \int d^4x\, \overline{\psi}(x)\big(i\gamma^\mu\partial_\mu - m\big)\psi(x).$$

We require the action to be real, and so we need the integral to be the same under complex conjugation. Well the result of the integral is just a number, so taking the transpose does

nothing (it's a 1×1 matrix), so we need the integrand to be hermitian. This is where the i comes in. When we take the hermitian conjugate, we will get the derivative acting on $\overline{\psi}(x)$, so we need to integrate by parts to get it to act on the $\psi(x)$. This will come with a minus sign, and so the complex conjugation of the i cancels this minus sign.

Notation. There is a common, and very handy, notation when considering the contraction of gamma matrices, it is a slash. For example we write

$$\not{\partial} := \gamma^\mu \partial_\mu.$$

We don't just do this for derivatives, but for general gamma matrix contraction, for example we use

$$\not{p} := \gamma^\mu p_\mu$$

for the contracted momentum. In this notation the Dirac equation reads

$$(i\not{\partial} - m)\psi(x) = 0.$$

10.2.1 Obtaining The Klein-Gordan Equation

So we have seen that the Dirac equation is linear in derivatives, while the Klein-Gordan equation is quadratic in derivatives. A reasonable question to ask is "does the Dirac equation imply the Klein-Gordan equation?" By which we mean can be obtain the Klein-Gordan equation from Equation (10.18)? Well consider

$$0 = (i\gamma^\mu \partial_\mu + m)(i\gamma^\mu \partial_\mu - m)\psi(x) = -\big(\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu + m^2\big)\psi(x)$$

Now we can split the product of two gamma matrices into the symmetric and antisymmetric parts,

$$\gamma^\mu \gamma^\nu = \frac{1}{2}\gamma^{(\mu}\gamma^{\nu)} + \frac{1}{2}\gamma^{[\mu}\gamma^{\nu]} = \frac{1}{2}\{\gamma^\mu, \gamma^\nu\} + \frac{1}{2}[\gamma^\mu, \gamma^\nu],$$

where the second line follows simply from the definition of the commutator and anticommutator. Now we know that $\partial_\mu \partial_\nu$ is symmetric in its indices and its true that

$$A^{[\mu\nu]} B_{(\mu\nu)} = 0,$$

and so we are just left with the anticommutator term. Finally use Equation (10.8) to get

$$\left(\frac{1}{2} 2\eta^{\mu\nu} \partial_\mu \partial_\nu + m^2\right)\psi(x) = \big(\partial^2 + m^2\big)\psi(x) = 0,$$

where the last line follows from the fact that the thing we started with being 0. This is exactly the Klein-Gordan equation.

10.2.2 γ^5 & Dirac Bilinears

We have seen that the Dirac bilinear $\overline{\psi}\gamma^\mu\psi$ is transforms as a scalar under Lorentz transformations. A question we could ask is "what if we put other matrices in between $\overline{\psi}$ and ψ?" Well these are 4×4 matrices, so if we can find 16 linearly independent matrices are can study

how a general matrix sandwiched between $\overline{\psi}/\psi$ transforms. We have 5 linearly independent matrices given by $\{\mathbb{1}, \gamma^\mu\}$, and the question is "can we get the other 11 using only the gammas?" The answer is yes. We have already found 6 others in $S^{\rho\sigma}$, so so far we have 11 out of 16. What else could we do? Well we could multiply all 4 of the gamma matrices together:[9]

$$\gamma^5 := i\gamma^0\gamma^1\gamma^2\gamma^3. \tag{10.19}$$

It follows from this definition that

$$\left(\gamma^5\right)^\dagger = \gamma^5, \qquad \left(\gamma^5\right)^2 = \mathbb{1}, \quad \text{and} \quad \{\gamma^5, \gamma^\mu\} = 0. \tag{10.20}$$

This is actually a very important combination, and we shall discuss it further in a moment. In the Dirac basis this takes the form

$$\gamma^5 = \begin{pmatrix} 0 & \mathbb{1}_{2\times 2} \\ \mathbb{1}_{2\times 2} & 0 \end{pmatrix}. \tag{10.21}$$

Remark 10.2.3. In the Weyl/chiral basis we get

$$\gamma^5 = \begin{pmatrix} \mathbb{1}_{2\times 2} & 0 \\ 0 & -\mathbb{1}_{2\times 2} \end{pmatrix}.$$

Comparing these two expressions with our definition of γ^0 (and recalling Remark 10.1.4) we see that the essential different between the Dirac and Weyl basis is $\gamma^0 \longleftrightarrow \gamma^5$.

Equation (10.21) shows us that γ^5 is linearly independent from the other matrices in the Dirac basis (and indeed this is true in an arbitrary representation), so we have $5+6+1 = 12$ of the 16. What about the remaining 4? Well so far we have considered products of 2 gammas ($S^{\rho\sigma}$) and 2 gammas (γ^5), what about a product of three?

Exercise

Show that

$$\gamma^\mu\gamma^\nu\gamma^\rho \sim \gamma^\sigma\gamma^5$$

for $\mu \neq \nu \neq \rho \neq \sigma$. *Hint: Multiply the left-hand side by* $(\gamma^\rho)^2 \sim \mathbb{1}$.

The result on the right-hand side of the above exercise has a single spacetime index and so corresponds to the final 4 linearly independent matrices.[10] So we can write any 4×4 matrix as a linear combination of the set

$$\Gamma = \{\mathbb{1}, \gamma^\mu, \gamma^5, \gamma^5\gamma^\mu, S^{\rho\sigma}\}.$$

So we can write a general Dirac bilinear in the form $\overline{\psi}\Gamma\psi$. Extending the argument made after Corollary 10.2.1, about treating the gamma index as a spacetime index, we can conclude that these things transform as follows

[9] Note we don't call it γ^4, despite it being the fourth gamma. This is just because in Euclidean space we would have $\mu = 1, 2, 3, 4$ and so 5 would make sense there.

[10] We haven't actually showed this is linearly independent, but you can check it in the Dirac and Weyl bases at least.

$\overline{\psi}\psi$	Scalar
$\overline{\psi}\gamma^\mu\psi$	Vector
$\overline{\psi}S^{\mu\nu}\psi$	Tensor
$\overline{\psi}\gamma^5\psi$	...Scalar
$\overline{\psi}\gamma^5\gamma^\mu\psi$	... Vector

As you probably noticed, the last two terms have ellipses before them, the reason why is set as an exercise.[11]

Exercise

Parity.
Parity is the transformation $x^\mu \to x'^\mu = (x^0, -\vec{x})$, which can be thought of as mapping the world onto its mirror image.

(a) Show that is $\psi(x)$ is a solution to the Dirac equation, then so is $\psi'(x') := \gamma^0\psi(x)$ in the parity transformed world. In other words, start with $(i\gamma^\mu\partial_\mu - m)\psi(x) = 0$, and manipulate it into the form

$$(i\gamma^\mu\partial'_\mu - m)\psi'(x') = 0,$$

where $\partial'_\mu = \partial/\partial x'^\mu$.

(b) Compute the transformation laws for the following bilinears under parity, i.e. calculate

$$\begin{aligned}
\overline{\psi}(x)\psi(x) &\to \overline{\psi}'(x')\psi'(x') = ?\\
\overline{\psi}(x)\gamma^5\psi(x) &\to \overline{\psi}'(x')\gamma^5\psi'(x') = ?\\
\overline{\psi}(x)\gamma^\mu\psi(x) &\to \overline{\psi}'(x')\gamma^\mu\psi'(x') = ?\\
\overline{\psi}(x)\gamma^\mu\gamma^5\psi(x) &\to \overline{\psi}'(x')\gamma^\mu\gamma^5\psi'(x') = ?
\end{aligned}$$

You should find that the bilinears transform as a scalar, pseudoscalar, vector and axial vector, respectively.

Remark 10.2.4. The last sentence in the exercise above is exactly the reason we put the ellipses in the table above for the Lorentz transformation of Dirac bilinears.

10.2.3 Chiral Spinors & Projections

So far we haven't actually said anything about the form of $\psi(x)$ apart from that its a 4-component column vector. We can look further into its construction by using a particular representation for our Clifford algebra. Let's consider the Weyl/chiral basis, i.e.

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \quad \text{and} \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}.$$

[11] This was an exercise from the course, so I don't want to type the results here. As always feel free to ask me any questions for further clarity.

In this basis the representations take the form

$$S[\Lambda_R] = \begin{pmatrix} e^{i\vec{\varphi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{i\vec{\varphi}\cdot\vec{\sigma}/2} \end{pmatrix}, \quad \text{and} \quad S[\Lambda_B] = \begin{pmatrix} e^{\vec{\chi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{-\vec{\chi}\cdot\vec{\sigma}/2} \end{pmatrix} \tag{10.22}$$

where the R/B stand for rotations and boosts respctively, and where the notation $\vec{\varphi}\cdot\vec{\sigma} = \varphi^1\sigma^1 + \varphi^2\sigma^2 + \varphi^3\sigma^3$ and similarly for $\vec{\chi}\cdot\vec{\sigma}$. The numbers φ^i/χ^i are the parameters that tell us how far to rotate/how fast to boost respectively.

So we see that our representations are in block diagonal form and so our representation is *reducible*,[12] and so we can write our $\psi(x)$s as

$$\psi(x) = \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix}. \tag{10.23}$$

where each $\psi_{R/L}$ is a 2-component column matrix, we call these *Weyl/Chiral spinors*. Note, from Equation (10.22) we see that ψ_R and ψ_L transform the same under rotations but *oppositely* under boosts!

We can actually use γ^5 to project our the ψ_R/ψ_L part of the field. We define

$$P_L := \frac{1}{2}(\mathbb{1} - \gamma^5), \quad \text{and} \quad P_R := \frac{1}{2}(\mathbb{1} + \gamma^5), \tag{10.24}$$

so that

$$\psi_L = P_L\psi, \quad \text{and} \quad \psi_R = P_R\psi.$$

We can see this is true for the Weyl basis explicitly, but we can actually *define* $\psi_{L/R}$ in a general representation by the above formulas. We call these the *left-handed* and *right-handed* spinors (hence the subscripts). We will see why this is the case in the following exercise.

