Quantum Field Theory

Damon Binder

19 June 2017

Books and Other Resources

Many books exist on quantum field theory and its applications. Here I mention just the ones that are most helpful as an introduction to the field, and which will best supplement these lecture notes.

Peskin & Schroeder, *An Introduction to Quantum Field Theory*

This is the gold standard for learning quantum field theory. In spite of its name, it contains enough material for three semesters worth of courses, and covers many advanced topics. It is well written and computationally focused. Unfortunately it introduces the path-integral rather late, instead focussing on canonical quantization.

Zee, *Quantum Field Theory in a Nutshell*

A fun introduction to the subject, with emphasis on physical insight and conceptual understanding rather than on computational mastery. The author does not let the truth get in the way of a good story. Uses the path-integral throughout.

Srednicki, *Quantum Field Theory*

Another path-integral oriented book, this text aims to get to loop diagrams and the mathematical structure of QFT quickly, with little emphasis on physical applications until much later in the book.

Schwartz, *Quantum Field Theory and the Standard Model*

A recent text using the canonical approach and focussed on applications of quantum field theory to particle physics.

Online Resources

A lot of online resources exist on quantum field theory. Three good course notes for other introductory QFT courses are:

- David Tong, `http://www.damtp.cam.ac.uk/user/tong/qft.html`.

- Michael Luke, `https://www.physics.utoronto.ca/~luke/PHY2403F/References.html`.

- Sidney Coleman, `https://arxiv.org/abs/1110.5013`.

Other online resources worth checking out:

- Quantum Condensed Matter Field Theory, by Ben Simons
  `http://www.tcm.phy.cam.ac.uk/~bds10/tp3.html`.

- Fields, by Warren Siegal
  `http://insti.physics.sunysb.edu/~siegel/errata.shtml`.

Acknowledgements

Most of my quantum field theory knowledge came from the textbooks and course notes listed on the previous page. These notes are heavily indebted the those resources. My main job has been to steal the good explanations from each text, and then synthesize into a coherent form.

I wrote these notes for Cédric Simenel and Joseph Hope, to form the basis of their quantum field theory course. I would like to thank them for their help with this project, and also for teaching me much of what I now know about physics.

Contents

Introduction

A field is something that takes a value at every point in space and time. For instance, if you are studying a metal rod, you may consider a 'temperature field' $T(x, t)$ which tells you the temperature at a given location x and time t. This field evolves according to the *heat equation:*

$$\frac{\partial T}{\partial t} = -\alpha \frac{\partial^2 T}{\partial x^2}.$$

(1.1)

We can likewise describe waves — whether they be acoustic waves, water waves, or light — via the wave equation:

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = \frac{\partial^2 \phi}{\partial x^2}.$$

(1.2)

Most of the fields you have encountered in your physics career thus far are phenomenological in nature. Water waves and temperature fields are not fundamental descriptions of reality. Instead they are useful approximations to describing macroscopic systems. So whilst a glass of water is made of $\sim 10^{23}$ particles, fluid mechanics is a very useful tool for describing this system.

When Maxwell first wrote down his famous equations, he thought of the electric and magnetic fields as propagating through an 'aether', analogous to sound waves through the air. It would take almost fifty years for this mistake to be rectified, but physicists would slowly realise that the 'aether' had to be incompressible, transparent, non-dispersive — almost like it wasn't there at all!

As you are no doubt aware, the nail in the coffin was special relativity. In our modern understanding, the electromagnetic field is a fundamental field — there really is an electric and magnetic field at every point in spacetime. As we will learn in this course, there are actually quite a number of fields permeating the universe. But these fields behave not classically but quantum mechanically. When we quantize a field, we discover that they give rise to discrete excitations called 'particles'. The electromagnetic field gives rise to the photon. Similarly, there is an electron field for electrons, proton fields for protons (which in turn is composed of quark fields), and so on for all particles in the universe.

