Introduction to Quantum Field Theory

Tongji University

Jinsu Kim

Last updated: June 21, 2024

Contents

※ Lecture 1 (2024.03.01)

1.1 Introduction

Welcome to Introduction to Quantum Field Theory (QFT) course. QFT is usually taught over 3 semesters, QFT I, QFT II, and QFT III (sometimes called Advanced QFT). It is obviously impossible to cover everything about QFT in just one semester. Our goal is to cover some basics that would correspond to QFT I contents.

1.1.1 What and Why

QFT is a study about our universe by using quantum fields. QFT consists of three words: Quantum, Field, and Theory. Let's have a closer look at these words. First, Theory: Simply speaking, a theory is a type of abstract thinking about a phenomenon or the results of such thinking. For us (physicists), a theory is an explanation of an aspect of nature and universe. Now, let's look at the word Quantum: a quantum is the minimum amount of any physical entity (physical property) involved in an interaction. This is the same Quantum you have encountered in Quantum Mechanics. What about the word Field? A field is something (normally represented by a scalar, vector, or tensor) that has a value for each point in space and time. Now, you may have some feelings as to what QFT is.

You have learnt Quantum Mechanics. You have learnt Classical Field Theory. At this stage, you may wonder why we need to learn QFT. According to the name QFT, it seems like it is just a combination of Quantum Mechanics and Classical Field Theory. The crucial aspect is the relativity, specifically Special Relativity. When Quantum Mechanics and Special Relativity meet each other, one observes that the particle number is not conserved. In other words, particles may pop up out of nowhere, or particles may disappear, which cannot be captured by, say, Quantum Mechanics alone.

So, in summary, QFT is the child from the marriage of Quantum Mechanics and Special Relativity, and it is the language for the description of the interactions of elementary particles in our universe, which are the products of field quantisations. QFT is essential for almost every branches of physics, including Condensed Matter Physics, High Energy Physics, Cosmology, Quantum Gravity, and Mathematical Physics.

1.1.2 Prerequisites

It is necessary that you have taken the following courses (or at least studied them):

- Classical Mechanics,

- Electrodynamics / Electromagnetism,
- Quantum Mechanics, and
- Special Relativity.

You must be comfortable with terms like the Langrangian formulation, the Hamiltonian formulation, locality, number operators, quantisation, Lorentz transformation, *etc*.

1.1.3 Useful Resources

There are many excellent textbooks on QFT:

- M. Peskin and D. Schroeder, ”An Introduction to Quantum Field Theory”
- M. Schwartz, ”Quantum Field Theory and the Standard Model”
- S. Weinberg, ”The Quantum Theory of Fields”
- M. Srednicki, ”Quantum Field Theory”
- A. Zee, ”Quantum Field Theory in a Nutshell”
- and many more

We will not be following any specific textbook, but the contents of this lecture will be similar to the first part of the Peskin and Schroeder.

There are also numerous lecture notes on QFT available online:

- D. Tong, ”Quantum Field Theory”
- C. Scrucca, ”Advanced Quantum Field Theory”
- and many more

The current lecture note is largely based on D. Tong’s beautiful lecture note.

1.2 Notations

We will work in the so-called “natural” units:

$$c = \hbar = 1\,.$$

In other words, the speed of light c ($= 3\times 10^8\,\mathrm{m/s}$) is set to unity, and reduced Planck's constant (Planck's constant divided by 2π) $\hbar$ ($= 1\times 10^{-34}\,\mathrm{J\cdot s}$) is also set to unity. What does it mean? To answer this question, let's have a look at their dimensions (recall that $\mathrm{J} = \mathrm{kg\cdot m^2/s^2}$):

$$\begin{aligned}\text{dim. of } c &= LT^{-1}\,,\\ \text{dim. of } \hbar &= L^2MT^{-1}\,.\end{aligned}$$

So, $c = 1$ allows us to view the space and time as the same thing; 1 metre is same as 1 second. Furthermore, $\hbar = 1$ (together with $c = 1$) allows us to equate mass and space; 1 kilogram is same as 1 inverse metre. We see the usefulness here: We can now express every quantities in terms of just one single scale, which we choose to be mass or, equivalently, energy, because Einstein said that $E = mc^2$ which has become $E = m$ as $c = 1$.

We will always work in the $(1+3)$-dimensional spacetime.

We will use Greek letters for space and time indices, *e.g.*, $\mu = 0, 1, 2, 3$, with 0 corresponding to time. For example, $x^\mu = \{t, x, y, z\}$.

We will use Latin letters for space indices, *e.g.*, $i = 1, 2, 3$. For example, $x^i = \{x, y, z\}$. Note that $x^\mu = \{t, x^i\}$.

We will use the metric convention of $(+,-,-,-)$. In other words, the Minkowski metric is given by

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

(I personally use $(-,+,+,+)$, but $(+,-,-,-)$ is more common in the Particle Physics / High Energy Physics community.)

We will use the Einstein summation convention:

$$A^\mu B_\mu = A^0B_0 + A^1B_1 + A^2B_2 + A^3B_3\,.$$

We will also use the notation $AB \equiv A\cdot B \equiv A^\mu B_\mu = \eta^{\mu\nu}A_\mu B_\nu = \eta_{\mu\nu}A^\mu B^\nu$. For example, $x^2 = x^\mu x_\mu = \eta_{\mu\nu}x^\mu x^\nu = t^2 - (x^2 + y^2 + z^2)$.

Many more as we go along.

1.3 Recap: Classical Field Theory and Quantum Mechanics

Before we talk about QFT, let us first review some basics of Classical Field Theory and Quantum Mechanics. Most of the things presented in this section should be familiar to you.

1.3.1 Action and Equation of Motion

Let ϕ be a (classical) field. Do you still remember what a field is? As we saw in the beginning of this lecture, a field is something (normally represented by a scalar, vector, or tensor) that has a value for each point in space and time. The value of the field ϕ at a specific spacetime point $(t, \mathbf{x})$ is $\phi(t, \mathbf{x})$. We shall use x to denote both the time and the space coordinates, $x = (t, \mathbf{x})$, and thus, we write $\phi(x)$. It is important to note that the position $\mathbf{x}$ is a mere label, not a dynamical variable.

Now, we would like to describe the dynamics of the field ϕ. In Classical Mechanics, we had something called Lagrangian. In Field Theory, since a field is a quantity at each point in space and time, we have a density, the Lagrangian density $\mathcal{L}$[1],

$$\mathcal{L} = \mathcal{L}(\phi, \partial_\mu \phi) \,,$$

where

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} \,.$$

Note that a Lagrangian density is a function of the field and its derivative. For a given time stamp, we can get the Lagrangian by integrating over space:

$$L = L(t) = \int d^3\mathbf{x}\, \mathcal{L}(\phi, \partial_\mu \phi) \,.$$

We rarely use the Lagrangian L in this course. We almost always use the Lagrangian density $\mathcal{L}$. So, $\mathcal{L}$ is sometimes simply called the Lagrangian. The action is then defined by

$$S = \int_{t_1}^{t_2} dt\, L(t) = \int d^4x\, \mathcal{L} \,.$$

Let us pause here for a moment and have a look at the dimensions; we use the notation $[A]$ to denote the mass dimension of A. We have $[d^4x] = -4$. So, in order to make the action dimensionless[2], we must have $[\mathcal{L}] = 4$, *i.e.*, the mass dimension of a Lagrangian density should be four (in four-dimensional spacetime, of course).

From the action S, we can obtain the equation of motion by the least action principle; $\delta S = 0$. Under $\phi \to \phi + \delta\phi$, we get

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

[1] Why only first derivative and no higher derivatives? Locality; $\phi(t, \mathbf{x})$ has no coupling to $\phi(t, \mathbf{y})$ if $\mathbf{x} \neq \mathbf{y}$. Why no explicit x-dependence in $\mathcal{L}$? Translational invariance.

[2] Recall the path integral from your Quantum Mechanics course where $\exp(-iS/\hbar)$ appears a lot.

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

The last term, which is the boundary term, vanishes for anything that decays at spatial infinity and obeys $\delta\phi(t_1, \mathbf{x}) = \delta\phi(t_2, \mathbf{x}) = 0$. Thus, requiring $\delta S = 0$ gives

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

This is called the Euler-Lagrange equation, and it is the equation of motion for the field ϕ.

Let us take a look at some examples. The simplest example would be a (free) real scalar field whose mass is m. The Lagrangian density for such a field (called the Klein-Gordon field) is given by

$$\mathcal{L} = \frac{1}{2}\partial^\mu \phi \partial_\mu \phi - \frac{1}{2} m^2 \phi^2 = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\nabla\phi)^2 - \frac{1}{2}m^2\phi^2 .$$

Notice that we have the kinetic energy $\dot{\phi}^2/2$ and the potential energy $(\nabla\phi)^2/2 + m^2\phi^2/2$; we have $\mathcal{L} = T - V$. Let us try to find the equation of motion for the field ϕ. It is straightforward to see that

$$\frac{\partial \mathcal{L}}{\partial \phi} = -m^2 \phi ,$$
$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = \partial^\mu \phi = (\dot{\phi}, -\nabla\phi) .$$

Therefore, the Euler-Lagrange equation tells us that

$$\partial_\mu \partial^\mu \phi + m^2 \phi = 0 ,$$

or, more explicitly,

$$\ddot{\phi} - \nabla^2 \phi + m^2 \phi = 0 .$$

This is called the Klein-Gordon equation.

We can easily generalise this simple case to the system of

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

with a generic potential function V. In this case, the Euler-Lagrange equation says that

$$\partial_\mu \partial^\mu \phi + \frac{\partial V}{\partial \phi} = 0 .$$

Exercise 1.1: Consider the following Lagrangian for a vector field A^μ:

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu A_\nu)(\partial^\mu A^\nu) + \frac{1}{2}(\partial_\mu A^\mu)^2 ,$$

or

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

where

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \,.$$

It is an exercise for you to work out the Euler-Lagrange equations:

$$\partial_i E_i = 0 \,, \quad \dot{E}_i = \epsilon_{ijk} \partial_j B_k \,.$$

Do you recognise them? They are the Maxwell equations.

We will not be able to cover how one may come up with all these Lagrangians, but, in general, one is guided by two principles, namely symmetry and renormalisability.

1.3.2 Lorentz Invariance

Now that we talked a little bit about (classical) fields, let us talk about the relativistic aspects. Recall that we want relativistic field theories. The time and space are on an equal footing, and the theory should remain the same whichever reference frame we choose. Thus, the Lagrangian should be Lorentz-invariant. In other words, the Lagrangian needs to be left unchanged under the Lorentz transformation,

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

with $\Lambda^\mu{}_\rho \eta^{\rho\sigma} \Lambda^\nu{}_\sigma = \eta^{\mu\nu}$. You may remember that a rotation by an angle θ around the x-axis is described by

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

and that a boost by v along the x-axis is described by

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

where $\gamma = 1/\sqrt{1-v^2}$ is the Lorentz gamma factor.

How do fields transform under the Lorentz transformation? The Lorentz transformations have a representation on the fields. A scalar field $\phi(x)$ transforms under the Lorentz transformation as[3]

$$\phi(x) \longrightarrow \phi'(x) = \phi(\Lambda^{-1}x)\,,$$

where $\Lambda^{-1}x \equiv (\Lambda^{-1})^{\mu}{}_{\nu}x^{\nu}$. This transformation tells us that the transformed field, evaluated at the transformed point, gives the same value as the original field evaluated at the point before the transformation[4]. If a theory is Lorentz invariant, and if $\phi(x)$ is a solution to the equation of motion, then ϕ' should also be a solution. This is equivalent to imposing the Lorentz invariance on the action S.

Exercise 1.2: Check that the Klein-Gordon theory is Lorentz-invariant.

1.3.3 Noether's Theorem

Just as in Classical Mechanics, we have Noether's theorem in the field theory: Symmetries of a Lagrangian give us conserved quantities. Let us be more specific. Noether's theorem states that every continuous symmetry of a Lagrangian gives rise to a conserved current $j^{\mu}(x)$ such that the equation of motion implies that

$$\partial_{\mu}j^{\mu} = 0\,.$$

A conserved current gives rise to a conserved charge,

$$Q = \int d^3\mathbf{x}\, j^0\,.$$

Note that

$$\frac{dQ}{dt} = \int d^3\mathbf{x}\,\frac{dj^0}{dt} = -\int d^3\mathbf{x}\,\nabla\cdot\mathbf{j} = 0\,,$$

given that $\mathbf{j} \longrightarrow 0$ as $|\mathbf{x}| \longrightarrow \infty$.

Let us prove the theorem. If a transformation $\phi \longrightarrow \phi + X$ is a symmetry, then we would have $\delta\mathcal{L} = 0$, *i.e.*, the Lagrangian density is left unchanged. Next, we note that, if we have a

[3] This is an active transformation. Alternatively, one may use a passive transformation where the coordinates would get transformed. In this case, we have $\phi(x) \longrightarrow \phi(\Lambda x)$. Note that it does not matter whether we choose the active one or the passive one. If Λ is a Lorentz transformation, then so is Λ^{-1}. So, being invariant under active transformations is the same as being invariant under passive transformations.

[4] Consider a temperature field, for example. Imagine that the original field $\phi(\mathbf{x})$ has a hotspot at $\mathbf{x} = (1, 0, 0)$. After a rotation, $\mathbf{x} \longrightarrow R\mathbf{x}$ about the z-axis, the new field $\phi'(\mathbf{x})$ will have the hotspot at $R\mathbf{x}$. If we want to express $\phi'(\mathbf{x})$ in terms of the original field ϕ, we need to place ourselves at $R\mathbf{x}$ and ask what the original field looked like where we have come from at $R^{-1}(R\mathbf{x}) = \mathbf{x}$. This is why we have the inverse Λ^{-1} in the argument.

continuous symmetry, we can always consider an infinitesimal change, *i.e.*, $X = \delta\phi$. Under the transformation,

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

When ϕ satisfies the equation of motion, the first term goes away, and we are left with

$$\delta\mathcal{L} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\right) .$$

For the symmetry transformation $\delta\phi = X$, we have $\delta\mathcal{L} = 0$. So,

$$\partial_\mu j^\mu = 0\,, \quad \text{with} \quad j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}X\,.$$

We can actually generalise it a bit: Instead of demanding $\delta\mathcal{L} = 0$ for a symmetry, we can say that X is a symmetry if $\delta\mathcal{L} = \partial_\mu F^\mu(\phi)$ for some $F^\mu(\phi)$, *i.e.*, a total derivative. Following the same steps presented above gives

$$\partial_\mu j^\mu = 0\,, \quad \text{with} \quad j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}X - F^\mu\,.$$

Let us now have some examples. Consider the following transformation:

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

Then, we have (note the plus sign as we are doing an active transformation)

$$\phi \longrightarrow \phi + \epsilon^\mu\partial_\mu\phi\,.$$

So, in the notation we used above, $X = \epsilon^\mu\partial_\mu\phi$. Similarly,

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

We see that the change in the Lagrangian is a total derivative, $\delta\mathcal{L} = \partial_\mu F^\mu$ with $F^\mu = \epsilon^\mu\mathcal{L}$. We thus have the following Noether current:

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

The ϵ^μ is just an overall constant, so we may define the current as follows:

$$(j^\mu)_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\partial_\nu \phi - \delta^\mu_\nu \mathcal{L}\,.$$

Note that we have 4 conserved currents, one for each of the translations. This current is the energy-momentum tensor $T^\mu{}_\nu$. It satisfies

$$\partial_\mu T^\mu{}_\nu = 0\,.$$

The four conserved quantities are given by

$$E = \int d^3\mathbf{x}\, T^{00}\,,$$

which is the total energy of the field configuration, and

$$P^i = \int d^3\mathbf{x}\, T^{0i}\,,$$

which is the total momentum of the field configuration.

Exercise 1.3: Work out the energy-momentum tensor for the Klein-Gordon theory.

Here is another example. Consider a complex (rather than real) scalar field

$$\psi(x) = \frac{\phi_1(x) + i\phi_2(x)}{\sqrt{2}}\,,$$

where $\phi_{1,2}$ are real scalar fields. Consider the following Lagrangian:

$$\mathcal{L} = \partial_\mu \psi^* \partial^\mu \psi - V(\psi^*\psi)\,,$$

where V is some potential that is a general polynomial in $\psi^*\psi$, *e.g.*,

$$V(\psi^*\psi) = m^2\psi^*\psi + \frac{\lambda}{2}(\psi^*\psi)^2\,.$$

Expanding ψ in terms of ϕ_1 and ϕ_2, we can find the equations of motion. Or, equivalently, we can treat ψ and ψ^* as independent variables and vary the action with respect to both of them to find the equations of motion. For example, $\delta S/\delta\psi^* = 0$ gives us

$$\partial_\mu \partial^\mu \psi + \frac{\partial V}{\partial \psi^*} = 0\,.$$

The Lagrangian has a symmetry, *i.e.*, $\delta\mathcal{L} = 0$, under

$$\psi \to e^{i\alpha}\psi\,.$$

If we view this transformation as an infinitesimal one, we have

$$\delta\psi = i\alpha\psi\,, \quad \delta\psi^* = -i\alpha\psi^*\,.$$

The associated conserved current is

$$j^\mu = i(\partial^\mu\psi^*)\psi - i\psi^*(\partial^\mu\psi)\,.$$

This type of symmetry is called an internal symmetry.

Exercise 1.4: Work out the current $j^\mu = i(\partial^\mu\psi^*)\psi - i\psi^*(\partial^\mu\psi)$.

❊ Lecture 2 (2024.03.08)

2.1 Recap: Classical Field Theory and Quantum Mechanics – continued

2.1.1 Hamiltonian Formalism

In Quantum Mechanics, we encountered a quantity called Hamiltonian. The Hamiltonian is related to the Lagrangian. In Classical Mechanics, for example, we learnt that we can define the so-called conjugate momentum and construct a Hamiltonian from a Lagrangian. We can do the same thing in field theories. We first define the conjugate momentum $\pi(x)$:

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

Note that this conjugate momentum is a function of x, and it is not the same thing as the total momentum P^i, which is a single number characterising the whole field configuration, we saw earlier. The Hamiltonian density is then given by

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

The Hamiltonian is

$$H = \int d^3\mathbf{x}\, \mathcal{H}\,.$$

Just like we were able to find equations of motion for the field ϕ from its Lagrangian, *i.e.*, the Euler-Lagrange equations, we can find equations of motion from the Hamiltonian by using Hamilton's equations:

$$\dot{\phi} = \frac{\partial \mathcal{H}}{\partial \pi}\,, \quad \dot{\pi} = -\frac{\partial \mathcal{H}}{\partial \phi}\,.$$

As an example, consider the Klein-Gordon theory,

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi) = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\nabla\phi)^2 - V(\phi)\,.$$

The conjugate momentum is

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

So, the Hamiltonian is given by

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

Do you recognise this? This agrees with the total energy that we get from the Lagrangian formulation.

2.1.2 Quantum Mechanics

We still have one more thing to recap: Quantum Mechanics. Two important concepts are the canonical quantisation and the harmonic oscillator. These will play an essential role in QFT.

Recall the simple harmonic oscillator from Classical Mechanics, which is given by the Hamiltonian

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

where p is the momentum, and q is the position. We want to quantise it and obtain the corresponding quantum system. We can do so via the canonical quantisation: We promote the p and q into complex operators, $\hat{p}$ and $\hat{q}$, and use the same formula for the Hamiltonian,

$$\hat{H} = \frac{1}{2}\hat{p}^2 + \frac{1}{2}\omega^2 \hat{q}^2 \,.$$

In the classical system, the p and q satisfy the Poisson brackets,

$$\{q, p\} = 1 \,.$$

In the quantum system, *i.e.*, after the canonical quantisation, they satisfy the commutation relation,

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

(If you need a refresh on the Poisson brackets and the commutators, see **Appendix – Poisson Brackets and Commutators**.)

Let us take a moment and think about this process. (We shall drop the hats, $\hat{\cdot}$, as we will be dealing with operators a lot.)

- We promoted p and q to operators, but what do they actually operate on? We will analyse these operators formally and show that there is a space the operators naturally act on. We will then take that space as our state space. In other words, we will suppose that there are some states our operators act on, and then try to figure out what properties the states must have.

- Consider, for example, a term p^2q. How should we promote this term? Classically, p^2q, pqp, and qp^2 are no different, but the operator p^2q is not the same thing as, say, the operator pqp. In other words, how we decide to order terms in the Hamiltonian actually matters. We shall see later the notion of "normal ordering" (or you may already have heard of this or learnt this) which gives a (slightly) more consistent way of doing the quantisation.

Once we are done with the quantisation, the time evolution of states is governed by the Schrödinger equation:

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

In practice, instead of trying to solve this thing, we try to find eigenstates $|E\rangle$ such that

$$H|E\rangle = E|E\rangle\,.$$

Then, defining

$$|\psi\rangle = e^{-iEt}|E\rangle$$

would give a nice, stable, solution to the Schrödinger equation.

Notice that, in the classical case, we can factorise the Hamiltonian as

$$H = \omega\left(\sqrt{\frac{\omega}{2}}q + \frac{i}{\sqrt{2\omega}}p\right)\left(\sqrt{\frac{\omega}{2}}q - \frac{i}{\sqrt{2\omega}}p\right)\,.$$

The Hamiltonian is now a product of two terms that are complex conjugates to each other; in terms of operators, they are adjoints. Thus, instead of working with two real objects, q and p, we can work with a single complex object, $\sqrt{\omega/2}q + ip/\sqrt{2\omega}$.

Let us do the exact same factorisation in the quantum case. Of course, we would not expect the result to be exactly the same as above due to the commutator relations. Still, we can try and define the operators

$$a = \sqrt{\frac{\omega}{2}}q + \frac{i}{\sqrt{2\omega}}p\,, \quad a^\dagger = \sqrt{\frac{\omega}{2}}q - \frac{i}{\sqrt{2\omega}}p\,.$$

These are known as the annihilation and creation operators as you already know. We may invert these and find

$$q = \frac{1}{\sqrt{2\omega}}\left(a + a^\dagger\right)\,, \quad p = -i\sqrt{\frac{\omega}{2}}\left(a - a^\dagger\right)\,.$$

From $[q, p] = i$, we can find that

$$[a, a^\dagger] = 1\,.$$

We further note that

$$a^\dagger a = \frac{\omega}{2}q^2 + \frac{p^2}{2\omega} - \frac{1}{2}\,,$$

and thus, $H = \omega(a^\dagger a + 1/2)$. (Equivalently, we could have started from the Hamiltonian expression and put p and q in terms of a and $a^\dagger$.) Now, we can compute

$$[H, a^\dagger] = \omega a^\dagger\,, \quad [H, a] = -\omega a\,.$$

From these, we note that a and $a^\dagger$ take us to different energy eigenstates. For example, if $H|E\rangle = E|E\rangle$, then

$$Ha^\dagger|E\rangle = (a^\dagger H + [H, a^\dagger])|E\rangle = (E+\omega)a^\dagger|E\rangle \,.$$

Similarly, we have

$$Ha|E\rangle = (E-\omega)a|E\rangle \,.$$

Thus, assuming that we have some energy eigenstate $|E\rangle$, the operators a and $a^\dagger$ give us loads more with eigenvalues

$$\cdots, E-2\omega, E-\omega, E, E+\omega, E+2\omega, \cdots$$

If the energy is bounded below, then there must be a ground state $|0\rangle$ such that $a|0\rangle = 0$. The other excited states would then come from applying $a^\dagger$ repeatedly, labelled by

$$|n\rangle = (a^\dagger)^n|0\rangle \,,$$

with

$$H|n\rangle = \left(n + \frac{1}{2}\right)\omega|n\rangle \,.$$

Note that we simply ignored the normalisation, so we do not have $\langle n|n\rangle = 1$.

Let us have a closer look at the ground state $|0\rangle$. The energy eigenvalue is given by

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

In other words, the ground-state energy is non-zero. Furthermore, we did not mention anything about having a particular state space, and yet, we managed to figure out what the eigenvalues of H must be. Having done that, we now know what the appropriate space is to work on: the right space is the Hilbert space generated by the orthonormal (again, we did not normalise the states before, but it is straightforward to do so) basis

$$\{|0\rangle, |1\rangle, |2\rangle, \cdots\} \,.$$

We are done with the recap. Time to explore QFT!

2.2 Free Real Scalar Quantum Fields

Similar to what we did in Quantum Mechanics, we use canonical quantisation to promote our classical fields to quantum fields. The plan is to repeat what we did for the harmonic oscillator. We know the commutation relations the creation and annihilation operators satisfy, so we will write the field and its conjugate momentum in terms of those operators. Then, we will be able to find the corresponding commutation relations for the field and its conjugate momentum.

Let us consider first a (free) real scalar field ϕ. At this stage, the field is still a classical field. The equation of motion is given by

$$\partial^2\phi + m^2\phi = 0\,, \tag{1}$$

where $\partial^2 \equiv \partial^\mu\partial_\mu$. We now take the Fourier transform,

$$\phi(t,\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}}\tilde{\phi}(t,\mathbf{p})\,. \tag{2}$$

From the Klein-Gordon equation, we see that $\tilde{\phi}$ satisfies

$$\left[\frac{\partial^2}{\partial t^2} + \left(\mathbf{p}^2 + m^2\right)\right]\tilde{\phi}(t,\mathbf{p}) = 0\,. \tag{3}$$

You may recognise this: This is the very equation for the simple harmonic oscillator with the frequency of $\omega_\mathbf{p} = \sqrt{\mathbf{p}^2 + m^2}$. Therefore, the solution to the Klein-Gordon equation is just a superposition of (infinite) simple harmonic oscillators.

Exercise 2.1: Verify Eq. (3).

Remember, we have been dealing with a classical field so far. It is time to quantise the field ϕ. We just saw that $\tilde{\phi}$ behaves like the position of a harmonic oscillator. Thus, its conjugate momentum $\tilde{\pi}$ would correspond to the momentum of a harmonic oscillator. We also noted that the position and the momentum can be written in terms of the creation and annihilation operators. Thus, we write, in analogy with the simple harmonic oscillators from Quantum Mechanics[5],

$$\phi(\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_\mathbf{p}}}\left(a_\mathbf{p} + a^\dagger_{-\mathbf{p}}\right)e^{i\mathbf{p}\cdot\mathbf{x}}\,, \tag{4}$$

$$\pi(\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^3}(-i)\sqrt{\frac{\omega_\mathbf{p}}{2}}\left(a_\mathbf{p} - a^\dagger_{-\mathbf{p}}\right)e^{i\mathbf{p}\cdot\mathbf{x}}\,. \tag{5}$$

Equivalently, we may write as follows:

$$\phi(\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_\mathbf{p}}}\left(a_\mathbf{p}e^{i\mathbf{p}\cdot\mathbf{x}} + a^\dagger_\mathbf{p}e^{-i\mathbf{p}\cdot\mathbf{x}}\right)\,, \tag{6}$$

$$\pi(\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^3}(-i)\sqrt{\frac{\omega_\mathbf{p}}{2}}\left(a_\mathbf{p}e^{i\mathbf{p}\cdot\mathbf{x}} - a^\dagger_\mathbf{p}e^{-i\mathbf{p}\cdot\mathbf{x}}\right)\,. \tag{7}$$

[5]You may wonder why the field and its conjugate momentum depend only on space, but not on time. This is because we are specifying a preferred time coordinate here. The state depends on time, while the field and the momentum depend on space as they are operators; this is just like the Schrödinger picture in Quantum Mechanics. We will get back to this later.

Note that ϕ is real. Similar to the way the position and momentum operators satisfy commutation relations, the field ϕ and its conjugate momentum π satisfy the following commutation relations:

$$\begin{aligned} [\phi(\mathbf{x}), \phi(\mathbf{y})] &= 0 = [\pi(\mathbf{x}), \pi(\mathbf{y})] \,, \\ [\phi(\mathbf{x}), \pi(\mathbf{y})] &= i\delta^{(3)}(\mathbf{x}-\mathbf{y}) \,. \end{aligned} \tag{8}$$

Exercise 2.2: Show that the commutation relations of Eq. (8) are equivalent to $[a_{\mathbf{p}}, a_{\mathbf{q}}] = 0 = [a^\dagger_{\mathbf{p}}, a^\dagger_{\mathbf{q}}]$ and $[a_{\mathbf{p}}, a^\dagger_{\mathbf{q}}] = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q})$.

The next step is to find the Hamiltonian in terms of the creation and annihilation operators, just like we did in Quantum Mechanics. All we have to do is to substitute Eq. (6) and Eq. (7) into

$$H = \int d^3\mathbf{x} \left[\frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2\right] \,. \tag{9}$$

Before we actually do the calculation, let us try to imagine what we will be seeing. We saw earlier that the system is just an infinite set of (decoupled) harmonic oscillators. So, the Hamiltonian will be, again, just an integral of the Hamiltonian of those infinite harmonic oscillators. But, we also saw earlier that a harmonic oscillator has a non-zero vacuum, ground-state energy, which is proportional to the frequency. Thus, if we sum up all, infinitely many, ground-state energies, we would get the ground-state energy of infinity. Furthermore, if you recall the steps for getting the ground-state energy, we will have something like $\omega_{\mathbf{p}}[a_{\mathbf{p}}, a^\dagger_{\mathbf{p}}]$, which is itself infinity. Therefore, we will be having an infinite sum of infinities.