Remark 10.2.5. Again the understanding of the remaining part of this material was set as an exercise on the course, so I have just inserted the question here.

[12] For more info see, for example, my notes on Dr. Dorigoni's Group theory course.

Exercise

(a) Use Equation (10.24) to show that P_L and P_R are projection operators, e.g. $P_L^2 = P_L$, $P_R^2 = P_R$ and $P_L P_R = 0$.

(b) Rewrite the Dirac Lagrangian Equation (10.17) in terms of ψ_L and ψ_R.

(c) Show how ψ_L and ψ_R look in the parity transformed world (see previous exercise). That is perform the transformation and then express everything in terms of $\psi'_L(x')$ and $\psi'_R(x)$. Show that the Lagrangian is invariant under parity transformation.

(d) Consider the Weyl spinor equation, Equation (10.23). Set $m = 0$ and write out the explicit form of the Dirac equation in this basis.

(e) Plug in plane wave solutions $\psi(x) = u(p)e^{-ip\cdot x}$ into the massless Dirac equation (note $p^0 = |\vec{p}|$ when $m = 0$) and show that u_L and u_R are eigenvectors of the helicity operator

$$h := \frac{1}{2}\begin{pmatrix} \hat{p}\cdot\vec{\sigma} & 0 \\ 0 & \hat{p}\cdot\vec{\sigma} \end{pmatrix}.$$

Here $\hat{p}$ is a unit vector in the direction of p, which you can always choose to be, say, in the z-direction. List the eigenvalues of the left and right-handed spinors. (After quantisation these correspond to the situation where the particle spin is either aligned or anti-aligned with the direction of motion.)

Remark 10.2.6. Part (e) of the exercise above is the motivation for the names left/right-handed. The result should say that the helicity of ψ_L and ψ_R are opposite. Helicity basically tells you the projection of the spin of a massless particle onto its momentum, we call the two options left- and right-handed. These names come from our hands: make a thumbs up but don't curl your fingers all the way in, now imagine your thumb points in the direction of momentum, then your fingers tell you about the spin direction. A right-handed spinor has spin-momentum projection like your right hand looks, and similarly for a left-handed spinor.

Left-Handed Right-Handed

Remark 10.2.7. As a final remark to this lecture, we started it by saying that we were going to find quantum fields corresponding to spin-1/2 particles. To emphasise again, it is not trivial to see at this stage that this is what we have done. We will return to this next lecture and compute the angular momentum and show we do indeed have spin-1/2 particles.

11 | Quantising The Dirac Field

We now want to proceed and try and quantise our Dirac Lagrangian and arrive at some QFT. In order to do this, we are going to draw comparisons to the complex scalar field and use the results from that to guide us.

11.1 Preparing For Quantisation

For the complex scalar field we could expand the classical theory as an integral over Fourier modes:

$$\phi(x) = \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{E_{\vec{p}}}}\left(b_{\vec{p}}\, e^{-ip\cdot x} + c^\dagger_{\vec{p}}\, e^{ip\cdot x}\right)\Big|_{p^0=E_{\vec{p}}}.$$

When we quantised this, promoting $\phi(x)$ to an operator, we saw it was built out of the classical plane wave solutions to the Klein-Gordan equation with the creation/annihilation operators being the Fourier coefficients. We now want to start from a similar anstaz for the Dirac field, but now we need to take into account the fact that the Dirac equation is a matrix equation (it has gamma matrices in it) and so we need to include some 4-component column matrices in there too. So we propose the anstaz

$$\psi(x) \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{E_{\vec{p}}}}\left(u(\vec{p},s)b_{\vec{p}}\, e^{-ip\cdot x} + v(\vec{p},s)c^\dagger_{\vec{p}}\, e^{ip\cdot x}\right)\Big|_{p^0=E_{\vec{p}}}. \tag{11.1}$$

where, as we will see, the s will label the different spin states, i.e. $s = 1, 2$. We then define $\overline{\psi}(x)$ from this, giving us

$$\overline{\psi}(x) \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{\sqrt{E_{\vec{p}}}}\left(\overline{u}(\vec{p},s)c_{\vec{p}}\, e^{-ip\cdot x} + \overline{v}(\vec{p},s)b^\dagger_{\vec{p}}\, e^{ip\cdot x}\right)\Big|_{p^0=E_{\vec{p}}}. \tag{11.2}$$

11.1.1 Positive/Negative Frequency Modes

Exercise

Use the Dirac equation, Equation (10.18), to show that

$$(\not{p} - m)u(\vec{p},s) = 0, \quad \text{and} \quad (\not{p} + m)v(\vec{p},s) = 0 \tag{11.3}$$

where we note the difference in sign of the m term.

We can use the result of this exercise to obtain the explicit form of $u(\vec{p},s)$ and $v(\vec{p},s)$ in a given representation. We will choose to use the Dirac basis. We also note that the Dirac equation is Lorentz invariant and so we can pick whichever frame we like, and so we choose the rest frame. We therefore get

$$\not{p} = \big(m\gamma^0, 0, 0, 0\big),$$

and so

$$0 = m\big(\gamma^0 - \mathbb{1}_{4\times4}\big)u(\vec{p},s) = \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{1}_{2\times2} \end{pmatrix}$$

so we get two independent solutions, i.e. the two s values:

$$u(\vec{p},1) = \sqrt{2m}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad \text{and} \quad u(\vec{p},2) = \sqrt{2m}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}.$$

where we have normalised the states with $\sqrt{2m}$. Similarly we get

$$v(\vec{p},1) = \sqrt{2m}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad \text{and} \quad u(\vec{p},2) = \sqrt{2m}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}.$$

Remark 11.1.1. A nice justification for the fact that we normalised the states with a $\sqrt{2m}$ above can be found in Prof. Tong's notes on page 100-102. However, it is important to note that he is using (and so does Peskin and Schroeder) the Weyl basis. This is related to the Dirac basis via the similarity transformation

$$S = \frac{1}{\sqrt{2}}\begin{pmatrix} \mathbb{1}_{2\times2} & -\mathbb{1}_{2\times2} \\ \mathbb{1}_{2\times2} & \mathbb{1}_{2\times2} \end{pmatrix},$$

and so the expressions he obtains (once you set $\vec{p}=0$) are

$$u^{\text{Tong}}(\vec{p},1) = Su^{\text{Us}}(\vec{p},1) = \sqrt{m}\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \quad \text{and} \quad u^{\text{Tong}}(\vec{p},2) = Su^{\text{Us}}(\vec{p},2) = \sqrt{m}\begin{pmatrix}0\\1\\0\\1\end{pmatrix}.$$

Claim 11.1.2. The normalisation given for our u/v results in the normalisation

$$\overline{u}(\vec{p},s)u(\vec{p},s) = 2m, \quad \text{and} \quad \overline{v}(\vec{p},s)v(\vec{p},s) = -2m$$

Proof. We do this by noting that $u(\vec{p},s)$ transforms the same as $\psi(x)$ and so we know that the products $\overline{u}(\vec{p},s)u(\vec{p},s)$ and $\overline{u}(\vec{p},s)\gamma^\mu p_\mu u(\vec{p},s)$ are Lorentz invariant so for the following expression

$$\overline{u}(\vec{p},s)\not{p}u(\vec{p},s) - m\overline{u}(\vec{p},s) = 0$$

we can use the rest frame again, which tells us

$$\begin{aligned} m\overline{u}(\vec{p},s)u(\vec{p},s) &= mu^\dagger(\vec{p},s)\left(\gamma^0\right)^2 u(\vec{p},s) \\ &= mu^\dagger(\vec{p},s)u(\vec{p},s) \\ &= 2m^2 \\ \implies \overline{u}(\vec{p},s)u(\vec{p},s) &= 2m, \end{aligned}$$

where we have used $(\gamma^0)^2 = \mathbb{1}$ and $u^\dagger(\vec{p},s)u(\vec{p},s) = 2m$. The result for $\overline{v}(\vec{p},s)u(\vec{p},s)$ follows similarly. We can also easily show that $\overline{v}(\vec{p},s)u(\vec{p},s) = 0$. □

The result of a above claim generalises to the following relation

$$\overline{u}(\vec{p},s)u(\vec{p},r) = 2m\delta^{rs}, \quad \text{and} \quad \overline{v}(\vec{p},s)v(\vec{p},r) = -2m\delta^{rs}. \tag{11.4}$$

It is important to note that this result is Lorentz invariant and basis independent and so will hold in general.

11.1.2 Spin Sums

Recall that when we were talking about cross sections we said we needed to consider all the different possible final states and sum over them. Here this will correspond to summing over the different spin values, i.e. over the s, r values, in the final states. We will also need to average over the initial spins. So we need to know what the result of the summing is, the answer is its a 4×4 matrix given by the following claim.

Claim 11.1.3. The spin sums satisfy

$$\sum_{s=1}^{2} u(\vec{p},s)\overline{u}(\vec{p},s) = \not{p} + m, \quad \text{and} \quad \sum_{s=1}^{2} v(\vec{p},s)\overline{v}(\vec{p},s) = \not{p} - m, \tag{11.5}$$

where the two spinors are not contracted but placed back to back to give a 4×4 matrix.

Proof. The easiest way to prove this, we would be to go back and look at the solution of Equation (11.3) in the Weyl basis and obtain an expression for $u(\vec{p},s)$ and $v(\vec{p},s)$ and then use these results to show Equation (11.5). These are done in Prof. Tong's notes on pages 100-101 ad 104-105, respectively, and I don't see the point in just copying this out when I can just point the reader there. □

11.2 The Quantisation

11.2.1 Fermions

The study of the atomic spectra of fermions showed that fermionic wavefunctions are antisymmetric under exchange of quantum numbers. It was shown by Jordan and Wigner that this is equivalent to creation/annihilation operators obeying *anti*commutator relations, in contrast to the commutator relations in the case of bosons. Let's say this matematically.