1.1 Why QFT?

You may be wondering why we need quantum field theory. After all, we have been using particles throughout our physics life, first in classically and then quantum mechanically. It turns out however that the particle view of the world is simply incompatible with special relativity.

In special relativity, physics must be local. Particles simply cannot talk to each other unless they are on top of each other, which never happens because particles are infinitely

small points.[1] In electrodynamics this problem is fixed by introducing the electromagnetic field: electrons talk to the field, which in turn propagates to the other electrons. If we want a non-trivial theory, we need to introduce a field.

The problems with particles in electrodynamics persist however:

1. Electrodynamics is plagued by pathologies, such as pre-acceleration solutions where a particle accelerates *before* a force is applied to it. These problems persist if we attempt to quantize, leading to violations of causality in the quantum theory.

2. In electrodynamics, the number of electrons cannot change. This persists if we attempt to quantize the theory, as no matter how hard you try, the Schrödinger equation cannot allow particle creation or destruction. This is in contradiction to observations such as electron-positron annihilation: $e^+ + e^- \to 2\gamma$. Theoretically, we see this relates to the famous equation $E = mc^2$ — if we have enough energy, nothing stops us creating or destroying electron-positron pairs.

3. Electrodynamics treats electrons and the electromagnetic field as fundamentally different. Yet from the wave-particle duality, we know that electrons have wave-like properties, and that photons can act like particles. It therefore seems perverse to treat them differently.

By introducing a field to describe the electrons, quantum field theory is able to overcome all of these problems.

1.2   What is QFT good for?

Everything. No seriously, everything. As we have seen, for relativistic systems, quantum field theory is necessary. Such is the case for particle physics, where we can summarise our knowledge of the universe (except, famously, gravity) in a single QFT — the Standard Model. To calculating anything, from the decay of the Higgs boson to the mass of the proton, requires quantum field theory. Deep down, everything is made of quantum fields.

But quantum field theory is also a useful tool in the non-relativistic world. In condensed matter physics, rather than solving the Schrödinger equation for 10^{23} particles, we are able to simplify problems through the introduction of fields. This is critical in the understanding of such phenomena as superconductivity, Bose-Einstein condensation, and the behaviour of metals.

Furthermore, we shall see later in this course, quantum field theory and statistical mechanics can be seen as different facets of the same subject. If you want to understand the behaviour of magnets, or surface growth, or the liquid-gas phase transition, that too requires quantum field theory.

1.3   Notation and Conventions

1.3.1   Natural Units

As grown ups, we work in natural units with $\hbar = c = k = 1$. In this system

$$[\text{time}] = [\text{length}] = [\text{energy}]^{-1} = [\text{mass}]^{-1} = [\text{temperature}]^{-1}. \tag{1.3}$$

---

[1] Rigid bodies do not exist in special relativity, so any particle must be infinitely small.

Quantity	Conversion to eV	Results
Height	$l \times (\hbar c)^{-1}$	1.1×10^{-7} eV^{-1}
Age	$t \times \hbar^{-1}$	9.1×10^{23} eV^{-1}
Mass	$m \times c^2$	4.0×10^{37} eV
Temperature	$T \times k$	2.6×10^{-2} eV

Table 1.1: The author, in natural units.

A particle of mass m therefore has energy $mc^2 = m$, and Compton wavelength $\hbar/mc = m^{-1}$.

All dimensionful quantities can be measured on a single scale, which we choose to be the energy. In particle physics this is measured in the electron volt (eV). For everyday quantities the electrovolt is not the most useful unit, as you can see in *Table 1.1*. When considering the typical scales involved in particle physics though, the electron volt is the unit of choice.

It is very useful to know the energy scales associated with various natural phenomena. The biggest scale we ever use is the distance to the cosmological horizon ~ 45 billion light years; in natural units this is $\sim 10^{33}$ eV^{-1}. On the other hand, the smallest scale we talk about is the length scale where quantum gravity becomes important; this is known as the Planck scale $\sim 10^{28}$ eV. The associated length scale, 10^{-28} eV^{-1} is often suggested to be the "smallest" length that it makes sense to talk about, due to quantum gravitational effects. This, however, is speculative, as we cannot yet observe the required energies experimentally.