Let us do the actual calculation and see if that indeed happens. For example, the first term gives

$$\begin{aligned} \int d^3\mathbf{x}\, \pi^2 &= -\int d^3\mathbf{x} \frac{d^3\mathbf{p} d^3\mathbf{q}}{(2\pi)^6} \frac{\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{q}}}}{2} \left(a_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} - a^\dagger_{\mathbf{p}} e^{-i\mathbf{p}\cdot\mathbf{x}}\right) \left(a_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{x}} - a^\dagger_{\mathbf{q}} e^{-i\mathbf{q}\cdot\mathbf{x}}\right) \\ &= -\int d^3\mathbf{x} \frac{d^3\mathbf{p} d^3\mathbf{q}}{(2\pi)^6} \frac{\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{q}}}}{2} \\ &\quad \times \left(a_{\mathbf{p}} a_{\mathbf{q}} e^{i(\mathbf{p}+\mathbf{q})\cdot\mathbf{x}} - a^\dagger_{\mathbf{p}} a_{\mathbf{q}} e^{i(\mathbf{q}-\mathbf{p})\cdot\mathbf{x}} - a_{\mathbf{p}} a^\dagger_{\mathbf{q}} e^{i(\mathbf{p}-\mathbf{q})\cdot\mathbf{x}} + a^\dagger_{\mathbf{p}} a^\dagger_{\mathbf{q}} e^{-i(\mathbf{p}+\mathbf{q})\cdot\mathbf{x}}\right) \\ &= -\int \frac{d^3\mathbf{p} d^3\mathbf{q}}{(2\pi)^3} \frac{\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{q}}}}{2} \\ &\quad \times \left(a_{\mathbf{p}} a_{\mathbf{q}} \delta^{(3)}(\mathbf{p}+\mathbf{q}) - a^\dagger_{\mathbf{p}} a_{\mathbf{q}} \delta^{(3)}(\mathbf{p}-\mathbf{q}) - a_{\mathbf{p}} a^\dagger_{\mathbf{q}} \delta^{(3)}(\mathbf{p}-\mathbf{q}) + a^\dagger_{\mathbf{p}} a^\dagger_{\mathbf{q}} \delta^{(3)}(\mathbf{p}+\mathbf{q})\right) \\ &= \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{\omega_{\mathbf{p}}}{2} \left[\left(a^\dagger_{\mathbf{p}} a_{\mathbf{p}} + a_{\mathbf{p}} a^\dagger_{\mathbf{p}}\right) - \left(a_{\mathbf{p}} a_{-\mathbf{p}} + a^\dagger_{\mathbf{p}} a^\dagger_{-\mathbf{p}}\right)\right] \,, \end{aligned} \tag{10}$$

where we have used

$$\int \frac{d^3\mathbf{p}}{(2\pi)^3} e^{-i\mathbf{p}\cdot\mathbf{x}} = \delta^{(3)}(\mathbf{x}) \,, \tag{11}$$

and the fact that $\omega_{\mathbf{p}} = \omega_{-\mathbf{p}}$. We repeat the calculation for the second and the third terms:

$$\int d^3\mathbf{x}\,(\nabla\phi)^2 = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{\mathbf{p}^2}{2\omega_{\mathbf{p}}}\left[\left(a^\dagger_{\mathbf{p}}a_{\mathbf{p}} + a_{\mathbf{p}}a^\dagger_{\mathbf{p}}\right) + \left(a_{\mathbf{p}}a_{-\mathbf{p}} + a^\dagger_{\mathbf{p}}a^\dagger_{-\mathbf{p}}\right)\right], \tag{12}$$

$$\int d^3\mathbf{x}\,m^2\phi^2 = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{m^2}{2\omega_{\mathbf{p}}}\left[\left(a^\dagger_{\mathbf{p}}a_{\mathbf{p}} + a_{\mathbf{p}}a^\dagger_{\mathbf{p}}\right) + \left(a_{\mathbf{p}}a_{-\mathbf{p}} + a^\dagger_{\mathbf{p}}a^\dagger_{-\mathbf{p}}\right)\right]. \tag{13}$$

Therefore, we get the Hamiltonian as

$$H = \frac{1}{4}\int \frac{d^3\mathbf{p}}{(2\pi)^3}\left[\left(-\omega_{\mathbf{p}} + \frac{\mathbf{p}^2}{\omega_{\mathbf{p}}} + \frac{m^2}{\omega_{\mathbf{p}}}\right)\left(a_{\mathbf{p}}a_{-\mathbf{p}} + a^\dagger_{\mathbf{p}}a^\dagger_{-\mathbf{p}}\right) + \left(\omega_{\mathbf{p}} + \frac{\mathbf{p}^2}{\omega_{\mathbf{p}}} + \frac{m^2}{\omega_{\mathbf{p}}}\right)\left(a^\dagger_{\mathbf{p}}a_{\mathbf{p}} + a_{\mathbf{p}}a^\dagger_{\mathbf{p}}\right)\right]. \tag{14}$$

The first term is zero because $\omega_{\mathbf{p}}^2 = \mathbf{p}^2 + m^2$. So,

$$\begin{aligned} H &= \frac{1}{2}\int \frac{d^3\mathbf{p}}{(2\pi)^3}\omega_{\mathbf{p}}\left(a^\dagger_{\mathbf{p}}a_{\mathbf{p}} + a_{\mathbf{p}}a^\dagger_{\mathbf{p}}\right) \\ &= \int \frac{d^3\mathbf{p}}{(2\pi)^3}\omega_{\mathbf{p}}\left(a^\dagger_{\mathbf{p}}a_{\mathbf{p}} + \frac{1}{2}[a_{\mathbf{p}}, a^\dagger_{\mathbf{p}}]\right) \\ &= \int \frac{d^3\mathbf{p}}{(2\pi)^3}\omega_{\mathbf{p}}\left(a^\dagger_{\mathbf{p}}a_{\mathbf{p}} + \frac{1}{2}(2\pi)^3\delta^{(3)}(\mathbf{0})\right). \end{aligned} \tag{15}$$

We indeed see the infinite sum of infinities. Let us just ignore the infinities for the moment and move on. Similar to what we did in Quantum Mechanics, we define a vacuum state $|0\rangle$ to be a state that satisfies

$$a_{\mathbf{p}}|0\rangle = 0\,, \tag{16}$$

for all $\mathbf{p}$. The ground-state energy is then given by

$$H|0\rangle = \frac{1}{2}\int d^3\mathbf{p}\,\omega_{\mathbf{p}}\delta^{(3)}(\mathbf{0})|0\rangle = \infty|0\rangle\,. \tag{17}$$

Hence, the ground-state energy is also infinity.

So, we tried to quantise the classical Klein-Gordon theory, computed the Hamiltonian and the ground-state energy, and obtained infinities. Let us now take a moment and think about those infinities more seriously. We have two kinds of infinity. The first one is the one coming from the delta function $\delta^3(\mathbf{0})$, which can be thought of as

$$(2\pi)^3\delta^{(3)}(\mathbf{p}) = \int d^3\mathbf{x}\,e^{i\mathbf{p}\cdot\mathbf{x}}\,. \tag{18}$$

When $\mathbf{p} = \mathbf{0}$, we are integrating 1 over all space. Therefore, we get the result of infinity. We get infinity because we are integrating over the whole space, and the space is infinitely big. This kind of infinity is called the infrared (or IR for short) divergence. The idea to get around this infinity is to consider the energy density, which is the energy per unit volume.

Imagine that the universe is in a box of volume V. Then, we can think of $(2\pi)^3\delta^{(3)}(\mathbf{0})$ as the volume V. This trick is known as the infrared cutoff. Then, the energy density would be given by

$$\mathcal{E}_0 = \frac{E}{V} = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{2}\omega_{\mathbf{p}}\,. \tag{19}$$

Later, we can then send V to infinity. Note that this is still infinity as $\omega_{\mathbf{p}}$ becomes infinitely large when we integrate over the momentum, which can again go up to infinity. This is the second kind of infinity we encountered.

The second infinity is the divergence at short distances or large momentum. This kind of infinity is called the ultraviolet (or UV) divergence. We (physicists) do not actually think that our current theory is valid all the way up to the infinite energy scale. In other words, at some point, the theory should be replaced or supplemented by new physics. Thinking like this, we want to cut off the integral at high momentum; we just put a bound on the momentum integral and do not go higher than that. It is important to note that the energy difference is not dependent of the cutoff. This is one way to think about this ultraviolet divergence, and there are other ways.

In the end, if we are interested in energy differences[6], we can simply throw away infinities. The Hamiltonian would then be given by

$$H = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\omega_{\mathbf{p}} a^\dagger_{\mathbf{p}} a_{\mathbf{p}}\,, \tag{20}$$

with a cutoff in the momentum integral. With this Hamiltonian, we get

$$H|0\rangle = 0\,. \tag{21}$$

If you think about it, the difference between the Hamiltonian (20) and the Hamiltonian before the removal of infinities is just a matter of ordering terms. To see the meaning of this, let us go back to Quantum Mechanics. Instead of using

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2\,, \tag{22}$$

if we were to use

$$H = \frac{1}{2}\left(\omega q - ip\right)\left(\omega q + ip\right)\,, \tag{23}$$

we would have obtained

$$H = \omega a^\dagger a\,. \tag{24}$$

[6]It is important to remember that we cannot always shift the energy, saying that only the difference matters as it is only the one that can be measured. This is not always the case; check out, *e.g.*, the cosmological constant and the Casimir effect. In this course, we will not talk much about these cases. For now, let us keep this issue in mind and move on.

Note that, classically, we can move around terms without a problem as they are commutative. In quantum theories, on the other hand, we cannot do that due to the commutation relations. In other words, a different way of writing terms leads to a different theory after quantisation.

Therefore, it is important to set some sort of standard. In fact, there is something called normal ordering: Given a product of operators,

$$\phi_1(\mathbf{x}_1)\phi_2(\mathbf{x}_2)\cdots\phi_n(\mathbf{x}_n)\,, \tag{25}$$

a normal ordering means that we put all the annihilation operators to the right of all the creation operators. The normal ordering is denoted by the symbol $:\cdots:$; for example,

$$:\phi_1(\mathbf{x}_1)\phi_2(\mathbf{x}_2)\cdots\phi_n(\mathbf{x}_n):\;. \tag{26}$$

For our case, we can thus say

$$:H := \int \frac{d^3\mathbf{p}}{(2\pi)^3}\omega_{\mathbf{p}}a^\dagger_{\mathbf{p}}a_{\mathbf{p}}\,. \tag{27}$$

The normal ordering will be employed when we quantise our theory.[7]

[7] I will drop the “$:\cdots:$” notation from time to time, but it should be clear from the context.

※ Lecture 3 (2024.03.15)

3.1 Free Real Scalar Quantum Fields – continued

Now that we have dealt with the Hamiltonian and the vacuum (the nothing), it is time to talk about particles (the something). The operators $a_{\mathbf{p}}$ and $a_{\mathbf{p}}^\dagger$ are called the annihilation and creation operators. Let's see if they still deserve the names. Note first that

$$\begin{aligned}[H, a_{\mathbf{p}}^\dagger] &= \int \frac{d^3\mathbf{q}}{(2\pi)^3}\omega_{\mathbf{q}}[a_{\mathbf{q}}^\dagger a_{\mathbf{q}}, a_{\mathbf{p}}^\dagger] \\ &= \int \frac{d^3\mathbf{q}}{(2\pi)^3}\omega_{\mathbf{q}} a_{\mathbf{q}}^\dagger (2\pi)^3 \delta^{(3)}(\mathbf{p}-\mathbf{q}) \\ &= \omega_{\mathbf{p}} a_{\mathbf{p}}^\dagger \,. \end{aligned} \tag{28}$$

Similarly, we get

$$[H, a_{\mathbf{p}}] = -\omega_{\mathbf{p}} a_{\mathbf{p}} \,. \tag{29}$$

Therefore, just like what we did for the simple harmonic oscillators in Quantum Mechanics, we can apply, say, the creation operators to construct energy eigenstates. Letting

$$|\mathbf{p}\rangle = a_{\mathbf{p}}^\dagger |0\rangle \,, \tag{30}$$

we get

$$H|\mathbf{p}\rangle = \omega_{\mathbf{p}}|\mathbf{p}\rangle \,, \tag{31}$$

where the eigenvalue is given by

$$\omega_{\mathbf{p}}^2 = \mathbf{p}^2 + m^2 \,. \tag{32}$$

From Special Relativity, we know that the energy of a particle of mass m and momentum $\mathbf{p}$ is

$$E_{\mathbf{p}}^2 = \mathbf{p}^2 + m^2 \,. \tag{33}$$

Thus, one may interpret the state $|\mathbf{p}\rangle$ as the momentum eigenstate of a particle of mass m and momentum $\mathbf{p}$, and one can identify m with the mass of the quantised particle. Let us see if this interpretation makes sense. Take the momentum operator, for example.[8] After the normal ordering, we get

$$:\mathbf{P} := -: \int d^3\mathbf{x}\, \pi(\mathbf{x})\nabla\phi(\mathbf{x}) := \int \frac{d^3\mathbf{k}}{(2\pi)^3}\mathbf{k} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \,. \tag{34}$$

[8]Recall from Lecture 1 that $P^i = \int d^3\mathbf{x}\dot{\phi}\partial^i\phi$.

Applying the momentum operator to the one-particle state $|\mathbf{p}\rangle$ gives

$$\mathbf{P}|\mathbf{p}\rangle = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\mathbf{k}a^\dagger_{\mathbf{k}}a_{\mathbf{k}}a^\dagger_{\mathbf{p}}|0\rangle = \mathbf{p}|\mathbf{p}\rangle\,. \tag{35}$$

Thus, the state has total momentum $\mathbf{p}$.

In a similar manner, we can build multi-particle states; all we have to do is to apply a bunch of the creation operators. For example, the n-particle state can be constructed by

$$|\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_n\rangle = a^\dagger_{\mathbf{p}_1}a^\dagger_{\mathbf{p}_2}\cdots a^\dagger_{\mathbf{p}_n}|0\rangle\,. \tag{36}$$

Note that, since the creation operators commute amongst themselves, $|\mathbf{p},\mathbf{q}\rangle = |\mathbf{q},\mathbf{p}\rangle$. In other words, the state is symmetric under exchange of any two particles, which means that the particles are bosons.

Our state space, or the full Hilbert space, would be given by the span of particles of the form

$$|0\rangle, a^\dagger_{\mathbf{p}_1}|0\rangle, a^\dagger_{\mathbf{p}_1}a^\dagger_{\mathbf{p}_2}|0\rangle, \cdots \tag{37}$$

This space is known as the Fock space; the Fock space is simply the sum of the n-particle Hilbert spaces. As in Quantum Mechanics, there is an operator that counts the particle number,

$$N = \int \frac{d^3\mathbf{p}}{(2\pi)^3}a^\dagger_{\mathbf{p}}a_{\mathbf{p}}\,. \tag{38}$$

It is quite easy to see that $[N, H] = 0$, which indicates that the particle number is conserved. It is important to note that we have been considering a free theory, *i.e.*, no interactions. Once we allow particles to interact, the particle number will no longer be conserved, and particles may be destroyed or created.

Exercise 3.1: Verify Eqs. (28) and (29).

Exercise 3.2: Check the second equality in Eq. (34), and verify Eq. (35).

It seems like we have successfully quantised the Klein-Gordon theory. We took a classical free real scalar field and quantised it by using the canonical quantisation. Are we done now? There is one very important aspect we forgot to talk about: Is our theory Lorentz invariant? It is not a trivial question. We started with a vacuum state $|0\rangle$, which we can normalise it as $\langle 0|0\rangle = 1$. We can reasonably assume that the vacuum state is Lorentz invariant. We then applied the creation operator to get 1-particle states $|\mathbf{p}\rangle$. They satisfy

$$\langle\mathbf{p}|\mathbf{q}\rangle = \langle 0|a_{\mathbf{p}}a^\dagger_{\mathbf{q}}|0\rangle = \langle 0|a^\dagger_{\mathbf{q}}a_{\mathbf{p}} + (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q})|0\rangle = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q})\,. \tag{39}$$

It is not so obvious whether this is Lorentz invariant or not, because we are only dealing with 3-vectors. In fact, it is not a Lorentz-invariant quantity; we will soon see why. Although it is not a Lorentz-invariant quantity, it is a scalar. So, we may come up with something like

$$|p\rangle = A_p|\mathbf{p}\rangle \tag{40}$$

so that

$$\langle p|q\rangle = (2\pi)^3 A_p^* A_q \delta^{(3)}(\mathbf{p}-\mathbf{q}) \tag{41}$$

becomes a Lorentz-invariant quantity. Notice the difference between p and $\mathbf{p}$; we think of the p ($\mathbf{p}$) in $|p\rangle$ ($|\mathbf{p}\rangle$) as the 4-vector (3-vector) p ($\mathbf{p}$). Now, we need to work out the normalisation factor A_p.

We know that $\int d^4p$ and p^2, or m^2, are Lorentz invariant. Thus,

$$\int d^4p\, \delta(p^2 - m^2) \tag{42}$$

must be a Lorentz-invariant quantity. Integrating over p_0, we see that

$$\int d^4p\, \delta(p^2 - m^2) = \int dp_0\, d^3\mathbf{p}\, \delta(p_0^2 - \mathbf{p}^2 - m^2) = \int d^3\mathbf{p} \frac{1}{2E_\mathbf{p}}\,, \tag{43}$$

where $E_\mathbf{p} \equiv \sqrt{\mathbf{p}^2 + m^2}$, and we have used

$$\int dx\, \delta(g(x)) = \sum_i \frac{1}{|g'(x_i)|}\,, \tag{44}$$

with $g(x)$ being a smooth function with simple zeros at $x = x_i$. Note that we have chosen $p_0 = +E_\mathbf{p}$ out of two branches, $p_0 = +E_\mathbf{p}$ and $p_0 = -E_\mathbf{p}$; to make this choice clear, it is sometimes useful to put the step function, $\int d^4p\, \delta(p_0^2 - \mathbf{p}^2 - m^2)\,\theta(p_0)$. Therefore, $\int d^3\mathbf{p}/(2E_\mathbf{p})$ is Lorentz invariant.

We also know that

$$1 = \int d^3\mathbf{p}\, \delta^{(3)}(\mathbf{p}-\mathbf{q}) = \int \frac{d^3\mathbf{p}}{2E_\mathbf{p}}\, 2E_\mathbf{p}\delta^{(3)}(\mathbf{p}-\mathbf{q})\,. \tag{45}$$

Since

$$\int \frac{d^3\mathbf{p}}{2E_\mathbf{p}} \tag{46}$$

and 1 are Lorentz invariant, we conclude that $2E_\mathbf{p}\delta^{(3)}(\mathbf{p}-\mathbf{q})$ is a Lorentz-invariant quantity.

Thus, our new state,

$$|p\rangle = \sqrt{2E_\mathbf{p}}|\mathbf{p}\rangle\,, \tag{47}$$

satisfies

$$\langle p|q\rangle = (2\pi)^3 (2E_\mathbf{p})\delta^{(3)}(\mathbf{p}-\mathbf{q})\,, \tag{48}$$

which is correctly Lorentz invariant. Similarly, one may define a new creation operator as

$$a^\dagger(p) = \sqrt{2E_\mathbf{p}}a_\mathbf{p}^\dagger\,, \tag{49}$$

in terms of which our field can be written as

$$\phi = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{2E_\mathbf{p}}\left[a(p)e^{i\mathbf{p}\cdot\mathbf{x}} + a^\dagger(p)e^{-i\mathbf{p}\cdot\mathbf{x}}\right]\,. \tag{50}$$

3.2 The Heisenberg Picture

Our field and its conjugate momentum are operators, and they depend only on space, while the time evolution is encoded in the states. Due to this, it was not so obvious to check whether our theory is Lorentz invariant or not. If we can instead somehow encode the time evolution in the operators themselves, the Lorentz invariance would become much clearer. Thankfully, we already know the way: the Heisenberg picture.

Let us remind ourselves what we learnt in Quantum Mechanics. The Schrödinger equation, which is the equation of motion, is given by

$$i\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle\,, \tag{51}$$

and the solution can formally be written as

$$|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle\,. \tag{52}$$

In other words, the time evolution is governed by the operator e^{-iHt}. Let's say that we are interested in the expectation value of an operator O at time t. We would then compute

$$\langle\psi(t)|O|\,\psi(t)\rangle\,. \tag{53}$$

The state at time t can be obtained by applying the time evolution operator, e^{-iHt}, to the initial state $|\psi(0)\rangle$, so we get

$$\langle\psi(t)|O|\,\psi(t)\rangle = \langle\psi(0)|e^{iHt}Oe^{-iHt}|\psi(0)\rangle\,. \tag{54}$$

Now, we see that, if we just interpret the whole situation as the operator being evolved in time, we can assign the time dependence to the operator; this is the Heisenberg picture in which the operator is time dependent,

$$O_H = e^{iHt}O_S e^{-iHt}\,. \tag{55}$$

Here, we have put the subscripts H (for the Heisenberg picture) and S (for the Schrödinger picture) to denote which picture we are in. The equation of motion is then given by

$$\frac{dO_H}{dt} = \frac{d}{dt}\left(e^{iHt}O_S e^{-iHt}\right) = iHe^{iHt}O_S e^{-iHt} + e^{iHt}O_S(-iHe^{-iHt}) = i[H,O_H]\,. \tag{56}$$

Let us try to evaluate the equation of motion for the free real scalar field ϕ (now the field depends on both time and space, $\phi(t,\mathbf{x})$):

$$\begin{aligned}\dot{\phi}(t,\mathbf{x}) &= i[H,\phi(t,\mathbf{x})]\\ &= \frac{i}{2}\int d^3\mathbf{y}\left[\pi^2(t,\mathbf{y}) + (\nabla_y\phi(t,\mathbf{y}))^2 + m^2\phi^2(t,\mathbf{y}),\phi(t,\mathbf{x})\right]\,.\end{aligned} \tag{57}$$

To proceed, we need to know the commutation relations. It is easy to show that our operators (the field and the conjugate momentum) satisfy the following equal-time commutation relations:

$$[\phi(t,\mathbf{x}),\phi(t,\mathbf{y})] = 0 = [\pi(t,\mathbf{x}),\pi(t,\mathbf{y})]\,, \quad [\phi(t,\mathbf{x}),\pi(t,\mathbf{y})] = i\delta^{(3)}(\mathbf{x}-\mathbf{y})\,. \tag{58}$$

Thus, we get

$$\begin{aligned}
\dot\phi(t,\mathbf{x}) &= i[H,\phi(t,\mathbf{x})] \\
&= \frac{i}{2}\int d^3\mathbf{y}\,\left[\pi^2(t,\mathbf{y}) + (\nabla_y\phi(t,\mathbf{y}))^2 + m^2\phi^2(t,\mathbf{y}),\phi(t,\mathbf{x})\right] \\
&= \frac{i}{2}\int d^3\mathbf{y}\,\left[\pi^2(t,\mathbf{y}),\phi(t,\mathbf{x})\right] \\
&= \frac{i}{2}\int d^3\mathbf{y}\,\left(\pi(t,\mathbf{y})\left[\pi(t,\mathbf{y}),\phi(t,\mathbf{x})\right] + \left[\pi(t,\mathbf{y}),\phi(t,\mathbf{x})\right]\pi(t,\mathbf{y})\right) \\
&= i\int d^3\mathbf{y}\,\pi(t,\mathbf{y})(-i\delta^{(3)}(\mathbf{x}-\mathbf{y})) \\
&= \pi(t,\mathbf{x})\,.
\end{aligned} \tag{59}$$

What about $\pi(t,\mathbf{x})$?

$$\begin{aligned}
\dot\pi(t,\mathbf{x}) &= i[H,\pi(t,\mathbf{x})] \\
&= \frac{i}{2}\int d^3\mathbf{y}\,\left[\pi^2(t,\mathbf{y}) + (\nabla_y\phi(t,\mathbf{y}))^2 + m^2\phi^2(t,\mathbf{y}),\pi(t,\mathbf{x})\right] \\
&= \frac{i}{2}\int d^3\mathbf{y}\,\left[(\nabla_y\phi(t,\mathbf{y}))^2 + m^2\phi^2(t,\mathbf{y}),\pi(t,\mathbf{x})\right] \\
&= \frac{i}{2}\int d^3\mathbf{y}\,\left(\left[(\nabla_y\phi(t,\mathbf{y}))^2,\pi(t,\mathbf{x})\right] + m^2\left[\phi^2(t,\mathbf{y}),\pi(t,\mathbf{x})\right]\right)\,.
\end{aligned} \tag{60}$$

The second term is easy to work out; it is given by $-m^2\phi(t,\mathbf{x})$. Let us work out the first term:

$$\begin{aligned}
&\frac{i}{2}\int d^3\mathbf{y}\,[(\nabla_y\phi(t,\mathbf{y}))^2,\pi(t,\mathbf{x})] \\
&= \frac{i}{2}\int d^3\mathbf{y}\,\left(\nabla_y\phi(t,\mathbf{y})\cdot[\nabla_y\phi(t,\mathbf{y}),\pi(t,\mathbf{x})] + [\nabla_y\phi(t,\mathbf{y}),\pi(t,\mathbf{x})]\cdot\nabla_y\phi(t,\mathbf{y})\right) \\
&= \frac{i}{2}\int d^3\mathbf{y}\,\left(\nabla_y\phi(t,\mathbf{y})\cdot\nabla_y[\phi(t,\mathbf{y}),\pi(t,\mathbf{x})] + \nabla_y[\phi(t,\mathbf{y}),\pi(t,\mathbf{x})]\cdot\nabla_y\phi(t,\mathbf{y})\right) \\
&= \frac{i}{2}\int d^3\mathbf{y}\,\left(\nabla_y\phi(t,\mathbf{y})\cdot\nabla_y(i\delta^{(3)}(\mathbf{y}-\mathbf{x})) + \nabla_y(i\delta^{(3)}(\mathbf{y}-\mathbf{x}))\cdot\nabla_y\phi(t,\mathbf{y})\right) \\
&= -\int d^3\mathbf{y}\,\nabla_y\phi(t,\mathbf{y})\cdot\nabla_y(\delta^{(3)}(\mathbf{y}-\mathbf{x})) \\
&= -\int d^3\mathbf{y}\,\nabla_y\cdot\left[\nabla_y(\phi(t,\mathbf{y}))\delta^{(3)}(\mathbf{y}-\mathbf{x})\right] + \int d^3\mathbf{y}\,\delta^{(3)}(\mathbf{y}-\mathbf{x})\nabla_y^2\phi(t,\mathbf{y}) = \nabla^2\phi(t,\mathbf{x})\,.
\end{aligned} \tag{61}$$

Thus, we get

$$\dot{\pi}(t,\mathbf{x}) = \nabla^2\phi(t,\mathbf{x}) - m^2\phi(t,\mathbf{x})\,. \tag{62}$$

Now, noting that $\dot{\pi}$ is nothing but $\ddot{\phi}$, we see that

$$\ddot{\phi}(t,\mathbf{x}) = \nabla^2\phi(t,\mathbf{x}) - m^2\phi(t,\mathbf{x})\,, \tag{63}$$

or, equivalently,

$$\partial^2\phi(t,\mathbf{x}) + m^2\phi(t,\mathbf{x}) = 0\,, \tag{64}$$

which is the usual Klein-Gordon equation, and this equation of motion is Lorentz invariant.

Exercise 3.3: Verify the equal-time commutation relations (58).

Previously, we wrote the field and momentum operators in terms of the annihilation and creation operators. Can we do the same thing when the field and momentum operators have the time dependence? We first note that

$$e^{iHt}a_{\mathbf{p}}e^{-iHt} = e^{-iE_{\mathbf{p}}t}a_{\mathbf{p}}\,, \quad e^{iHt}a^\dagger_{\mathbf{p}}e^{-iHt} = e^{+iE_{\mathbf{p}}t}a^\dagger_{\mathbf{p}}\,. \tag{65}$$

Therefore, we have

$$\begin{aligned}\phi(x) &= \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\left(a_{\mathbf{p}}e^{-iE_{\mathbf{p}}t}e^{i\mathbf{p}\cdot\mathbf{x}} + a^\dagger_{\mathbf{p}}e^{+iE_{\mathbf{p}}t}e^{-i\mathbf{p}\cdot\mathbf{x}}\right)\\ &= \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\left(a_{\mathbf{p}}e^{-ip\cdot x} + a^\dagger_{\mathbf{p}}e^{+ip\cdot x}\right)\,.\end{aligned} \tag{66}$$

Note that all the parts appearing in this expression are Lorentz invariant. Thus, finally, we are now completely Lorentz invariant. Are we, though?

Exercise 3.4: Derive Eq. (65) by noting that $e^ABe^{-A} = B+[A,B]+[A,[A,B]]/2!+[A,[A,[A,B]]]/3!+\cdots$ for operators A and B.

We are actually not completely Lorentz invariant, due to the equal-time commutation relations $[\phi(t,\mathbf{x}),\pi(t,\mathbf{y})] = i\delta^{(3)}(\mathbf{x}-\mathbf{y})$. The time and the space are treated separately here, so we are not completely Lorentz invariant. Can we do something more general? Can we say something about $[O_1(x), O_2(y)]$ for arbitrary space and time?

First of all, our theory should be causal; if two spacetime points x and y are spacelike separated, a measurement at x should not affect the measurement at y. To ensure the causality, all spacelike-separated operators should commute,

$$[O_1(x), O_2(y)] = 0 \text{ for all spacelike separations, } (x-y)^2 < 0\,. \tag{67}$$

Let us check if our theory is causal. For convenience, we define

$$\Delta(x-y) = [\phi(x), \phi(y)]\,. \tag{68}$$

We now assume that x and y are spacelike separated. Our field is Lorentz invariant, so we can go to a frame where their times are aligned, *i.e.*,

$$x = (t, \mathbf{x}) \quad \text{and} \quad y = (t, \mathbf{y})\,. \tag{69}$$

Then, we see that

$$\Delta(x-y) = [\phi(x), \phi(y)] = [\phi(t,\mathbf{x}), \phi(t,\mathbf{y})] = 0\,. \tag{70}$$

Thus, we conclude that our theory is indeed causal.

What about timelike separations? Choosing a frame where $x = (0, \mathbf{x})$ and $y = (t, \mathbf{x})$, we can obtain

$$\Delta(x-y) \sim e^{-imt} - e^{imt}\,, \tag{71}$$

which does not vanish. To see this, we can just compute the commutation $[\phi(x), \phi(y)]$:

$$\begin{aligned}
\Delta(x-y) &= \int \frac{d^3\mathbf{p}d^3\mathbf{q}}{(2\pi)^6\sqrt{4E_\mathbf{p}E_\mathbf{q}}}\Big([a_\mathbf{p}, a_\mathbf{q}]e^{i(-p\cdot x - q\cdot y)} + [a_\mathbf{p}, a_\mathbf{q}^\dagger]e^{i(-p\cdot x + q\cdot y)} \\
&\qquad\qquad + [a_\mathbf{p}^\dagger, a_\mathbf{q}]e^{i(p\cdot x - q\cdot y)} + [a_\mathbf{p}^\dagger, a_\mathbf{q}^\dagger]e^{i(p\cdot x + q\cdot y)}\Big) \\
&= \int \frac{d^3\mathbf{p}d^3\mathbf{q}}{(2\pi)^6\sqrt{4E_\mathbf{p}E_\mathbf{q}}}(2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q})\Big(e^{i(-p\cdot x + q\cdot y)} - e^{i(p\cdot x - q\cdot y)}\Big) \\
&= \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_\mathbf{p}}\Big(e^{ip\cdot(-x+y)} - e^{ip\cdot(x-y)}\Big)\,.
\end{aligned} \tag{72}$$

Note that it is Lorentz invariant.

※ Lecture 4 (2024.03.22)

4.1 Propagators

We have accomplished many things so far. We started with a classical free real scalar field theory, quantised it using the canonical quantisation, constructed the Hamiltonian and the ground state as well as multi-particle states, examined the energy eigenvalues, talked about how to deal with the infinities, made our theory Lorentz invariant, and tested the causality. We are now ready to do some fun stuffs, such as moving particles around and making them collide. To do so, we need to introduce something called propagator.

The propagator of a real scalar field ϕ is defined as

$$D(x-y) \equiv \langle 0|\phi(x)\phi(y)|0\rangle \,. \tag{73}$$

As the name suggests, the propagator tells us the amplitude for a particle to propagate from y to x; this will become much clearer later when we study interactions. It is straightforward to show that

$$\begin{aligned} D(x-y) &= \int \frac{d^3\mathbf{p}d^3\mathbf{q}}{(2\pi)^6}\frac{1}{\sqrt{4E_\mathbf{p}E_\mathbf{q}}}\langle 0|a_\mathbf{p}a_\mathbf{q}^\dagger|0\rangle e^{-ip\cdot x+iq\cdot y} \\ &= \int \frac{d^3\mathbf{p}d^3\mathbf{q}}{(2\pi)^6}\frac{1}{\sqrt{4E_\mathbf{p}E_\mathbf{q}}}\langle 0|[a_\mathbf{p},a_\mathbf{q}^\dagger]|0\rangle e^{-ip\cdot x+iq\cdot y} \\ &= \int \frac{d^3\mathbf{p}d^3\mathbf{q}}{(2\pi)^6}\frac{1}{\sqrt{4E_\mathbf{p}E_\mathbf{q}}}(2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q})e^{-ip\cdot x+iq\cdot y} \\ &= \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{2E_\mathbf{p}}e^{-ip\cdot(x-y)} \,. \end{aligned} \tag{74}$$

Compared to Eq. (72), we see that

$$\Delta(x-y) = D(x-y) - D(y-x) \,. \tag{75}$$

<u>**Interlude:**</u> Contour integral; see **Appendix – Contour Integral**.