Our states are defined in the same way to the bosonic case, i.e.

$$|\alpha\rangle := b_\alpha^\dagger |0\rangle ,$$

where α is some general quantum number, with multiparticle states then given by

$$|\alpha, \beta\rangle = b_\alpha^\dagger b_\beta^\dagger |0\rangle .$$

In the bosonic case it didn't matter what way we ordered the creation operators as they commuted, however, as we have just said, Fermi statistics tell us that the wavefunction changes sign if we switch the quantum numbers $\alpha \leftrightarrow \beta$, i.e. we require

$$|\alpha, \beta\rangle = - |\beta, \alpha\rangle .$$

This tells us that we require

$$\{b_\alpha^\dagger, b_\beta^\dagger\} := b_\alpha^\dagger b_\beta^\dagger + b_\beta^\dagger b_\alpha^\dagger = 0.$$

Taking the complex conjugate of this tells us that we also require

$$\{b_\alpha, b_\beta\} = 0.$$

Now note that the antisymmetric nature tells us that we cannot create two particles with the same quantum numbers, and the same for annihiliations i.e.

$$b_\alpha^\dagger b_\alpha^\dagger = 0, \quad \text{and} \quad b_\alpha b_\alpha = 0.$$

This essentially tells us we can label the elements of our Hilbert space by stating whether the slot is filled or not, e.g. $|0, 0\rangle$ and $|1, 0\rangle$ etc. From this we can write the action of the creation/annihilation operators as

$$\begin{aligned} b_\alpha^\dagger |..., 0, ...\rangle &= |..., 1, ...\rangle \quad \text{and} \quad b_\alpha^\dagger |..., 1, ...\rangle = 0 \\ b_\alpha |..., 1, ...\rangle &= |..., 0, ...\rangle \quad \text{and} \quad b_\alpha |..., 0, ...\rangle = 0, \end{aligned}$$

where the shown slot is the α-th slot.

We can use this to show that for Fermionic creation/annihilation operators we also require

$$\{b_\alpha, b_\beta^\dagger\} = \delta_{\alpha\beta}.$$

The fact that this is true is meant to be motivated by the next exercise.

Exercise

Show that for a system that only contains two particles, i.e. the only states available are $|0, 0\rangle , |0, 1\rangle , |1, 0\rangle$ and $|1, 1\rangle$, that $\{b_\alpha, b_\beta^\dagger\} = \delta_{\alpha\beta}$ for $\alpha, \beta = 1, 2$. *Hint: You will want to use the antisymmetric property of the states to make statements like* $|0, 1\rangle + |1, 0\rangle = 0$.

11.2.2 Quantising The Fields

So we now promote the classical expressions Equations (11.1) and (11.2) into operator equations, and we interpret $b^\dagger_{\vec{p},s}$ as an operator which creates the particles associated to the spinor $u(\vec{p},s)$ and similarly for $c^\dagger_{\vec{p},s}$ with $v(\vec{p},s)$. With the scalar field in mind, we can anticipate that these will correspond to particle and antiparticles, respectively. However it is important to note that we have no proof of this statement at this point.

For a reason we will explain shortly (see Remark 11.2.4), it turns out we need to consider anticommutation relations for our Dirac fields, not commutation relations like we did for the scalar fields. That is we want the *equal time* anticommutation relations

$$\begin{aligned} \{\psi_\alpha(\vec{x}),\psi_\beta(\vec{y})\} = 0 = \{\psi^\dagger_\alpha(\vec{x}),\psi^\dagger_\beta(\vec{y})\} \\ \{\psi_\alpha(\vec{x}),\psi^\dagger_\beta(\vec{y})\} = \delta_{\alpha\beta}\delta^{(3)}\big(\vec{x}-\vec{y}\big) \end{aligned} \tag{11.6}$$

Exercise

Show that Equation (11.6) is equivalent to

$$\begin{aligned} \{b_{\vec{p},r},b^\dagger_{\vec{q},s}\} = (2\pi)^3\delta_{rs}\delta^{(3)}\big(\vec{p}-\vec{q}\big) \\ \{c_{\vec{p},r},c^\dagger_{\vec{q},s}\} = (2\pi)^3\delta_{rs}\delta^{(3)}\big(\vec{p}-\vec{q}\big) \end{aligned} \tag{11.7}$$

and all other anticommutators vanishing. *Hint: Use Equations (11.1) and (11.2) and then use $(\gamma^0)^2 = \mathbb{1}$ so that you can use the spin sums Equation (11.5). If you cant work it out, Prof. Tong does it for commutation relations in his notes on page 107 which should give you the outline of the proof.*

Remark 11.2.1. Note it is easy to see from the above exercise that

$$\big\{\overline{\psi}_\alpha(\vec{x}),\overline{\psi}_\beta(\vec{y})\big\} = 0$$

as we see in Equation (11.2) that we will only get anticommutators between c and $b^\dagger$.

Remark 11.2.2. Note it is not a fully convincing argument to say "the quantisation of Dirac fields are spin-1/2 particles and we know spin-1/2 particles are Fermions so we should use Fermi statistics." Firstly, we haven't yet shown that Dirac fields correspond to spin-1/2 particles. We actually could have done this already, but even still I don't believe its a convincing argument to say we therefore use Fermi statistics for the fields. As we will see there is a much more physically crucial[1] reason why we have to use anticommutation relations.

11.2.3 The Hamiltonian

As before, we are ultimately interested in the spectrum of the theorem, that is the eigenvalues of the Hamiltonian. In order to do that, of course we need to know what the Hamiltonian

[1] As a 'spoiler' its because otherwise the energy is unbounded from below.

actually is. Recall that the Hamiltonian density is given by

$$\mathcal{H} = \pi\dot{\psi} - \mathcal{L}, \qquad \text{with} \qquad \pi := \frac{\partial \mathcal{L}}{\partial \dot{\psi}}.$$

Using the Lagrangian Equation (10.17) we get

$$\pi = i\overline{\psi}\gamma^0 = i\psi^\dagger\gamma^0\gamma^0 = i\psi^\dagger, \tag{11.8}$$

from which we get that the Hamiltonian density is

$$\mathcal{H} = \overline{\psi}\big(-i\gamma^i\partial_i + m\big)\psi.$$

Integrating over space will give us the Hamiltonian,

$$H = \int d^3\vec{x}\,\overline{\psi}\big(-i\gamma^i\partial_i + m\big)\psi.$$

This is the classical result, we get the quantum version by promoting everything to operators and inserting our operator expressions Equations (11.1) and (11.2). Doing this we obtain[2]

$$\begin{aligned} H &= \int \frac{d^3\vec{p}}{(2\pi)^3} \sum_{s=1}^{2} \Big(b^\dagger_{\vec{p},s} b_{\vec{p},s} - c_{\vec{p},s} c^\dagger_{\vec{p},s}\Big) E_{\vec{p}} \\ &= \int \frac{d^3\vec{p}}{(2\pi)^3} \sum_{s=1}^{2} \Big(b^\dagger_{\vec{p},s} b_{\vec{p},s} + c^\dagger_{\vec{p},s} c_{\vec{p},s} - (2\pi)^3\delta^{(3)}(0)\Big) E_{\vec{p}} \\ \therefore\ H &= \int \frac{d^3\vec{p}}{(2\pi)^3} \sum_{s=1}^{2} \Big(b^\dagger_{\vec{p},s} b_{\vec{p},s} + c^\dagger_{\vec{p},s} c_{\vec{p},s}\Big) E_{\vec{p}} \end{aligned} \tag{11.9}$$

where we have used the anticommutation relation Equation (11.7), and where the therefore symbol contains the information about us dropping the infinity term by a redefinition of the ground state energy, as we did for the scalar case.

So we see that the theory contains two types of particles, one created by $b^\dagger$ and the other by $c^\dagger$. These particles have the same mass (as they give the same energy contribution, $E_{\vec{p}}$, and each carry one of two spins. This is exactly what we expect of spin-1/2 Fermions, and we shall indeed provide further support of this claim. We use this interpration to write the Hamiltonian in terms of the number operators

$$H = \int \frac{d^3\vec{p}}{(2\pi)^3} \sum_{s=1}^{2} \big(N_{b,s} + N_{c,s}\big) E_{\vec{p}}.$$

Remark 11.2.3. Notice that in Equation (11.9) the ground state energy has the opposite sign to the scalar case. That is for the scalar field we have the ground state energy being $+\infty$ while here we have it being $-\infty$.

[2] To show this we need the same relations we need for the proof of the spin sums, Equation (11.5). Seeing as these aren't presented here, we do not prove this formula. Again it is done in Prof. Tong's notes, page 108. Note to self: Maybe after the courses are finished and you have more time come back and include all this stuff.

Remark 11.2.4. We can now wee why we needed to consider anticommutation relations and not commutation relations. If we had used the commutation relation

$$[c_{\vec{p},s}, c^\dagger_{\vec{q},r}] = \delta_{sr}\delta^{(3)}(\vec{p}-\vec{q}),$$

then when we switched the $c_{\vec{p},s}c^\dagger_{\vec{p},s}$ term in the first line of Equation (11.9) we would obtain

$$H = \int \frac{d^3\vec{p}}{(2\pi)^3}\sum_{s=1}^{2}\left(b^\dagger_{\vec{p},s}b_{\vec{p},s} - c^\dagger_{\vec{p},s}c_{\vec{p},s} + (2\pi)^3\delta^{(3)}(0)\right)E_{\vec{p}}.$$

Again we can drop the delta function however the negative sign infront of the $cc^\dagger$ term is a dagger[3] to the heart of our theory; it tells us that we our spectrum is unbounded from below! That is we could continually lower the energy of our system by repeated creation of $c^\dagger$ particles. This is obviously not good, and it is for this reason that we had to use anticommutation relations for our fields.