All other energy scales fall somewhere in between these two extremes. In *Table 1.2*, you can find representative energies from a range of fields in physics. For thermal and atomic physics, the electron volt is a logical unit to use. For nuclear physics, we would usually work in MeVs instead, and in particle physics, GeVs. The LHC, the most powerful particle accelerator ever built, collides protons together with total energy 13 TeV.

Field	Quantity	Value (eV)
Cosmology	(Age of Universe)$^{-1}$	1.5×10^{-33}
Thermal	Room Temperature	2.6×10^{-2}
Atomic	Ground state energy of H	1.3×10^1
Nuclear	Energy per nucleon of ^{238}U	7.6×10^6
Strong Force	Pion Mass	1.4×10^8
Weak Force	Higgs Mass	1.3×10^{11}
Gravity	Planck Scale	10^{28}

Table 1.2: Energy Scales of the Universe

1.3.2   Other Conventions

As a general rule, we will use the notations and conventions use in *Peskin & Schroeder*. We will use the metric tensor

$$g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{1.4}$$

Greek letters, $\mu, \nu, \lambda, ...$ will be used for space-time indices, where as Latin letters $i, j, k, ...$ will be used only for spatial indices. Four-vectors will be written as p^μ, and we will bold three-vectors; for example, $p^\mu = (p^0, \mathbf{p})$.

For Fourier transforms, we will use the convention

$$f(x) = \int \frac{d^4 k}{(2\pi)^4} e^{-ikx} f(k)$$

$$f(k) = \int d^4 x \; e^{ikx} f(x). \tag{1.5}$$

This convention is different to the one commonly used in quantum mechanics, so be careful!

Classical Field Theory

In classical mechanics, we describe a system by a series of variables $q_a(t)$. Here a is an index, labelling the coordinates $q_1, q_2, ..., q_n$. The fundamental aim of classical mechanics is to compute the evolution of a system over time, and this can be achieved using the principle of least action. Every system has a Lagrangian $L(q_a, \dot{q}_a, t)$, and system follows the trajectory that minimizes the action

$$S[q_a] = \int L(q_a, \dot{q}_a, t) \ dt. \tag{2.1}$$

Using variational calculus, we can derive the Euler-Lagrange equations:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_a} - \frac{\partial L}{\partial q_a} = 0, \tag{2.2}$$

giving us the equations of motion for our system.

Classical field theory is a simple generalization of classical mechanics. The main difference is that now we are interested not in coordinates $q_a(t)$ which depend only on time, but instead fields $\phi_a(\vec{x}, t)$ which are functions of both space and time. The action is now given

$$S[\phi_a] = \int \ dt \ L[\phi_a, \partial_\mu \phi_a] = \int \ dt \ d^3x \ \mathcal{L}(\phi_a(x, t), \partial_\mu \phi_a). \tag{2.3}$$

where $\mathcal{L}$ is technically the *Lagrangian density*, but almost always simply called the Lagrangian. The dynamics of our fields can depend not only on $\phi_a(x, t)$ and its time derivative, $\dot{\phi}_a(x, t)$, but also on the spatial derivatives $\nabla \phi_a(x, t)$.

In classical mechanics we usually restrict to Lagrangians which depend only on q and $\dot{q}$: no $\ddot{q}$ terms allowed. For the same reason, we restrict to Lagrangian densities which may depend on $\dot{\phi}_a(x, t)$ but cannot depend on higher derivatives.[1] No such problems arise for higher spatial derivatives, such as $\nabla^2 \phi_a$, but as we will mainly be working in relativistic theories, we can ignore these. This is also the reason we ignore an explicit x^μ dependence in the Lagrangian.