4.1.1 The Feynman Propagator

There are many other propagators. One of the most useful propagators is the Feynman propagator:

$$\Delta_F(x-y) = \langle 0|T\phi(x)\phi(y)|0\rangle = \begin{cases} D(x-y) & \text{for} \quad x^0 > y^0 \\ D(y-x) & \text{for} \quad x^0 < y^0 \end{cases} \tag{76}$$

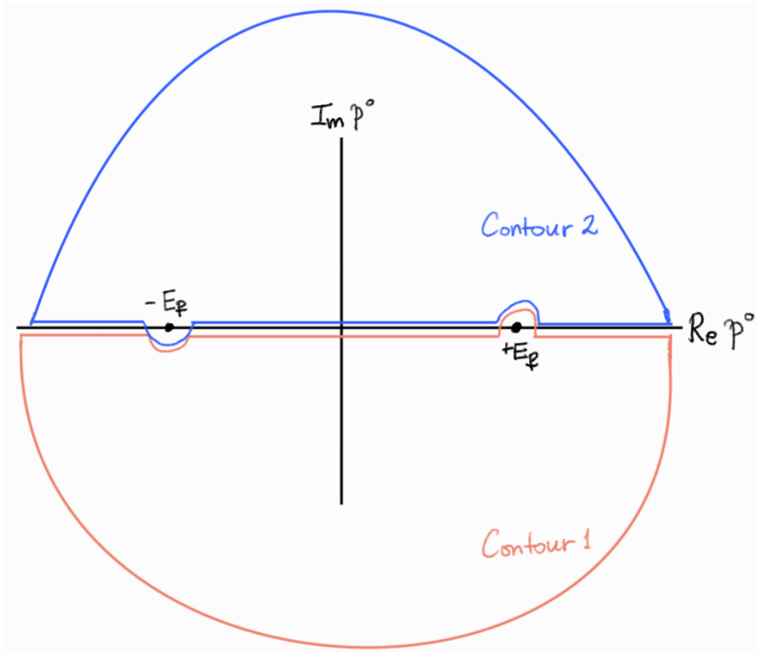


Figure 1: (Sorry about the bad drawing.) Contour integral for the Feynman propagator. For $x^0 > y^0$, we choose Contour 1. For $x^0 < y^0$, we choose Contour 2.

where T means time ordering. The time-ordering operator T places all operators evaluated at later times to the left, *e.g.*,

$$T\phi(x)\phi(y) = \begin{cases} \phi(x)\phi(y) & \text{for} \quad x^0 > y^0 \\ \phi(y)\phi(x) & \text{for} \quad x^0 < y^0 \end{cases} \tag{77}$$

We now claim that the Feynman propagator (76) can be written as

$$\Delta_F(x-y) = \int \frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2}e^{-ip\cdot(x-y)}\,. \tag{78}$$

Note that the denominator, $p^2 - m^2 = (p^0)^2 - \mathbf{p}^2 - m^2 = (p^0)^2 - E_\mathbf{p}^2$ contains a pole at $p^0 = \pm E_\mathbf{p}$. Thus, we need a prescription for avoiding these singularities in the p^0 integral. For this, we will choose the contour depicted in Fig. 1.[9]

Let us show Eq. (78). Note first that the residue of the pole at $p^0 = \pm E_\mathbf{p}$ is $\pm 1/(2E_\mathbf{p})$. For $x^0 > y^0$, we choose Contour 1 shown in Fig. 1, *i.e.*, $p^0 \to -i\infty$. Then, $e^{-ip^0(x^0-y^0)} \to 0$, and the p^0-integral will pick up the residue at the pole $p^0 = E_\mathbf{p}$. Thus, we get

$$\begin{aligned}\int_{\text{Contour 1}} \frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2}e^{-ip\cdot(x-y)} &= \int \frac{d^3\mathbf{p}}{(2\pi)^4}\frac{(-2\pi i)}{2E_\mathbf{p}}ie^{-iE_\mathbf{p}(x^0-y^0)+i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} \\ &= \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{2E_\mathbf{p}}e^{-ip\cdot(x-y)} = D(x-y)\,. \end{aligned} \tag{79}$$

[9] As you may have guessed, if we make another choice for the contours, we get another propagator.

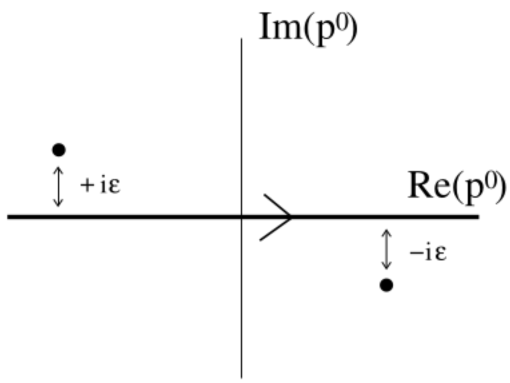


Figure 2: The $i\epsilon$ prescription for the Feynman propagator. (Figure from D. Tong's lecture note.)

For $x^0 < y^0$, we choose Contour 2 shown in Fig. 1, *i.e.*, $p^0 \to +i\infty$. Then, $e^{-ip^0(x^0-y^0)} \to 0$, and the p^0-integral will pick up the residue at the pole $p^0 = -E_{\mathbf{p}}$. Thus, we get

$$\begin{aligned}\int_{\text{Contour 2}} \frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2}e^{-ip\cdot(x-y)} &= \int \frac{d^3\mathbf{p}}{(2\pi)^4}\frac{(2\pi i)}{-2E_{\mathbf{p}}} i e^{iE_{\mathbf{p}}(x^0-y^0)+i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} \\ &= \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}} e^{-iE_{\mathbf{p}}(y^0-x^0)-i\mathbf{p}\cdot(\mathbf{y}-\mathbf{x})} \\ &= \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}} e^{-iE_{\mathbf{p}}(y^0-x^0)+i\mathbf{p}\cdot(\mathbf{y}-\mathbf{x})} \\ &= \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}} e^{-ip\cdot(y-x)} = D(y-x)\,. \end{aligned} \tag{80}$$

Here, we have flipped the sign of $\mathbf{p}$. Thus, we have shown that the Feynman propagator (76) can be written as

$$\Delta_F(x-y) = \int \frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2}e^{-ip\cdot(x-y)}\,. \tag{81}$$

Another way of presenting the Feynman propagator is

$$\Delta_F(x-y) = \int \frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2+i\epsilon}e^{-ip\cdot(x-y)}\,, \tag{82}$$

where $\epsilon > 0$ is infinitesimal. The $i\epsilon$ term has the effect of moving the pole $-E_{\mathbf{p}}$ ($+E_{\mathbf{p}}$) to the positive-imaginary (negative-imaginary) p^0 side; see Fig. 2. Thus, integrating along the real p^0 axis is equivalent to what we did before. Inserting the $i\epsilon$ term is called the "$i\epsilon$ prescription".

Note that the propagator is, in fact, Green's function of the Klein-Gordon operator:

$$\left(\partial^2 + m^2\right)\Delta_F(x-y) = -i\delta^{(4)}(x-y)\,. \tag{83}$$

Exercise 4.1: Verify Eq. (83).

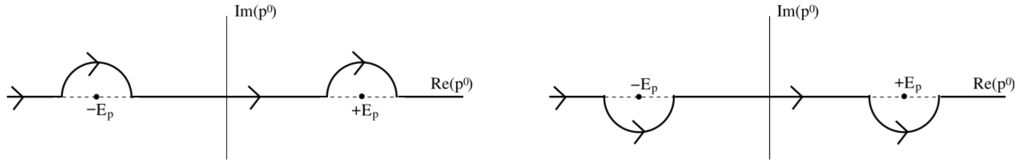


Figure 3: The retarded (left) and advanced (right) contours. (Figure from D. Tong's lecture note.)

4.1.2 Other Propagators

It is useful to consider other choices for the contour. Let us consider the contour depicted in the left panel of Fig. 3; this is called the retarded contour. This choice gives rise to the so-called retarded Green's function:

$$\Delta_R(x-y) = \begin{cases} D(x-y) - D(y-x) & x^0 > y^0 \\ 0 & y^0 > x^0 \end{cases} \tag{84}$$

Instead, if we consider the contour depicted in the right panel of Fig. 3, which is called the advanced contour, we get the so-called advanced Green's function:

$$\Delta_A(x-y) = \begin{cases} 0 & x^0 > y^0 \\ D(y-x) - D(x-y) & y^0 > x^0 \end{cases} \tag{85}$$

The retarded Green's function is useful if we know the initial field configuration, and we want to know its evolution in the presence of a source. The advanced Green's function is useful if we know the final field configuration, and we want to know its initial set-up. In this course, we will mostly be dealing with the Feynman propagator.

❊ Lecture 5 (2024.03.29)

5.1 Interacting Real Scalar Quantum Fields

So far, we have discussed free theories, especially the free real scalar field theory. We have learnt a lot, and we have studied many important things. However, it was not so fun. Although we were able to determine the exact spectrum, nothing interesting happened. We could create particle excitations, but those particles did nothing.

Now is the time to move away from free theories and start to include interaction terms. One obvious question to ask is: "How can one introduce interactions?" Let us look back at the free theories. The Lagrangian contained at most quadratic terms. In turn, the equations of motion became linear. Thus, if we want to introduce interactions, we can put some higher-order (cubic, quartic, *etc.*) terms in the Lagrangian. These extra terms can then act as the potential term in the Lagrangian.

We shall start with the simplest case: one real scalar field interacting with itself. We can write down the general Lagrangian with interaction terms as follows:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \sum_{n=3}^{\infty}\frac{\lambda_n}{n!}\phi^n\,, \tag{86}$$

where the $\lambda_n\phi^n$ terms are the interaction terms; λ_n are called the coupling constants[10]. We will regard the added interaction terms as "small" perturbations to the free theory[11]. But, what do we mean by "small" perturbations? If λ_n were dimensionless parameters, we may impose $\lambda_n \ll 1$. However, are λ_n dimensionless? Since $[S] = 0$ and $[d^4x] = -4$, and because $S = \int d^4x\mathcal{L}$, we see that

$$[\mathcal{L}] = 4\,. \tag{87}$$

From the kinetic term, using $[\partial_\mu] = 1$, we find that

$$[\phi] = 1\,. \tag{88}$$

Using $[\phi] = 1$, we get (from the mass term)

$$[m] = 1\,, \tag{89}$$

and (from the interaction terms)

$$[\lambda_n] = 4 - n\,. \tag{90}$$

[10]For the moment, let us restrict ourselves to the $\lambda_n \geq 0$ case.

[11]It is basically impossible to do anything with such a generic Lagrangian.

Thus, λ_4 is the only parameter that is dimensionless, and $\lambda_{n\neq 4}$ are not dimensionless. Thus, the statement $\lambda_n \ll 1$ would not make sense.

What we can do instead is to compare the coupling parameters with some relevant energy scale E[12]. For a given energy scale E, we can then look at $\lambda_n E^{n-4}$ as it is a dimensionless parameter; note that $[E] = 1$. In this way, the coupling parameters λ_n can be categorised as follows:

- $\underline{n = 3}$: In this case, we have $\lambda_n E^{n-4} \to \lambda_3/E$, which decreases with the energy. In other words, λ_3/E is small (large) at high (low) energy scales. Such terms are called "relevant" as they are most relevant at low energy scales.

  $\to$ Note that, in a relativistic theory, where $E > m$, we can impose $\lambda_3 \ll m$ to make the cubic term a small perturbation.

- $\underline{n = 4}$: In this case, λ_4 is a dimensionless parameter, and thus we can impose $\lambda_4 \ll 1$ to make the quartic term a small perturbation. Such terms are called "marginal".

- $\underline{n > 4}$: In this case, we have $\lambda_n E^{n-4}$ which increases with the energy. In other words, it becomes small (large) at low (high) energy scales. Such terms are called "irrelevant" as they become small at low energy scales.

  $\to$ Note that, despite the name "irrelevant", they play a crucial role in QFT, as QFT typically involves high energies. These terms may lead to the so-called non-renormalisable field theories; in this course, we do not discuss such cases.

In this course, we will deal with interaction terms that can truly be considered as small perturbations of the free field theory at all energies. As an example, let us consider the following theory (called the ϕ^4 theory):

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4\,. \tag{91}$$

In order to treat the quartic term as a small perturbation, we need to impose $\lambda \ll 1$. Another example is the so-called scalar Yukawa theory:

$$\mathcal{L} = \partial_\mu\psi^*\partial^\mu\psi + \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - M^2\psi^*\psi - \frac{1}{2}m^2\phi^2 - g\psi^*\psi\phi\,. \tag{92}$$

The first and the second terms are the kinetic terms, and the third and the fourth terms are the mass terms. The last term is the interaction term with the coupling parameter g whose mass dimension is 1, *i.e.*, $[g] = 1$. In order to treat the interaction term as a small perturbation, we impose $g \ll M, m$; another way to put it is to impose $g/M \ll 1$ and $g/m \ll 1$.

[12] For example, if we are interested in particle collisions at a collider experiment, then the energy of the colliding particles could be used as the relevant energy scale.

5.2 The Interaction Picture

Let us study interacting fields. We have added interaction terms, which are small enough to be considered as tiny perturbations, to the Lagrangian. Thus, the Hamiltonian would then be split into two parts: the free part H_0 and the interaction part H_{int}. We already discussed how we can incorporate the time dependence for the free theory by going to the Heisenberg picture. There, the operators contain the simple time evolutions. When the interaction part is a small perturbation to the well-understood Hamiltonian H_0, we can leave the time dependence for H_{int} in the states, just like in the Schrödinger picture. This viewpoint, a mixture of the Heisenberg picture and the Schrödinger picture, is called the interaction picture; we actually encountered this picture in Quantum Mechanics.

Let us quickly review the pictures again. In the Schrödinger picture, a state $|\psi(t)\rangle_S$ evolves in time, and operators O_S do not. The equation that governs the time evolution is given by the Schrödinger equation,

$$i\frac{d|\psi\rangle_S}{dt} = H|\psi\rangle_S\,. \tag{93}$$

If H is constant in time, we can write down the solution as

$$|\psi(t)\rangle_S = e^{-iHt}|\psi(0)\rangle_S\,, \tag{94}$$

or, more generally,

$$|\psi(t)\rangle_S = e^{-iH(t-t_0)}|\psi(t_0)\rangle_S = U_S(t,t_0)|\psi(t_0)\rangle_S\,, \tag{95}$$

where $U_S(t,t_0) = e^{-iH(t-t_0)}$ is the time evolution operator, which is unitary.

In the Heisenberg picture, the states are fixed, and the operators evolve in time,

$$O_H(t) = e^{iHt}O_S e^{-iHt}\,, \tag{96}$$
$$|\psi\rangle_H = e^{iHt}|\psi(t)\rangle_S\,. \tag{97}$$

The equation of motion is then given by

$$\frac{dO_H}{dt} = i[H,O_H]\,. \tag{98}$$

In the interaction picture, we do a bit of both: We send the free part H_0 to the Heisenberg picture and leave the interaction part H_{int} in the Schrödinger picture. In other words, we do

$$O_I(t) = e^{iH_0t}O_S e^{-iH_0t}\,, \tag{99}$$
$$|\psi(t)\rangle_I = e^{iH_0t}|\psi(t)\rangle_S\,. \tag{100}$$

Note that, in the interaction picture, the interaction part of the Hamiltonian $H_{\rm int}$ is given by

$$H_I \equiv H_{{\rm int},I} = e^{iH_0 t} H_{{\rm int},S} e^{-iH_0 t} \,. \tag{101}$$

Note also that, since

$$\begin{aligned} i\frac{d|\psi\rangle_S}{dt} = H_S|\psi\rangle_S \Rightarrow i\frac{d}{dt}\left(e^{-iH_0 t}|\psi\rangle_I\right) &= (H_0 + H_{\rm int})_S \, e^{-iH_0 t}|\psi\rangle_I \\ \Rightarrow H_0 e^{-iH_0 t}|\psi\rangle_I + ie^{-iH_0 t}\frac{d|\psi\rangle_I}{dt} &= H_0 e^{-iH_0 t}|\psi\rangle_I + H_{{\rm int},S} e^{-iH_0 t}|\psi\rangle_I \\ \Rightarrow ie^{-iH_0 t}\frac{d|\psi\rangle_I}{dt} &= H_{{\rm int},S} e^{-iH_0 t}|\psi\rangle_I \,, \end{aligned} \tag{102}$$

we get

$$i\frac{d|\psi\rangle_I}{dt} = e^{iH_0 t} H_{{\rm int},S} e^{-iH_0 t}|\psi\rangle_I = H_I|\psi\rangle_I \,. \tag{103}$$

Exercise 5.1: Show that $d\mathcal{O}_I/dt = i[H_0, \mathcal{O}_I]$ in the interaction picture.

Why do we consider the interaction picture? Why is the interaction picture a good picture to work in? One reason is that we already know how to deal with the free theory in the Heisenberg picture. Another reason is that considering the time evolution of states generated by the interaction part is much simpler than considering the time evolution of states generated by the whole Hamiltonian.

Once we move to the interaction picture, we need to solve Eq. (103). As we did in the Schrödinger picture, we may try to say

$$|\psi(t)\rangle_I = e^{-iH_I(t-t_0)}|\psi(t_0)\rangle_I \,. \tag{104}$$

However, this would be wrong, because H_I is now time dependent. Let us write down the formal solution as

$$|\psi(t)\rangle_I = U(t,t_0)|\psi(t_0)\rangle_I \,. \tag{105}$$

Our job is then to find $U(t,t_0)$ which is called the time-evolution operator. Note that we can say a few things about this operator right away. When $t = t_0$, we should get the same state, so $U(t_0,t_0) = 1$. Also, if we evolve from t_0 to t_1, and then from t_1 to t_2, this should be identical as evolving from t_0 to t_2. Thus, $U(t_2,t_1)U(t_1,t_0) = U(t_2,t_0)$. Furthermore, by substituting Eq. (105) into Eq. (103), we find that the time-evolution operator satisfies

$$i\frac{dU}{dt} = H_I U \,. \tag{106}$$

Can we say that the solution is

$$U(t,t_0) = \exp\left(-i\int_{t_0}^{t} H_I(t')dt'\right) ? \tag{107}$$

This would be the correct answer if H_I were a function. The problem is that H_I is an operator. The above solution is thus incorrect. We can explicitly verify that this is an incorrect solution; we can just take the time derivative. Noting that the exponential of an operator is defined as a series,

$$\exp\left(-i\int_{t_0}^{t} H_I(t')dt'\right) = 1 + \left(-i\int_{t_0}^{t} H_I(t')dt'\right) + \frac{1}{2!}\left(-i\int_{t_0}^{t} H_I(t')dt'\right)^2 + \cdots , \tag{108}$$

if we take the time derivative, we get, from the quadratic term, for example,

$$-\frac{1}{2}H_I(t)\int_{t_0}^{t} H_I(t')dt' - \frac{1}{2}\left(\int_{t_0}^{t} H_I(t')dt'\right)H_I(t)\,, \tag{109}$$

which is not what we want. The first term looks good as the terms are in the correct order, but the second term is not good. Remember that we cannot simply interchange terms here because $H_I(t')$ do not commute with $H_I(t)$ when $t' \neq t$.

What is the answer then? The answer is

$$U(t,t_0) = T\exp\left(-i\int_{t_0}^{t} H_I(t')dt'\right)\,. \tag{110}$$

This is called the Dyson formula. We can easily show that this is indeed the correct solution. Thanks to the time-ordering operator T, we can change orders of operators inside T, because the time-ordering operator will do the ordering. In order words, under T, all operators commute. Thus, we find that

$$i\frac{\partial}{\partial t}\left[T\exp\left(-i\int_{t_0}^{t} dt' H_I(t')\right)\right] = T\left[H_I(t)\exp\left(-i\int_{t_0}^{t} dt' H_I(t')\right)\right] = H_I(t)T\exp\left(-i\int_{t_0}^{t} dt' H_I(t')\right)\,. \tag{111}$$

Note that t is the latest time.

※ Lecture 6 (2024.04.12)

6.1 The Interaction Picture – continued

In the last lecture, we learnt that the solution to the equation of motion in the interaction picture,

$$i\frac{d|\psi\rangle_I}{dt} = e^{iH_0 t} H_{\text{int},S} e^{-iH_0 t}|\psi\rangle_I = H_I|\psi\rangle_I \,, \tag{112}$$

reads

$$|\psi(t)\rangle_I = U(t,t_0)|\psi(t_0)\rangle_I \,, \tag{113}$$

where the time-evolution operator is given by the Dyson formula,

$$U(t,t_0) = T\exp\left(-i\int_{t_0}^{t} H_I(t')dt'\right) . \tag{114}$$

It is generally hard to directly apply the Dyson formula as computing time-ordered exponentials is typically quite difficult in practice. However, when H_I is small so that it could be regarded as a small perturbation, we may use an expanded form of the formula. Let us examine the expanded expression of the Dyson formula:

$$\begin{aligned} U(t,t_0) &= T\left[1 + \left(-i\int_{t_0}^{t} H_I(t')dt'\right) + \frac{1}{2!}\left(-i\int_{t_0}^{t} H_I(t')dt'\right)^2 + \cdots\right] \\ &= T\left[1 - i\int_{t_0}^{t} dt' H_I(t') + \frac{(-i)^2}{2!}\int_{t_0}^{t} dt'\int_{t_0}^{t} dt'' H_I(t')H_I(t'') + \cdots\right] \\ &= T\Bigg[1 - i\int_{t_0}^{t} dt' H_I(t') \\ &\qquad + \frac{(-i)^2}{2!}\left(\int_{t_0}^{t} dt'\int_{t_0}^{t'} dt'' H_I(t')H_I(t'') + \int_{t_0}^{t} dt'\int_{t'}^{t} dt'' H_I(t')H_I(t'')\right) + \cdots\Bigg] \\ &= 1 - i\int_{t_0}^{t} dt' H_I(t') \\ &\qquad + \frac{(-i)^2}{2!}\left(\int_{t_0}^{t} dt'\int_{t_0}^{t'} dt'' H_I(t')H_I(t'') + \int_{t_0}^{t} dt'\int_{t'}^{t} dt'' H_I(t'')H_I(t')\right) + \cdots \end{aligned} \tag{115}$$

Note that

$$\int_{t_0}^{t} dt'\int_{t'}^{t} dt'' H_I(t'')H_I(t') = \int_{t_0}^{t} dt''\int_{t_0}^{t''} dt' H_I(t'')H_I(t') = \int_{t_0}^{t} dt'\int_{t_0}^{t'} dt'' H_I(t')H_I(t'') \,, \tag{116}$$

where, in the last expression, we have switched the dummy indices $t' \leftrightarrow t''$. Thus, we obtain

$$U(t,t_0) = 1 - i\int_{t_0}^{t} dt' H_I(t') + (-i)^2\int_{t_0}^{t} dt'\int_{t_0}^{t'} dt'' H_I(t')H_I(t'') + \cdots \tag{117}$$

(We may understand Eq. (117) from Fig. 4.) When the interactions are small, we can thus take the first two or three terms.

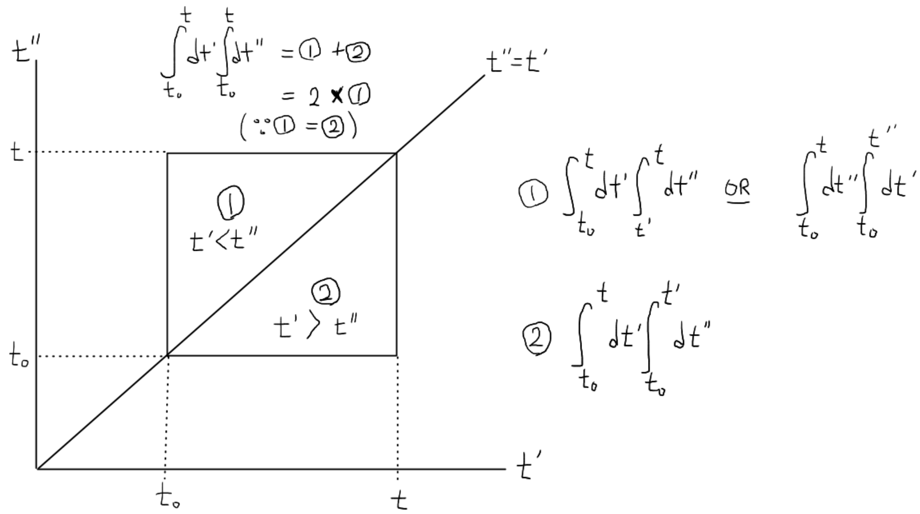


Figure 4: The third term in the expanded expression of the Dyson formula (117).

6.2 Example: Scalar Yukawa Theory

Let us take the scalar Yukawa theory as an example and try to apply what we have learnt. As we briefly mentioned in the previous lecture, the scalar Yukawa theory is described by

$$\mathcal{L} = \partial_\mu \psi^* \partial^\mu \psi + \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - M^2 \psi^* \psi - \frac{1}{2} m^2 \phi^2 - g\psi^* \psi \phi \,, \tag{118}$$

where ϕ is a real scalar field, ψ is a complex scalar field, and the last term, $g\psi^*\psi\phi$, is the interaction term.

We have not talked about a complex scalar field in detail. Now seems to be a good time to talk about it. Let us start from a classical free theory with the following Lagrangian:

$$\mathcal{L} = \partial_\mu \psi^* \partial^\mu \psi - \mu^2 \psi^* \psi \,. \tag{119}$$

We would like to quantise the theory by using the canonical quantisation. Just like the way we expressed the quantised real scalar field in terms of annihilation and creation operators,

$$\phi(\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_\mathbf{p}}} \left(a_\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{x}} + a^\dagger_\mathbf{p} e^{-i\mathbf{p}\cdot\mathbf{x}} \right) \,, \tag{120}$$

we can write down the quantised ψ as

$$\psi(\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_\mathbf{p}}} \left(b_\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{x}} + c^\dagger_\mathbf{p} e^{-i\mathbf{p}\cdot\mathbf{x}} \right) \,, \tag{121}$$

$$\psi^\dagger(\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_\mathbf{p}}} \left(b^\dagger_\mathbf{p} e^{-i\mathbf{p}\cdot\mathbf{x}} + c_\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{x}} \right) \,. \tag{122}$$

Note that, in the case of a real scalar field, we had to have $a_{\mathbf{p}}$ and $a_{\mathbf{p}}^\dagger$ in order to ensure that ϕ is real. However, in the case of ψ, there is no reason to impose such a condition. The conjugate momentum operator (recall that $\pi = \partial\mathcal{L}/\partial\dot{\psi} = \dot{\psi}^*$ classically) is given by

$$\pi(\mathbf{x}) = i\int \frac{d^3\mathbf{p}}{(2\pi)^3}\sqrt{\frac{E_{\mathbf{p}}}{2}}\left(b_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}} - c_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}}\right), \tag{123}$$

$$\pi^\dagger(\mathbf{x}) = -i\int \frac{d^3\mathbf{p}}{(2\pi)^3}\sqrt{\frac{E_{\mathbf{p}}}{2}}\left(b_{\mathbf{p}} e^{i\mathbf{p}\cdot\mathbf{x}} - c_{\mathbf{p}}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}}\right). \tag{124}$$

The commutation relations are

$$[\psi(\mathbf{x}), \pi(\mathbf{y})] = i\delta^{(3)}(\mathbf{x}-\mathbf{y}), \quad [\psi(\mathbf{x}), \pi^\dagger(\mathbf{y})] = 0, \tag{125}$$

with others related by complex conjugation; all the other commutation relations are zero. We may easily show the followings:

$$[b_{\mathbf{p}}, b_{\mathbf{q}}^\dagger] = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q}), \quad [c_{\mathbf{p}}, c_{\mathbf{q}}^\dagger] = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{q}), \tag{126}$$

with others being zero. For a complex scalar field, we have two creation operators, $b_{\mathbf{p}}^\dagger$ and $c_{\mathbf{p}}^\dagger$. They create two types of particle (with the same mass) which are interpreted as particles and anti-particles. (In the case of a real scalar field, the interpretation is that the particle is its own anti-particle.) Finally, similar to the real scalar field, we can go to the Heisenberg picture where we have $e^{ip\cdot x}$, not $e^{-i\mathbf{p}\cdot\mathbf{x}}$.

Exercise 6.1: Show Eq. (126).

Let us now get back to our example. The interaction Hamiltonian is given by

$$H_{\text{int}} = g\int d^3\mathbf{x}\,\psi^\dagger\psi\phi\,. \tag{127}$$

We have three types of particle here:

- $\phi \sim a + a^\dagger$:
  - This operator can create or destroy ϕ particles.
  - We shall call them mesons.
- $\psi \sim b + c^\dagger$:
  - This operator can create anti-particles (via $c^\dagger$) and destroy particles (via b).
  - We shall call them anti-nucleons and nucleons.
- $\psi^\dagger \sim b^\dagger + c$:

– This operator can create nucleons (via $b^\dagger$) and destroy anti-nucleons (via c).

Expanding H_{int} in terms of the creation and annihilation operators, we end up with a bunch of terms. For example, H_{int} contains $b^\dagger c^\dagger a$. This operator will destroy a meson and create a pair of nucleon and anti-nucleon; it would thus contribute to a meson decay. Of course, we also have H_{int}^2 in the Dyson formula. One term in H_{int}^2 is $b^\dagger c^\dagger acba^\dagger$. This term creates a meson by destroying a pair of nucleon and anti-nucleon, and then destroys the meson by creating a pair of nucleon and anti-nucleon.

In this section, we will try to compute the decay of a meson. To proceed, we assume that initial states and final states are eigenstates of the free theory. In other words, we take the initial state $|i\rangle$ at $t \to -\infty$ and the final states $|f\rangle$ at $t \to +\infty$ to be eigenstates of the free Hamiltonian H_0.[13] Picture the following scenario: At $t \to -\infty$, the particles are far away from each other, so they do not sense the presence of each other. As time flows, the particles approach each other, and they interact with each other. After the interaction, they depart, each going on its own way. In this sense, we may treat the initial and final states as eigenstates of the free theory. The amplitude to go from $|i\rangle$ to $|f\rangle$ is

$$\lim_{t_\pm \to \pm\infty} \langle f|U(t_+, t_-)|i\rangle \equiv \langle f|S|i\rangle \,, \tag{128}$$

where the unitary operator S is known as the S-matrix (scattering matrix).

For the example of the meson decay, we may consider the following initial and final states:

$$\begin{aligned} |i\rangle &= \sqrt{2E_{\mathbf{p}}}\, a_{\mathbf{p}}^\dagger|0\rangle \,, \\ |f\rangle &= \sqrt{4E_{\mathbf{q}_1}E_{\mathbf{q}_2}}\, b_{\mathbf{q}_1}^\dagger c_{\mathbf{q}_2}^\dagger|0\rangle \,. \end{aligned} \tag{129}$$

In other words, we are preparing an initial state of a single meson and a final state of a nucleon–anti-nucleon pair. Note that the states are relativistically normalised. The amplitude for the meson decay is then given by

$$\begin{aligned} \langle f|S|i\rangle &= \langle f|\left[1 - i\int_{-\infty}^{+\infty} dt\, H_I(t) + \cdots\right]|i\rangle \\ &= \langle f|1|i\rangle - i\langle f|\int_{-\infty}^{+\infty} dt\, H_I(t)|i\rangle + \cdots \,, \end{aligned} \tag{130}$$

where 1 should be understood as the identity; some often use $\mathbf{1}$ for the identity. The first term becomes zero as the initial state and the final state are distinct eigenstates which are orthogonal. To leading order in the coupling constant g, we get

$$\langle f|S|i\rangle = -ig\langle f|\int d^4x\, \psi^\dagger(x)\psi(x)\phi(x)|i\rangle \,. \tag{131}$$

[13] Whether this assumption is solid or not is beyond the scope of this course. In QFT-II and Advanced QFT, we will discuss more on this.