11.2.4 Conserved Currents/Charges

Let's look at the Noether currents for this theory. The first obvious one to consider is the stress tensor, which we recall is given by

$$T^\mu{}_\nu = \frac{\partial\mathcal{L}}{\partial\partial_\mu\psi}\partial_\nu\psi - \delta^\mu{}_\nu\mathcal{L}.$$

Using the Dirac Lagrangian, Equation (10.17), we get

$$\begin{aligned} T^{\mu\nu} &= i\overline{\psi}\gamma^\mu\partial^\nu\psi - \eta^{\mu\nu}\mathcal{L} \\ &= i\overline{\psi}\gamma^\mu\partial^\nu\psi - \eta^{\mu\nu}\overline{\psi}(i\not{\partial} - m)\psi. \end{aligned}$$

This doesn't look particularly nice because of the second term. However we can make this much nicer by noting that the a current is conserved only when the equations of motion are satisfied, so we can freely impose Equation (10.18) to set the second term to zero. We therefore have

$$T^{\mu\nu} = i\overline{\psi}\gamma^\mu\partial^\nu\psi. \tag{11.10}$$

From this we can find the physical momentum via

$$\begin{aligned} P^i &:= \int d^3\vec{x}\, T^{0i} \\ &= \int \frac{d^3\vec{p}}{(2\pi)^3}\sum_{s=1}^{2}\vec{p}\left(b^\dagger_{\vec{p},s}b_{\vec{p},s} + c^\dagger_{\vec{p},s}c_{\vec{p},s}\right) \\ &= \int \frac{d^3\vec{p}}{(2\pi)^3}\sum_{s=1}^{2}\vec{p}\left(N_{b,s} + N_{c,s}\right), \end{aligned}$$

which just tells us that the total momentum (not including spin!) in the system is given by the sum of momentum from the $b^\dagger$ particles and the $c^\dagger$ particles.

[3] Awful pun, I know...

Remark 11.2.5. Note that we could have also used the equations of motion when considering the scalar fields earlier in the course, however they weren't really any help because the stress-tensor was linear in derivatives whereas the equations of motion were quadratic. This is one of the beauties of the Dirac equation being linear in derivatives.

Now let's look for the conserved charge corresponding to the internal $U(1)$ symmetry:

$$\psi \to e^{-i\alpha}\psi$$

for some fixed $\alpha \in \mathbb{R}$. For the scalar field case this gave rise to conservation of particle number for the real scalar field and conservation of the difference of particle numbers for the complex scalar field. We can show this gives rise to the current

$$j^\mu = \overline{\psi}\gamma^\mu\psi,$$

and so Equation (2.6) tells us our conserved charge is

$$\begin{aligned} Q &= \int d^3\vec{x}\,\overline{\psi}\gamma^0\psi \\ &= \int \frac{d^3\vec{p}}{(2\pi)^3}\sum_{s=1}^{2}\left(b^\dagger_{\vec{p},s}b_{\vec{p},s} - c^\dagger_{\vec{p},s}c_{\vec{p},s}\right) \\ &= \int \frac{d^3\vec{p}}{(2\pi)^3}\sum_{s=1}^{2}\left(N_{b,s} - N_{c,s}\right). \end{aligned} \tag{11.11}$$

It is at *this point* that we get our particle antiparticle interpretation. This is because of the relative minus sign above which tells us that the $c^\dagger$ particle carries the opposite charge to the $b^\dagger$ particle.

Exercise

Show that the conserved current $j^\mu = \overline{\psi}\gamma^\mu\psi$ is indeed conserved. That is show

$$\nabla_\mu j^\mu = 0.$$

11.2.5 Spectrum

We can now look at the spectrum of the system. As always we define the vacuum to be the state that is annihilated by all annihilation operators,

$$b_{\vec{p},s}\,|0\rangle = 0 = c_{\vec{p},s}\,|0\rangle$$

for all $\vec{p}$ and s. It follows immediately from the previous expressions that the vacuum obeys

$$H\,|0\rangle = \vec{P}\,|0\rangle = Q\,|0\rangle = 0,$$

which tells us the vacuum has zero energy, zero momentum and no particles.[4]

[4]Note the condition $Q\,|0\rangle = 0$ only tells us that there is an equal number of particles and antiparticles, but because the momentum and energy are zero, it follows that we cannot have any particles.

We have two types of one particle state, we label these by

$$|\vec{p},s\rangle_\psi := \sqrt{2E_{\vec{p}}}\, b^\dagger_{\vec{p},s}\,|0\rangle\,, \quad \text{and} \quad |\vec{p},s\rangle_{\overline{\psi}} := \sqrt{2E_{\vec{p}}}\, c^\dagger_{\vec{p},s}\,|0\rangle \tag{11.12}$$

These have energy, momentum and charge given by

$$\begin{aligned} H\,|\vec{p},s\rangle_\psi &= E_{\vec{p}}\,|\vec{p},s\rangle_\psi \quad \text{and} \quad H\,|\vec{p},s\rangle_{\overline{\psi}} = E_{\vec{p}}\,|\vec{p},s\rangle_{\overline{\psi}} \\ \vec{P}\,|\vec{p},s\rangle_\psi &= \vec{p}\,|\vec{p},s\rangle_\psi \quad \text{and} \quad \vec{P}\,|\vec{p},s\rangle_{\overline{\psi}} = \vec{p}\,|\vec{p},s\rangle_{\overline{\psi}} \\ Q\,|\vec{p},s\rangle_\psi &= +\,|\vec{p},s\rangle_\psi \quad \text{and} \quad Q\,|\vec{p},s\rangle_{\overline{\psi}} = -\,|\vec{p},s\rangle_{\overline{\psi}}\,, \end{aligned}$$

which again supports our statement that $b^\dagger$ makes particles while $c^\dagger$ makes antiparticles.

Exercise

Find the energy, momentum and charge of multiparticle states. *Hint: Be careful about the states you allow, as we have Fermions so* $|\vec{p},s;\vec{p},s\rangle_\psi = 0$ *by antisymmetry, and similarly for the antiparticle states.*

11.2.6 Spin

So what about the spin? Well, recalling what we did in lectures 2 and 3, we get this by considering an infinitesimal Lorentz transformation on our system. We have already seen, Equation (10.12), that the Dirac spinor transforms under the Lorentz group as

$$\psi^\alpha(x) \to S[\Lambda]^\alpha{}_\beta \psi^\beta\big(\Lambda^{-1}x\big),$$

so we just need to consider the infinitesimal version of this. We have already seen

$$\Lambda^\mu{}_\nu \approx \delta^\mu{}_\nu + \omega^\mu{}_\nu$$

and from Equation (10.13) we have

$$S[\Lambda]^\alpha{}_\beta \approx \delta^\alpha{}_\beta + \frac{1}{2}\Omega_{\rho\sigma}\big(S^{\rho\sigma}\big)^\alpha{}_\beta,$$

which gives us

$$\delta\psi^\alpha(x) = -\omega^\mu{}_\nu x^\nu \partial_\mu \psi^\alpha(x) + \frac{1}{2}\Omega_{\rho\sigma}\big(S^{\rho\sigma}\big)^\alpha{}_\beta \psi^\beta(x),$$

where the minus sign on the first term comes from the fact that we have Λ^{-1}.[5] Now we use Equation (10.5), which says

$$\omega^{\mu\nu} = \Omega^{\mu\nu},$$

to obtain

$$\delta\psi^\alpha(x) = -\omega^{\mu\nu}\Big[x_\nu \partial_\mu \psi^\alpha(x) - \frac{1}{2}\big(S^{\mu\nu}\big)^\alpha{}_\beta \psi^\beta(x)\Big].$$

If we compare this to the derivation of Equation (2.12) we see that the first term above is exactly the same, while the second term above is new so will give an additional contribution to the Lorentz current. In total we have

$$(\mathcal{J}^\mu)^{\rho\sigma} = x^\rho T^{\mu\sigma} - x^\sigma T^{\mu\rho} - i\overline{\psi}\gamma^\mu S^{\rho\sigma}\psi.$$

[5] See Section 2.1.1 if this doesn't make sense.

So we have the angular momentum operator

$$J^i := \epsilon^{ijk} \int d^3\vec{x}\, (\mathcal{J}^0)^{jk}, \tag{11.13}$$

which, using our expressions for the stress tensor, gives

$$\vec{J} = \int d^3\vec{x}\, \psi^\dagger \Big[\vec{x} \times (-i\vec{\nabla}) + \frac{1}{2}\vec{S}\Big]\psi,$$

where[6]

$$\vec{S} = \big(S^{23}, S^{31}, S^{12}\big), \qquad S^{ij} = -\frac{i}{2}\epsilon^{ijk} \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix},$$

where σ^k are the Pauli matrices. This expression can be used to show that our particles do indeed have spin-1/2. We do not do the calculation here but instead point the interested reader to pages 60-61 of Peskin and Schroeder.

Exercise

Show the Equation (11.13) holds. *Hint: Do it for one component J^i and notice how it extends to this formula.*

[6] I think the entries for $\vec{S}$ are correct. If you disagree please feel free to email me a correction.

12 | Dirac Propagators & Perturbation Theory

12.1 Propagators

12.1.1 Causality

So far we have quantised the Dirac equation and showed (or at least claimed it can be shown) that they correspond to spin-1/2 Fermions. The next thing we need to check is causality, that is we need the field operators to not communicate outside the lightcone. For the complex scalar field we saw that this boiled down to the condition

$$[\psi(x), \psi^\dagger(y)] = 0, \qquad \forall (x-y)^2 < 0.$$

We therefore expect to have something like

$$[\psi(x), \overline{\psi}(y)] = 0 \qquad \forall (x-y)^2 < 0.$$

However there is a problem here: we can expand $\psi/\overline{\psi}$ in terms of the creation and annihilation operators $b, b^\dagger, c$ and $c^\dagger$, however we do not know what they commutation relationships are. That is

$$[b_{\vec{p},s}, b^\dagger_{\vec{q},r}] = ?$$

Now we can relate these to the anticommutation relations using the identity

$$[A, B] = \{A, B\} - 2BA, \tag{12.1}$$

which is easily verified. However this raises a problem as, using Equations (11.1) and (11.2), we see that the commutator will contain terms that have creation operators to the right. For example

$$[c^\dagger_{\vec{p},s} c_{\vec{q},r}] = \{c^\dagger_{\vec{p},s} c_{\vec{q},r}\} - 2c_{\vec{q},r} c^\dagger_{\vec{q},s},$$

and while the first term will just give us a delta function, as per Equation (11.7), the second term will give a non-zero expectation value. Equally we will get terms arising containing both b and c operators, e.g.

$$[c^\dagger_{\vec{p},s}, b^\dagger_{\vec{q},r}] = -2b^\dagger_{\vec{q},r} c^\dagger_{\vec{p},s},$$

where we have used the fact that the anticommutator vanishes. It is not trivial to see that all these contirbutions will somehow cancel and give us a vanishing commutator for spacelike separations.