Classical field theory still obeys the principle of least action. To derive the equations of motion, we simply perturb our fields by $\delta\phi_a$, requiring that the change in action, δS, is

---

[1] The technical reason for this is the *Ostrogradsky Instability*. This generically arises in higher-derivative theories, and precludes the existence of a stable ground state.

zero:

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

$$= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi_a} \delta \phi_a - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \right) \delta \phi_a + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \delta \phi_a \right) \right) \tag{2.4}$$

This last term is a total divergence, and so is equal to a surface integral over the boundary of the integration region. If we assume that our variation, $\delta \phi_a$ vanishes on this boundary, we can simply ignore the term, and so

$$\delta S = \int d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi_a} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \right) \right) \delta \phi_a = 0. \tag{2.5}$$

For this to be true for all possible deviations $\delta \phi_a$, we then find the Euler-Lagrange equations:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \right) - \frac{\partial \mathcal{L}}{\partial \phi_a} = 0. \tag{2.6}$$

These equations, one for each field ϕ_a, are the equations of motion for our fields.

2.1   Examples

The Wave Equation

Take the Lagrangian

$$\mathcal{L}(\phi, \partial_\mu \phi) = \frac{1}{2}(\partial^\mu \phi)(\partial_\mu \phi) = \frac{1}{2}(\dot{\phi})^2 - (\nabla \phi)^2, \tag{2.7}$$

governing the behaviour of a single scalar field ϕ. We can use the Euler-Lagrange equations to deduce the equation of motion

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = \partial^2 \phi = 0. \tag{2.8}$$

Restoring factors of c, we find that the equation of motion is simply:

$$\partial_t^2 \phi = c^2 \nabla^2 \phi. \tag{2.9}$$

This is the wave equation for a wave propagating at the speed of light. As we shall see, in quantum field theory, such fields will give rise to massless particles, such as the photon.

The Klein-Gordon Equation

The *Klein-Gordon Lagrangian* is a generalization of the above Lagrangian:

$$\mathcal{L}(\phi, \partial_\mu \phi) = \frac{1}{2}(\partial^\mu \phi)(\partial_\mu \phi) - \frac{1}{2}m^2 \phi^2. \tag{2.10}$$

We now find an extra contribution to the Euler-Lagrange equations given by:

$$\frac{\partial \mathcal{L}}{\partial \phi} = m^2 \phi$$

and this give us the *Klein-Gordon equation*:

$$\partial^2 \phi + m^2 \phi = 0. \tag{2.11}$$

The Klein-Gordon equation is of central importance in quantum field theory. The m^2 is called the *mass term*, as when we quantize we find that the theory gives rise to spinless particles of mass m.

We can generalize our example still further:

$$\mathcal{L}(\phi, \partial_\mu \phi) = \frac{1}{2}(\partial^\mu \phi)(\partial_\mu \phi) - V(\phi), \tag{2.12}$$

where $V(\phi)$ is an arbitrary function of ϕ. We call $V(\phi)$ the *potential*, in analogy to the potential in classical mechanics. Our equation of motion is now

$$\partial^2 \phi + V'(\phi) = 0. \tag{2.13}$$

The Electromagnetic Field

As you may have seen in previous courses, the Maxwell Lagrangian is given by:

$$\mathcal{L} = -\frac{1}{2}(\partial^\nu A^\mu)(\partial_\nu A_\mu) + \frac{1}{2}(\partial^\nu A_\nu)^2. \tag{2.14}$$

The first term looks very similar to the Klein-Gordon Lagrangian, although we have a minus sign now to make sure that the $(\dot{A}_i)^2$ are positive. The equations of motion are

$$\partial^2 A^\nu - \partial^\nu(\partial_\mu A^\mu) = \partial_\mu(\partial^\mu A^\nu - \partial^\nu A_\mu) = \partial_\mu F^{\mu\nu} = 0 \tag{2.15}$$

where $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A_\mu$ is the electromagnetic field strength. The tensor $F^{\mu\nu}$ can be utilized to write the Maxwell Lagrangian more concisely as:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{2.16}$$