Note that the fields are in the interaction picture (or the Heisenberg picture of the free theory). We can express the fields in terms of the creation and annihilation operators,

$$\phi(x) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{k}}}}\left(a_{\mathbf{k}}e^{-ik\cdot x} + a^\dagger_{\mathbf{k}}e^{ik\cdot x}\right), \tag{132}$$

$$\psi(x) = \int \frac{d^3\mathbf{k_1}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{k_1}}}}\left(b_{\mathbf{k_1}}e^{-ik_1\cdot x} + c^\dagger_{\mathbf{k_1}}e^{ik_1\cdot x}\right), \tag{133}$$

$$\psi^\dagger(x) = \int \frac{d^3\mathbf{k_2}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{k_2}}}}\left(b^\dagger_{\mathbf{k_2}}e^{ik_2\cdot x} + c_{\mathbf{k_2}}e^{-ik_2\cdot x}\right), \tag{134}$$

and do the calculation. Before we blindly substitute these expressions, let us observe, from

$$\langle f|\psi^\dagger\psi\phi|i\rangle \sim \langle 0|cb(b^\dagger + c)(b + c^\dagger)(a + a^\dagger)a^\dagger|0\rangle\,, \tag{135}$$

that (i) the $a^\dagger a^\dagger$ term will be zero due to $\langle 0|$ (recall that a commute with b and c), and thus, we are left with $\langle f|\psi^\dagger\psi\phi|i\rangle \sim \langle 0|cb(b^\dagger + c)(b + c^\dagger)aa^\dagger|0\rangle$, and (ii) amongst $cb(b^\dagger + c)(b + c^\dagger) = cbb^\dagger b + cbb^\dagger c^\dagger + cbcb + cbcc^\dagger$, the first, third, and fourth terms will be zero due to $|0\rangle$ (recall that b and c commute). Thus, we are left with

$$\langle f|\psi^\dagger\psi\phi|i\rangle \sim \langle 0|cbb^\dagger c^\dagger aa^\dagger|0\rangle\,. \tag{136}$$

Another way to understand this is by looking at the followings:

$$\begin{aligned}
\langle f|\psi^\dagger\psi\phi|i\rangle &\sim \langle 0|cb(b^\dagger + c)(b + c^\dagger)(a + a^\dagger)a^\dagger|0\rangle \\
&\sim \left[\langle 0|cb(b^\dagger + c)(b + c^\dagger)\right]\left[(a + a^\dagger)a^\dagger|0\rangle\right] \\
&\sim \left[(b^\dagger + c)(b + c^\dagger)b^\dagger c^\dagger|0\rangle\right]^\dagger\left[(a + a^\dagger)a^\dagger|0\rangle\right] \\
&\sim \left[(b^\dagger bb^\dagger c^\dagger + b^\dagger c^\dagger b^\dagger c^\dagger + cbb^\dagger c^\dagger + cc^\dagger b^\dagger c^\dagger)|0\rangle\right]^\dagger\left[(aa^\dagger + a^\dagger a^\dagger)|0\rangle\right] \\
&\sim \left[|1b1c\rangle + |2b2c\rangle + |0\rangle + |1b1c\rangle\right]^\dagger\left[|0\rangle + |2a\rangle\right] \\
&\sim \left[\langle 1b1c| + \langle 2b2c| + \langle 0| + \langle 1b1c|\right]\left[|0\rangle + |2a\rangle\right]\,.
\end{aligned} \tag{137}$$

The only non-zero contribution comes from the product of the third term in the first square bracket (which is originated from the $cbb^\dagger c^\dagger$ term) and the first term in the second square bracket (which is originated from the $aa^\dagger$ term).

We may understand this from physics point of view. We have an initial state of one meson, and we are interested in the meson decay. Thus, we first need to kill the meson; this corresponds to the a term in ϕ. Next, we need to create a pair of nucleon and anti-nucleon; this corresponds to picking up $b^\dagger c^\dagger$ term from $\psi^\dagger\psi$.

Therefore, we only need to consider

$$\langle f|S|i\rangle = -ig\langle f|\int d^4x\, \psi^\dagger(x)\psi(x)\phi(x)|i\rangle$$

$$= -ig\langle 0|\sqrt{4E_{\mathbf{q_1}}E_{\mathbf{q_2}}}c_{\mathbf{q_2}}b_{\mathbf{q_1}}\int d^4x\int\frac{d^3\mathbf{k}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{k}}}}\int\frac{d^3\mathbf{k_1}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{k_1}}}}\int\frac{d^3\mathbf{k_2}}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{k_2}}}}$$
$$\times b^\dagger_{\mathbf{k_2}}e^{ik_2\cdot x}c^\dagger_{\mathbf{k_1}}e^{ik_1\cdot x}a_{\mathbf{k}}e^{-ik\cdot x}\sqrt{2E_{\mathbf{p}}}a^\dagger_{\mathbf{p}}|0\rangle\,. \tag{138}$$

Organising terms, we have

$$\langle f|S|i\rangle = -ig\langle 0|\int d^4x\int\frac{d^3\mathbf{k}d^3\mathbf{k_1}d^3\mathbf{k_2}}{(2\pi)^9}\frac{\sqrt{E_{\mathbf{p}}E_{\mathbf{q_1}}E_{\mathbf{q_2}}}}{\sqrt{E_{\mathbf{k}}E_{\mathbf{k_1}}E_{\mathbf{k_2}}}}e^{i(k_1+k_2)\cdot x}e^{-ik\cdot x}c_{\mathbf{q_2}}b_{\mathbf{q_1}}b^\dagger_{\mathbf{k_2}}c^\dagger_{\mathbf{k_1}}a_{\mathbf{k}}a^\dagger_{\mathbf{p}}|0\rangle\,. \tag{139}$$

Using $a_{\mathbf{k}}a^\dagger_{\mathbf{p}}|0\rangle = [a_{\mathbf{k}}, a^\dagger_{\mathbf{p}}]|0\rangle = (2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{k})|0\rangle$, we get

$$\langle f|S|i\rangle = -ig\langle 0|\int d^4x\int\frac{d^3\mathbf{k_1}d^3\mathbf{k_2}}{(2\pi)^6}\frac{\sqrt{E_{\mathbf{q_1}}E_{\mathbf{q_2}}}}{\sqrt{E_{\mathbf{k_1}}E_{\mathbf{k_2}}}}e^{i(k_1+k_2)\cdot x}e^{-ip\cdot x}c_{\mathbf{q_2}}b_{\mathbf{q_1}}b^\dagger_{\mathbf{k_2}}c^\dagger_{\mathbf{k_1}}|0\rangle\,. \tag{140}$$

Similarly, using $b_{\mathbf{q_1}}b^\dagger_{\mathbf{k_2}}|0\rangle = [b_{\mathbf{q_1}}, b^\dagger_{\mathbf{k_2}}]|0\rangle = (2\pi)^3\delta^{(3)}(\mathbf{k_2}-\mathbf{q_1})|0\rangle$ and $c_{\mathbf{q_2}}c^\dagger_{\mathbf{k_1}}|0\rangle = [c_{\mathbf{q_2}}, c^\dagger_{\mathbf{k_1}}]|0\rangle = (2\pi)^3\delta^{(3)}(\mathbf{k_1}-\mathbf{q_2})|0\rangle$, we get

$$\begin{aligned}\langle f|S|i\rangle &= -ig\langle 0|\int d^4x e^{i(q_1+q_2)\cdot x}e^{-ip\cdot x}|0\rangle\\ &= -ig(2\pi)^4\delta^{(4)}(q_1+q_2-p)\,.\end{aligned} \tag{141}$$

What the result, *i.e.*, the delta function, tells us is that the decay happens when the constraint $p = q_1 + q_2$ is satisfied.

※ Lecture 7 (2024.04.26)

7.1 Wick's Theorem

In Lecture 6, we discussed the Dyson formula and showed its application by considering the decay of a meson in the scalar Yukawa theory; the process was described by the first-order term in the interaction Hamiltonian. We can go on and consider the second order, third order, and so on. We just need to expand the fields in terms of the creation and annihilation operators. We then move them around by using the commutation relations and remove terms if annihilation operators sit right to creation operators (as they act on $|0\rangle$). It would be great if we can somehow start with a configuration where all the annihilation operators sit right to all the creation operators. Recall that it is the definition of normal ordering. Wick's theorem tells us how to go from time-ordered products to normal-ordered products.

We start with defining something called contraction. The contraction of two fields, say ϕ and ψ, is defined as

$$\contraction{}{\phi}{}{\psi}\phi\psi = T(\phi\psi) - \,:\phi\psi:\,. \tag{142}$$

Let us consider the contraction of two real scalar fields,

$$\contraction{}{\phi}{(x)}{\phi}\phi(x)\phi(y)\,. \tag{143}$$

Writing $\phi = \phi^+ + \phi^-$, where

$$\phi^+(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_\mathbf{p}}}a_\mathbf{p}e^{-ip\cdot x}\,,\quad \phi^-(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2E_\mathbf{p}}}a^\dagger_\mathbf{p}e^{ip\cdot x}\,, \tag{144}$$

we obtain, for the case of $x^0 > y^0$,

$$\begin{aligned} T\phi(x)\phi(y) &= \phi(x)\phi(y) \\ &= \left[\phi^+(x) + \phi^-(x)\right]\left[\phi^+(y) + \phi^-(y)\right] \\ &= \phi^+(x)\phi^+(y) + \phi^-(x)\phi^+(y) + \phi^+(x)\phi^-(y) + \phi^-(x)\phi^-(y) \\ &= \phi^+(x)\phi^+(y) + \phi^-(x)\phi^+(y) + [\phi^+(x), \phi^-(y)] + \phi^-(y)\phi^+(x) + \phi^-(x)\phi^-(y)\,, \end{aligned} \tag{145}$$

and

$$\begin{aligned} :\phi(x)\phi(y): &=: \phi^+(x)\phi^+(y) + \phi^-(x)\phi^+(y) + \phi^+(x)\phi^-(y) + \phi^-(x)\phi^-(y): \\ &= \phi^+(x)\phi^+(y) + \phi^-(x)\phi^+(y) + \phi^-(y)\phi^+(x) + \phi^-(x)\phi^-(y)\,. \end{aligned} \tag{146}$$

Thus, we see that

$$\contraction{}{\phi}{(x)}{\phi}\phi(x)\phi(y) = [\phi^+(x), \phi^-(y)]\,. \tag{147}$$

Recall that

$$D(x-y) \equiv \langle 0|\phi(x)\phi(y)|0\rangle = \langle 0|\phi^+(x)\phi^-(y)|0\rangle = \langle 0|[\phi^+(x), \phi^-(y)]|0\rangle\,. \tag{148}$$

Therefore, we find that

$$\wick{\c{\phi}(x)\c{\phi}(y)} = D(x-y)\,. \tag{149}$$

We can repeat the calculation for the case of $x^0 < y^0$; we get

$$\wick{\c{\phi}(x)\c{\phi}(y)} = D(y-x)\,. \tag{150}$$

Finally, recalling that

$$\Delta_F(x-y) = \begin{cases} D(x-y) & \text{for} \quad x^0 > y^0 \\ D(y-x) & \text{for} \quad x^0 < y^0 \end{cases}, \tag{151}$$

we conclude that

$$\wick{\c{\phi}(x)\c{\phi}(y)} = \Delta_F(x-y)\,. \tag{152}$$

Note that, although $\phi(x)\phi(y)$, $T\phi(x)\phi(y)$, and $:\phi(x)\phi(y):$ are operators, the contraction gives us a c-number function. As a consequence, we can move around the contracted ones without the worry of commutations.

Similarly, we have, for complex scalar fields,

$$\wick{\c{\psi}(x)\c{\psi}^\dagger(y)} = \Delta_F(x-y)\,, \quad \wick{\c{\psi}^\dagger(x)\c{\psi}(y)} = \Delta_F(y-x)\,, \quad \wick{\c{\psi}(x)\c{\psi}(y)} = 0 = \wick{\c{\psi}^\dagger(x)\c{\psi}^\dagger(y)}\,, \tag{153}$$

where $\Delta_F(x-y) = \langle 0|T\psi(x)\psi^\dagger(y)|0\rangle$ for the complex scalar field.

Exercise 7.1: Show Eq. (150).

Exercise 7.2: Show Eq. (153).

Wick's theorem states that, for any collection of fields, we have

$$T(\phi_1\cdots\phi_n) =: \phi_1\cdots\phi_n : + : \text{all possible contractions} :\,, \tag{154}$$

where $\phi_1 = \phi(x_1)$, $\phi_2 = \phi(x_2)$, *etc.* For example,

$$\begin{aligned} T(\phi_1\phi_2\phi_3\phi_4) =&: \phi_1\phi_2\phi_3\phi_4 : + \wick{\c{\phi}_1\c{\phi}_2} : \phi_3\phi_4 : + \wick{\c{\phi}_1\c{\phi}_3} : \phi_2\phi_4 : + \wick{\c{\phi}_1\c{\phi}_4} : \phi_2\phi_3 : \\ &+ \wick{\c{\phi}_2\c{\phi}_3} : \phi_1\phi_4 : + \wick{\c{\phi}_2\c{\phi}_4} : \phi_1\phi_3 : + \wick{\c{\phi}_3\c{\phi}_4} : \phi_1\phi_2 : \\ &+ \wick{\c1{\phi}_1\c1{\phi}_2\c2{\phi}_3\c2{\phi}_4} + \wick{\c1{\phi}_1\c2{\phi}_2\c1{\phi}_3\c2{\phi}_4} + \wick{\c1{\phi}_1\c2{\phi}_2\c2{\phi}_3\c1{\phi}_4}\,. \end{aligned} \tag{155}$$

We have already proven Wick's theorem for $n = 2$. Let us suppose that Wick's theorem is true for $\phi_2 \cdots \phi_n$. Let us now add ϕ_1. Consider the case where $x_1^0 > x_m^0$ where $m = 2, 3, \cdots, n$. Then,

$$T(\phi_1\phi_2\cdots\phi_n) = (\phi_1^+ + \phi_1^-)(: \phi_2\cdots\phi_n : + : \text{all possible contractions} :)\,. \tag{156}$$

The ϕ_1^- term is fine as it is already normal ordered. The ϕ_1^+ term needs to be moved, passing all the ϕ_m^- operators. Every time we pass ϕ_m^-, we will get a contraction $\overline{\phi_1\phi_m}$. This is how Wick's theorem could be proved.

7.2 Example: Nucleon Scattering

As an example, let's consider the nucleon-nucleon scattering in the scalar Yukawa theory. The initial and final states are

$$|i\rangle = \sqrt{2E_{\mathbf{p}_1}}\sqrt{2E_{\mathbf{p}_2}}b^\dagger_{\mathbf{p}_1}b^\dagger_{\mathbf{p}_2}|0\rangle \equiv |p_1, p_2\rangle\,, \tag{157}$$

$$|f\rangle = \sqrt{2E_{\mathbf{p}'_1}}\sqrt{2E_{\mathbf{p}'_2}}b^\dagger_{\mathbf{p}'_1}b^\dagger_{\mathbf{p}'_2}|0\rangle \equiv |p'_1, p'_2\rangle\,. \tag{158}$$

The scattering amplitude is given by

$$\begin{aligned}\langle f|S|i\rangle &= \langle f|T\exp\left(-i\int_{-\infty}^{+\infty}H_I(t)dt\right)|i\rangle \\ &= \langle f|1|i\rangle - i\langle f|\int_{-\infty}^{+\infty}H_I(t)dt|i\rangle - \frac{1}{2}\langle f|T\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}H_I(t_1)H_I(t_2)dt_1dt_2|i\rangle + \cdots \\ &= \text{no scattering} + O(g)\text{ term} + O(g^2)\text{ term} + \cdots\,.\end{aligned} \tag{159}$$

We are not interested in the no-scattering case. The $O(g)$ term becomes zero because the $\phi \sim a + a^\dagger$ in $H_I \sim \psi^\dagger\psi\phi$ will kill $|0\rangle$ and $\langle 0|$. Thus, the leading-order term is the $O(g^2)$ term:

$$\begin{aligned}O(g^2)\text{ term} &= -\frac{1}{2}\langle f|T\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}H_I(t_1)H_I(t_2)dt_1dt_2|i\rangle \\ &= -\frac{g^2}{2}\int d^4x_1d^4x_2\langle f|T\psi^\dagger(x_1)\psi(x_1)\phi(x_1)\psi^\dagger(x_2)\psi(x_2)\phi(x_2)|i\rangle\,.\end{aligned} \tag{160}$$

We can use Wick's theorem to convert the time-ordered product into a sum of normal-ordered products. Amongst many terms, we have

$$\overline{\phi(x_1)\phi(x_2)} : \psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2) : \tag{161}$$

which will contribute to the scattering process; note that $: \psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2) :$ contains $b^\dagger b^\dagger bb$.

Exercise 7.3: Convince yourself that Eq. (161) is the only relevant term.

Thus, we only need to compute

$$\langle f| :\psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2): |i\rangle \,. \tag{162}$$

It is given by

$$\begin{aligned}\langle f| :\psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2): |i\rangle = \int \frac{d^3\mathbf{k_1} d^3\mathbf{k_2} d^3\mathbf{k_3} d^3\mathbf{k_4}}{(2\pi)^{12}} \frac{\sqrt{16E_{\mathbf{p}_1}E_{\mathbf{p}_2}E_{\mathbf{p}'_1}E_{\mathbf{p}'_2}}}{\sqrt{16E_{\mathbf{k}_1}E_{\mathbf{k}_2}E_{\mathbf{k}_3}E_{\mathbf{k}_4}}} \\ \times \langle 0|b_{\mathbf{p}'_2}b_{\mathbf{p}'_1}b^\dagger_{\mathbf{k}_1}b^\dagger_{\mathbf{k}_2}b_{\mathbf{k}_3}b_{\mathbf{k}_4}b^\dagger_{\mathbf{p}_1}b^\dagger_{\mathbf{p}_2}|0\rangle \\ \times e^{i(k_1\cdot x_1 + k_2\cdot x_2 - k_3\cdot x_1 - k_4\cdot x_2)} \,.\end{aligned} \tag{163}$$

Noting that

$$\begin{aligned}
&\langle 0|b_{\mathbf{p}'_2}b_{\mathbf{p}'_1}b^\dagger_{\mathbf{k}_1}b^\dagger_{\mathbf{k}_2}b_{\mathbf{k}_3}b_{\mathbf{k}_4}b^\dagger_{\mathbf{p}_1}b^\dagger_{\mathbf{p}_2}|0\rangle \\
&= \langle 0|b_{\mathbf{p}'_2}b_{\mathbf{p}'_1}b^\dagger_{\mathbf{k}_1}b^\dagger_{\mathbf{k}_2}b_{\mathbf{k}_3}[b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_1}]b^\dagger_{\mathbf{p}_2}|0\rangle + \langle 0|b_{\mathbf{p}'_2}b_{\mathbf{p}'_1}b^\dagger_{\mathbf{k}_1}b^\dagger_{\mathbf{k}_2}b_{\mathbf{k}_3}b^\dagger_{\mathbf{p}_1}b_{\mathbf{k}_4}b^\dagger_{\mathbf{p}_2}|0\rangle \\
&= [b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_1}]\langle 0|b_{\mathbf{p}'_2}b_{\mathbf{p}'_1}b^\dagger_{\mathbf{k}_1}b^\dagger_{\mathbf{k}_2}b_{\mathbf{k}_3}b^\dagger_{\mathbf{p}_2}|0\rangle + \langle 0|b_{\mathbf{p}'_2}b_{\mathbf{p}'_1}b^\dagger_{\mathbf{k}_1}b^\dagger_{\mathbf{k}_2}b_{\mathbf{k}_3}b^\dagger_{\mathbf{p}_1}[b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_2}]|0\rangle \\
&= [b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_1}]\langle 0|b_{\mathbf{p}'_2}b_{\mathbf{p}'_1}b^\dagger_{\mathbf{k}_1}b^\dagger_{\mathbf{k}_2}[b_{\mathbf{k}_3}, b^\dagger_{\mathbf{p}_2}]|0\rangle + [b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_2}]\langle 0|b_{\mathbf{p}'_2}b_{\mathbf{p}'_1}b^\dagger_{\mathbf{k}_1}b^\dagger_{\mathbf{k}_2}[b_{\mathbf{k}_3}, b^\dagger_{\mathbf{p}_1}]|0\rangle \\
&= [b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_1}][b_{\mathbf{k}_3}, b^\dagger_{\mathbf{p}_2}]\langle 0|b_{\mathbf{p}'_2}[b_{\mathbf{p}'_1}, b^\dagger_{\mathbf{k}_1}]b^\dagger_{\mathbf{k}_2}|0\rangle + [b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_1}][b_{\mathbf{k}_3}, b^\dagger_{\mathbf{p}_2}]\langle 0|b_{\mathbf{p}'_2}b^\dagger_{\mathbf{k}_1}b_{\mathbf{p}'_1}b^\dagger_{\mathbf{k}_2}|0\rangle \\
&\quad + [b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_2}][b_{\mathbf{k}_3}, b^\dagger_{\mathbf{p}_1}]\langle 0|b_{\mathbf{p}'_2}[b_{\mathbf{p}'_1}, b^\dagger_{\mathbf{k}_1}]b^\dagger_{\mathbf{k}_2}|0\rangle + [b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_2}][b_{\mathbf{k}_3}, b^\dagger_{\mathbf{p}_1}]\langle 0|b_{\mathbf{p}'_2}b^\dagger_{\mathbf{k}_1}b_{\mathbf{p}'_1}b^\dagger_{\mathbf{k}_2}|0\rangle \\
&= [b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_1}][b_{\mathbf{k}_3}, b^\dagger_{\mathbf{p}_2}][b_{\mathbf{p}'_1}, b^\dagger_{\mathbf{k}_1}][b_{\mathbf{p}'_2}, b^\dagger_{\mathbf{k}_2}] + [b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_1}][b_{\mathbf{k}_3}, b^\dagger_{\mathbf{p}_2}]\langle 0|[b_{\mathbf{p}'_2}, b^\dagger_{\mathbf{k}_1}]b_{\mathbf{p}'_1}b^\dagger_{\mathbf{k}_2}|0\rangle \\
&\quad + [b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_2}][b_{\mathbf{k}_3}, b^\dagger_{\mathbf{p}_1}][b_{\mathbf{p}'_1}, b^\dagger_{\mathbf{k}_1}][b_{\mathbf{p}'_2}, b^\dagger_{\mathbf{k}_2}] + [b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_2}][b_{\mathbf{k}_3}, b^\dagger_{\mathbf{p}_1}]\langle 0|[b_{\mathbf{p}'_2}, b^\dagger_{\mathbf{k}_1}]b_{\mathbf{p}'_1}b^\dagger_{\mathbf{k}_2}|0\rangle \\
&= [b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_1}][b_{\mathbf{k}_3}, b^\dagger_{\mathbf{p}_2}][b_{\mathbf{p}'_1}, b^\dagger_{\mathbf{k}_1}][b_{\mathbf{p}'_2}, b^\dagger_{\mathbf{k}_2}] + [b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_1}][b_{\mathbf{k}_3}, b^\dagger_{\mathbf{p}_2}][b_{\mathbf{p}'_2}, b^\dagger_{\mathbf{k}_1}][b_{\mathbf{p}'_1}, b^\dagger_{\mathbf{k}_2}] \\
&\quad + [b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_2}][b_{\mathbf{k}_3}, b^\dagger_{\mathbf{p}_1}][b_{\mathbf{p}'_1}, b^\dagger_{\mathbf{k}_1}][b_{\mathbf{p}'_2}, b^\dagger_{\mathbf{k}_2}] + [b_{\mathbf{k}_4}, b^\dagger_{\mathbf{p}_2}][b_{\mathbf{k}_3}, b^\dagger_{\mathbf{p}_1}][b_{\mathbf{p}'_2}, b^\dagger_{\mathbf{k}_1}][b_{\mathbf{p}'_1}, b^\dagger_{\mathbf{k}_2}] \\
&= (2\pi)^{12}\delta^{(3)}(\mathbf{k_4}-\mathbf{p_1})\delta^{(3)}(\mathbf{k_3}-\mathbf{p_2})\delta^{(3)}(\mathbf{k_1}-\mathbf{p}'_1)\delta^{(3)}(\mathbf{k_2}-\mathbf{p}'_2) \\
&\quad + (2\pi)^{12}\delta^{(3)}(\mathbf{k_4}-\mathbf{p_1})\delta^{(3)}(\mathbf{k_3}-\mathbf{p_2})\delta^{(3)}(\mathbf{k_1}-\mathbf{p}'_2)\delta^{(3)}(\mathbf{k_2}-\mathbf{p}'_1) \\
&\quad + (2\pi)^{12}\delta^{(3)}(\mathbf{k_4}-\mathbf{p_2})\delta^{(3)}(\mathbf{k_3}-\mathbf{p_1})\delta^{(3)}(\mathbf{k_1}-\mathbf{p}'_1)\delta^{(3)}(\mathbf{k_2}-\mathbf{p}'_2) \\
&\quad + (2\pi)^{12}\delta^{(3)}(\mathbf{k_4}-\mathbf{p_2})\delta^{(3)}(\mathbf{k_3}-\mathbf{p_1})\delta^{(3)}(\mathbf{k_1}-\mathbf{p}'_2)\delta^{(3)}(\mathbf{k_2}-\mathbf{p}'_1) \,,
\end{aligned} \tag{164}$$

we get

$$\begin{aligned}
\langle f| :\psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2): |i\rangle &= e^{i(p'_1\cdot x_1 + p'_2\cdot x_2 - p_2\cdot x_1 - p_1\cdot x_2)} + e^{i(p'_2\cdot x_1 + p'_1\cdot x_2 - p_2\cdot x_1 - p_1\cdot x_2)} \\
&\quad + e^{i(p'_1\cdot x_1 + p'_2\cdot x_2 - p_1\cdot x_1 - p_2\cdot x_2)} + e^{i(p'_2\cdot x_1 + p'_1\cdot x_2 - p_1\cdot x_1 - p_2\cdot x_2)} \\
&= e^{i(p'_1-p_2)\cdot x_1}e^{i(p'_2-p_1)\cdot x_2} + e^{i(p'_2-p_2)\cdot x_1}e^{i(p'_1-p_1)\cdot x_2} \\
&\quad + e^{i(p'_1-p_1)\cdot x_1}e^{i(p'_2-p_2)\cdot x_2} + e^{i(p'_2-p_1)\cdot x_1}e^{i(p'_1-p_2)\cdot x_2} \,.
\end{aligned} \tag{165}$$

Another way to arrive at Eq. (165) is by noting the fact that, for relativistically normalised states,

$$\langle 0|\psi(x)|p\rangle = e^{-ip\cdot x} \,. \tag{166}$$

So, we get,

$$
\begin{aligned}
&\langle p_1', p_2'| : \psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2) : |p_1, p_2\rangle \\
&= \langle p_1', p_2'|\psi^\dagger(x_1)\psi^\dagger(x_2)|0\rangle\langle 0|\psi(x_1)\psi(x_2)|p_1, p_2\rangle \\
&= \left(e^{ip_1'\cdot x_1 + ip_2'\cdot x_2} + e^{ip_1'\cdot x_2 + ip_2'\cdot x_1}\right)\left(e^{-ip_1\cdot x_1 - ip_2\cdot x_2} + e^{-ip_1\cdot x_2 - ip_2\cdot x_1}\right) \\
&= e^{i(p_1'-p_2)\cdot x_1}e^{i(p_2'-p_1)\cdot x_2} + e^{i(p_2'-p_2)\cdot x_1}e^{i(p_1'-p_1)\cdot x_2} \\
&\quad + e^{i(p_1'-p_1)\cdot x_1}e^{i(p_2'-p_2)\cdot x_2} + e^{i(p_2'-p_1)\cdot x_1}e^{i(p_1'-p_2)\cdot x_2}\,.
\end{aligned}
\tag{167}
$$

Exercise 7.4: Show Eq. (166).

Thus,

$$
\begin{aligned}
O(g^2)\text{ term} &= -\frac{g^2}{2}\int d^4x_1 d^4x_2\Big[e^{i(p_1'-p_2)\cdot x_1}e^{i(p_2'-p_1)\cdot x_2} + e^{i(p_2'-p_2)\cdot x_1}e^{i(p_1'-p_1)\cdot x_2} \\
&\quad + e^{i(p_1'-p_1)\cdot x_1}e^{i(p_2'-p_2)\cdot x_2} + e^{i(p_2'-p_1)\cdot x_1}e^{i(p_1'-p_2)\cdot x_2}\Big]\int\frac{d^4k}{(2\pi)^4}\frac{i}{k^2-m^2+i\epsilon}e^{ik\cdot(x_1-x_2)} \\
&= -g^2\int\frac{d^4k}{(2\pi)^4}\int d^4x_1 d^4x_2\,\frac{i}{k^2-m^2+i\epsilon} \\
&\quad\times\left[e^{i(p_1'-p_1)\cdot x_1}e^{i(p_2'-p_2)\cdot x_2} + e^{i(p_2'-p_1)\cdot x_1}e^{i(p_1'-p_2)\cdot x_2}\right]e^{ik\cdot(x_1-x_2)} \\
&= -g^2\int\frac{d^4k}{(2\pi)^4}\int d^4x_1 d^4x_2\,\frac{i}{k^2-m^2+i\epsilon} \\
&\quad\times\left[+e^{i(p_1'-p_1+k)\cdot x_1}e^{i(p_2'-p_2-k)\cdot x_2} + e^{i(p_2'-p_1+k)\cdot x_1}e^{i(p_1'-p_2-k)\cdot x_2}\right] \\
&= -g^2\int\frac{d^4k}{(2\pi)^4}\,\frac{i}{k^2-m^2+i\epsilon} \\
&\quad\times(2\pi)^8\left[\delta^{(4)}\left(p_1'-p_1+k\right)\delta^{(4)}\left(p_2'-p_2-k\right) + \delta^{(4)}\left(p_2'-p_1+k\right)\delta^{(4)}\left(p_1'-p_2-k\right)\right] \\
&= -g^2(2\pi)^4\delta^{(4)}\left(p_1+p_2-p_1'-p_2'\right) \\
&\quad\times\left[\frac{i}{(p_1-p_1')^2-m^2+i\epsilon} + \frac{i}{(p_1-p_2')^2-m^2+i\epsilon}\right]\,.
\end{aligned}
\tag{168}
$$

We again see a delta function.

Thanks to Wick's theorem, we were able to do the computation in a simpler manner. However, the computation is still rather tedious. There is a much better way; the Feynman diagrams.

7.3 Feynman Diagrams

As we are not interested in no-scattering processes, what we really want to compute is the object $\langle f|S-1|i\rangle$. The perturbative expansion resulted in multiple terms. Each term can be represented pictorially as a diagram. The rules, called the Feynman rules, are as follows (note that we are considering the scalar Yukawa theory):

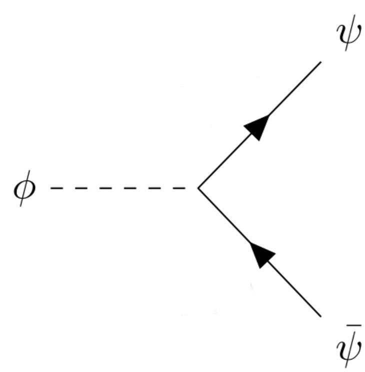
- The initial state $|i\rangle$ and the final state $|f\rangle$ can be depicted as an external line.
  - We will choose dashed lines for mesons (or ϕ-particles):

  - We will choose solid lines for nucleons (or $\psi/\bar{\psi}$-particles; $\bar{\psi}$ is to denote the anti-particle):

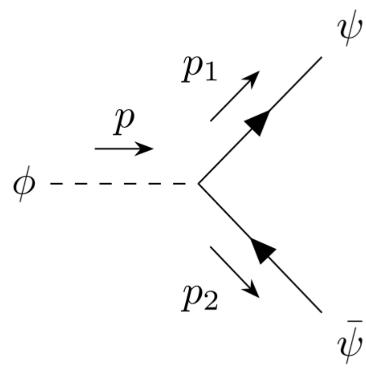
- The charge of the particles can be represented as an arrow.
  - We will choose an incoming arrow in the initial state for ψ:

  - We will choose an outgoing arrow in the initial state for $\bar{\psi}$:

  - For the final state, we reverse the convention; outgoing arrow in the final state for ψ and incoming arrow in the final state for $\bar{\psi}$.
- We join the external lines together with trivalent vertices:



  - This represents the interaction $\psi^\dagger\psi\phi$ with the coupling constant g.
- We then assign a directed momentum to each line:
  - For the internal lines, we assign dummy momenta.



- To each vertex, we write down a factor of

$$(-ig)(2\pi)^4\delta^{(4)}\left(\sum_i p_i\right) \tag{169}$$

where $\sum_i p_i$ is the sum of all momenta flowing into the vertex.

- For each internal dashed line, we write down a factor of

$$\int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2 - m^2 + i\epsilon}, \tag{170}$$

where we have assumed that momentum k is assigned to the ϕ particle. For the internal solid lines, we replace the mass.