So what do we do? Well we note that every single observable quantity we have written for the Dirac field has contained bilinears, $\overline{\psi}\Gamma\psi$. Note this is obviously true because what we measure are numbers but $\psi/\overline{\psi}$ are matrices, so we need to contract them to get a number. We therefore propose that what we need for our causality condition is

$$\Big[\overline{\psi}(x)\psi(x), \overline{\psi}(y)\psi(y)\Big] = 0 \qquad \forall (x-y)^2 < 0. \tag{12.2}$$

Claim 12.1.1. We can satisfy Equation (12.2) if

$$iS_{\alpha\beta}(x-y) := \big\{\psi_\alpha(x), \overline{\psi}_\beta(y)\big\} = 0 \qquad \forall (x-y)^2 < 0. \tag{12.3}$$

Proof. We just need to show the two relations are related, and then of course they agree under the given conditions. That is we need to show

$$\Big[\overline{\psi}(x)\psi(x), \overline{\psi}(y)\psi(y)\Big] \sim \big\{\psi_\alpha(x), \overline{\psi}_\beta(y)\big\}.$$

We do this by repeated use of Equation (12.1) along with

$$[AB, C] = A[B, C] + [A, C]B, \quad \text{and} \quad [A, CD] = C[A, D] + [A, C]D.$$

We will also need to use the anticommutation relations Equation (11.6) along with

$$\overline{\psi}\psi = \overline{\psi}_\alpha\psi_\alpha,$$

where there is an implicit sum. To save notation we shall use α for the fields as a function of x and β for the y fields, we therefore have

$$\begin{aligned}
\Big[\overline{\psi}_\alpha\psi_\alpha, \overline{\psi}_\beta\psi_\beta\Big] &= \overline{\psi}_\alpha\Big[\psi_\alpha, \overline{\psi}_\beta\psi_\beta\Big] + \Big[\overline{\psi}_\alpha, \overline{\psi}_\beta\psi_\beta\Big]\psi_\alpha \\
&= \overline{\psi}_\alpha\Big[\psi_\alpha, \overline{\psi}_\beta\Big]\psi_\beta + \overline{\psi}_\alpha\overline{\psi}_\beta\Big[\psi_\alpha, \psi_\beta\Big] + \Big[\overline{\psi}_\alpha, \overline{\psi}_\beta\Big]\psi_\beta\psi_\alpha + \overline{\psi}_\beta\Big[\overline{\psi}_\alpha, \psi_\beta\Big]\psi_\alpha \\
&= \overline{\psi}_\alpha\big\{\psi_\alpha, \overline{\psi}_\beta\big\}\psi_\beta - 2\overline{\psi}_\alpha\overline{\psi}_\beta\psi_\alpha\psi_\beta + \overline{\psi}_\alpha\overline{\psi}_\beta\big\{\psi_\alpha, \psi_\beta\big\} - 2\overline{\psi}_\alpha\overline{\psi}_\beta\psi_\beta\psi_\alpha \\
&\quad + \big\{\overline{\psi}_\alpha, \overline{\psi}_\beta\big\}\psi_\beta\psi_\alpha - 2\overline{\psi}_\beta\overline{\psi}_\alpha\psi_\beta\psi_\alpha + \overline{\psi}_\beta\big\{\overline{\psi}_\alpha, \psi_\beta\big\}\psi_\alpha - 2\overline{\psi}_\beta\psi_\beta\overline{\psi}_\alpha\psi_\alpha \\
&= \overline{\psi}_\alpha\big\{\psi_\alpha, \overline{\psi}_\beta\big\}\psi_\beta + \overline{\psi}_\beta\big\{\overline{\psi}_\alpha, \psi_\beta\big\}\psi_\alpha - 2\overline{\psi}_\alpha\overline{\psi}_\beta\big\{\psi_\alpha, \psi_\beta\big\} - 2\overline{\psi}_\beta\overline{\psi}_\alpha\psi_\beta\psi_\alpha \\
&\quad - 2\overline{\psi}_\beta\psi_\beta\overline{\psi}_\alpha\psi_\alpha \\
&= \overline{\psi}_\alpha\big\{\psi_\alpha, \overline{\psi}_\beta\big\}\psi_\beta + \overline{\psi}_\beta\big\{\overline{\psi}_\alpha, \psi_\beta\big\}\psi_\alpha - 2\underline{\overline{\psi}_\beta\overline{\psi}_\alpha\psi_\beta\psi_\alpha} - 2\overline{\psi}_\beta\big\{\psi_\beta, \overline{\psi}_\alpha\big\}\psi_\alpha \\
&\quad + 2\underline{\overline{\psi}_\beta\overline{\psi}_\alpha\psi_\beta\psi_\alpha} \\
&= \overline{\psi}_\alpha\big\{\psi_\alpha, \overline{\psi}_\beta\big\}\psi_\beta + \overline{\psi}_\beta\big\{\psi_\beta, \overline{\psi}_\alpha\big\}\psi_\alpha - 2\overline{\psi}_\beta\big\{\psi_\beta, \overline{\psi}_\alpha\big\}\psi_\alpha \\
&= \overline{\psi}_\alpha\big\{\psi_\alpha, \overline{\psi}_\beta\big\}\psi_\beta - \overline{\psi}_\beta\big\{\psi_\beta, \overline{\psi}_\alpha\big\}\psi_\alpha,
\end{aligned}$$

where I have tried to colour/underline the terms that carry over from line to line. For clarity, we have used

$$\big\{\psi_\alpha, \psi_\beta\big\} = 0 = \big\{\overline{\psi}_\alpha, \overline{\psi}_\beta\big\}$$

to drop all terms of that form. □

Ok so we can impose our causality condition by requiring Equation (12.3) holds. This is left as an exercise.

Exercise

Show that

$$iS_{\alpha\beta}(x-y) = \left(i\not{\partial}_x + m\right)_{\alpha\beta}\big(D(x-y) - D(y-x)\big), \tag{12.4}$$

where $D(x,y)$ is the propagator from the scalar theory, i.e.

$$D(x-y) = \int \frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{2E_{\vec{p}}}e^{-ip\cdot(x-y)}.$$

Hint: You will want to use the spin sums, Equation (11.5). If you get stuck, Prof. Tong does this calculation on page 113.

Remark 12.1.2. Recall that the propagator $D(x-y)$ is in the kernel of the Klein-Gordan equation if we are on shell, $p^2 = m^2$, that is

$$(\partial_x^2 + m^2)D(x-y)\big|_{p^2=m^2} = 0.$$

From this we see that

$$(i\not{\partial}_x - m)S(x-y)\big|_{p^2=m^2} = (\partial_x^2 + m^2)\big(D(x-y) - D(y-x)\big)\big|_{p^2=m^2} = 0$$

Remark 12.1.3. For some of the expressions that follow, we shall drop the $\alpha\beta$ labels for the matrices, however it is important to remember that our results are matrices. This should generally be clear given that gamma matrices will appear.

12.1.2 Time & Normal Ordering For Fermions

So we have found the propagator for the theory, Equation (12.4), now we want to find the Feynman propagator for the theory. As with the scalar field, we define this as the expectation value of the time ordered product of the fields $\psi(x)\overline{\psi}(y)$. However in order to write this expression down, we need to make a small correction to our formula for the time ordered product. We will also need this for Dyson's formula. Similarly we need to tweak the normal ordering operator for Wick's theorem.

The thing we need to account for is the fact that Fermion fields anticommute, and so we pick up minus signs when moving them past each other within T. We can, however, move *pairs* of fields freely as they always come with a $(-1)^2 = +1$, and because of this Dyson's formula is unaffected, as the Hamiltonian always contains bilinears.

Example 12.1.4. As usual let's denote $\psi_i := \psi(x_i)$. Then, for $x_3^0 > x_1^0 > x_4^0 > x_2^0$, we have

$$\begin{aligned}\mathcal{T}\left[\psi_1\psi_2\psi_3\psi_4\right] &= \mathcal{T}\left[\psi_3\psi_4\psi_1\psi_2\right] \\ &= -\mathcal{T}\left[\psi_3\psi_1\psi_4\psi_2\right] \\ &= -\psi_3\psi_1\psi_4\psi_2\end{aligned}$$

We have to do a similar thing for normal ordering as the creation/annihilation operators now anticommute.

Example 12.1.5. For Fermions we have the following

$$: b_{\vec{p},s} b_{\vec{q},r} b^\dagger_{\vec{\ell},t} : \; = (-1)^2 : b^\dagger_{\vec{\ell},t} b_{\vec{p},s} b_{\vec{q},r} : \; = b^\dagger_{\vec{\ell},t} b_{\vec{p},s} b_{\vec{q},r}.$$

A similar type of thing holds for c operators.

12.1.3 Feynman Propagator For Fermions

We can now construct the Feynman propagator for Fermions:

$$S_F(x-y) := \langle 0| \mathcal{T}\psi(x)\overline{\psi}(y) |0\rangle = \begin{cases} \langle 0| \psi(x)\overline{\psi}(y) |0\rangle & x^0 > y^0 \\ \langle 0| -\overline{\psi}(y)\psi(x) |0\rangle & y^0 > x^0 \end{cases} \tag{12.5}$$

where we note the minus sign in the cases. This minus sign is necessary to preserve Lorentz invariance: recall that when $(x-y)^2 < 0$ there is no Lorentz invariant way to decided if $x^0 > y^0$ or $y^0 > x^0$, putting this together with Equation (12.3) we see the minus sign is needed. That is, Equation (12.3) tells us

$$\psi(x)\overline{\psi}(y) = -\overline{\psi}(y)\psi(x),$$

for spacelike separation, but in here we can transition from $x^0 > y^0$ to $y^0 > x^0$ using a Lorentz transformation, so we need a minus sign in the time ordering expression.