A Non-Relativistic Example

In condensed matter physics, it is useful to consider the complex field ψ, with Lagrangian:

$$\mathcal{L} = \frac{i}{2}(\Psi^*\dot{\Psi} - \dot{\Psi}^*\Psi) - \frac{1}{2M}\nabla\Psi^* \cdot \nabla\Psi - U(x)|\Psi|^2. \tag{2.17}$$

It is easy to check that the Lagrangian is real, even though the field is complex. The $U(x)|\Psi|^2$ term explicitly breaks translational invariance — very useful when considering atoms trapped in a potential well, or electrons in a crystal lattice. To derive the equations of motion, we treat Ψ and Ψ^* as independent fields:

$$\frac{\partial\mathcal{L}}{\partial\dot{\Psi}^*} = -\frac{i}{2}\Psi, \quad \frac{\partial\mathcal{L}}{\partial\nabla\Psi^*} = -\frac{1}{2M}\nabla\Psi, \quad \frac{\partial\mathcal{L}}{\partial\Psi^*} = \frac{i}{2}\dot{\Psi} - U(x)\Psi \tag{2.18}$$

and so:

$$i\dot{\Psi} = -\frac{1}{2M}\nabla^2\Psi + U(x)\Psi. \tag{2.19}$$

Other than some factors of $\hbar$, this equation looks like the Schrödinger equation! But it isn't quite: our field Ψ is purely classical, and has none of the probabilistic interpretation

of the Schrödinger equation.

2.2   Noether's Theorem

In classical mechanics, you will have learnt that continuous symmetries correspond to conservation laws. In classical field theory, this generalizes. A continuous symmetry corresponds to a *conserved current* $j^\mu(x)$, which satisfies the equation

$$\partial_\mu j^\nu = 0. \tag{2.20}$$

If we rewrite this condition in components, we recognise this as the continuity equation

$$\frac{\partial j^0}{\partial t} = \nabla \cdot \mathbf{j}. \tag{2.21}$$

You have probably met this in electromagnetism, where it is used to express conservation of charge.

For our current j^μ, the corresponding conserved charge is

$$Q = \int d^3x \; j^0. \tag{2.22}$$

To show that this quantity is in fact conserved, we can compute:

$$\frac{dQ}{dt} = \int d^3x \; \frac{\partial j^0}{\partial x^0} = \int d^3x \; \nabla \cdot \mathbf{j} = 0. \tag{2.23}$$

The last inequalities follows because we are integrating over a total divergence. Notice that (2.20) is stronger than simply demanding the existence of a conserved quantity: it implies that the charge is conserved *locally*. For if we define the charge in a finite volume V to equal

$$Q_V = \int_V d^3x \; j^0 \tag{2.24}$$

we then find that

$$\frac{dQ_V}{dt} = \int_V d^3x \; \nabla \cdot \mathbf{j} = \int_{\partial V} \mathbf{j} \cdot d\mathbf{S} \tag{2.25}$$

where the last integral is a surface integral over the boundary ∂V. This equation says that the change of charge in a region is equal to the amount of charge flowing through the regions boundary.

Proving Noether's Theorem

The proof of Noether's theorem is very similar to the proof in classical mechanics. Under an infinitesimal symmetry, the fields will transform as

$$\delta \phi_a = G_a(\phi), \tag{2.26}$$

whereas the Lagrangian will change by a total derivative:

$$\delta \mathcal{L} = \partial_\mu F^\mu(\phi). \tag{2.27}$$

As in our derivation of the Euler-Lagrange equations, now the transformation of $\mathcal{L}$

under the infinitesimal symmetry $\delta\phi_a$:

$$\begin{aligned}
\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi_a}\delta\phi_a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\partial_\mu(\delta\phi_a) \\
&= \left[\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right)\right]\delta\phi_a + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a\right).
\end{aligned} \tag{2.28}$$

For fields satisfying the equations of motion, the first term is zero, and this leaves

$$\delta\mathcal{L} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a\right). \tag{2.29}$$

Now specializing to the case of a symmetry transformation, we find that

$$\partial_\mu F^\mu(\phi) = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}G_a(\phi)\right) \tag{2.30}$$

and hence, the current

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}G_a(\phi) - F^\mu(\phi) \tag{2.31}$$

is conserved.