❊ Lecture 8 (2024.05.10)

8.1 Feynman Diagrams – continued

Let us recap the Feynman rules for the scalar Yukawa theory:

- The initial state $|i\rangle$ and the final state $|f\rangle$ can be depicted as an external line.
  - We choose dashed lines for mesons (or ϕ-particles).
  - We choose solid lines for nucleons (or $\psi/\bar{\psi}$-particles).
- The charge of the particles can be represented as an arrow.
  - We choose an incoming arrow in the initial state for ψ.
  - We choose an outgoing arrow in the initial state for $\bar{\psi}$.
  - For the final state, we reverse the convention.
- We join the external lines together with trivalent vertices (for the interaction $\psi^\dagger\psi\phi$).
- We then assign a directed momentum to each line.
  - For the internal lines, we assign dummy momenta.
- To each vertex, we write down a factor of

$$(-ig)(2\pi)^4\delta^{(4)}\left(\sum_i p_i\right) \tag{171}$$

  where $\sum_i p_i$ is the sum of all momenta flowing into the vertex.
- For each internal dashed line, we write down a factor of

$$\int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2 - m^2 + i\epsilon}\,, \tag{172}$$

  where we have assumed that momentum k is assigned to the ϕ particle. For the internal solid lines, we replace the mass.

Note that the delta function is the manifestation of the 4-momentum conservation. As this factor is common to all S-matrix elements, we may define the amplitude $\mathcal{A}_{fi}$ as

$$\langle f|S - 1|i\rangle = i\mathcal{A}_{fi}(2\pi)^4\delta^{(4)}(p_F - p_I)\,, \tag{173}$$

where p_I (p_F) is the sum of the initial (final) 4-momenta. The factor of i in front is just a convention. We can then re-write the Feynman rules for the amplitude $i\mathcal{A}_{fi}$ itself as follows:

- Draw possible diagrams with appropriate external legs.
- Impose 4-momentum conservation at each vertex.
- Write down a factor of $(-ig)$ at each vertex.
- Write down propagator for each internal line.
- Integrate over the internal momentum.

8.1.1 Example: Nucleon Scattering

Let us consider the nucleon–nucleon scattering process again. We can draw two diagrams as shown in Fig. 5. For the left diagram, we see that $k = p_1 - p_1'$ (or $k = p_2' - p_2$), while for the right

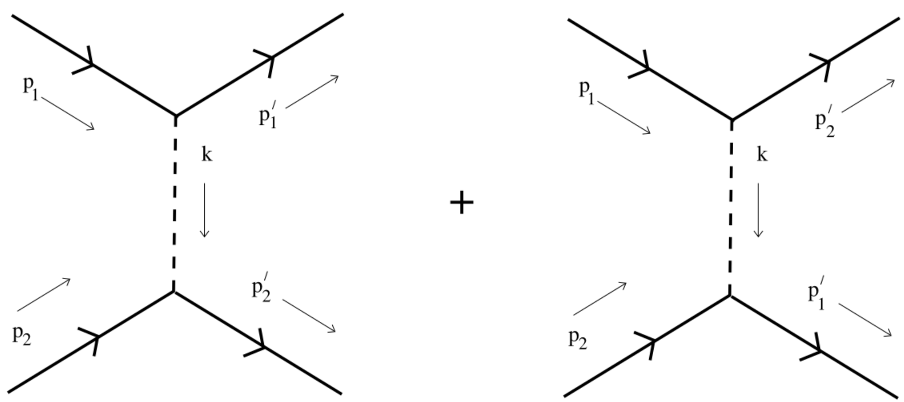


Figure 5: Feynman diagrams for nucleon–nucleon scattering

one, we have $k = p_1 - p_2'$ (or $k = p_1' - p_2$). Thus, applying the Feynman rules, we obtain

$$i\mathcal{A}_{fi} = (-ig)^2 \left[\frac{i}{(p_1 - p_1')^2 - m^2 + i\epsilon} + \frac{i}{(p_1 - p_2')^2 - m^2 + i\epsilon} \right] . \tag{174}$$

This is exactly the same as the result we obtained earlier. Note that the meson, whose momentum is k, does not satisfy the usual energy dispersion relation as $k^2 \neq m^2$[14]; this type of particles is called a virtual particle and is said to be off-shell. On the other hand, the external legs, namely the nucleon in this case, satisfy the usual energy dispersion relation; they are on-shell.

Exercise 8.1: Draw Feynman diagrams for nucleon–anti-nucleon scattering into meson–meson and compute the amplitude.

[14]One way to see this is by considering the centre-of-mass frame where $k = (0, \mathbf{p_1} - \mathbf{p}_1')$. We see that $k^2 \leq 0$.

8.1.2 Example: ϕ^4 Theory

For the ϕ^4 theory, the interaction Hamiltonian is

$$H_{\text{int}} = \frac{\lambda}{4!}\phi^4 \,. \tag{175}$$

In this case, instead of the trivalent vertex, we have a 4-point vertex as depicted in Fig. 6. This

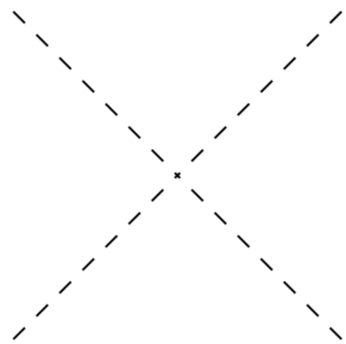


Figure 6: 4-point vertex in the ϕ^4 theory

vertex is associated with a factor of $(-i\lambda)$. One may ask what happened to the factor of $1/4!$. To understand the absence of the $1/4!$ factor, let us consider the $\phi\phi \to \phi\phi$ scattering. At $O(\lambda)$, we have

$$-i\frac{\lambda}{4!}\langle p_1', p_2'| : \phi(x)\phi(x)\phi(x)\phi(x) : |p_1, p_2\rangle \,. \tag{176}$$

There are 4! ways of doing the contraction. This cancels the $1/4!$ factor.[15]

8.2 Amplitude to Observables

The amplitude $\mathcal{A}_{fi}$ can be related to measurable quantities, such as cross sections and decay rates. In this course, we simply present expressions for the decay rate and cross section without derivations.

Let us first consider a single particle $|i\rangle$ of momentum p_I decaying into some number of particles $|f\rangle$ with momentum p_i; we denote the total final momentum by $p_F = \sum_i p_i$. The decay probability per unit time is given by

$$\Gamma = \frac{1}{2m} \sum_{\text{final states}} \int |\mathcal{A}_{fi}|^2 \, d\Pi \,, \tag{177}$$

[15] Sometimes, extra factors, known as symmetry factors, are attached to Feynman diagrams.

where

$$d\Pi = (2\pi)^4 \delta^{(4)}(p_F - p_I) \prod_{\text{final states}} \frac{d^3 p_i}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}_i}} \,. \tag{178}$$

The object Γ is called the decay rate or width of the particle. The inverse is equal to the half-life, $\tau = \Gamma^{-1}$.

Next, consider two particles[16], fired towards each other. Sometimes, they will collide and bounce off each other. Sometimes, they will miss and pass through. The fraction of the time they collide is called the cross section σ. In fact, we can calculate a more refined quantity known as the differential cross section $d\sigma$; it is the probability for a scattering process to occur in the solid angle. Imagine that we are sitting in the centre-of-mass frame of the collision. Let us denote the energies of the incoming particles by E_1 and E_2. The differential cross section is then given by

$$d\sigma = \frac{1}{4E_1 E_2} \frac{1}{|\mathbf{v}_1 - \mathbf{v}_2|} |\mathcal{A}_{fi}|^2 \, d\Pi \,, \tag{179}$$

where $\mathbf{v}_i$ is the 3-velocity.

8.3 Spinors

We have covered a lot about scalar fields. Quantisation of scalar fields led to spin-0 particles. Spin-0, or scalar, particles are, however, not the only particles in the universe. In fact, most particles in the universe have spin. In QFT, these particles can be represented as quantisations of fields which transform non-trivially under the Lorentz group.

In general, one may consider vector-valued fields, ϕ^a. For a Lorentz transformation Λ, the field would have a corresponding transformation; let us call it $D(\Lambda)$. The transformation rule would be written as

$$x \longrightarrow \Lambda x \,, \quad \phi \longrightarrow D(\Lambda)\phi \,, \tag{180}$$

or, in components,

$$x^\mu \longrightarrow \Lambda^\mu{}_\nu x^\nu \,, \quad \phi^a \longrightarrow D(\Lambda)^a{}_b \phi^b \,. \tag{181}$$

We can immediately say a few things about $D(\Lambda)$. For the Lorentz transformation $\Lambda = 1$, *i.e.*, if we do nothing, then the field should stay the same, *i.e.*, $D(1) = 1$. Next, let us consider applying two consecutive Lorentz transformations, say Λ_1 and Λ_2. Then, the field would transform by $D(\Lambda_1)$ and $D(\Lambda_2)$. However, we know that applying Λ_1 and Λ_2 is same as applying $\Lambda_3 = \Lambda_2\Lambda_1$

[16] In reality, we collide beams, or collections, of particles.

once. Under the transformation Λ_3, the field would transform by $D(\Lambda_3) = D(\Lambda_2\Lambda_1)$. We thus need

$$D(\Lambda_2)D(\Lambda_1) = D(\Lambda_2\Lambda_1)\,. \tag{182}$$

This is called a representation of the Lorentz group. The Lorentz group, $O(1,3)$, is a group of all Lorentz transformations,

$$O(1,3) = \{\Lambda \in M_{4\times 4} \ : \ \Lambda^T \eta \Lambda = \eta\}\,, \tag{183}$$

where $\eta_{\mu\nu}$ is the Minkowski metric. A representation of the Lorentz group is a vector space V (which is called the representation space) and a map $D(\Lambda) \ : \ V \longrightarrow V$ for each $\Lambda \in O(1,3)$ such that

$$D(1) = 1\,, \quad D(\Lambda_2)D(\Lambda_1) = D(\Lambda_2\Lambda_1)\,, \tag{184}$$

for any $\Lambda_1, \Lambda_2 \in O(1,3)$.

Exercise 8.2: Show that the Lorentz transformations $\Lambda^T\eta\Lambda = \eta$ truly leave x^2 invariant.

Different fields have different representations. The question is then: How do we find different representations? As a first step, let us consider infinitesimal Lorentz transformations,

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \epsilon\omega^\mu{}_\nu\,, \tag{185}$$

where ϵ is infinitesimal. As Λ is a Lorentz transformation, it should satisfy $\Lambda^T\eta\Lambda = \eta$ or, in components,

$$(\Lambda^T)_\mu{}^\rho \eta_{\rho\sigma}\Lambda^\sigma{}_\nu = \eta_{\mu\nu} \Longleftrightarrow \Lambda^\rho{}_\mu\Lambda^\sigma{}_\nu\eta_{\rho\sigma} = \eta_{\mu\nu}\,. \tag{186}$$

Equivalently, we may say, since $\Lambda\eta\Lambda^T = \eta$,

$$\Lambda^\mu{}_\rho\eta^{\rho\sigma}(\Lambda^T)_\sigma{}^\nu = \eta^{\mu\nu} \Longleftrightarrow \Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma\eta^{\rho\sigma} = \eta^{\mu\nu}\,. \tag{187}$$

For the infinitesimal transformation (185), the condition requires that

$$\omega^{\mu\nu} + \omega^{\nu\mu} = 0\,, \tag{188}$$

i.e., ω needs to be anti-symmetric. Thus, an infinitesimal Lorentz transformation is an anti-symmetric matrix. This is known as the Lie algebra of the Lorentz group,

$$\mathfrak{o}(1,3) = \{\omega \in M_{4\times 4} \ : \ \omega^{\mu\nu} + \omega^{\nu\mu} = 0\}\,. \tag{189}$$

Exercise 8.3: Show Eq. (188).

An infinitesimal Lorentz transformation is an anti-symmetric matrix. A 4×4 anti-symmetric matrix has 6 independent components. These correspond to the 6 transformations of the Lorentz group, namely 3 rotations and 3 boosts. It proves to be useful to introduce a basis of these six 4×4 anti-symmetric matrices, $(\mathcal{M}^A)^{\mu\nu}$ with $A = 1, 2, \cdots, 6$. If fact, it is better to use a pair of anti-symmetric indices $[\rho\sigma]$ $(\rho, \sigma = 0, \cdots, 3)$ instead of A. Note that the anti-symmetry on ρ and σ leads to 6 different matrices. It is important not to forget that the matrices are also anti-symmetric on μ and ν. We can then write down a basis of six 4×4 anti-symmetric matrices as

$$(\mathcal{M}^{\rho\sigma})^{\mu\nu} = \eta^{\rho\mu}\eta^{\sigma\nu} - \eta^{\sigma\mu}\eta^{\rho\nu}\,, \tag{190}$$

where μ and ν are for the 4×4 matrix, and ρ and σ are for the basis element. It is easy to show that

$$(\mathcal{M}^{\rho\sigma})^{\mu}{}_{\nu} = \eta^{\rho\mu}\delta^{\sigma}{}_{\nu} - \eta^{\sigma\mu}\delta^{\rho}{}_{\nu}\,, \tag{191}$$

which would be more useful if we want to do any practical calculation with these matrices, such as multiplications, as we need to contract indices. Notice that $(\mathcal{M}^{\rho\sigma})^{\mu}{}_{\nu}$, with one index lowered, are no longer necessarily anti-symmetric.

Now that we have a basis, we can express any $\omega^{\mu}{}_{\nu}$ as a linear combination of $\mathcal{M}^{\rho\sigma}$:

$$\omega^{\mu}{}_{\nu} = \frac{1}{2}\Omega_{\rho\sigma}(\mathcal{M}^{\rho\sigma})^{\mu}{}_{\nu}\,, \tag{192}$$

where $\Omega_{\rho\sigma} = -\Omega_{\sigma\rho}$ are 6 numbers. The 6 basis matrices $\mathcal{M}^{\rho\sigma}$ are called the generators of the Lorentz transformations, and they obey

$$[\mathcal{M}^{\rho\sigma}, \mathcal{M}^{\mu\nu}] = \eta^{\sigma\mu}\mathcal{M}^{\rho\nu} - \eta^{\rho\mu}\mathcal{M}^{\sigma\nu} + \eta^{\rho\nu}\mathcal{M}^{\sigma\mu} - \eta^{\sigma\nu}\mathcal{M}^{\rho\mu}\,, \tag{193}$$

which is called the Lorentz Lie algebra relations. A finite Lorentz transformation can then be expressed as

$$\Lambda = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma}\right)\,. \tag{194}$$

Exercise 8.4: Show Eq. (193).

※ Lecture 9 (2024.05.17)

9.1 The Clifford Algebra

Last time, we saw that an infinitesimal Lorentz transformation is an anti-symmetric matrix. We discussed the generators of the Lorentz transformations $\mathcal{M}^{\rho\sigma}$ which satisfy the Lorentz Lie algebra relations.

There is in fact a useful way of finding a representation of the Lorentz algebra. To talk about this, we first need to talk about the Clifford algebra:

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}1\,, \tag{195}$$

where γ^μ ($\mu = 0, \cdots, 3$) are a set of 4 matrices, and 1 is the unit matrix. Note that, for $\mu \neq \nu$,

$$\gamma^\mu\gamma^\nu = -\gamma^\nu\gamma^\mu\,, \tag{196}$$

while, for $\mu = \nu$,

$$(\gamma^0)^2 = 1\,, \quad (\gamma^i)^2 = -1 \;\; (i = 1, 2, 3)\,. \tag{197}$$

We would like to find a collection of matrices γ^μ that satisfies the Clifford algebra; it is called a representation. One representation is

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \tag{198}$$

where σ^i are the Pauli matrices,

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{199}$$

which satisfy

$$\{\sigma^i, \sigma^j\} = 2\delta^{ij}1\,. \tag{200}$$

This representation is called the chiral (or Weyl) representation. Note that the gamma matrices here are 4×4 matrices, and the Pauli matrices are 2×2 matrices; sometimes people use $\{\mathbb{1}, \mathbb{0}\}$ or $\{\mathbf{1}, \mathbf{0}\}$ to stress that these are unit and null (2×2 in this case) matrices.

The chiral representation is not the only possibility. For any invertible matrix V, a new representation can be constructed by $\gamma^\mu \longrightarrow V\gamma^\mu V^{-1}$. It turns out that any 4-dimensional representation of the Clifford algebra comes from the chiral representation.

The reason we are talking about the Clifford algebra is because every representation of the Clifford algebra gives us a representation of the Lorentz algebra. To be specific, the matrices

$$S^{\rho\sigma} = \frac{1}{4}[\gamma^\rho, \gamma^\sigma]\,, \tag{201}$$

where γ^ρ is a representation of the Clifford algebra, define a representation of the Lorentz algebra (193). It means that the matrices $S^{\rho\sigma}$ satisfy

$$[S^{\mu\nu}, S^{\rho\sigma}] = \eta^{\nu\rho}S^{\mu\sigma} - \eta^{\mu\rho}S^{\nu\sigma} + \eta^{\mu\sigma}S^{\nu\rho} - \eta^{\nu\sigma}S^{\mu\rho}\,. \tag{202}$$

Before we prove this, note first that $S^{\rho\sigma} = 0$ for $\rho = \sigma$ and $\gamma^\rho\gamma^\sigma/2$ for $\rho \neq \sigma$. Thus, we can write

$$S^{\rho\sigma} = \frac{1}{2}\gamma^\rho\gamma^\sigma - \frac{1}{2}\eta^{\rho\sigma}\,. \tag{203}$$

Note also that

$$[S^{\rho\sigma}, \gamma^\mu] = \gamma^\rho\eta^{\sigma\mu} - \gamma^\sigma\eta^{\mu\rho}\,. \tag{204}$$

Exercise 9.1: Prove Eq. (204).

We are now ready to prove Eq. (202). It is straightforward to show that

$$\begin{aligned}
[S^{\mu\nu}, S^{\rho\sigma}] &= \frac{1}{2}[S^{\mu\nu}, \gamma^\rho\gamma^\sigma - \eta^{\rho\sigma}] \\
&= \frac{1}{2}[S^{\mu\nu}, \gamma^\rho\gamma^\sigma] \\
&= \frac{1}{2}\left([S^{\mu\nu}, \gamma^\rho]\gamma^\sigma + \gamma^\rho[S^{\mu\nu}, \gamma^\sigma]\right) \\
&= \frac{1}{2}\left(\gamma^\mu\eta^{\nu\rho}\gamma^\sigma - \gamma^\nu\eta^{\mu\rho}\gamma^\sigma + \gamma^\rho\gamma^\mu\eta^{\nu\sigma} - \gamma^\rho\gamma^\nu\eta^{\mu\sigma}\right) \\
&= \frac{1}{2}\left(\gamma^\mu\gamma^\sigma\eta^{\nu\rho} - \gamma^\nu\gamma^\sigma\eta^{\mu\rho} + \gamma^\rho\gamma^\mu\eta^{\nu\sigma} - \gamma^\rho\gamma^\nu\eta^{\mu\sigma}\right)\,,
\end{aligned} \tag{205}$$

where we have used Eq. (204). Now, since $\gamma^\mu\gamma^\sigma = 2S^{\mu\sigma} + \eta^{\mu\sigma}$, we have

$$\begin{aligned}
[S^{\mu\nu}, S^{\rho\sigma}] &= \frac{1}{2}\left[(2S^{\mu\sigma} + \eta^{\mu\sigma})\eta^{\nu\rho} - (2S^{\nu\sigma} + \eta^{\nu\sigma})\eta^{\mu\rho} + (2S^{\rho\mu} + \eta^{\rho\mu})\eta^{\nu\sigma} - (2S^{\rho\nu} + \eta^{\rho\nu})\eta^{\mu\sigma}\right] \\
&= S^{\mu\sigma}\eta^{\nu\rho} - S^{\nu\sigma}\eta^{\mu\rho} + S^{\rho\mu}\eta^{\nu\sigma} - S^{\rho\nu}\eta^{\mu\sigma}\,.
\end{aligned} \tag{206}$$

As $S^{\mu\nu}$ are 4×4 matrices, let us introduce α and β which run from 1 to 4 to denote their rows and columns; thus, we have $(S^{\mu\nu})^\alpha{}_\beta$. The matrices $(S^{\mu\nu})^\alpha{}_\beta$ would act on a field, $\psi^\alpha(x)$. The field $\psi^\alpha(x)$ transforms under Lorentz transformations as

$$\psi^\alpha(x) \to S(\Lambda)^\alpha{}_\beta\psi^\beta(\Lambda^{-1}x)\,, \tag{207}$$

where

$$\Lambda = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma}\right), \quad S(\Lambda) = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right). \tag{208}$$

Note that the same 6 numbers $\Omega_{\rho\sigma}$ are used in both Λ and $S(\Lambda)$ to ensure that we are applying the same Lorentz transformation on x and ψ. We call the field $\psi^\alpha(x)$ the Dirac spinor field.

One important property of the spinor representation $S(\Lambda)$ is that it is not necessarily unitary. In fact, there are no finite-dimensional unitary representations of the Lorentz group. As

$$S(\Lambda) = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma}\right), \tag{209}$$

the representation would be unitary if $S^{\rho\sigma}$ are anti-hermitian, *i.e.*, $(S^{\rho\sigma})^\dagger = -S^{\rho\sigma}$. However, we have

$$(S^{\rho\sigma})^\dagger = -\frac{1}{4}[(\gamma^\rho)^\dagger, (\gamma^\sigma)^\dagger], \tag{210}$$

which would be anti-hermitian if either all γ^μ are hermitian or all γ^μ are anti-hermitian. The problem is that we do not have this because

$$(\gamma^0)^2 = 1, \quad (\gamma^i)^2 = -1. \tag{211}$$

We can explicitly see that, in the chiral representation, $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^i)^\dagger = -\gamma^i$. In general, there is no way to choose γ^μ in such a way that $S^{\mu\nu}$ become anti-hermitian.

9.2 The Dirac Spinor Field

Now that we have a new field, the Dirac spinor ψ, we would like to construct its Lorentz-invariant action. The equation of motion would then follow from the action.

We need a Lorentz scalar to construct an action. As a first candidate, we may consider

$$\psi^\dagger\psi, \tag{212}$$

with the spinor indices summed over. Here, $\psi^\dagger = (\psi^*)^T$ is the adjoint. Let us see how this transforms under the Lorentz transformations. Since

$$\psi \to S(\Lambda)\psi, \quad \psi^\dagger \to \psi^\dagger S(\Lambda)^\dagger, \tag{213}$$

we have

$$\psi^\dagger\psi \to \psi^\dagger S(\Lambda)^\dagger S(\Lambda)\psi. \tag{214}$$

In general, $\psi^\dagger S(\Lambda)^\dagger S(\Lambda)\psi \neq \psi^\dagger\psi$ since the representation is not unitary as we discussed. Therefore, we conclude that $\psi^\dagger\psi$ is not the correct answer.

The reason we failed is due to the fact that $S(\Lambda)$ is not unitary, in general. Earlier, we saw that, for a representation of the Clifford algebra like the chiral representation, we have $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^i)^\dagger = -\gamma^i$. Then, we may write

$$\gamma^0\gamma^\mu\gamma^0 = (\gamma^\mu)^\dagger\,. \tag{215}$$

In turn, we find

$$(S^{\mu\nu})^\dagger = -\frac{1}{4}[(\gamma^\mu)^\dagger, (\gamma^\nu)^\dagger] = -\frac{1}{4}\gamma^0[\gamma^\mu, \gamma^\nu]\gamma^0 = -\gamma^0 S^{\mu\nu}\gamma^0\,. \tag{216}$$

Thus,

$$S(\Lambda)^\dagger = \exp\left(\frac{1}{2}\Omega_{\rho\sigma}(S^{\rho\sigma})^\dagger\right) = \gamma^0 S(\Lambda)^{-1}\gamma^0\,. \tag{217}$$

Writing in this way, the transformation of our first attempt $\psi^\dagger\psi$ looks like

$$\psi^\dagger\psi \to \psi^\dagger\gamma^0 S(\Lambda)^{-1}\gamma^0 S(\Lambda)\psi\,. \tag{218}$$

From this expression, we may guess that the quantity $\psi^\dagger\gamma^0\psi$ is a Lorentz scalar. Let us see if this is true. Under a Lorentz transformation, we have

$$\begin{aligned}\psi^\dagger\gamma^0\psi &\to \psi^\dagger S(\Lambda)^\dagger\gamma^0 S(\Lambda)\psi \\ &= \psi^\dagger\gamma^0 S(\Lambda)^{-1}\gamma^0\gamma^0 S(\Lambda)\psi \\ &= \psi^\dagger\gamma^0 S(\Lambda)^{-1} S(\Lambda)\psi \\ &= \psi^\dagger\gamma^0\psi\,. \end{aligned} \tag{219}$$

Indeed, we obtain the correct transformation for a Lorentz scalar. The object $\psi^\dagger\gamma^0$ is called the Dirac adjoint and denoted by $\bar\psi$.

Exercise 9.2: Check Eq. (215).

With the Dirac spinor ψ and its adjoint $\bar\psi$, we can construct several Lorentz-covariant objects other than the scalar $\bar\psi\psi$. For example, $\bar\psi\gamma^\mu\psi$ is a Lorentz vector, *i.e.*,

$$\bar\psi\gamma^\mu\psi \to \Lambda^\mu{}_\nu\bar\psi\gamma^\nu\psi\,. \tag{220}$$

It indicates that we can treat the index μ on γ^μ as a true vector index. In particular, we can construct Lorentz scalars by contracting it with other Lorentz indices. To prove Eq. (220), we need to show

$$S(\Lambda)^{-1}\gamma^\mu S(\Lambda) = \Lambda^\mu{}_\nu\gamma^\nu\,, \tag{221}$$

because

$$\bar\psi\gamma^\mu\psi = \psi^\dagger\gamma^0\gamma^\mu\psi \to \psi^\dagger S(\Lambda)^\dagger\gamma^0\gamma^\mu S(\Lambda)\psi$$

$$
\begin{aligned}
&= \psi^\dagger \gamma^0 S(\Lambda)^{-1} \gamma^0 \gamma^0 \gamma^\mu S(\Lambda) \psi \\
&= \psi^\dagger \gamma^0 S(\Lambda)^{-1} \gamma^\mu S(\Lambda) \psi \\
&= \bar{\psi} S(\Lambda)^{-1} \gamma^\mu S(\Lambda) \psi \,. \qquad (222)
\end{aligned}
$$

To show Eq. (221), let us consider an infinitesimal transformation,

$$
\Lambda \approx 1 + \frac{1}{2}\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma} \,, \quad S(\Lambda) \approx 1 + \frac{1}{2}\Omega_{\rho\sigma} S^{\rho\sigma} \,. \tag{223}
$$

Then, Eq. (221) is same as

$$
\begin{aligned}
&\left(1 - \frac{1}{2}\Omega_{\rho\sigma} S^{\rho\sigma}\right) \gamma^\mu \left(1 + \frac{1}{2}\Omega_{\rho\sigma} S^{\rho\sigma}\right) = \left(1 + \frac{1}{2}\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma}\right)^\mu{}_\nu \gamma^\nu \\
\Leftrightarrow\, &\gamma^\mu \Omega_{\rho\sigma} S^{\rho\sigma} - \Omega_{\rho\sigma} S^{\rho\sigma} \gamma^\mu = \Omega_{\rho\sigma} (\mathcal{M}^{\rho\sigma})^\mu{}_\nu \gamma^\nu \\
\Leftrightarrow\, &\gamma^\mu S^{\rho\sigma} - S^{\rho\sigma} \gamma^\mu = (\mathcal{M}^{\rho\sigma})^\mu{}_\nu \gamma^\nu \\
\Leftrightarrow\, &- [S^{\rho\sigma}, \gamma^\mu] = (\mathcal{M}^{\rho\sigma})^\mu{}_\nu \gamma^\nu \,. \qquad (224)
\end{aligned}
$$

Thus, if we show Eq. (224), we prove Eq. (220). Using

$$
(\mathcal{M}^{\rho\sigma})^\mu{}_\nu = \eta^{\rho\mu} \delta^\sigma_\nu - \eta^{\sigma\mu} \delta^\rho_\nu \,, \tag{225}
$$

we find

$$
(\mathcal{M}^{\rho\sigma})^\mu{}_\nu \gamma^\nu = \eta^{\rho\mu} \gamma^\sigma - \eta^{\sigma\mu} \gamma^\rho \,, \tag{226}
$$

and thus, Eq. (224) becomes

$$
-[S^{\rho\sigma}, \gamma^\mu] = \eta^{\rho\mu} \gamma^\sigma - \eta^{\sigma\mu} \gamma^\rho \,, \tag{227}
$$

which is nothing but Eq. (204). Therefore, we conclude that $\bar{\psi}\gamma^\mu\psi$ is indeed a Lorentz vector.

In a similar manner, one can show that $\bar{\psi}\gamma^\mu\gamma^\nu\psi$ is a Lorentz tensor. One can see that the symmetric part, which is proportional to $\eta^{\mu\nu}\bar{\psi}\psi$, and the anti-symmetric part, which is proportional to $\bar{\psi}S^{\mu\nu}\psi$, are both Lorentz tensors.

Exercise 9.3: Show that $\bar{\psi}\gamma^\mu\gamma^\nu\psi$ is a Lorentz tensor.

❊ Lecture 10 (2024.05.24)

10.1 The Dirac Action and Equation

Now that we know how to build Lorentz-covariant objects, we can write down a Lorentz-invariant action from these objects. We choose

$$S = \int d^4x\, \bar{\psi}(x) \left(i\gamma^\mu \partial_\mu - m\right) \psi(x) \,. \tag{228}$$

This is the Dirac action.

Exercise 10.1: Can you tell why there is a factor of i in the Dirac action?

We will see later that the theory describes particles and anti-particles of mass m and spin $1/2$ after quantisation. Note that, unlike the scalar field Lagrangian which is of the second order, the Dirac Lagrangian is of the first order. Furthermore, the mass appears as m rather than m^2.

From the Dirac action (228), we can obtain the equation of motion for the Dirac field. Varying the action with respect to $\bar{\psi}$ gives

$$\left(i\gamma^\mu \partial_\mu - m\right) \psi = 0 \,. \tag{229}$$

This is the Dirac equation. Varying the action with respect to ψ gives the conjugate equation,

$$i\partial_\mu \bar{\psi}\gamma^\mu + m\bar{\psi} = 0 \,. \tag{230}$$

Exercise 10.2: Check Eqs. (229) and (230).

Exercise 10.3: Check that the Dirac equation is Lorentz invariant.

We will see a lot of γ^μ-contractions. Thus, it is useful to introduce the following notation:

$$\not{A} \equiv \gamma^\mu A_\mu \,. \tag{231}$$

This is called the slash. With this notation, the Dirac equation reads

$$\left(i\not{\partial} - m\right) \psi = 0 \,. \tag{232}$$

Note that the γ^μ matrices are not diagonal. Due to this fact, different components of the Dirac spinor get mixed up. Interestingly, it turns out that each component satisfies the Klein-Gordon equation. To see this, let us multiply the Dirac equation by $(i\not{\partial} + m)$:

$$0 = (i\gamma^\nu\partial_\nu + m)(i\gamma^\mu\partial_\mu - m)\psi = -(\gamma^\nu\gamma^\mu\partial_\nu\partial_\mu + m^2)\psi \,. \tag{233}$$

Using

$$\gamma^\nu\gamma^\mu\partial_\nu\partial_\mu = \frac{1}{2}\{\gamma^\mu, \gamma^\nu\}\partial_\mu\partial_\nu = \partial_\mu\partial^\mu \,, \tag{234}$$

we find

$$(\partial^2 + m^2)\psi = 0\,. \tag{235}$$

This final equation has no γ^μ matrices, and nothing mixes up different components; each component of ψ satisfies the Klein-Gordon equation. In this sense, it is often said that the Dirac equation is the "square root" of the Klein-Gordon equation.