Exercise

Use the Fourier expansions Equations (11.1) and (11.2) to show that

$$\langle 0| \psi_\alpha(x)\overline{\psi}_\beta(y) |0\rangle = \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} e^{-ip\cdot(x-y)} \big(\not{p} + m\big)_{\alpha\beta}$$
$$-\langle 0| \overline{\psi}_\beta(y)\psi_\alpha(x) |0\rangle = -\int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} e^{ip\cdot(x-y)} \big(\not{p} - m\big)_{\alpha\beta}.$$

Hint: If you have done the previous exercise this one is basically done.

Claim 12.1.6. We can combine the above two relations to write the Feynman propagator as a 4-momentum integral

$$S_F(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} e^{-ip\cdot(x-y)}. \tag{12.6}$$

Proof. You do a contour integral. The calculation is similar to the one we did for the Feynman propagator for the scalar field, Equation (5.10). We 'prove' it here by noting the similarity to that expression. □

Remark 12.1.7. Note in extension of Remark 12.1.2, we see that the Feynman propagator for Ferimons is a Green's function of the Dirac equation, that is

$$\big(i\not{\partial}_x - m\big) S_F(x-y) = i\delta^{(4)}(x-y).$$

12.2 Perturbation Theory

We now want to look at the interacting theory and derive the Feynman rules. As before we do this by using Wick's theorem and considering the contractions. As we mentioned above, we need to make a small tweak to Wick contractions in order to account for the anticommutation relations.

12.2.1 Wicks Theorem For Fermions

The contractions actually adapt from the scalar case very straightforwardly.[1] We have

$$\overbrace{\psi(x)\psi(y)} = 0 = \overbrace{\overline{\psi}(x)\overline{\psi}(y)}, \quad \text{and} \quad \overbrace{\psi(x)\overline{\psi}(y)} = S_F(x-y). \tag{12.7}$$

Then if we are taking a contraction within a string of normal ordered objects we obtain the result by using the anticommutation relations to put the contracted fields next to each other and then normal ordering the rest.

Example 12.2.1. Let's say we want to contract ψ_1 with ψ_5 and ψ_2 with ψ_3 in the following normal ordering. Then we move the ψ_5 so that it's next to the ψ_1 and remove the two contractions. That is

$$\begin{aligned} :\overbrace{\psi_1\psi_2\overline{\psi}_3\psi_4\overline{\psi}_5}\psi_6: &= (-1)^3 :\overbrace{\psi_1\overline{\psi}_5}\overbrace{\psi_2\overline{\psi}_3}\psi_4\psi_6: \\ &= -S_F(x_1-x_5)S_F(x_2-x_3):\psi_4\psi_6: \end{aligned}$$

We then define the contraction of the operators on states. However we need to account for the fact that we have two different kinds of states, and so in total we have 4 kinds of contractions. These correspond to the two types of particle (i.e. particle and antiparticle) and whether they are ingoing or outgoing. We list them in the table below.

Particle/Antiparticle	Ingoing/Outgoing	Contraction Expression
Particle	Ingoing	$\overbrace{\psi(x)\,\|\vec{p},s\rangle}_{\psi} = e^{-ip\cdot x}u(\vec{p},s)$
Antiarticle	Ingoing	$\overbrace{\overline{\psi}(x)\,\|\vec{p},s\rangle}_{\overline{\psi}} = e^{-ip\cdot x}\overline{v}(\vec{p},s)$
Particle	Outgoing	$\overbrace{\langle\vec{p},s\,\|\,\overline{\psi}}_{\overline{\psi}} = e^{ip\cdot x}\overline{u}(\vec{p},s)$
Antiarticle	Outgoing	$\overbrace{\langle\vec{p},s\,\|\,\psi}_{\psi} = e^{ip\cdot x}v(\vec{p},s)$

There's four things to notice here, which we list in bullet points below.

[1] See Section 7.2 if you need a reminder on how we obtain this.

- We only contract ψ with the ψ states (i.e. have a subscript ψ). Similarly for contracting $\overline{\psi}$.
- Incoming particles come with a factor of $e^{-ip\cdot x}$ while outgoing particles come with $e^{+ip\cdot x}$.
- The particles are given by the letter u while the antiparticles are given by v.
- If the contraction is obtained using a $\overline{\psi}$ then the resulting symbol has a bar over it, e.g. an outgoing particle has a $\overline{u}$.

Remark 12.2.2. Note that if we had a multiparticle state, as per the method of putting the contracted objects next to each other, we need to switch the ordering of the ket so that we contract the field with the first entry. With the Fermion states discussion above, we see this gives us potential minus signs. For clarity, if we want to contract $\psi(x)$ with the second entry in a two particle state, we have

$$\overbrace{\psi(x)\,|\vec{p},s;\vec{q},r\rangle} = -\overbrace{\psi(x)\,|\vec{q},r;\vec{p},s\rangle} = -e^{-iq\cdot x}u(\vec{q},r)\,|\vec{p},s\rangle\,.$$

Note that if we then contracted this result with another field operator, say $\psi(y)$, we wouldn't get any more minus signs as the prefactor on the right-hand side expression doesn't have the anticommutation relation with $\psi(y)$. Recalling that the operators $\psi(x)$ and $\psi(y)$ anticommute, allows us to reformulate this procedure by saying "put the operator with the $(\vec{p},s)$ closest to the ket-vector", i.e.

$$\overbrace{\psi(y)\underbrace{\psi(x)\,|\vec{p},s;\vec{q},r\rangle}} = -\underbrace{\psi(x)\overbrace{\psi(y)\,|\vec{p},s;\vec{q},r\rangle}}$$

Wick's theorem is then exactly the same, namely

$$\mathcal{T}(\psi_1\overline{\psi}_2...) =: \psi_1\overline{\psi}_2...: \ + \text{all contractions}$$

Ok so let's do some examples to clear up any confusion and also derive some Feynman rules.

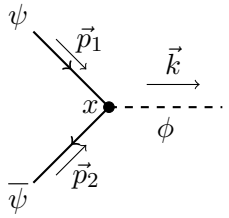
12.2.2 Yakawa Theory

Yakawa theory carries over[2] to the Dirac fields, and the Lagrangian is given by

$$\mathcal{L} = \overline{\psi}(i\partial\!\!\!/ - m)\psi + \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}\mu^2\phi^2 - g\overline{\psi}\psi\phi,$$

where μ^2 is the mass term for the real scalar field ϕ.

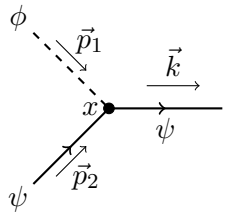
Before we even derive the Feynman rules we can sort of guess what the diagrams will look like. We see that we have three fields $\psi, \overline{\psi}$ and ϕ and the interaction term couples one of each together with coupling strength g. So we expect the diagram to have a vertex something like



[2] In fact this is the actual Yakawa theory. The scalar Yakawa theory was just studied because its useful for the study now.

This represents the process $\psi(p_1)\overline{\psi}(p_2) \to \phi(k)$, which is particle-antiparticle annihilation.

Now it follows from the fact that we can contract the $\overline{\psi}$ with an outgoing state and produce a particle, that we can rotate this Feynman diagram and obtain the following one:

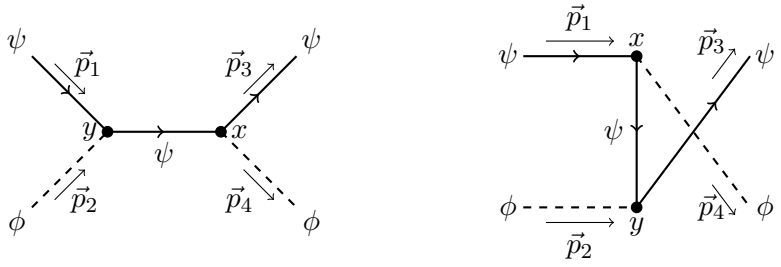


which represents the scattering process $\phi(p_1)\psi(p_2) \to \psi(k)$. Note there are no antiparticles present here. Indeed it we can always rotate a Feynman diagram and obtain new ones, provided the particles in the new one are valid. That is, both of the above diagrams are fine because we have particles and antiparticles in this system.[3] Note that we always have one Fermion arrow (i.e. the arrows on the lines) flowing into a vertex and one flowing out. This corresponds to our charge Equation (11.11) being conserved locally at the vertex.

Ok let's find some Feynman rules. We'll start by looking at the scattering process

$$\psi(p_1)\phi(p_2) \to \psi(p_3)\phi(p_4).$$

There are two distinct types of contractions we can do here. We shall use our guess work above to draw these diagrams now for pictorial clarity, however it is important to note that it could be entirely possible that what we draw doesn't correspond to the Wick contraction result at all.[4] The two diagrams are:



Let's just focus on the left-hand one. This corresponds to a second order term in Dyson's formula and is given by the following contraction (note we have included a α/β index on the $\psi/\overline{\psi}$s to show how the matrices are contracted. Also note that we have used $|\psi(p,s)\rangle$ instead of $|p,s\rangle_\psi$ to help keep track of what comes from where.)

$$\frac{(-ig)^2}{2!}\int d^4x d^4y\, \langle \underbrace{\phi(p_4)\overline{\psi(p_3,s)|}\,\overline{\psi}^\alpha_x \psi^\alpha_x \phi_x \overline{\psi}^\beta_y \psi^\beta_y \phi_y}\, |\psi(p_1,r)\phi(p_2)\rangle\,,$$

[3] I have never actually met a case where this doesn't work out, but I thought it was worth clarifying in case such cases do exist.

[4] Of course it will, but I just mean to highlight that just because you think it should look like this, doesn't mean it has to.

which using our contraction rules gives us

$$\frac{(-ig)^2}{2!}\int d^4x\, d^4y\, e^{i(p_3+p_4)\cdot x}\,\overline{u}(\vec{p}_3,s)_\alpha\, S_F(x-y)_{\alpha\beta}\, u(\vec{p}_1,r)_\beta\, e^{-i(p_1+p_2)\cdot y}.$$

We can then insert the integral expression for $S_F(x-y)$, Equation (12.6), as

$$S_F(x-y) = \int \frac{d^4q}{(2\pi)^4}\frac{i(\not{q}+m)}{q^2-m^2+i\epsilon}e^{-iq\cdot(x-y)},$$

and then do the integral over y and q to set $q = p_1+p_2$, and then the integral over x gives a delta function $\delta^{(4)}(p_1+p_2-p_3-p_4)$, which gives us the expression

$$\frac{(-ig)^2}{2!}\delta^{(4)}(p_1+p_2-p_3-p_4)\overline{u}(\vec{p}_3,s)_\alpha\frac{i\left[\not{p}_1+\not{p}_2+m\right]_{\alpha\beta}}{(p_1+p_2)^2-m^2+i\epsilon}u(\vec{p}_1,r)_\beta.$$

Finally we cancel the factor of 2 by the symmetry $x \longleftrightarrow y$.