2.2.1 Internal Symmetries

An *internal symmetry* is a symmetry that only depend the fields at a given location in spacetime, acting the same at every point. For instance, our Lagrangian for the wave equation:

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 \tag{2.32}$$

is invariant under the translation symmetry

$$\phi(x) \rightarrow \phi(x) + \alpha, \tag{2.33}$$

where α is a constant. Using (2.31), we can compute the corresponding conserved quantity:

$$j^\mu = \partial^\mu\phi. \tag{2.34}$$

As a less trivial example, consider a complex scalar field:

$$\mathcal{L} = \partial^\mu\Phi^*\partial_\mu\Phi - \frac{1}{2}|\Phi|^2. \tag{2.35}$$

This is invariant under phase rotation:

$$\Phi \rightarrow e^{i\alpha}\Phi. \tag{2.36}$$

The associated conservation law is:

$$j^\mu = i(\partial^\mu\Phi^*)\Phi - i\Phi^*(\partial^\mu\Phi). \tag{2.37}$$

Later we shall see that conservation of electric charge arises from a symmetry of this form.

2.2.2   Translations and Energy-Momentum Conservation

As you will remember from your classical mechanics course, conservation of momentum arises from spatial-translational symmetry, and conservations of energy arises from time-translational symmetry. In a relativistic theory, these symmetries are combined into space-time translation, under which:

$$x^\nu \to x^\nu - \varepsilon^\nu, \quad \phi_a(x) \to \phi_a(x + \varepsilon) = \phi_a(x) + \varepsilon^\nu \partial_\nu \phi_a(x). \tag{2.38}$$

Since the Lagrangian is itself a scalar, we find that

$$\mathcal{L}(x) \to \mathcal{L}(x) + \varepsilon^\nu \partial_\nu \mathcal{L} \tag{2.39}$$

once we have substituted our fields in. We now invoke Noether's theorem to conclude that we have four conserved currents $(j^\mu)_\nu$, for $\nu = 0, 1, 2$ and 3:

$$(j^\mu)_\nu = \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi_a} \partial_\mu \phi_a - \delta^\mu_\nu \mathcal{L} = T^\mu_\nu. \tag{2.40}$$

The tensor T^μ_ν is called the *energy-momentum* tensor, and satisfies

$$\partial_\mu T^\mu_\nu = 0. \tag{2.41}$$

The total energy H and momentum P_i of a system is then given by

$$H = \int d^3x \, T^0_0, \quad P_i = \int d^3x \, T^0_i. \tag{2.42}$$

The Klein-Gordon Energy-Momentum Tensor

For the Klein-Gordon Lagrangian, we find that

$$T^{\mu\nu} = \partial^\mu \phi \partial^\nu \phi - \frac{1}{2} \delta^\mu_\nu \partial^\rho \phi \partial_\rho \phi + \frac{1}{2} m^2 \phi^2. \tag{2.43}$$

The total energy of the system is then:

$$H = \int d^3x \, T^{00} = \int d^3x \, \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}(\nabla \phi)^2 + \frac{1}{2}m^2\phi^2, \tag{2.44}$$

while the total momentum is

$$P_i = -\int dx^3 \, \dot{\phi}\nabla_i\phi. \tag{2.45}$$

2.3   Hamiltonian Field Theory

Just as classical mechanics can be formulated using either the Lagrangian or the Hamiltonian, so to can we use both Lagrangians and Hamiltonians for classical field theory. In the Lagrangian approach symmetries, and in particular Lorentz invariance, are manifest. This is why we began by discussing Lagrangians. Quantizing a theory from the Lagrangian is however difficult: this is the Feynman many-path formulation of quantum mechanics. Rather than begin with path-integrals, we will first use canonical quantization. This is analogous to the approach we usually use in quantum mechanics. In this approach, the Hamiltonian is central.