10.2 Chiral/Weyl Spinors

For the γ^μ matrices, we discussed one explicit form, namely the chiral representation,

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \tag{236}$$

with σ^i being the Pauli matrices. Using this explicit form, we can work out the explicit form of the $S^{\mu\nu} = \gamma^\mu\gamma^\nu/2 - \eta^{\mu\nu}/2$ matrices. Let us first look at S^{ij} $(i \neq j)$:

$$S^{ij} = \frac{1}{2}\begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -\sigma^i\sigma^j & 0 \\ 0 & -\sigma^i\sigma^j \end{pmatrix} = -\frac{i}{2}\sum_k \epsilon^{ijk}\begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}. \tag{237}$$

If we write $\Omega_{ij} = -\epsilon_{ijk}\varphi^k$, then

$$S(\Lambda) = \exp\left(\frac{1}{2}\Omega_{\mu\nu}S^{\mu\nu}\right) = \begin{pmatrix} e^{i\boldsymbol{\varphi}\cdot\boldsymbol{\sigma}/2} & 0 \\ 0 & e^{i\boldsymbol{\varphi}\cdot\boldsymbol{\sigma}/2} \end{pmatrix}. \tag{238}$$

This represents the spinor rotation. It is important to note that it is different from the usual vector rotation done by Λ. To see that they are indeed different, let us consider a rotation by 2π about the z-axis, which is achieved by $\boldsymbol{\varphi} = (0, 0, 2\pi)$. Then, the spinor rotation matrix becomes

$$S(\Lambda) = \begin{pmatrix} e^{i\pi\sigma^3} & 0 \\ 0 & e^{i\pi\sigma^3} \end{pmatrix} = -1\,. \tag{239}$$

In other words, under a 2π rotation,

$$\psi^\alpha \to -\psi^\alpha\,. \tag{240}$$

This is clearly different from the usual vector rotation by Λ in which case we would get $\Lambda = 1$.

For the boosts, we find

$$S^{0i} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -\sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}. \tag{241}$$

If we write $\Omega_{i0} = \chi_i$, then

$$S(\Lambda) = \begin{pmatrix} e^{\chi\cdot\sigma/2} & 0 \\ 0 & e^{-\chi\cdot\sigma/2} \end{pmatrix} . \tag{242}$$

Note that, in the chiral representation, the spinor rotation and boost transformations are block diagonal. It indicates that the Dirac spinor representation is reducible. We may decompose this into two irreducible representations. Motivated by this, we write, in the chiral representation,

$$\psi = \begin{pmatrix} u_+ \\ u_- \end{pmatrix} , \tag{243}$$

where $u_\pm$ are two-component spinors and called Weyl (or chiral) spinors. Under rotations, they transform as

$$u_\pm \longrightarrow e^{i\varphi\cdot\sigma/2} u_\pm \, , \tag{244}$$

while, under boosts,

$$u_\pm \longrightarrow e^{\pm\chi\cdot\sigma/2} u_\pm \, . \tag{245}$$

We can re-write the Dirac Lagrangian and the Dirac equation in terms of the Weyl spinors. The Dirac Lagrangian can be decomposed into

$$\begin{aligned}
\mathcal{L} = \bar\psi(i\partial\!\!\!/ - m)\psi &= \begin{pmatrix} u_+^\dagger & u_-^\dagger \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \left[\begin{pmatrix} 0 & i\partial_0 \\ i\partial_0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & i\sigma^i\partial_i \\ -i\sigma^i\partial_i & 0 \end{pmatrix} - \begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix} \right] \begin{pmatrix} u_+ \\ u_- \end{pmatrix} \\
&= \begin{pmatrix} u_-^\dagger & u_+^\dagger \end{pmatrix} \begin{pmatrix} -m & i\partial_0 + i\sigma^i\partial_i \\ i\partial_0 - i\sigma^i\partial_i & -m \end{pmatrix} \begin{pmatrix} u_+ \\ u_- \end{pmatrix} \\
&= \begin{pmatrix} u_-^\dagger & u_+^\dagger \end{pmatrix} \begin{pmatrix} -mu_+ + i(\partial_0 + \sigma^i\partial_i)u_- \\ i(\partial_0 - \sigma^i\partial_i)u_+ - mu_- \end{pmatrix} \\
&= -mu_-^\dagger u_+ + iu_-^\dagger(\partial_0 + \sigma^i\partial_i)u_- + iu_+^\dagger(\partial_0 - \sigma^i\partial_i)u_+ - mu_+^\dagger u_- \\
&= iu_-^\dagger(\partial_0 + \sigma^i\partial_i)u_- + iu_+^\dagger(\partial_0 - \sigma^i\partial_i)u_+ - m(u_+^\dagger u_- + u_-^\dagger u_+) \, .
\end{aligned} \tag{246}$$

Defining

$$\sigma^\mu = (1, \sigma^i) \, , \quad \bar\sigma^\mu = (1, -\sigma^i) \, , \tag{247}$$

we find

$$\mathcal{L} = \bar\psi(i\partial\!\!\!/ - m)\psi = iu_-^\dagger\sigma^\mu\partial_\mu u_- + iu_+^\dagger\bar\sigma^\mu\partial_\mu u_+ - m(u_+^\dagger u_- + u_-^\dagger u_+) \, . \tag{248}$$

We see that u_+ and u_- are coupled through the mass term, and a massless field can be described by u_+ or u_- alone. The equation of motion is then given by

$$i\bar{\sigma}^\mu \partial_\mu u_+ = 0\,, \quad \text{or} \quad i\sigma^\mu \partial_\mu u_- = 0\,. \tag{249}$$

These are called the Weyl equations.

Note that the Dirac fermion has 4 components. Each component is complex, so we have in total 8 real components. In the case of a real scalar field, we had 1 real component, and we obtained one type of particle. In the case of a complex scalar field, we had 2 real components, and we obtained two types of particle, namely the particle and the anti-particle. Then, do we have 8 types of particle in the case of the Dirac fermion? The answer is no, because, unlike the scalar field cases, the equation of motion is first order rather than second order. As a result, the momentum conjugate to the spinor ψ becomes

$$\pi_\psi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\psi^\dagger\,. \tag{250}$$

In other words, the phase space for a spinor is parametrised by ψ and $\psi^\dagger$; recall that, for a scalar field, it is parametrised by ϕ and $\dot{\phi}$. Therefore, the phase space of the Dirac spinor has 8 real dimensions, which means that the dimension of the configuration space, *i.e.*, the number of degrees of freedom, is 4. We will see later that this 4 manifests as the spin-up particle, spin-down particle, spin-up anti-particle, and spin-down anti-particle.

Exercise 10.4: Argue that the Weyl fermion has 2 degrees of freedom.

So far, we have discussed that, in the chiral representation, the Lorentz group matrices $S(\Lambda)$ become block diagonal, and the Dirac spinor gets decomposed into the Weyl spinors. What would happen if we choose a different representation, so that $\gamma^\mu \to V\gamma^\mu V^{-1}$ and $\psi \to V\psi$? The $S(\Lambda)$ will not be block diagonal. Does it mean that we cannot define chiral spinors?

There is, in fact, an invariant way to define the chiral spinors. To do so, we first introduce the γ^5 as follows:

$$\gamma^5 = -i\gamma^0\gamma^1\gamma^2\gamma^3\,. \tag{251}$$

It is easy to show that the γ^5 satisfies

$$\{\gamma^5, \gamma^\mu\} = 0\,, \quad (\gamma^5)^2 = 1\,. \tag{252}$$

With this γ^5, the set of γ-matrices, $\tilde{\gamma}^A = (\gamma^\mu, i\gamma^5)$, satisfies the 5-dimensional Clifford algebra $\{\tilde{\gamma}^A, \tilde{\gamma}^B\} = 2\eta^{AB}$. One may also show that $[S^{\mu\nu}, \gamma^5] = 0$, which indicates that γ^5 is a scalar under rotations and boosts.

Exercise 10.5: Check Eq. (252).

We now introduce the so-called Lorentz-invariant projection operators:

$$P_\pm = \frac{1}{2}(1 \pm \gamma^5)\,. \tag{253}$$

The projection operators have the following properties:

$$P_+^2 = P_+\,, \quad P_-^2 = P_-\,, \quad P_+P_- = 0\,. \tag{254}$$

Let us see what they do in the chiral representation. In the chiral representation,

$$\gamma^5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{255}$$

and thus,

$$P_+ = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad P_- = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{256}$$

In other words, the projection operators project onto the Weyl spinors $u_\pm$ (hence the name "projection" operators).

For an arbitrary representation of the Clifford algebra, we may use γ^5 to define the chiral spinors,

$$\psi_\pm = P_\pm \psi\,, \tag{257}$$

which form the irreducible representations of the Lorentz group. The ψ_+ (ψ_-) is often called a left-handed (right-handed) spinor.[17]

The spinors $\psi_\pm$ are related to each other by parity. What is parity? Let us recall that the Lorentz group is defined in such a way that, under $x^\mu \longrightarrow \Lambda^\mu{}_\nu x^\nu$, $\Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \eta^{\rho\sigma} = \eta^{\mu\nu}$. So far, we have discussed transformations Λ which are continuously connected to the identity; this is why we were able to consider infinitesimal transformations. However, there are two discrete symmetries that are also part of the Lorentz group, namely the time reversal,

$$T\ :\ (x^0, x^i) \longrightarrow (-x^0, x^i)\,, \tag{258}$$

and the parity,

$$P\ :\ (x^0, x^i) \longrightarrow (x^0, -x^i)\,. \tag{259}$$

To see that the parity exchanges the left- and right-handed spinors, let us consider the chiral representation where the Weyl spinors transform as

$$u_\pm \longrightarrow e^{i\varphi\cdot\sigma/2} u_\pm\,, \tag{260}$$

[17] The word "chirality" is derived from the Greek word $\chi\varepsilon\iota\rho$ which means "hand".

under rotations, while, under boosts,

$$u_\pm \longrightarrow e^{\pm\chi\cdot\sigma/2}u_\pm\,. \tag{261}$$

Rotations do not change sign under parity. On the other hand, boosts flip sign. Thus, we see that the parity exchanges the left- and right-handed spinors, $P\;:\;u_\pm \longrightarrow u_\mp$. In other words, the upper (lower) two components of the Dirac spinor become the lower (upper) two components, which is exactly what γ^0 does, $P\;:\;\psi(t,\mathbf{x}) \longrightarrow \gamma^0\psi(t,-\mathbf{x})$.

Notice that if $\psi(t,\mathbf{x})$ satisfies the Dirac equation, then the parity-transformed spinor, $\gamma^0\psi(t,-\mathbf{x})$, also satisfies the Dirac equation:

$$\begin{aligned}(i\partial\!\!\!/ - m)\gamma^0\psi(t,-\mathbf{x}) &= (i\gamma^0\partial_t + i\gamma^i\partial_i - m)\gamma^0\psi(t,-\mathbf{x})\\ &= (i\gamma^0\gamma^0\partial_t + i\gamma^i\gamma^0\partial_i - m\gamma^0)\psi(t,-\mathbf{x})\\ &= (i\gamma^0\gamma^0\partial_t - i\gamma^0\gamma^i\partial_i - \gamma^0 m)\psi(t,-\mathbf{x})\\ &= \gamma^0(i\gamma^0\partial_t - i\gamma^i\partial_i - m)\psi(t,-\mathbf{x})\,,\end{aligned} \tag{262}$$

which is 0 because the derivative is now acting on $-\mathbf{x}$ instead of $\mathbf{x}$.

Now that we know how the spinor transforms under parity, let us examine how interaction terms would change under parity. We start from $\bar\psi\psi$. Recalling that $\bar\psi = \psi^\dagger\gamma^0$, we see that

$$P\;:\;\bar\psi\psi \longrightarrow \bar\psi\psi\,, \tag{263}$$

where we have omitted the spacetime arguments for brevity. It indicates that $\bar\psi\psi$ transforms as a scalar. Next, let us consider $\bar\psi\gamma^\mu\psi$. The temporal component transforms as

$$P\;:\;\bar\psi\gamma^0\psi \longrightarrow \bar\psi\gamma^0\psi\,, \tag{264}$$

while the spatial components transform as

$$P\;:\;\bar\psi\gamma^i\psi \longrightarrow -\bar\psi\gamma^i\psi\,. \tag{265}$$

In other words, $\bar\psi\gamma^\mu\psi$ transforms as a vector with the spatial part changing sign. Similarly, one can show that $\bar\psi S^{\mu\nu}\psi$ transforms as a tensor.

How about $\bar\psi\gamma^5\psi$ and $\bar\psi\gamma^5\gamma^\mu\psi$? Under parity, we see that they transform as

$$P\;:\;\bar\psi\gamma^5\psi \longrightarrow -\bar\psi\gamma^5\psi\,, \tag{266}$$

$$P\;:\;\bar\psi\gamma^5\gamma^\mu\psi \longrightarrow \begin{cases} -\bar\psi\gamma^5\gamma^0\psi & (\mu=0)\,,\\ \bar\psi\gamma^5\gamma^i\psi & (\mu=i)\,.\end{cases} \tag{267}$$

Objects that transform as $\bar\psi\gamma^5\psi$ are called pseudoscalars, and objects that transform as $\bar\psi\gamma^5\gamma^\mu\psi$ are called axial vectors.

Exercise 10.6: Show Eqs. (266) and (267).

In summary, we have the following spinor bilinears:

$$\bar{\psi}\psi\,, \quad \bar{\psi}\gamma^\mu\psi\,, \quad \bar{\psi}S^{\mu\nu}\psi\,, \quad \bar{\psi}\gamma^5\psi\,, \quad \bar{\psi}\gamma^5\gamma^\mu\psi\,. \tag{268}$$

Exercise 10.7: Convince yourself that these are all the bilinears we need.

Armed with new terms involving γ^5, we can construct new Lagrangians, *i.e.*, new theories. Sometimes, these new terms break parity invariance. Sometimes, they don't. If a theory treats $\psi_\pm$ in the same manner, such a theory is called a vector-like theory. On the other hand, a theory where ψ_+ and ψ_- appear differently is called a chiral theory.

※ Lecture 11 (2024.05.31)

11.1 Majorana Fermions

The spinor ψ^α is a complex object. Even if we make the spinor real by imposing $\psi^* = \psi$, it will not stay that way once we perform a Lorentz transformation; note that the representation $S(\Lambda)$ is normally complex as well. However, there is a way to impose a reality condition.

To discuss the possibility, let us look at the following basis for the Clifford algebra:

$$\gamma^0 = \begin{pmatrix} 0 & \sigma^2 \\ \sigma^2 & 0 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} i\sigma^3 & 0 \\ 0 & i\sigma^3 \end{pmatrix}, \quad \gamma^2 = \begin{pmatrix} 0 & -\sigma^2 \\ \sigma^2 & 0 \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} -i\sigma^1 & 0 \\ 0 & -i\sigma^1 \end{pmatrix}. \tag{269}$$

They are all pure imaginary, $(\gamma^\mu)^* = -\gamma^\mu$. This means that the generators $S^{\mu\nu} = [\gamma^\mu, \gamma^\nu]/4$, and hence $S(\Lambda)$, are real. Therefore, with this basis, we can impose the condition $\psi^* = \psi$ and work with a real spinor; this reality condition is preserved under the Lorentz transformation. We call such spinors Majorana spinors and the basis the Majorana basis.

Exercise 11.1: Show that the gamma matrices shown in Eq. (269) satisfy the Clifford algebra.

What if we want to play with the Majorana spinors in a different basis? In other words, is there a way to impose the reality condition in a general way? We will only focus on the basis that satisfies $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^i)^\dagger = -\gamma^i$. Let us then define:

$$\psi^{(c)} = C\psi^*, \tag{270}$$

where C, which is a 4×4 matrix, satisfies

$$C^\dagger C = 1, \quad C^\dagger \gamma^\mu C = -(\gamma^\mu)^*. \tag{271}$$

We see that $\psi^{(c)}$ transforms as

$$\psi^{(c)} \to CS(\Lambda)^*\psi^* = CS(\Lambda)^* C^\dagger C\psi^* = S(\Lambda)C\psi^* = S(\Lambda)\psi^{(c)}, \tag{272}$$

where we have used $CS(\Lambda)^*C^\dagger = S(\Lambda)$ because $C^\dagger[\gamma^\mu, \gamma^\nu]C = [(\gamma^\mu)^*, (\gamma^\nu)^*]$. We also see that $\psi^{(c)}$ satisfies the Dirac equation, if ψ satisfies the Dirac equation, because if we take the complex conjugate and apply C to the Dirac equation of ψ, we get

$$C(-i\partial\!\!\!/^* - m)\psi^* = (i\partial\!\!\!/ - m)C\psi^* = (i\partial\!\!\!/ - m)\psi^{(c)} = 0. \tag{273}$$

Thus, with C, we can now impose the Lorentz-invariant reality condition on the Dirac spinor,

$$\psi^{(c)} = \psi, \tag{274}$$

which will give us a Majorana spinor. After quantisation, the Majorana spinor gives rise to a fermion that is its own anti-particle.

The matrix C is called the charge conjugation. In the Majorana basis, we have $C_{\text{Maj}} = 1$, and the reality condition (also known as the Majorana condition) becomes $\psi = \psi^*$. In the chiral basis, we may take $C_{\text{chiral}} = i\gamma^2$. Before we move on, let us see how the Majorana condition would look in terms of the Weyl spinors. We find that $u_+ = i\sigma^2 u_-^*$ and $u_- = -i\sigma^2 u_+^*$; the Majorana condition relates u_- and u_+. In other words, a Majorana spinor can be written in terms of Weyl spinors as

$$\psi = \begin{pmatrix} u_+ \\ -i\sigma^2 u_+^* \end{pmatrix} . \tag{275}$$

11.2 Symmetries of the Dirac Lagrangian

The Dirac Lagrangian has a number of symmetries:

- Spacetime translations:
  The spinor transforms under spacetime translations as

  $$\delta\psi = \epsilon^\mu \partial_\mu \psi , \tag{276}$$

  and the associated current is the energy-momentum tensor, given by

  $$T^{\mu\nu} = i\bar{\psi}\gamma^\mu \partial^\nu \psi - \eta^{\mu\nu}\mathcal{L} . \tag{277}$$

  Note that a current is conserved only when the equations of motion are obeyed. Thus, we can impose the equations of motion already at the level of $T^{\mu\nu}$. For a spinor field, this means that

  $$T^{\mu\nu} = i\bar{\psi}\gamma^\mu \partial^\nu \psi , \tag{278}$$

  as $\mathcal{L} = 0$ after the use of equations of motion.

- Lorentz transformations:
  Under an infinitesimal Lorentz transformation,

  $$\delta\psi^\alpha = -\omega^\mu{}_\nu x^\nu \partial_\mu \psi^\alpha + \frac{1}{2}\Omega_{\rho\sigma}(S^{\rho\sigma})^\alpha{}_\beta \psi^\beta . \tag{279}$$

  Noting that $\omega^\mu{}_\nu = \Omega_{\rho\sigma}(\mathcal{M}^{\rho\sigma})^\mu{}_\nu/2$, we get

  $$\delta\psi^\alpha = -\omega^{\mu\nu}\left[x_\nu \partial_\mu \psi^\alpha - \frac{1}{2}(S_{\mu\nu})^\alpha{}_\beta \psi^\beta\right] . \tag{280}$$

  The associated conserved current is given by

  $$(\mathcal{J}^\mu)^{\rho\sigma} = x^\rho T^{\mu\sigma} - x^\sigma T^{\mu\rho} - i\bar{\psi}\gamma^\mu S^{\rho\sigma}\psi . \tag{281}$$

- Internal vector symmetry:
  The Dirac Lagrangian is invariant under $\psi \longrightarrow e^{-i\alpha}\psi$. The associated current is
  $$j^\mu_V = \bar{\psi}\gamma^\mu\psi\,, \tag{282}$$
  and the conserved quantity is
  $$Q = \int d^3x\, \bar{\psi}\gamma^0\psi = \int d^3x\, \psi^\dagger\psi\,, \tag{283}$$
  which, as we shall see later, has the interpretation of electric charge (or particle number) for fermions.

- Axial symmetry:
  The Dirac Lagrangian develops one extra internal symmetry when $m = 0$:
  $$\psi \longrightarrow e^{i\alpha\gamma^5}\psi\,, \qquad \bar{\psi} \longrightarrow \bar{\psi}e^{i\alpha\gamma^5}\,. \tag{284}$$
  The associated current is
  $$j^\mu_A = \bar{\psi}\gamma^\mu\gamma^5\psi\,. \tag{285}$$
  It is crucial that this is conserved only when $m = 0$. With the full Dirac Lagrangian, we find
  $$\partial_\mu j^\mu_A = 2im\bar{\psi}\gamma^5\psi\,, \tag{286}$$
  which vanishes only for $m = 0$. The current j^μ_A is called the axial current, and this is particularly interesting as it does not survive the quantisation process when the theory is coupled to gauge fields. In other words, while the axial transformation is a good symmetry of the classical Lagrangian, after the quantisation, the symmetry gets broken. This is called an anomaly.

11.3 Solutions of the Dirac Equation

Let us now discuss the solutions to the Dirac equation,
$$(i\partial\!\!\!/ - m)\psi = 0\,. \tag{287}$$
We start with a simple ansatz:
$$\psi = u(\mathbf{p})e^{-ip\cdot x}\,, \tag{288}$$
where $u(\mathbf{p})$ is a 4-component spinor that can depend on the 3-momentum $\mathbf{p}$. With this ansatz, the Dirac equation becomes
$$0 = (p\!\!\!/ - m)u(\mathbf{p}) = \begin{pmatrix} -m & p\cdot\sigma \\ p\cdot\bar{\sigma} & -m \end{pmatrix} u(\mathbf{p})\,, \tag{289}$$

which has the solution

$$u(\mathbf{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi \\ \sqrt{p\cdot\bar{\sigma}}\,\xi \end{pmatrix} . \tag{290}$$

Here, ξ is a 2-component spinor normalised to $\xi^\dagger\xi = 1$. To prove that this is indeed the solution, let us first write

$$u(\mathbf{p}) = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} . \tag{291}$$

Then, the Dirac equation, $(\not{p} - m)u(\mathbf{p}) = 0$, tells us that

$$(p\cdot\sigma)u_2 = mu_1 \,, \quad (p\cdot\bar{\sigma})u_1 = mu_2 \,. \tag{292}$$

Note that $(p\cdot\sigma)(p\cdot\bar{\sigma}) = p_\mu\sigma^\mu p_\nu\bar{\sigma}^\nu = (p_0+p_i\sigma^i)(p_0-p_j\sigma^j) = p_0^2-p_ip_j\sigma^i\sigma^j = p_0^2-p_ip_j\delta^{ij} = p_\mu p^\mu = m^2$. It means that, from the first equation, $(p\cdot\sigma)u_2 = mu_1$, we read

$$\begin{aligned}
&(p\cdot\bar{\sigma})(p\cdot\sigma)u_2 = m(p\cdot\bar{\sigma})u_1 \\
&\Leftrightarrow m^2u_2 = m(p\cdot\bar{\sigma})u_1 \\
&\Leftrightarrow mu_2 = (p\cdot\bar{\sigma})u_1 \,,
\end{aligned} \tag{293}$$

which is the second equation. Now, let us try the ansatz $u_1 = (p\cdot\sigma)\xi'$ for some spinor ξ'. Then, the second equation, $(p\cdot\bar{\sigma})u_1 = mu_2$, tells us that $u_2 = m\xi'$. Thus, any spinor of the form

$$u(\mathbf{p}) = A\begin{pmatrix} (p\cdot\sigma)\xi' \\ m\xi' \end{pmatrix} \tag{294}$$

with a constant A is a solution. Finally, choosing $A = 1/m$ and $\xi' = \sqrt{p\cdot\bar{\sigma}}\,\xi$, we get

$$u(\mathbf{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi \\ \sqrt{p\cdot\bar{\sigma}}\,\xi \end{pmatrix} . \tag{295}$$

What if we start with the ansatz

$$\psi = v(\mathbf{p})e^{+ip\cdot x} \,, \tag{296}$$

instead of $\psi = u(\mathbf{p})e^{-ip\cdot x}$? The Dirac equation then tells us that $v(\mathbf{p})$ satisfies

$$0 = (\not{p} + m)v(\mathbf{p}) = \begin{pmatrix} m & p\cdot\sigma \\ p\cdot\bar{\sigma} & m \end{pmatrix} v(\mathbf{p}) \,, \tag{297}$$

whose solution is given by

$$v(\mathbf{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\eta \\ -\sqrt{p\cdot\bar{\sigma}}\,\eta \end{pmatrix} . \tag{298}$$

Here, η is a 2-component spinor normalised to $\eta^\dagger\eta = 1$.

Exercise 11.2: Show that the solution (298) is indeed a solution.

Solutions of the form $\psi = u(\mathbf{p})e^{-ip\cdot x} \sim e^{-iEt}$ are called positive frequency solutions because if we compute the energy (see Eq. (278)),

$$E = \int d^3x\, T^{00} = \int d^3x\, i\bar{\psi}\gamma^0\dot{\psi} = \int d^3x\, \psi^\dagger\gamma^0(-i\gamma^i\partial_i + m)\psi\,, \tag{299}$$

then we see that it is positive. On the other hand, solutions of the form $\psi = v(\mathbf{p})e^{+ip\cdot x} \sim e^{+iEt}$ are called negative frequency solutions because we get negative energy if we compute the energy.

Exercise 11.3: Compute the energies for $\psi = u(\mathbf{p})e^{-ip\cdot x}$ and $\psi = v(\mathbf{p})e^{+ip\cdot x}$.

As an example, let us consider the positive frequency solution with mass m and zero 3-momentum, $\mathbf{p} = 0$,

$$u(\mathbf{p}) = \sqrt{m}\begin{pmatrix} \xi \\ \xi \end{pmatrix} . \tag{300}$$

Under rotations, ξ transforms as $\xi \to e^{i\varphi\cdot\sigma/2}\xi$. The 2-component spinor ξ defines the spin of the field. Just like Quantum Mechanics, a field with spin up (down) along a given direction is described by the eigenvector of the corresponding Pauli matrix with eigenvalue 1 (-1). For example, $\xi^T = (1\ \ 0)$ ($\xi^T = (0\ \ 1)$) describes a field with spin up (down) along the z-axis. After quantisation, this will become the spin of the associated particle.

Let us now consider boosting the field with spin $\xi^T = (1\ \ 0)$ along the z direction with $p^\mu = (E, 0, 0, p^3)$. The solution then becomes

$$u(\mathbf{p}) = \begin{pmatrix} \sqrt{E-p^3}\,\binom{1}{0} \\ \sqrt{E+p^3}\,\binom{1}{0} \end{pmatrix} . \tag{301}$$

In particular, for a massless field, we get

$$u(\mathbf{p}) = \sqrt{2E}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} , \tag{302}$$

as $E = p^3$. Similarly, for a boosted solution of the spin down field, $\xi^T = (0 \;\; 1)$, we have

$$u(\mathbf{p}) = \begin{pmatrix} \sqrt{E + p^3} \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \\ \sqrt{E - p^3} \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \end{pmatrix} , \tag{303}$$

which, in the massless case, becomes

$$u(\mathbf{p}) = \sqrt{2E} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} . \tag{304}$$

11.3.1 Helicity

One of the important concept is the helicity which is the projection of the angular momentum along the direction of momentum,

$$h = \frac{i}{2} \epsilon_{ijk} \hat{p}^i S^{jk} = \frac{1}{2} \hat{p}_i \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix} . \tag{305}$$

Let us work out the helicity for the examples we considered above. The massless field with spin $\xi^T = (1 \;\; 0)$, we find $h = 1/2$; we say that it is "right-handed". For the massless field with spin $\xi^T = (0 \;\; 1)$, we find $h = -1/2$; we say that it is "left-handed".

Exercise 11.4: Can you tell how the helicity is different from the chirality? Are both of them invariant under a Lorentz transformation? Do they become identical in a special case?

11.3.2 Spinor Manipulations

Later, we will be encountering products of the spinors u and v multiple times. So, let us discuss some useful formulae here.

First of all, it is convenient to introduce a basis ξ^s and η^s, where $s = 1, 2$, for the 2-component spinors such that

$$\xi^{r\dagger} \xi^s = \delta^{rs} , \quad \eta^{r\dagger} \eta^s = \delta^{rs} . \tag{306}$$

An example would be

$$\xi^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \quad \xi^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} , \tag{307}$$

and similarly for η^s. With this basis, we can write, *e.g.*, the two independent positive frequency solutions as

$$u^s(\mathbf{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\,\xi^s \end{pmatrix} . \tag{308}$$

We can compute the inner product as follows:

$$\begin{aligned} u^{r\dagger}(\mathbf{p})\cdot u^s(\mathbf{p}) &= \begin{pmatrix} \xi^{r\dagger}\sqrt{p\cdot\sigma} & \xi^{r\dagger}\sqrt{p\cdot\bar{\sigma}} \end{pmatrix} \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\,\xi^s \end{pmatrix} \\ &= \xi^{r\dagger} p\cdot\sigma\xi^s + \xi^{r\dagger} p\cdot\bar{\sigma}\xi^s \\ &= 2p_0\xi^{r\dagger}\xi^s = 2p_0\delta^{rs} . \end{aligned} \tag{309}$$

Of course, $u^\dagger\cdot u$ is not Lorentz invariant as we already discussed. The Lorentz-invariant inner product would be $\bar{u}\cdot u$, which is given by

$$\begin{aligned} \bar{u}^r(\mathbf{p})\cdot u^s(\mathbf{p}) &= \begin{pmatrix} \xi^{r\dagger}\sqrt{p\cdot\sigma} & \xi^{r\dagger}\sqrt{p\cdot\bar{\sigma}} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\,\xi^s \end{pmatrix} \\ &= 2m\delta^{rs} . \end{aligned} \tag{310}$$

We can do the same thing for the negative frequency solutions,

$$v^s(\mathbf{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\eta^s \\ -\sqrt{p\cdot\bar{\sigma}}\,\eta^s \end{pmatrix} . \tag{311}$$

It is straightforward to show that

$$v^{r\dagger}(\mathbf{p})\cdot v^s(\mathbf{p}) = 2p_0\delta^{rs} , \quad \bar{v}^r(\mathbf{p})\cdot v^s(\mathbf{p}) = -2m\delta^{rs} . \tag{312}$$

We may also consider the inner product between u and v. It is easy to show that

$$\bar{u}^r(\mathbf{p})\cdot v^s(\mathbf{p}) = 0 , \quad \bar{v}^r(\mathbf{p})\cdot u^s(\mathbf{p}) = 0 , \tag{313}$$

and (note the minus sign in front of the second momentum)

$$u^{r\dagger}(\mathbf{p})\cdot v^s(-\mathbf{p}) = 0 , \quad v^{r\dagger}(\mathbf{p})\cdot u^s(-\mathbf{p}) = 0 . \tag{314}$$

Exercise 11.5: Show Eqs. (313) and (314).

Before we move on, let us talk about the outer products as well,

$$\sum_{s=1}^{2} u^s(\mathbf{p})\bar{u}^s(\mathbf{p}) , \quad \sum_{s=1}^{2} v^s(\mathbf{p})\bar{v}^s(\mathbf{p}) . \tag{315}$$

For the first one, we have

$$\sum_{s=1}^{2} u^s(\mathbf{p})\bar{u}^s(\mathbf{p}) = \sum_{s=1}^{2} \begin{pmatrix} \sqrt{p\cdot\sigma}\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\xi^s \end{pmatrix} \begin{pmatrix} \xi^{s\dagger}\sqrt{p\cdot\bar{\sigma}} & \xi^{s\dagger}\sqrt{p\cdot\sigma} \end{pmatrix} . \tag{316}$$

Since $\sum_s \xi^s \xi^{s\dagger} = 1$ (the 2×2 unit matrix), we find

$$\sum_{s=1}^{2} u^s(\mathbf{p})\bar{u}^s(\mathbf{p}) = \begin{pmatrix} m & p\cdot\sigma \\ p\cdot\bar{\sigma} & m \end{pmatrix} = \not{p} + m \, . \tag{317}$$

Exercise 11.6: Show that $\sum_s v^s(\mathbf{p})\bar{v}^s(\mathbf{p}) = \not{p} - m$.

※ Lecture 12 (2024.06.07)

12.1 Quantisation of the Dirac Field

We are now ready to quantise the Dirac theory. As a first attempt, it seems reasonable to do the quantisation as we did for the scalar field. We start by defining the conjugate momentum,

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\bar{\psi}\gamma^0 = i\psi^\dagger \,. \tag{318}$$

As we discussed earlier, the momentum conjugate to ψ does not involve the time derivative of ψ.