Exercise

Show that the leading order expression for the process

$$\overline{\psi}(p_1)\phi(p_2) \to \overline{\psi}(p_3)\phi(p_4)$$

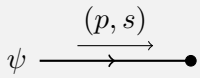
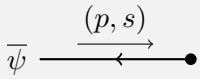
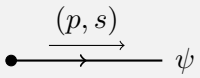
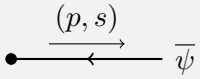
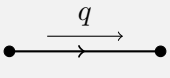
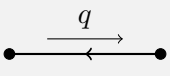
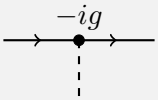
is given by the final result (before the symmetry factor)

$$\frac{(-ig)^2}{2!}\delta^{(4)}(p_1+p_2-p_3-p_4)\overline{v}(\vec{p}_1,r)_\alpha\frac{i\left[-(\not{p}_1+\not{p}_2)+m\right]_{\alpha\beta}}{(p_1+p_2)^2-m^2+i\epsilon}v(\vec{p}_3,s)_\beta.$$

Hint: Note the minus sign in the numerator, this should come from a delta function.

The results of the above exercises suggest the following *momentum space* Feynman rules.

All the usual stuff for the scalar field ϕ and then:

Type	Diagram	Maths Expression
Incoming Fermion		$u(\vec{p},s)$
Incoming Antiermion		$\overline{v}(\vec{p},s)$
Outgoing Fermion		$\overline{u}(\vec{p},s)$
Outgoing Antifermion		$v(\vec{p},s)$
Fermion Propagator		$\int \frac{d^4q}{(2\pi)^4} \frac{i(\not{q}+m)}{q^2-m^2+i\epsilon}$
Antifermion Propagator		$\int \frac{d^4q}{(2\pi)^4} \frac{i(-\not{q}+m)}{q^2-m^2+i\epsilon}$
Vertex		$-(ig)$

Note that we also label the spin s on the momentum arrow, and note again that the antifermion propagator comes with a $-\not{q}$ in the fraction.

Exercise

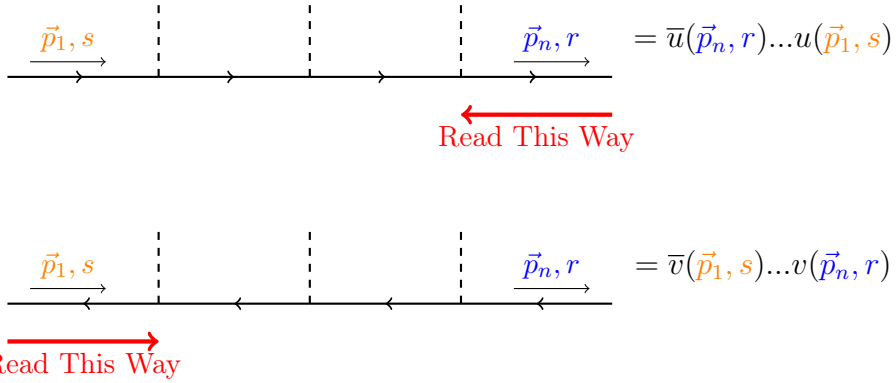
Draw the corresponding Feynman diagram to the previous exercise.

12.2.3 Following The Arrows

It takes some practice to get used to obtaining the mathematical expressions from the Feynman diagrams for Fermionic fields. This is because the $u/\overline{u}/v/\overline{v}$ are matrices and obviously we have to write them in certain order if we want to get a number as the answer. To clarify, the unbarred u and v are 4-component column matrices and the barred $\overline{u}$ and $\overline{v}$ are 4-component row matrices (as the bar contains a Hermitian conjugate). So if we want to get a number out at the end, then we have to put all the barred elements to the left, as we did in the two expressions above.

So how do you ensure that you always get this from the Feynman diagram? the answer is you start at the *end* of the Fermion flow and work backwards. That is, you work backwards

along the arrows that appear on the Fermion lines:[5]

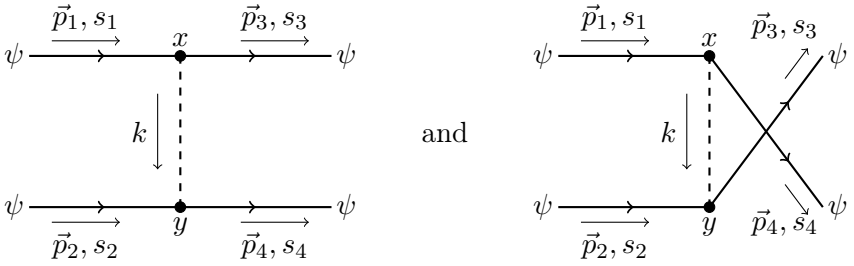


12.2.4 A Fermionic Field Theorists New Best/Worst Friend: Minus Signs

As we have tried to stress multiple times, when dealing with Fermionic fields, we cannot exchange particles willy-nilly to obtain new expressions. Of course we can do it, we just have to be careful about minus signs. These are both good and bad. They're bad because who likes to keep track of relative minus signs, but they're good because they lead to terms cancelling in our expressions that would otherwise cause problems.[6]

So what do we do? Let's look at where these minus signs manifest from Dyson's formula and Wick's theorem. As we explained above, the way you do the Wick contractions with Fermions is to move everything so that the contracted fields are next to each other, picking up the minus signs as you go, and then do the contractions. If we switch two outgoing particles, essentially what we're doing is swapping which vertices they come from (i.e. t channel vs. u channel), and so we need to move them so that they contract with the relevant propagator vertex point. Putting this together with Remark 12.2.2 we can see where these minus signs come from. This is perhaps easiest to see with an example.

Example 12.2.3. Let's consider the contractions for a t channel vs. a u channel for $\psi\psi \to \psi\psi$. The two diagrams are drawn below:



The relevant term from Dyson's formula is (forgetting about the integrals etc, its just the contractions we care about here)

$$\langle \psi(\vec{p}_4, s_4)\psi(\vec{p}_3, s_3)| \overline{\psi}_x\psi_x\phi_x\overline{\psi}_y\psi_y\phi_y |\psi(\vec{p}_1, s_1)\psi(\vec{p}_2, s_2)\rangle$$

[5] Note I have tried to colour coordinate the momentum/spin to make it more clear.
[6] See QED lectures for more information.

Now if we want to only have Fermions and no antifermions, we will need to contract the $\overline{\psi}$s with the bra vectors and the ψs with the ket vector. So firstly we move them, picking up the minus signs. Step by step, this gives us

$$\begin{aligned} &- \langle \psi(\vec{p}_4, s_4)\psi(\vec{p}_3, s_3)| \, \overline{\psi}_x \overline{\psi}_y \psi_x \phi_x \psi_y \phi_y \, |\psi(\vec{p}_1, s_1)\psi(\vec{p}_2, s_2)\rangle \\ = &- \langle \psi(\vec{p}_4, s_4)\psi(\vec{p}_3, s_3)| \, \overline{\psi}_x \overline{\psi}_y \phi_x \phi_y \psi_x \psi_y \, |\psi(\vec{p}_1, s_1)\psi(\vec{p}_2, s_2)\rangle \end{aligned}$$

were we have use the fact that we can move the ψ through the ϕs with no problems. We now use Remark 12.2.2, which basically tells us that we have to order our $\psi_x\psi_y/\overline{\psi}_x\overline{\psi}_y$ so that we have the correct momenta/spin going to the correct vertex, i.e. x or y. The convention is we fix what we call x and what we call y by where the initial momentum flows.[7] This means we essentially don't have a choice for the $\psi_x\psi_y$ contractions. As per the diagrams above, we choose to have $\psi(\vec{p}_1, s_1)$ flowing into vertex x,[8] so we need to move the ψ_x closest to the ket vector. Clearly we also have to contract the ϕ_x and the ϕ_y to get the propagator between the vertices. So we get

$$\langle \psi(\vec{p}_4, s_4)\psi(\vec{p}_3, s_3)| \, \overline{\psi}_x \overline{\psi}_y \overbrace{\phi_x \phi_y} \underbrace{\psi_y \overbrace{\psi_x \, |\psi(\vec{p}_1, s_1)}} \psi(\vec{p}_2, s_2)\rangle$$

for both the t and u channel. Seeing as this is the same for both we can drop it here and just look at the contractions between the $\overline{\psi}$s and the bra-vector. For the t channel we need to contract the $\overline{\psi}_x$ with the $\psi(\vec{p}_3, s_3)$ and the $\overline{\psi}_y$ with the $\psi(\vec{p}_4, s_4)$ (see the left diagram above). So, as per Remark 12.2.2, we want the $\overline{\psi}_x$ closest to the bra-vector, which is how our expression is already written. So for the t channel we don't pick up any more minus signs. However for the u channel we need flip the above contractions, that is contract $\overline{\psi}_x$ with $\psi(\vec{p}_4, s_4)$ and $\overline{\psi}_y$ with $\psi(\vec{p}_3, s_3)$. This means we need to place the $\overline{\psi}_y$ closest to the bra-vector and so we have to anticommute it past the $\overline{\psi}_x$. This is where our sign comes from!

For completeness, the full expressions for the t and u channel are

$$\frac{i(-ig)^2 \overline{u}(\vec{p}_3, s_3)_\alpha u(\vec{p}_1, s_1)_\alpha \overline{u}(\vec{p}_4, s_4)_\beta u(\vec{p}_2, s_2)_\beta}{(p_1 - p_3)^2 - \mu^2 + i\epsilon}$$

and

$$-\frac{i(-ig)^2 \overline{u}(\vec{p}_4, s_4)_\alpha u(\vec{p}_1, s_1)_\alpha \overline{u}(\vec{p}_3, s_3)_\beta u(\vec{p}_2, s_2)_\beta}{(p_1 - p_4)^2 - \mu^2 + i\epsilon}$$

respectively. So we don't just swap $p_3 \to p_4$ in the denominator, but we also include the minus sign in front of the second expression for the Fermion statistics.