We then promote the field and the conjugate momentum to operators with the following canonical commutation relations:

$$[\psi_\alpha(\mathbf{x}), \psi_\beta(\mathbf{y})] = 0 = [\psi^\dagger_\alpha(\mathbf{x}), \psi^\dagger_\beta(\mathbf{y})]\,, \quad [\psi_\alpha(\mathbf{x}), \psi^\dagger_\beta(\mathbf{y})] = \delta_{\alpha\beta}\delta^{(3)}(\mathbf{x}-\mathbf{y})\,, \tag{319}$$

just like we did in the case of a scalar field theory. The operators ψ and $\psi^\dagger$ may be expressed as a sum of plane waves (note that we are not considering any interactions at the moment),

$$\begin{aligned}
\psi(\mathbf{x}) &= \sum_{s=1}^{2} \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_\mathbf{p}}} \left[b^s_\mathbf{p} u^s(\mathbf{p}) e^{+i\mathbf{p}\cdot\mathbf{x}} + c^{s\dagger}_\mathbf{p} v^s(\mathbf{p}) e^{-i\mathbf{p}\cdot\mathbf{x}} \right] \,, \\
\psi^\dagger(\mathbf{x}) &= \sum_{s=1}^{2} \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_\mathbf{p}}} \left[b^{s\dagger}_\mathbf{p} u^{s\dagger}(\mathbf{p}) e^{-i\mathbf{p}\cdot\mathbf{x}} + c^{s}_\mathbf{p} v^{s\dagger}(\mathbf{p}) e^{+i\mathbf{p}\cdot\mathbf{x}} \right] \,.
\end{aligned} \tag{320}$$

The operators $b^{s\dagger}_\mathbf{p}$ ($c^{s\dagger}_\mathbf{p}$) would have the role of creating particles associated to the spinor $u^s(\mathbf{p})$ ($v^s(\mathbf{p})$). The commutation relations (319) imply that

$$[b^r_\mathbf{p}, b^{s\dagger}_\mathbf{q}] = (2\pi)^3\delta^{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q})\,, \quad [c^r_\mathbf{p}, c^{s\dagger}_\mathbf{q}] = -(2\pi)^3\delta^{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q})\,. \tag{321}$$

We start to sense that something is going wrong. Let us imagine the ground state $|0\rangle$ defined as $c^r_\mathbf{p}|0\rangle$, just like we did for the scalar case. Then, the excited, one-particle state, which would be $c^{s\dagger}_\mathbf{p}|0\rangle$, has negative norm; $\langle 0|cc^\dagger|0\rangle = \langle 0|[c, c^\dagger]|0\rangle < 0$. We will discuss how to avoid this problem later, but for now, let us carry on and see where we are led to.

<u>Exercise 12.1:</u> Verify Eq. (321).

The next step is to construct the Hamiltonian, $H = \int d^3\mathbf{x}\, \mathcal{H}$, for the Dirac theory. We find

$$\mathcal{H} = \pi\dot{\psi} - \mathcal{L} = \bar{\psi}(-i\gamma^i\partial_i + m)\psi\,, \tag{322}$$

which is to be turned into an operator. Let us first note that

$$(-i\gamma^i\partial_i + m)\psi = \sum_{s=1}^{2} \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_\mathbf{p}}} \left[b^s_\mathbf{p}(-\gamma^i p_i + m) u^s(\mathbf{p}) e^{+i\mathbf{p}\cdot\mathbf{x}} + c^{s\dagger}_\mathbf{p}(\gamma^i p_i + m) v^s(\mathbf{p}) e^{-i\mathbf{p}\cdot\mathbf{x}} \right] \,. \tag{323}$$

We have seen the equations that u and v satisfy, namely

$$(\not{p} - m)u^s(\mathbf{p}) = 0\,, \quad (\not{p} + m)v^s(\mathbf{p}) = 0\,, \tag{324}$$

from which we read

$$(-\gamma^i p_i + m)u^s(\mathbf{p}) = \gamma^0 p_0 u^s(\mathbf{p})\,, \quad (\gamma^i p_i + m)v^s(\mathbf{p}) = -\gamma^0 p_0 v^s(\mathbf{p})\,. \tag{325}$$

Thus, we can write

$$(-i\gamma^i\partial_i + m)\psi = \sum_{s=1}^{2}\int \frac{d^3\mathbf{p}}{(2\pi)^3}\sqrt{\frac{E_\mathbf{p}}{2}}\gamma^0\left[b^s_\mathbf{p}u^s(\mathbf{p})e^{+i\mathbf{p}\cdot\mathbf{x}} - c^{s\dagger}_\mathbf{p}v^s(\mathbf{p})e^{-i\mathbf{p}\cdot\mathbf{x}}\right]\,. \tag{326}$$

The Hamiltonian operator is, therefore, given by (noting that $\bar{\psi} = \psi^\dagger\gamma^0$ and $(\gamma^0)^2 = 1$)

$$\begin{aligned}
H &= \sum_{r=1}^{2}\sum_{s=1}^{2}\int d^3\mathbf{x}\frac{d^3\mathbf{p}d^3\mathbf{q}}{(2\pi)^6}\frac{1}{2}\sqrt{\frac{E_\mathbf{p}}{E_\mathbf{q}}} \\
&\qquad \left[b^{r\dagger}_\mathbf{q}u^{r\dagger}(\mathbf{q})e^{-i\mathbf{q}\cdot\mathbf{x}} + c^r_\mathbf{q}v^{r\dagger}(\mathbf{q})e^{+i\mathbf{q}\cdot\mathbf{x}}\right]\cdot\left[b^s_\mathbf{p}u^s(\mathbf{p})e^{+i\mathbf{p}\cdot\mathbf{x}} - c^{s\dagger}_\mathbf{p}v^s(\mathbf{p})e^{-i\mathbf{p}\cdot\mathbf{x}}\right] \\
&= \sum_{r=1}^{2}\sum_{s=1}^{2}\int d^3\mathbf{x}\frac{d^3\mathbf{p}d^3\mathbf{q}}{(2\pi)^6}\frac{1}{2}\sqrt{\frac{E_\mathbf{p}}{E_\mathbf{q}}} \\
&\qquad \Big\{b^{r\dagger}_\mathbf{q}b^s_\mathbf{p}u^{r\dagger}(\mathbf{q})\cdot u^s(\mathbf{p})e^{i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{x})} - b^{r\dagger}_\mathbf{q}c^{s\dagger}_\mathbf{p}u^{r\dagger}(\mathbf{q})\cdot v^s(\mathbf{p})e^{-i(\mathbf{p}\cdot\mathbf{x}+\mathbf{q}\cdot\mathbf{x})} \\
&\qquad\quad + c^r_\mathbf{q}b^s_\mathbf{p}v^{r\dagger}(\mathbf{q})\cdot u^s(\mathbf{p})e^{i(\mathbf{p}\cdot\mathbf{x}+\mathbf{q}\cdot\mathbf{x})} - c^r_\mathbf{q}c^{s\dagger}_\mathbf{p}v^{r\dagger}(\mathbf{q})\cdot v^s(\mathbf{p})e^{-i(\mathbf{p}\cdot\mathbf{x}-\mathbf{q}\cdot\mathbf{x})}\Big\} \\
&= \sum_{r=1}^{2}\sum_{s=1}^{2}\int \frac{d^3\mathbf{p}d^3\mathbf{q}}{(2\pi)^3}\frac{1}{2}\sqrt{\frac{E_\mathbf{p}}{E_\mathbf{q}}} \\
&\qquad \Big\{b^{r\dagger}_\mathbf{q}b^s_\mathbf{p}u^{r\dagger}(\mathbf{q})\cdot u^s(\mathbf{p})\delta^{(3)}(\mathbf{p}-\mathbf{q}) - b^{r\dagger}_\mathbf{q}c^{s\dagger}_\mathbf{p}u^{r\dagger}(\mathbf{q})\cdot v^s(\mathbf{p})\delta^{(3)}(\mathbf{p}+\mathbf{q}) \\
&\qquad\quad + c^r_\mathbf{q}b^s_\mathbf{p}v^{r\dagger}(\mathbf{q})\cdot u^s(\mathbf{p})\delta^{(3)}(\mathbf{p}+\mathbf{q}) - c^r_\mathbf{q}c^{s\dagger}_\mathbf{p}v^{r\dagger}(\mathbf{q})\cdot v^s(\mathbf{p})\delta^{(3)}(\mathbf{p}-\mathbf{q})\Big\} \\
&= \sum_{r=1}^{2}\sum_{s=1}^{2}\int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{2}\Big\{b^{r\dagger}_\mathbf{p}b^s_\mathbf{p}u^{r\dagger}(\mathbf{p})\cdot u^s(\mathbf{p}) - b^{r\dagger}_{-\mathbf{p}}c^{s\dagger}_\mathbf{p}u^{r\dagger}(-\mathbf{p})\cdot v^s(\mathbf{p}) \\
&\qquad\quad + c^r_{-\mathbf{p}}b^s_\mathbf{p}v^{r\dagger}(-\mathbf{p})\cdot u^s(\mathbf{p}) - c^r_\mathbf{p}c^{s\dagger}_\mathbf{p}v^{r\dagger}(\mathbf{p})\cdot v^s(\mathbf{p})\Big\}\,.
\end{aligned} \tag{327}$$

If we re-label $\mathbf{p} \longrightarrow -\mathbf{p}$ in the second and third terms, we get

$$H = \sum_{r=1}^{2}\sum_{s=1}^{2}\int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{2}\Big\{b^{r\dagger}_\mathbf{p}b^s_\mathbf{p}u^{r\dagger}(\mathbf{p})\cdot u^s(\mathbf{p}) - b^{r\dagger}_\mathbf{p}c^{s\dagger}_{-\mathbf{p}}u^{r\dagger}(\mathbf{p})\cdot v^s(-\mathbf{p})$$

$$+ c^r_{\mathbf{p}} b^s_{-\mathbf{p}} v^{r\dagger}(\mathbf{p}) \cdot u^s(-\mathbf{p}) - c^r_{\mathbf{p}} c^{s\dagger}_{\mathbf{p}} v^{r\dagger}(\mathbf{p}) \cdot v^s(\mathbf{p}) \Big\} . \tag{328}$$

Using

$$u^{r\dagger}(\mathbf{p}) \cdot u^s(\mathbf{p}) = 2p_0 \delta^{rs} = v^{r\dagger}(\mathbf{p}) \cdot v^s(\mathbf{p}) , \quad u^{r\dagger}(\mathbf{p}) \cdot v^s(-\mathbf{p}) = 0 = v^{r\dagger}(\mathbf{p}) \cdot u^s(-\mathbf{p}) , \tag{329}$$

we get

$$H = \sum_{s=1}^{2} \int \frac{d^3\mathbf{p}}{(2\pi)^3} E_{\mathbf{p}} \left(b^{s\dagger}_{\mathbf{p}} b^s_{\mathbf{p}} - c^s_{\mathbf{p}} c^{s\dagger}_{\mathbf{p}} \right) . \tag{330}$$

Since $[c^r_{\mathbf{p}}, c^{s\dagger}_{\mathbf{q}}] = -(2\pi)^3 \delta^{rs} \delta^{(3)}(\mathbf{p} - \mathbf{q})$, we can express the Hamiltonian as

$$H = \sum_{s=1}^{2} \int \frac{d^3\mathbf{p}}{(2\pi)^3} E_{\mathbf{p}} \left(b^{s\dagger}_{\mathbf{p}} b^s_{\mathbf{p}} - c^{s\dagger}_{\mathbf{p}} c^s_{\mathbf{p}} + (2\pi)^3 \delta^{(3)}(0) \right) . \tag{331}$$

The last term, $\delta^{(3)}(0)$, is the familiar infinity that we have already encountered in the scalar case; we know how to deal with it by using the normal ordering. The first term, $b^{s\dagger}_{\mathbf{p}} b^s_{\mathbf{p}}$, is also something we have seen in the scalar case; $b^{s\dagger}_{\mathbf{p}}$ creates positive-energy states,

$$[H, b^{s\dagger}_{\mathbf{p}}] = E_{\mathbf{p}} b^{s\dagger}_{\mathbf{p}} . \tag{332}$$

Now, the second term looks a bit weird. Firstly, the minus sign in front of $c^{s\dagger}_{\mathbf{p}} c^s_{\mathbf{p}}$ seems wrong. Nevertheless, if we think of $c^{s\dagger}_{\mathbf{p}}$ as a creation operator, we still see that $c^{s\dagger}_{\mathbf{p}}$ creates positive-energy states,

$$[H, c^{s\dagger}_{\mathbf{p}}] = E_{\mathbf{p}} c^{s\dagger}_{\mathbf{p}} . \tag{333}$$

However, the problem is that these states have negative norm. One may try to avoid this issue by interpreting c, rather that $c^\dagger$, as the creation operator. Then, we would get

$$[H, c^s_{\mathbf{p}}] = -E_{\mathbf{p}} c^s_{\mathbf{p}} , \tag{334}$$

i.e., the Hamiltonian is not bounded from below, which is not a good thing.

Exercise 12.2: Check Eqs. (333) and (334).

12.2 Quantisation of the Dirac Field – Again

Our first attempt at quantising the Dirac theory turned out to be a failure. The main problem was the minus sign; it seemed to be the case that we cannot avoid the minus sign no matter what we do. What did we do wrong? What we missed is the fact that spin-1/2 particles are fermions. They obey the Fermi-Dirac statistics; interchanging any two particles would give rise to a minus sign. Unlike the scalar fields, which must be quantised as bosons, spin-1/2 fields must be quantised as fermions.[18] Thus, let us discuss how we can quantise a field as a fermion.

Let us recall that when we quantised the scalar field, we saw that

$$a^\dagger_{\mathbf{p}} a^\dagger_{\mathbf{q}}|0\rangle \equiv |\mathbf{p},\mathbf{q}\rangle = |\mathbf{q},\mathbf{p}\rangle\,, \tag{335}$$

essentially because $[a^\dagger_{\mathbf{p}}, a^\dagger_{\mathbf{q}}] = 0$. So, the resulting particles obeyed the bosonic statistics. If we want states to obey fermionic statistics, the commutator is not the right choice because we do not get a minus sign upon the interchange of two particles. What gives a minus sign is the anti-commutation relations,

$$\{A, B\} \equiv AB + BA\,. \tag{336}$$

Therefore, instead of Eq. (319), let us impose

$$\{\psi_\alpha(\mathbf{x}), \psi_\beta(\mathbf{y})\} = 0 = \{\psi^\dagger_\alpha(\mathbf{x}), \psi^\dagger_\beta(\mathbf{y})\}\,, \quad \{\psi_\alpha(\mathbf{x}), \psi^\dagger_\beta(\mathbf{y})\} = \delta_{\alpha\beta}\delta^{(3)}(\mathbf{x}-\mathbf{y})\,. \tag{337}$$

We then obtain

$$\{b^r_{\mathbf{p}}, b^{s\dagger}_{\mathbf{q}}\} = (2\pi)^3\delta^{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q})\,, \quad \{c^r_{\mathbf{p}}, c^{s\dagger}_{\mathbf{q}}\} = (2\pi)^3\delta^{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q})\,. \tag{338}$$

Exercise 12.3: Verify Eq. (338).

With the anti-commutation relations, the Hamiltonian now becomes

$$H = \sum_{s=1}^{2}\int \frac{d^3\mathbf{p}}{(2\pi)^3} E_{\mathbf{p}}\left(b^{s\dagger}_{\mathbf{p}} b^s_{\mathbf{p}} - c^s_{\mathbf{p}} c^{s\dagger}_{\mathbf{p}}\right) = \sum_{s=1}^{2}\int \frac{d^3\mathbf{p}}{(2\pi)^3} E_{\mathbf{p}}\left(b^{s\dagger}_{\mathbf{p}} b^s_{\mathbf{p}} + c^{s\dagger}_{\mathbf{p}} c^s_{\mathbf{p}} - (2\pi)^3\delta^{(3)}(0)\right). \tag{339}$$

We no longer have the minus-sign problem.

The vacuum $|0\rangle$ can be defined as

$$b^s_{\mathbf{p}}|0\rangle = c^s_{\mathbf{p}}|0\rangle = 0\,, \tag{340}$$

just as in the scalar case. We can easily show that

$$[H, b^r_{\mathbf{p}}] = -E_{\mathbf{p}} b^r_{\mathbf{p}}\,, \quad [H, b^{r\dagger}_{\mathbf{p}}] = E_{\mathbf{p}} b^{r\dagger}_{\mathbf{p}}\,, \tag{341}$$

and

$$[H, c^r_{\mathbf{p}}] = -E_{\mathbf{p}} c^r_{\mathbf{p}}\,, \quad [H, c^{r\dagger}_{\mathbf{p}}] = E_{\mathbf{p}} c^{r\dagger}_{\mathbf{p}}\,. \tag{342}$$

[18] This is the so-called spin-statistics theorem: integer spin fields are quantised as bosons, and half-integer spin fields are quantised as fermions.

Thus, we can construct a tower of energy eigenstates just like how we did in the scalar case. For example, the one-particle states would be

$$|\mathbf{p}, r\rangle = b_{\mathbf{p}}^{r\dagger}|0\rangle\,, \tag{343}$$

and the two-particle states would be

$$|\mathbf{p}_1, r_1; \mathbf{p}_2, r_2\rangle = b_{\mathbf{p}_1}^{r_1\dagger} b_{\mathbf{p}_2}^{r_2\dagger}|0\rangle\,. \tag{344}$$

Note that $|\mathbf{p}_1, r_1; \mathbf{p}_2, r_2\rangle = -|\mathbf{p}_2, r_2; \mathbf{p}_1, r_1\rangle$; the particles now obey the Fermi-Dirac statistics. One important consequence of the Fermi-Dirac statistics is the Pauli-Exclusion principle, $|\mathbf{p}, r; \mathbf{p}, r\rangle = 0$.

Exercise 12.4: Show Eqs. (341) and (342).

Exercise 12.5: Confirm that a particle $|\mathbf{p} = 0, r\rangle$ carries spin $1/2$. (To do so, you can act with the angular momentum operator (281).)

12.3 Propagators for Dirac Fields

We now move to the Heisenberg picture, where the spinors,

$$\psi(t, \mathbf{x}) = \psi(x) = \sum_{s=1}^{2} \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \left[b_{\mathbf{p}}^{s} u^{s}(\mathbf{p}) e^{-ip\cdot x} + c_{\mathbf{p}}^{s\dagger} v^{s}(\mathbf{p}) e^{+ip\cdot x} \right]\,, \tag{345}$$

satisfy the operator equation,

$$\frac{\partial \psi}{\partial t} = i[H, \psi]\,. \tag{346}$$

We define the fermionic propagator as

$$iS_{\alpha\beta}(x-y) = \{\psi_\alpha(x), \bar{\psi}_\beta(y)\}\,, \tag{347}$$

or, in short, $iS(x-y) = \{\psi(x), \bar{\psi}(y)\}$. Explicitly, it is given by

$$\begin{aligned}
iS(x-y) &= \sum_{s=1}^{2}\sum_{r=1}^{2} \int \frac{d^3\mathbf{p}\, d^3\mathbf{q}}{(2\pi)^6} \frac{1}{\sqrt{4E_{\mathbf{p}}E_{\mathbf{q}}}} \\
&\qquad \times \left[\{b_{\mathbf{p}}^{s}, b_{\mathbf{q}}^{r\dagger}\} u^{s}(\mathbf{p}) \bar{u}^{r}(\mathbf{q}) e^{-i(p\cdot x - q\cdot y)} + \{c_{\mathbf{p}}^{s\dagger}, c_{\mathbf{q}}^{r}\} v^{s}(\mathbf{p}) \bar{v}^{r}(\mathbf{q}) e^{+i(p\cdot x - q\cdot y)} \right] \\
&= \sum_{s=1}^{2} \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} \left[u^{s}(\mathbf{p}) \bar{u}^{s}(\mathbf{p}) e^{-ip\cdot(x-y)} + v^{s}(\mathbf{p}) \bar{v}^{s}(\mathbf{p}) e^{+ip\cdot(x-y)} \right] \\
&= \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2E_{\mathbf{p}}} \left[(\not{p} + m) e^{-ip\cdot(x-y)} + (\not{p} - m) e^{+ip\cdot(x-y)} \right]\,.
\end{aligned} \tag{348}$$

Recalling that the propagator for a real scalar field can be written as

$$D(x-y) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2E_\mathbf{p}} e^{-ip\cdot(x-y)}\,, \tag{349}$$

we can write down the fermionic propagator as

$$iS(x-y) = (i\partial\!\!\!/_x + m)\left[D(x-y) - D(y-x)\right]\,. \tag{350}$$

We can also define the Feynman propagator $S_F(x-y)$:

$$S_F(x-y) = \langle 0|T\psi(x)\bar{\psi}(y)|0\rangle \equiv \begin{cases} \langle 0|\psi(x)\bar{\psi}(y)|0\rangle & \text{for} \quad x^0 > y^0\,, \\ -\langle 0|\bar{\psi}(y)\psi(x)|0\rangle & \text{for} \quad x^0 < y^0\,. \end{cases} \tag{351}$$

Equivalently, we can write

$$S_F(x-y) = i\int \frac{d^4p}{(2\pi)^4} \frac{p\!\!\!/ + m}{p^2 - m^2 + i\epsilon} e^{-ip\cdot(x-y)}\,. \tag{352}$$

The Feynman propagator S_F is a Green's function for the Dirac operator, *i.e.*, $(i\partial\!\!\!/_x - m)S_F(x-y) = i\delta^{(4)}(x-y)$.

Exercise 12.6: Convince yourself that the minus sign in the Feynman propagator (351) is necessary for Lorentz invariance.

Exercise 12.7: Show Eq. (352).

12.4 Feynman Rules for Fermions

Having established the quantisation of the Dirac field, let us now consider interactions. In particular, we will consider an interaction between a Dirac fermion and a real scalar field. The interaction is governed by the so-called Yukawa theory,

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}\mu^2\phi^2 + \bar{\psi}(i\partial\!\!\!/ - m)\psi - \lambda\phi\bar{\psi}\psi\,, \tag{353}$$

where the first two terms describe the real scalar field (of mass μ), the third term is for the Dirac fermion (of mass m), and the last term is the Yukawa interaction term. We will again refer to the ϕ particles as mesons and the ψ particles as nucleons; still, this is not exactly correct.

Before we consider any specific scattering processes, let us have a look at the dimensions of the fields/couplings. The scalar field has mass-dimension one, $[\phi] = 1$. From the second term, we see that the mass of ϕ is of the mass-dimension one, $[\mu] = 1$. The kinetic term for ψ tells us that $[\psi] = 3/2$, and thus, $[m] = 1$. Finally, from the interaction term, we see that the coupling λ is dimensionless, $[\lambda] = 0$, unlike in the scalar Yukawa case.

For the Yukawa theory, the Feynman rules for amplitude computations are as follows:

- Fermions (scalars) are denoted by a solid (dashed) line.
- To each incoming (outgoing) fermion with momentum p and spin s, we associate a spinor $u^s(\mathbf{p})$ ($\bar{u}^s(\mathbf{p})$) with an arrow directing from left to right.
- To each incoming (outgoing) anti-fermion with momentum p and spin s, we associate a spinor $\bar{v}^s(\mathbf{p})$ ($v^s(\mathbf{p})$) with an arrow directing from right to left.
- For each vertex, we put a factor of $-i\lambda$.
- For each internal line, we put a factor of the relevant propagator, namely

$$\frac{i}{p^2 - \mu^2 + i\epsilon}, \tag{354}$$

  for scalars, and

$$\frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon}, \tag{355}$$

  for fermions.

  – It is important to note that the fermionic propagator is a 4×4 matrix; the matrix indices are contracted at each vertex, either with further propagators or with external spinors.

- Finally, we add extra minus signs for statistics.

Let us consider, as an example, the nucleon scattering process and see how the Feynman rules are applied. For the $\psi\psi \to \psi\psi$ process, the leading-order Feynman diagrams are shown in Fig. 7.

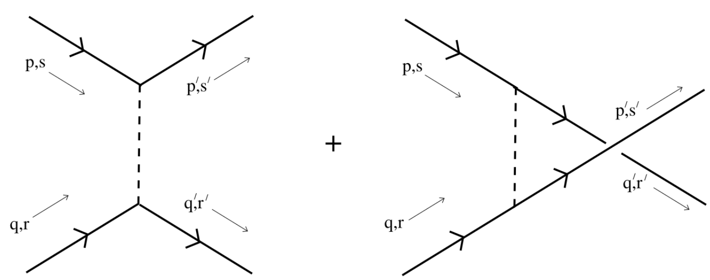


Figure 7: Feynman diagrams for nucleon-nucleon scattering

Using the aforementioned Feynman rules gives the amplitude,

$$\mathcal{A} = (-i\lambda)^2 \left(\frac{\left[\bar{u}^{s'}(\mathbf{p}') \cdot u^s(\mathbf{p})\right]\left[\bar{u}^{r'}(\mathbf{q}') \cdot u^r(\mathbf{q})\right]}{(p - p')^2 - \mu^2} - \frac{\left[\bar{u}^{s'}(\mathbf{p}') \cdot u^r(\mathbf{q})\right]\left[\bar{u}^{r'}(\mathbf{q}') \cdot u^s(\mathbf{p})\right]}{(p - q')^2 - \mu^2} \right), \tag{356}$$

where we see that the second diagram has picked up a minus sign due to statistics; note that the second diagram is the same as the first one except that the outgoing legs are crossed. We have used the momentum conservation at the vertex to find the momentum for the meson propagator.

Exercise 12.8: Compute the amplitude for the $\psi\bar{\psi} \to \psi\bar{\psi}$ process.

※ Lecture 13 (2024.06.21)

13.1 Electromagnetic Field

We have so far discussed real/complex scalar fields and the Dirac field which, after quantisation, give rise to spin-0 bosonic particles and spin-1/2 fermionic particles, respectively. However, none of them is fit for the description of light which we know exists in our Universe. We now discuss the theory of light and how light may interact with matter fields.

We already know the classical theory of light: Maxwell's equations. The Lagrangian is given, in the absence of any sources, by

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \,, \tag{357}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength. The equations of motion are then given by

$$\partial_\mu F^{\mu\nu} = 0 \,. \tag{358}$$

If we define $A^\mu = (\phi, \mathbf{A})$, then the electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$ are defined by

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \,, \quad \mathbf{B} = \nabla \times \mathbf{A} \,. \tag{359}$$

In these notations, the field strength can be written as

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix} . \tag{360}$$

From the definition of $F_{\mu\nu}$, we see that it satisfies

$$\partial_\rho F_{\mu\nu} + \partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} = 0 \,, \tag{361}$$

which is called the Bianchi identity. In terms of $\mathbf{E}$ and $\mathbf{B}$, we get

$$\nabla \cdot \mathbf{B} = 0 \,, \quad \frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} \,, \tag{362}$$

which are two of the familiar forms of Maxwell's equations. The remaining two Maxwell's equations,

$$\nabla \cdot \mathbf{E} = 0 \,, \quad \frac{\partial \mathbf{E}}{\partial t} = \nabla \times \mathbf{B} \,, \tag{363}$$

come from the equations of motion.

Exercise 13.1: Show Eq. (358).

Exercise 13.2: Verify Eqs. (362) and (363).

The vector field A_μ is called the gauge field, and it is massless in this case. As it has 4 components, we may feel tempted to say that the field has 4 degrees of freedom. However, light, or the photon, has only 2 degrees of freedom (their 2 polarisation states). To understand what is going on here, we first note that A_0 has no kinetic term, $\dot{A}_0$, in the Lagrangian. In other words, it is not dynamical. A_0 is not independent, and we are left with 3 degrees of freedom now.[19]

Next, let us consider the transformation

$$A_\mu(x) \to A_\mu(x) + \partial_\mu \lambda(x)\,, \tag{366}$$

where $\lambda(x)$ is any function that dies off reasonably quickly at $\mathbf{x} \to \infty$. Under the transformation,

$$F_{\mu\nu} \to \partial_\mu(A_\nu + \partial_\nu \lambda) - \partial_\nu(A_\mu + \partial_\mu \lambda) = F_{\mu\nu}\,. \tag{367}$$

Therefore, we see that under the transformation, the Lagrangian is invariant. This symmetry is called a gauge symmetry. Two states that are related by a gauge symmetry are the same physical states.[20] In other words, a gauge symmetry is to be viewed as a redundancy in our theory. This interpretation would have been problematic if A_μ were a physical, observable object. However, if A_μ and $A_\mu + \partial_\mu \lambda$ correspond to the same physical state, then we are okay. As the gauge symmetry is a redundancy, we can pick one representative, *i.e.*, we can pick one particular λ. We call this process choosing a gauge, and different representative configurations of a physical state are called different gauges. Two popular gauge choices are Lorenz (not Lorentz) gauge, $\partial_\mu A^\mu = 0$ and Coulomb gauge, $\nabla \cdot \mathbf{A} = 0$. Choosing a particular gauge will remove one more degree of freedom, leaving 2 degrees of freedom in the end.

13.2 Quantisation of the Electromagnetic Field

We are now ready to quantise the (free) Maxwell theory. We will work in the Lorenz gauge. In this gauge, the equations of motion, $\partial_\mu F^{\mu\nu} = 0$, become $\partial_\mu \partial^\mu A^\nu = 0$ since $\partial_\mu A^\mu = 0$. We can

[19] From the equation of motion $\nabla \cdot \mathbf{E} = 0$, we see that

$$\nabla^2 A_0 + \nabla \cdot \frac{\partial \mathbf{A}}{\partial t} = 0\,, \tag{364}$$

whose solution is given by

$$A_0(\mathbf{x}) = \int d^3\mathbf{x}' \, \frac{1}{4\pi|\mathbf{x} - \mathbf{x}'|} \nabla \cdot \dot{\mathbf{A}}(\mathbf{x}')\,. \tag{365}$$

Clearly, A_0 is not independent.

[20] This statement has a caveat. In electromagnetism, the gauge symmetry $\psi \to e^{ie\lambda(x)}\psi$ leads to the conservation of electric charge because amongst the infinite number of gauge symmetries, there is also a single global symmetry, $\lambda(x) = const$. This is a true symmetry of the system, and it takes us to another physical state.

tweak our Lagrangian in such a way that it gives rise to $\partial_\mu \partial^\mu A^\nu = 0$ directly as its equations of motion. We can do so with[21]

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} (\partial_\mu A^\mu)^2 \,. \tag{369}$$

The equations of motion are then

$$0 = \partial_\mu F^{\mu\nu} + \partial^\nu (\partial_\mu A^\mu) = \partial_\mu \partial^\mu A^\nu \,, \tag{370}$$

as desired.