Remark 12.2.4. Note that we have introduced labels α/β on the matrices so we can see what is contracted with what, e.g. for the first expression the incoming $\psi(\vec{p}_1, s_1)$ couples to the outgoing $\psi(\vec{p}_3, s_3)$.

Remark 12.2.5. Note it doesn't matter which term we put the minus sign in front of, as long as there is a relative minus sign between the two expressions. This is just because the matrix element for a given process is given by a sum over all the different contributions (this

[7] More accurately, we include a term $x \leftrightarrow y$ in our expressions, so we fix the x and y and just calculate the answer for this, then add the symmetry factors at the end.

[8] Note we do this so that we cancel the minus factor in our expression above.

just comes from the fact that each diagrams corresponds to a different contraction type, and we sum over the contractions). Now the two expressions above are just the individual contributions to the matrix element, but the individual terms are not measurable; all we can measure is the absolute value of the full matrix element squared. The difference between $\mathcal{M}_t - \mathcal{M}_u$ and $\mathcal{M}_u - \mathcal{M}_t$ is clearly just an *overall* minus sign, and which is completely redundant. That is

$$|\mathcal{M}_t - \mathcal{M}_u|^2 = |\mathcal{M}_u - \mathcal{M}_t|^2.$$

12.3 QED

We end the course with a very brief discussion of QED. We do not have the time here to discuss it in detail, but there is a separate QED course which will expand much further on the theory. The aim here is basically just to introduce the Lagrangian and state the Feynman rules.

Remark 12.3.1. I have tried to expand on some of the equations/ideas presented in the course, however there is still a significant amount of information that would be useful to know. As always, basically all of this information can be found in Prof. Tong's notes, chapter 6.

Ok so here's the information we need but don't derive:

- The photon field is denoted by A^μ.[9]
- The Lagrangian for the photon field is given by

$$\mathcal{L}_{\text{Maxwell}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{12.8}$$

  where

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

- The photon couples to electric charge, by which we mean the conserved charge for QED is given by local conservation of electric charge. So things like electrons participate in QED reactions, and, as the photon is electrically neutral, the only acceptable vertices for an electron in QED is scattering (i.e. one electron in one electron out) or annihilation (i.e. electron and positron in, photon out).

Remark 12.3.2. Note that the fact that the photon is electrically neutral means that we cannot have any photon-photon interactions. This is contrast to, say, ϕ^4 theory where we had a vertex with more than one ϕ. A better comparison to make would be to QCD (the QFT of the strong force) where the gauge boson,[10] called a *gluon*, is charged under the QCD charge (called colour). This means that we *can* get gluon self interactions. This is not discussed at all in this course, but just mentioned for interest.

[9] Note physical photons only have 2 polarisations, so we should only have 2 degrees of freedom in A^μ, however we naïvely have 4 (one for each $\mu = 0, .., 3$. It turns out you can remove two of these leaving only 2 degrees of freedom. We obtain this using gauge fixing, see Prof. Tong's notes for more details.

[10] Basically this just means the particle that describes the force. So the gauge boson for QED is the photon.

Ok so let's couple the Maxwell theory to some Fermions. The Lagrangian is given by putting together the Maxwell Lagrangian, Equation (12.8), and the Dirac Lagrangian, Equation (10.17), and then include a coupling between the fields. The result is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\psi}(i\not{D} - m)\psi \qquad D_\mu := \partial_\mu + ieA_\mu. \tag{12.9}$$

where we recall that $\not{D} := \gamma^\mu D_\mu$. We see that our coupling constant is $-ie\gamma^\mu$, and so we have

$$H_I = ie\int d^3\vec{x}\,\overline{\psi}\not{A}\psi \tag{12.10}$$

Remark 12.3.3. The definition of D_μ in Equation (12.9) might seem a bit strange, however we can show that this definition is required to preserve our internal symmetries. For an outline of this, we set the following exercise.

Exercise

Show that Equation (12.9) is invariant under the transformation

$$A_\mu \to A_\mu + \partial_\mu\alpha, \quad \text{and} \quad \psi \to e^{-ie\alpha}\psi$$

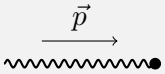
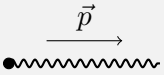
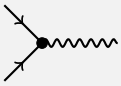
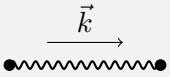
for some function $\alpha(x)$. *Hint: Show that*

$$D_\mu\psi \to e^{-ie\alpha}D_\mu\psi$$

to argue that $\overline{\psi}D_\mu\psi$ *is invariant.*

12.3.1 Feynman Rules

We already have all the Feynman rules for the Fermion fields, so all we need to include here are the vertex factor and the photon rules. We derive these by considering the quantisation of gauge fields, however there was no time in this course to cover that, so we just state the momentum space Feynman rules in the table below.

Type	Diagram	Maths Expression
Incoming Photon		$\varepsilon_\mu(p) = \overbrace{A_\mu \, \vert\vec{p},\varepsilon\rangle}$
Outgoing Photon		$\varepsilon^*_\mu(p) = \langle\vec{p},\overbrace{\varepsilon\vert \, A_\mu}$
Vertex		$-ie\gamma^\mu$
Photon Propagator[a]		$\dfrac{-i\eta_{\mu\nu}}{k^2+i\epsilon}$

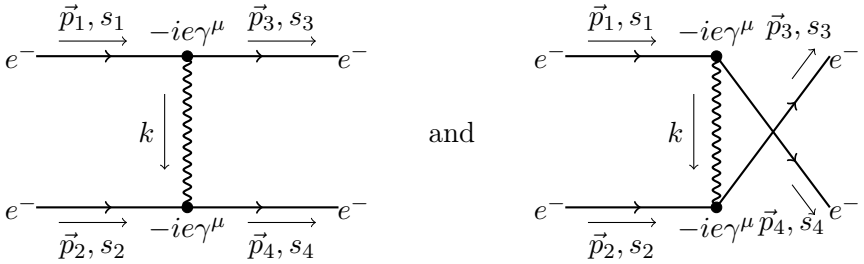
[a] In the Feynman gauge. We give a more general expression in the QED course.

We use these rules in the same way as for the previous theories to convert Feynman diagrams into mathematical expressions. Note that we are considering couplings to Ferimions (i.e. the Dirac fields) and so we have to include the factors of -1 between diagrams as per the end of the last section.

Remark 12.3.4. Note there is no mass term in the denominator of the propagator term. This is because the photon is massless.

We end the course by giving one example of a Feynman diagram in QED and its corresponding mathmatical expression.

Example 12.3.5. As we stated in the bullet points above, photons couple to electric charge, and so the scattering process $e^-e^- \to e^-e^-$ is valid. The only two valid diagrams are the t channel and u channel, i.e.[11]



[11] The right diagram is a bit cramped, apologies. I just wanted to save time drawing a whole new diagram so copied the previous one and made some small edits. I think its clear what labels what so decided not to spend ages tweaking it.

These two diagrams correspond to the expressions

$$\overline{u}(\vec{p}_3, s_3)_\alpha(-ie\gamma^\mu)u(\vec{p}_1, s_1)_\alpha \frac{(-i\eta_{\mu\nu})}{(p_1 - p_3)^2 + i\epsilon}\overline{u}(\vec{p}_4, s_4)_\beta(-ie\gamma^\mu)u(\vec{p}_2, s_2)_\beta$$

and

$$-\overline{u}(\vec{p}_4, s_4)_\alpha(-ie\gamma^\mu)u(\vec{p}_1, s_1)_\alpha \frac{(-i\eta_{\mu\nu})}{(p_1 - p_4)^2 + i\epsilon}\overline{u}(\vec{p}_3, s_3)_\beta(-ie\gamma^\mu)u(\vec{p}_2, s_2)_\beta$$

respectively. Note the minus sign in front of the second expression. This comes from the exchange of the end state particles.

For more experience with QED we end with some exercises.

Exercise

Explain why we don't get a s channel. *Hint: The answer to this is essentially given at the start of this section, so if you've properly read everything this is done...*

Exercise

Draw the Feynman diagrams and obtain the corresponding mathematical expression for the QED process $e^-e^+ \to e^-e^+$. *Hint: The answer to the last exercise might be useful.*

Exercise

There is something called a muon, which is roughly speaking just a heavier version of the electron. We denote it by μ^-. Importantly it is also electrically charged, and so couples to the photon. Draw the Feynman diagrams for the scattering process $e^-\mu^- \to e^-\mu^-$, and write down their mathematical expressions. If you like you can distinguish between electron and muon type particles by denoting electrons with u/vs and muons with an a/bs.

Useful Texts & Further Readings

Textbooks

- M. E. Peskin and D. V. Schroeder, "*An Introduction to QFT*" (Addison Wesley, 1995).
- C. Itzykson and J.-B. Zuber, "*Quantum Field Theory*" (McGraw-Hill,1980).
- L. H. Ryder, "*Quantum Field Theory*" (Cambridge University Press, 1996).
- P. Ramond, "*Field Theory, A Modern Primer*" (Benjamin, 1994).
- S. Weinberg, "*Quantum Theory of Fields*", vol I and II (Cambridge University Press, 1996).
- A. Zee, "*Quantum Field Theory in a Nutshell*" (Princeton University Press, 2010).
- F. Mandl and G. Shaw, "*Quantum Field Theory*" (Wiley, 1984.
- L. H. Ryder, "*Quantum Field Theory*" (Cambridge University Press, 1996).
- Schwartz, Matthew D. *Quantum field theory and the standard model* (Cambridge University Press, 2014).

Other Similar Courses Available Online

- Prof. David Tong, "Lectures on Quantum Field Theory", Cambridge University.
- Prof. Michael Luke, "PHY2403F Lecture Notes", University Of Toronto.
- Prof. Timo Weigand "Quantum Field Theory I + II", Institute for Theoretical Physics, Heidelberg University.