So, our plan is to quantise the theory (369) and later impose the constraint $\partial_\mu A^\mu = 0$. We start by computing the momenta π^μ conjugate to A_μ:

$$\pi^0 = \frac{\partial \mathcal{L}}{\partial \dot{A}_0} = -\partial_\mu A^\mu \,, \quad \pi^i = \frac{\partial \mathcal{L}}{\partial \dot{A}_i} = \partial^i A^0 - \dot{A}^i \,. \tag{371}$$

We then promote the classical fields A_μ and the conjugate momenta π^μ to operators with the usual commutation relations,

$$[A_\mu(\mathbf{x}), \pi_\nu(\mathbf{y})] = i \eta_{\mu\nu} \delta^{(3)}(\mathbf{x} - \mathbf{y}) \,, \quad \text{and} \quad 0 \quad \text{otherwise} \,. \tag{372}$$

Next, we express the fields and the momenta in terms of the creation and annihilation operators,

$$A_\mu(\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2|\mathbf{p}|}} \sum_{\lambda=0}^{3} \epsilon_\mu^\lambda(\mathbf{p}) \left[a_\mathbf{p}^\lambda e^{+i\mathbf{p}\cdot\mathbf{x}} + a_\mathbf{p}^{\lambda\dagger} e^{-i\mathbf{p}\cdot\mathbf{x}} \right] \,, \tag{373}$$

$$\pi^\mu(\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} i \sqrt{\frac{|\mathbf{p}|}{2}} \sum_{\lambda=0}^{3} (\epsilon^\lambda)^\mu(\mathbf{p}) \left[a_\mathbf{p}^\lambda e^{+i\mathbf{p}\cdot\mathbf{x}} - a_\mathbf{p}^{\lambda\dagger} e^{-i\mathbf{p}\cdot\mathbf{x}} \right] \,. \tag{374}$$

Here, $\epsilon^\lambda(\mathbf{p})$ are four polarisation 4-vectors; we pick ϵ^0 ($\epsilon^{1,2,3}$) to be timelike (spacelike) with the normalisation

$$\epsilon^\lambda \cdot \epsilon^{\lambda'} = \eta^{\lambda\lambda'} \,. \tag{375}$$

We will choose ϵ^1 and ϵ^2 to lie transverse to the momentum,

$$\epsilon^1 \cdot p = 0 = \epsilon^2 \cdot p \,, \tag{376}$$

[21] In fact, we can consider a bit more general Lagrangian,

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2\alpha} (\partial_\mu A^\mu)^2 \,, \tag{368}$$

with arbitrary α. The quantisation of the theory is independent of α, and different choices of α are sometimes also referred to as different gauges. We are using $\alpha = 1$, which is called the Feynman gauge.

and ϵ^3 to be the longitudinal polarisation. For example, if $p \sim (1,0,0,1)$, *i.e.*, the momentum lies along the x^3 direction, then

$$\epsilon^0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \epsilon^1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \epsilon^2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \epsilon^3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{377}$$

The commutation relations indicate that

$$[a^\lambda_{\mathbf{p}}, a^{\lambda'\dagger}_{\mathbf{p}'}] = -\eta^{\lambda\lambda'}(2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{p}'). \tag{378}$$

Note that, for spacelike $\lambda, \lambda' = 1, 2, 3$, we get

$$[a^\lambda_{\mathbf{p}}, a^{\lambda'\dagger}_{\mathbf{p}'}] = \delta^{\lambda\lambda'}(2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{p}'), \tag{379}$$

while for the timelike $\lambda, \lambda' = 0$, we have

$$[a^0_{\mathbf{p}}, a^{0\dagger}_{\mathbf{p}'}] = -(2\pi)^3\delta^{(3)}(\mathbf{p}-\mathbf{p}'), \tag{380}$$

with a minus sign.[22] This minus sign leads to the same problem of the negative norm we have already encountered when trying to quantise the Dirac theory.

Exercise 13.3: Show Eq. (378).

We should remember that we have not yet imposed the constraint equation, $\partial_\mu A^\mu = 0$. We will see that it will remove the timelike, negative-norm states, solving the problem. We will also see that this will reduce the degrees of freedom down to 2. Working in the Heisenberg picture, where $\partial_\mu A^\mu = 0$ makes sense as an operator equation, we could try to impose the condition on operators. The problem is that, in this case $\pi^0 = 0$, and thus, the commutation relation, $[A_\mu(\mathbf{x}), \pi_\nu(\mathbf{y})] = i\eta_{\mu\nu}\delta^{(3)}(\mathbf{x}-\mathbf{y})$ cannot be satisfied. Instead, we could try to impose the condition on the Hilbert space,

$$\partial_\mu A^\mu|\Psi\rangle = 0. \tag{381}$$

The problem is that, if we use this condition, not even the vacuum $|0\rangle$ becomes a physical state. To see what it means, let us write $A_\mu(x) = A^+_\mu(x) + A^-_\mu(x)$ where

$$A^+_\mu(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2|\mathbf{p}|}}\sum_{\lambda=0}^{3}\epsilon^\lambda_\mu(\mathbf{p})a^\lambda_{\mathbf{p}}e^{-ip\cdot x}, \tag{382}$$

[22] This minus sign can be traced back to the wrong sign of the kinetic term for A_0 in our Lagrangian: $\mathcal{L} = \frac{1}{2}\dot{\mathbf{A}}^2 - \frac{1}{2}\dot{A}_0^2 + \cdots$

$$A_\mu^-(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2|\mathbf{p}|}} \sum_{\lambda=0}^{3} \epsilon_\mu^\lambda(\mathbf{p}) a_\mathbf{p}^{\lambda\dagger} e^{+ip\cdot x} \,. \tag{383}$$

Then, $A_\mu^+|0\rangle = 0$ automatically, but $\partial^\mu A_\mu^-|0\rangle \neq 0$.

As a final attempt, let us ask that physical states $|\Psi\rangle$ are defined by

$$\partial^\mu A_\mu^+ |\Psi\rangle = 0 \,, \tag{384}$$

which is known as the Gupta-Bleuler condition. This ensures that

$$\langle \Psi' | \partial_\mu A^\mu | \Psi \rangle = 0 \,, \tag{385}$$

so that the operator $\partial_\mu A^\mu$ has vanishing matrix elements between physical states. The physical states $|\Psi\rangle$ span a physical Hilbert space $\mathcal{H}_{\text{phys}}$.

Did we solve the negative-norm problem then? Consider a basis of states for the Fock space. We can decompose any element of this basis as $|\Psi\rangle = |\psi_T\rangle|\phi\rangle$, where $|\psi_T\rangle$ contains only transverse photons, created by $a_\mathbf{p}^{1\dagger}$ and $a_\mathbf{p}^{2\dagger}$, while $|\phi\rangle$ contains the timelike photons created by $a_\mathbf{p}^{0\dagger}$ and longitudinal photons created by $a_\mathbf{p}^{3\dagger}$. The Gupta-Bleuler condition requires that

$$(a_\mathbf{p}^0 - a_\mathbf{p}^3)|\phi\rangle = 0 \,. \tag{386}$$

It means that the physical states must contain combinations of timelike and longitudinal photons. In general, $|\phi\rangle$ will be a linear combination of states $|\phi_n\rangle$ containing n pairs of timelike and longitudinal photons. So, we may write

$$|\phi\rangle = \sum_{n=0}^{\infty} C_n |\phi_n\rangle \,, \tag{387}$$

with $|\phi_0\rangle = |0\rangle$. Although the condition $(a_\mathbf{p}^0 - a_\mathbf{p}^3)|\phi\rangle = 0$ decouples the negative-norm states, all the remaining states involving timelike and longitudinal photons have zero norm, $\langle\phi_m|\phi_n\rangle = \delta_{n0}\delta_{m0}$. Thus, we still need to deal with all these zero-norm states.

We deal with the zero-norm states by treating them as gauge equivalent to the vacuum. Two states that differ in their timelike and longitudinal photon content, $|\phi_n\rangle$ with $n \geq 1$, are said to be physically equivalent. This treatment would be okay if they give the same expectation value for all physical observables. Let us see if this is true. In particular, we can check that this is true for the Hamiltonian,

$$H = \int \frac{d^3\mathbf{p}}{(2\pi)^3} |\mathbf{p}| \left(\sum_{i=1}^{3} a_\mathbf{p}^{i\dagger} a_\mathbf{p}^i - a_\mathbf{p}^{0\dagger} a_\mathbf{p}^0 \right) \,. \tag{388}$$

The Gupta-Bleuler condition ensures that $\langle\Psi|a_{\mathbf{p}}^{3\dagger}a_{\mathbf{p}}^{3}|\Psi\rangle = \langle\Psi|a_{\mathbf{p}}^{0\dagger}a_{\mathbf{p}}^{0}|\Psi\rangle$. Thus, the contributions from the timelike and longitudinal photons cancel amongst themselves in the Hamiltonian. Furthermore, this renders the Hamiltonian positive-definite, leaving us just with the contribution from the transverse photons.

Before we move on, let us talk about the propagator. In the Lorenz gauge, it is given by

$$\langle 0|TA_\mu(x)A_\nu(y)|0\rangle = \int \frac{d^4p}{(2\pi)^4}\frac{-i\eta_{\mu\nu}}{p^2+i\epsilon}e^{-ip\cdot(x-y)}\,. \tag{389}$$

Exercise 13.4: Show Eq. (389).

13.3 Quantum Electrodynamics

We now want to build an interacting theory of light and matter; we want to write down a Lagrangian which couples the gauge field A_μ to some scalars or spinors. In particular, the theory of electromagnetism coupled to fermions can be described by the Lagrangian,

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - e\bar{\psi}\gamma^\mu A_\mu\psi\,. \tag{390}$$

Here, e is a coupling constant which has the interpretation of the electric charge of the fermion ψ. We may write this Lagrangian as

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi\,, \tag{391}$$

where $D_\mu = \partial_\mu + ieA_\mu$ is called the covariant derivative. Note that this Lagrangian is invariant under the gauge transformations

$$A_\mu \longrightarrow A_\mu + \partial_\mu\lambda\,, \quad \psi \longrightarrow e^{-ie\lambda}\psi\,, \tag{392}$$

for an arbitrary function $\lambda(x)$. This theory, the theory of electrons interacting with light, is called quantum electrodynamics (QED).

Exercise 13.5: Show the invariance of the Lagrangian under the gauge transformation (392).

The Feynman rules for QED are as follows (note that we work in the Lorenz gauge):

- We use wiggle (solid) lines to describe photons (fermions).
- For each vertex (two solid lines and one wiggle line), we put $-ie\gamma^\mu$.
- For the photon propagator, we put

$$-\frac{i\eta_{\mu\nu}}{p^2+i\epsilon}\,, \tag{393}$$

and for the fermion propagator, we put

$$\frac{i(\not{p}+m)}{p^2-m^2+i\epsilon}\,. \tag{394}$$

- For incoming (outgoing) photon external lines, we add a polarisation vector $\epsilon^{\mu}_{\text{in}}$ ($\epsilon^{\mu}_{\text{out}}$).
- For incoming (outgoing) fermion external lines, we add a spinor u^r ($\bar{u}^r$).
- For incoming (outgoing) anti-fermion external lines, we add a spinor $\bar{v}^r$ (v^r).

As an example, let us consider the electron scattering process, $e^-e^- \to e^-e^-$. Two leading-order Feynman diagrams are shown in Fig. 8. Applying the QED Feynman rules, we find that

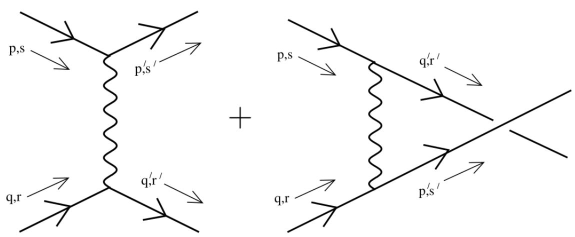


Figure 8: Feynman diagrams for electron-electron scattering

the amplitude is given by

$$-i(-ie)^2\left(\frac{[\bar{u}^{s'}(\mathbf{p}')\gamma^{\mu}u^{s}(\mathbf{p})][\bar{u}^{r'}(\mathbf{q}')\gamma_{\mu}u^{r}(\mathbf{q})]}{(p'-p)^2}-\frac{[\bar{u}^{s'}(\mathbf{p}')\gamma^{\mu}u^{r}(\mathbf{q})][\bar{u}^{r'}(\mathbf{q}')\gamma_{\mu}u^{s}(\mathbf{p})]}{(p-q')^2}\right)\,. \tag{395}$$

※ Appendix – Poisson Brackets and Commutators

In Classical Mechanics, the Poisson bracket of two variables, $A(q,p)$ and $B(q,p)$, which are functions of the canonical coordinates q and momenta p, is defined as

$$\{A,B\} \equiv \sum_i \left(\frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q_i} \right) .$$

In Quantum Mechanics, the commutator or Dirac bracket of two operators, A and B, is

$$[A,B] \equiv AB - BA .$$

Both types of brackets (denoted only by [] here) have similar algebraic properties:

- Linearity:

$$[a_1A_1 + a_2A_2, B] = a_1[A_1,B] + a_2[A_2,B]$$

$$[A, b_1B_1 + b_2B_2] = b_1[A,B_1] + b_2[A,B_2]$$

- Antisymmetry:

$$[A,B] = -[B,A]$$

- Leibniz rules:

$$[AB,C] = A[B,C] + [A,C]B$$

$$[A,BC] = B[A,C] + [A,B]C$$

- Jacobi identity:

$$[A,[B,C]] + [B,[C,A]] + [C,[A,B]] = 0$$

Furthermore, both types of brackets involving the Hamiltonian can be used to describe the time dependence of the classical/quantum variables. In Classical Mechanics,

$$\begin{aligned} \frac{d}{dt}A(q,p) &= \sum_i \left(\frac{\partial A}{\partial q_i}\frac{dq_i}{dt} + \frac{\partial A}{\partial p_i}\frac{dp_i}{dt} \right) \\ &= \sum_i \left(\frac{\partial A}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial H}{\partial q_i} \right) \end{aligned}$$

$$= \{A, H\} ,$$

where we have used the Hamilton equations of motion. In Quantum Mechanics, from the Heisenberg-Dirac equation, we have

$$i\hbar \frac{d}{dt} \langle\psi|\hat{A}|\psi\rangle = \langle\psi|[\hat{A}, \hat{H}]|\psi\rangle ,$$

or, in the Heisenberg picture,

$$i\hbar \frac{d}{dt} \hat{A} = [\hat{A}, \hat{H}] .$$

The similarity between the classical Poisson brackets and the quantum commutators can simply be stated as follows: If classically $\{A, B\} = C$, then the corresponding operators in Quantum Mechanics should obey $[\hat{A}, \hat{B}] = i\hbar\hat{C}$.

In particular, if we have classical canonical variables q_i and p_i, then

$$\{q_i, q_j\} = 0 , \quad \{p_i, p_j\} = 0 , \quad \{q_i, p_j\} = \delta_{ij} ,$$

and the corresponding quantum operators should obey the canonical commutation relations,

$$[\hat{q}_i, \hat{q}_j] = 0 , \quad [\hat{p}_i, \hat{p}_j] = 0 , \quad [\hat{q}_i, \hat{p}_j] = i\hbar\delta_{ij} .$$

※ Appendix – Contour Integral

This appendix is based on Chapter 14 of "Mathematical Methods in the Physical Sciences" by M. Boas.

Let us consider a function $f(z)$ with a complex variable z. The function $f(z)$ is called analytic in a region of the complex plane if it has a derivative at every point in the region. For example, the statement "The function $f(z)$ is analytic at a point $z = a$" means that $f(z)$ has a derivative at every point inside some small circle about $z = a$. Note that the derivative of $f(z)$ is defined as

$$f'(z) = \frac{df}{dz} = \lim_{\Delta z \to 0} \frac{\Delta f}{\Delta z},$$

where $\Delta f = f(z + \Delta z) - f(z)$ and $\Delta z = \Delta x + i\Delta y$.

- **The Cauchy-Riemann Conditions:** If $f(z) = u(x, y) + iv(x, y)$ is analytic in a region, then

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y} \tag{396}$$

  in that region.

  – **Proof:** Consider $f = f(z)$ with $z = x + iy$. Then,

$$\frac{\partial f}{\partial x} = \frac{df}{dz}\frac{\partial z}{\partial x} = \frac{df}{dz} \quad \text{and} \quad \frac{\partial f}{\partial y} = \frac{df}{dz}\frac{\partial z}{\partial y} = i\frac{df}{dz}.$$

  Since $f = u + iv$, we get

$$\frac{\partial f}{\partial x} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} \quad \text{and} \quad \frac{\partial f}{\partial y} = \frac{\partial u}{\partial y} + i\frac{\partial v}{\partial y}.$$

  Thus, combining the above two,

$$\frac{df}{dz} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} \quad \text{and} \quad \frac{df}{dz} = -i\frac{\partial u}{\partial y} + \frac{\partial v}{\partial y}.$$

  Hence,

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}.$$

A regular point of $f(z)$ is a point at which $f(z)$ is analytic, and a singular point (or singularity) of $f(z)$ is a point at which $f(z)$ is not analytic. It is called an isolated singular point if $f(z)$ is analytic everywhere else inside some small circle about the singular point.

- **The Cauchy Theorem:** Let C be a simple[23] closed curve with a continuously turning tangent except possibly at a finite number of points; that is, we allow a finite number of corners, but otherwise the curve must be smooth. If $f(z)$ is analytic on and inside C, then

$$\oint_C f(z)dz = 0\,. \tag{397}$$

  - **Proof:** Expanding $f(z)$ and z,

$$\oint_C f(z)dz = \oint_C (u+iv)(dx+idy) = \oint_C (udx - vdy) + i\oint_C (vdx+udy)\,.$$

    Green's theorem in the plane says that, if $P(x,y)$, $Q(x,y)$, and their partial derivatives are continuous in a simply-connected region R, then

$$\oint_C Pdx + Qdy = \iint_{A_C}\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)dxdy\,,$$

    where A_C stands for the area inside C. So, if we let $P = u$ and $Q = -v$, then

$$\oint_C (udx - vdy) = \iint_{A_C}\left(-\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)dxdy\,,$$

    which is zero by the Cauchy-Riemann conditions (396). Thus, we obtain

$$\oint_C f(z)dz = i\oint_C (vdx+udy)\,.$$

    Similarly, if we let $P = v$ and $Q = u$, and apply Green's theorem,

$$\oint_C (vdx+udy) = \iint_{A_C}\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right)dxdy\,,$$

    which is again zero by the Cauchy-Riemann conditions. Therefore,

$$\oint_C f(z)dz = 0\,.$$

    Note that we have assumed that $f'(z)$ is continuous so that u, v, and their derivatives are continuous; one may prove the theorem without this assumption as well which we do not cover here.

- **Cauchy's Integral Formula:** If $f(z)$ is analytic on and inside a simple closed curve C, the value of $f(z)$ at a point $z = a$ inside C is given by the following contour integral along C:

$$f(a) = \frac{1}{2\pi i}\oint_C \frac{f(z)}{z-a}dz\,. \tag{398}$$

[23] A simple curve is a curve that does not cross itself.

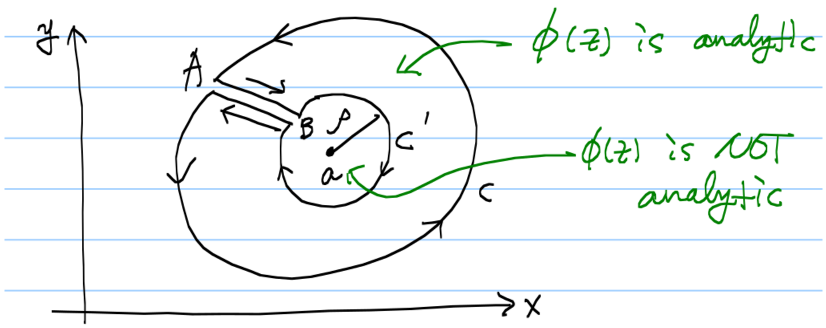


Figure 9: A cut AB between the curve C and the circle C′.

– **Proof:** Let a be a fixed point inside the simple closed curve C, and consider the function

$$\phi(z) = \frac{f(z)}{z-a}\,.$$

Let C' be a small circle inside C with centre at a and radius ρ. Now, we make a cut between C and C' along AB; see Fig. 9. We are going to integrate along the path from A, around C, to B, around C', and back to A. According to the Cauchy theorem,

$$\oint_{C_{ccw}} \phi(z)dz + \oint_{C'_{cw}} \phi(z)dz = 0\,,$$

where ccw and cw stand for counter-clockwise and clockwise, respectively. Changing C'_{cw} to C'_{ccw}, we get

$$\oint_C \phi(z)dz = \oint_{C'} \phi(z)dz\,,$$

now with counter-clockwise for both C and C'. Along the circle C', we have $z = a + \rho e^{i\theta}$ and $dz = i\rho e^{i\theta}d\theta$. So, the above equation becomes

$$\oint_C \phi(z)dz = \oint_{C'} \phi(z)dz = \oint_{C'} \frac{f(z)}{z-a}dz = \int_0^{2\pi} \frac{f(z)}{\rho e^{i\theta}} i\rho e^{i\theta}d\theta = \int_0^{2\pi} if(z)d\theta\,.$$

We now let $\rho \to 0$; that is, we let $z \to a$. Then, we have

$$\oint_C \phi(z)dz = 2\pi i f(a)\,.$$

Therefore,

$$f(a) = \frac{1}{2\pi i}\oint_C \frac{f(z)}{z-a}dz\,.$$

- **The Laurent Theorem:** Let C_1 and C_2 be two circles with centre at z_0. Let $f(z)$ be analytic in the region R between the circles. Then, $f(z)$ can be expanded in a series of the form

$$\begin{aligned} f(z) = a_0 + a_1(z - z_0) + a_2(z - z_0)^2 + \cdots \\ + \frac{b_1}{z - z_0} + \frac{b_2}{(z - z_0)^2} + \cdots \end{aligned} \tag{399}$$

which is convergent in R; the series is called Laurent's series.

  - The "a" series is a power series, and a power series converges inside some circle. The "b" series, called the principal part of the Laurent series, is a series of inverse powers of z and so converges for $|1/z| < c$, where c is some constant; thus, the "b" series converges outside some circle. Then, a Laurent series converges between two circles.
  - The coefficients are given by

$$a_n = \frac{1}{2\pi i} \oint_C \frac{f(z)}{(z - z_0)^{n+1}} dz \,, \qquad b_n = \frac{1}{2\pi i} \oint_C \frac{f(z)}{(z - z_0)^{-n+1}} dz \,, \tag{400}$$

    where C is any simple closed curve surrounding z_0 and lying in R. It is easy to verify this once using the residue theorem.
  - If all the bs are zero, $f(z)$ is analytic at $z = z_0$, and we call z_0 a regular point.
  - If $b_n \neq 0$, but all the bs after b_n are zero, $f(z)$ is said to have a pole of order n at $z = z_0$; if $n = 1$, we say that $f(z)$ has a simple pole.
  - If there are an infinite number of bs different from zero, $f(z)$ has an essential singularity at $z = z_0$.
  - The coefficient b_1 of $1/(z - z_0)$ is called the residue of $f(z)$ at $z = z_0$.

Note that most of the functions we encounter is analytic except for poles; such functions are called meromorphic functions.

- **The Residue Theorem:** The residue theorem states that, for a simple closed curve C,

$$\oint_C f(z) dz = 2\pi i \times \text{sum of the residues of } f(z) \text{ inside } C \,, \tag{401}$$

where the integral around C is in the counter-clockwise direction.

  - **Proof:** Let z_0 be an isolated singular point of $f(z)$. We would like to find

$$\oint_C f(z) dz$$

around a simple closed curve C surrounding z_0, but enclosing no other singularities. Let $f(z)$ be expanded in the Laurent series

$$f(z) = a_0 + a_1(z - z_0) + a_2(z - z_0)^2 + \cdots + \frac{b_1}{z - z_0} + \frac{b_2}{(z - z_0)^2} + \cdots$$

about z_0 that converges near $z = z_0$. By the Cauchy theorem, the integral of the "a" series is zero since this part is analytic. To evaluate the integrals of the terms in the "b" series, we use the same trick we used to derive Cauchy's integral formula. For the b_1 term, we can apply Cauchy's integral formula to get

$$\oint_C \frac{b_1}{z - z_0} dz = 2\pi i b_1 \,.$$

The integrals of all the other b_n terms are zero because, if we let

$$I \equiv \oint_C \frac{b_n}{(z - z_0)^n} dz \,,$$

and use the same trick we used to derive Cauchy's integral formula, we get

$$I = \int_0^{2\pi} \frac{b_n}{\rho^n e^{in\theta}} i\rho e^{i\theta} d\theta = i b_n \rho^{1-n} \int_0^{2\pi} e^{i\theta(1-n)} d\theta = i b_n \rho^{1-n} \frac{1}{i(1-n)} \left[e^{i(1-n)2\pi} - 1 \right] = 0 \,.$$

Therefore,

$$\oint_C f(z) dz = 2\pi i b_1 \,,$$

where b_1 is the residue of $f(z)$ at $z = z_0$. Thus, we arrive at

$$\oint_C f(z) dz = 2\pi i \times \text{residue of } f(z) \text{ at the singular point inside } C \,.$$

We can do the same thing if there are several isolated singularities inside C.

There are many ways to find residues. If it is easy to write down the Laurent series for $f(z)$ about $z = z_0$ that is valid near z_0, then the residue is just the coefficient b_1 of the term $1/(z - z_0)$. Take

$$f(z) = \frac{e^z}{z - 1}$$

as an example. Since

$$e^z = e^{z-1} e = e \left[1 + (z - 1) + \frac{(z-1)^2}{2!} + \cdots \right]$$

we get

$$\frac{e^z}{z-1} = \frac{e}{z-1} + e + \cdots$$

Thus, the residue of $f(z)$ at $z = 1$ is e.

If $f(z)$ has a simple pole at $z = z_0$, we may find the residue by multiplying $f(z)$ by $(z - z_0)$ and evaluating the result at $z = z_0$ because, in this way, all the terms in the Laurent series will vanish except b_1. Take

$$f(z) = \frac{z}{(2z+1)(5-z)}$$

as an example. We see that there are two simple poles: one at $z = 5$ and the other at $z = -1/2$. To find the residue at $z = 5$, we multiply $f(z)$ by $(z - 5)$,

$$-\frac{z}{2z+1}$$

and evaluate it at $z = 5$,

$$-\frac{5}{11}\,.$$

For $z = -1/2$, we get

$$\left[\frac{z}{(2z+1)(5-z)} \times \left(z + \frac{1}{2}\right)\right]\bigg|_{z=-1/2} = -\frac{1}{22}\,.$$

If $f(z)$ can be written as $g(z)/h(z)$, where $g(z)$ is analytic and not zero at z_0, while $h(z)$ becomes zero at z_0 and has the property of $h'(z_0) \neq 0$, then the residue at z_0 is given by

$$R(z_0) = \lim_{z\to z_0} \frac{g(z)}{h(z)}(z - z_0) = g(z_0) \lim_{z\to z_0} \frac{z - z_0}{h(z)} = g(z_0) \lim_{z\to z_0} \frac{1}{h'(z)} = \frac{g(z_0)}{h'(z_0)}\,,$$

where we have used L'Hôpital's rule.

When $f(z)$ has a pole of order n, we can use the following method: Multiply $f(z)$ by $(z-z_0)^m$, where m is an integer greater than or equal to the order n of the pole, differentiate the result $m - 1$ times, divide by $(m - 1)!$, and evaluate the resulting expression at $z = z_0$. As an example, consider

$$f(z) = \frac{z \sin z}{(z-\pi)^3}\,.$$

What is $R(\pi)$? Using the method stated above,

$$R(\pi) = \left[\frac{1}{2!}\frac{d^2}{dz^2}\left(f(z)(z-\pi)^3\right)\right]\bigg|_{z=\pi} = \left[\frac{1}{2!}\frac{d^2}{dz^2}(z \sin z)\right]\bigg|_{z=\pi} = \frac{1}{2}\left(2\cos z - z \sin z\right)\bigg|_{z=\pi} = -1\,.$$

Using the residue theorem, we can evaluate definite integrals of a function that has poles. Let us have a look at some examples. First, consider

$$I = \int_0^{2\pi} \frac{d\theta}{5 + 4\cos\theta} .$$

Letting $z = e^{i\theta}$, we get, since $dz = iz d\theta$ and $\cos\theta = (e^{i\theta} + e^{-i\theta})/2 = (z + 1/z)/2$,

$$I = \oint_C \frac{\frac{1}{iz} dz}{5 + 2z + 2/z} = \frac{1}{i}\oint_C \frac{dz}{2z^2 + 5z + 2} = \frac{1}{i}\oint_C \frac{dz}{(2z+1)(z+2)} .$$

The pole is at $z = -1/2$. Using the residue theorem, we get

$$I = \frac{1}{i} \times 2\pi i R\left(-\frac{1}{2}\right) .$$

The residue is given by

$$R\left(-\frac{1}{2}\right) = \frac{1}{(2z+1)(z+2)}\left(z + \frac{1}{2}\right)\bigg|_{z=-1/2} = \frac{1}{2(z+2)}\bigg|_{z=-1/2} = \frac{1}{3} .$$

Thus,

$$I = \frac{2\pi}{3} .$$

Next, consider the following integral:

$$I = \int_0^\infty \frac{r^{p-1}}{1+r} dr , \quad 0 < p < 1 .$$

Let us first try to evaluate

$$\oint_C \frac{z^{p-1}}{1+z} dz , \quad 0 < p < 1 ,$$

with $z = re^{i\theta}$. Consider the $p = 1/2$ case. Choose, for example, $\theta = \pi/4$. Then, we have

$$z^{-1/2} = r^{-1/2} e^{-i\pi/8} .$$

If we now increase θ by 2π, we get

$$z = re^{i(\pi/4 + 2\pi)} \Rightarrow z^{-1/2} = -r^{-1/2} e^{-i\pi/8} .$$

We see that, for any starting point (with $r \neq 0$), $z^{-1/2}$ comes back at a different value (*i.e.*, a different branch) when θ increases by 2π. This is same for z^{p-1}. Thus, we must decide on some

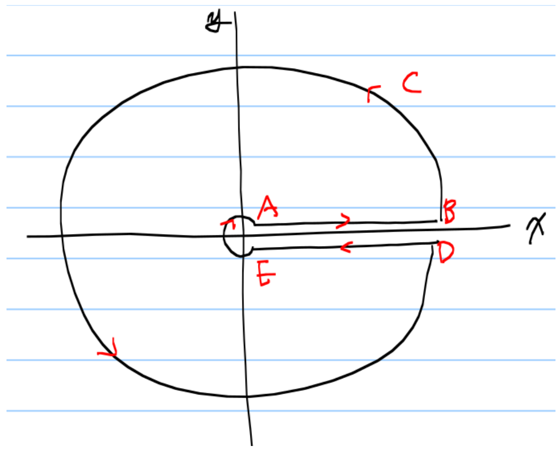


Figure 10: An example for a branch cut and a branch point.

interval of length 2π for θ in order to have a single-valued function. Let us agree to restrict θ to the value of 0 to 2π in evaluating the contour integral,

$$\oint_C \frac{z^{p-1}}{1+z} dz\,, \quad 0 < p < 1\,.$$

We may imagine an artificial barrier or cut along the positive x axis, which is called the branch cut; see Fig. 10. A point which we cannot encircle without crossing a branch cut is called a branch point. In our example, the origin is a branch point. Now, the factor z^{p-1} is a single-valued function. Inside the closed curve C, $z^{p-1}/(1+z)$ is an analytic function except for the pole at $z = -1 = e^{i\pi}$. The residue at the pole is given by

$$R(-1) = \frac{z^{p-1}}{1+z}(z+1)\bigg|_{z=-1} = (-1)^{p-1} = e^{i\pi(p-1)} = -e^{ip\pi}\,.$$

Thus, using the residue theorem, we get

$$\oint_C \frac{z^{p-1}}{1+z} dz = 2\pi i R(-1) = -2\pi i e^{i\pi p}\,,$$

for $0 < p < 1$. The contour C can be divided into 4 sections; the big circle, the small circle, the line AB, and the line DE. For both the big and small circles, we can write $z = re^{i\theta}$, and the integral is given by

$$\int \frac{r^{p-1}e^{i(p-1)\theta}}{1+re^{i\theta}} rie^{i\theta} d\theta = i\int \frac{r^p e^{ip\theta}}{1+re^{i\theta}} d\theta\,.$$

Note that this goes to 0 if we send r to 0. It also goes to 0 if we send r to ∞. Thus, the integrals along the circular parts of the contour tend to zero as the little circle shrinks to a point and the large circle expands indefinitely. We are then left with the two integrals along the positive x

axis with AB extending from 0 to ∞ and DE from ∞ to 0. Along AB, we have $\theta = 0$ and $z = r$. Thus,

$$\int_0^\infty \frac{r^{p-1}}{1+r} dr \,.$$

Along DE, we have $\theta = 2\pi$ and $z = re^{2\pi i}$. Thus,

$$\int_\infty^0 \frac{r^{p-1}e^{2\pi i p}}{1+re^{2\pi i}} dr = -\int_0^\infty \frac{r^{p-1}e^{2\pi i p}}{1+r} dr \,.$$

Adding everything gives

$$\left(1 - e^{2\pi i p}\right) \int_0^\infty \frac{r^{p-1}}{1+r} dr = -2\pi i e^{i\pi p} \,.$$

Then, the integral we wanted to evaluate is given by

$$\int_0^\infty \frac{r^{p-1}}{1+r} dr = -\frac{2\pi i e^{i\pi p}}{1-e^{2\pi i p}} = \frac{2\pi i}{e^{i\pi p} - e^{-\pi i p}} = \frac{\pi}{\sin(\pi p)} \,.$$