In classical mechanics, we define the conjugate momenta p_a via the formula:

$$p_a = \frac{\partial L}{\partial \dot{q}}, \tag{2.46}$$

and by analogy, we define the *conjugate field momenta* $\pi_a(x)$ via the formula:

$$\pi_a(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}_a(x)}. \tag{2.47}$$

We then define the *Hamiltonian density* as

$$\mathcal{H}(x) = \pi_a(x)\dot{\phi}_a(x) - \mathcal{L}(x), \tag{2.48}$$

eliminating all $\dot{\phi}_a$ just as in classical mechanics. The Hamiltonian is then given by

$$H = \int dx^3 \; \pi_a(x)\dot{\phi}_a(x) - \mathcal{L}(x). \tag{2.49}$$

The equations of motion are given by

$$\begin{aligned}
\dot{\phi}_a(x,t) &= \frac{\partial H}{\partial \pi_a(x,t)} \\
\dot{\pi}_a(x,t) &= -\frac{\partial H}{\partial \phi_a(x,t)},
\end{aligned} \tag{2.50}$$

analogous to the equations for particles. Unlike the Euler-Lagrange equations, the Lorentz invariance of these equations is not manifest.

The Klein-Gordon Hamiltonian

For the Klein-Gordon Lagrangian, the conjugate momentum field is

$$\pi_a(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}_a(x)} = \dot{\phi}(x) \tag{2.51}$$

and hence

$$H = \int dx^3 \; \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2. \tag{2.52}$$

Note that this is consistent with our formula for energy (2.44).

The Klein-Gordon Field

Just as in quantum mechanics we promote our coordinates q_a and their momenta p_a to operators, in quantum field theory our operators will be field operators $\phi_a(\mathbf{x})$ and $\pi_a(\mathbf{x})$. Because we are in the Schrödinger pictures, these on spatial coordinates but not on time. In quantum mechanics, our coordinates satisfy the *canonical commutation relations*

$$[q_a, q_b] = [p_a, p_b] = 0, \quad [q_a, p_b] = i\delta_{ab}. \tag{3.1}$$

The *canonical commutation relations* for a field theory are a slight generalization of this

$$[\phi_a(\mathbf{x}), \phi_b(\mathbf{y})] = [\pi_a(\mathbf{x}), \pi_b(\mathbf{y})] = 0, \quad [\phi_a(\mathbf{x}), \pi_b(\mathbf{y})] = \delta_{ab}\delta(\mathbf{x} - \mathbf{y}). \tag{3.2}$$

These do not look very Lorentz invariant, but this is because we are working in the Schrödinger picture. Our operators do not evolve in time at all. Instead, we evolve the states forward according to the Schrödinger equation

$$i\frac{d\,|\psi\rangle}{dt} = H\,|\psi\rangle. \tag{3.3}$$

This should all look very similar to quantum mechanics. We might now attempt to calculate the energy eigenstates of our Hamiltonian, allowing us to evolve states forward in time.

But this is where the similarities end. Whilst in quantum mechanics is is not hard to compute the spectrum of H for a system, at least in principle, quantum field theories are much trickier to study. Part of the difficulty is that $|\psi\rangle$ is not a function but a *functional*; it associates an complex amplitude to each possible field configuration $\phi(\mathbf{x})$. The Hilbert spaces required for quantum field theory are stupendously big, and we almost never can define the state of a quantum field.

There are however a few classes of quantum field theory which we can solve. The simplest and most important of these theories are *free field theories*. These are theories where the Lagrangian is quadratic, and the equations of motion are linear. This vastly simplifies them, and allows us to exactly solve the theories.

The simplest example of a free field is the Klein-Gordon Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{m^2}{2}\phi^2, \tag{3.4}$$

which we already met in the previous chapter. It governs the behaviour of spin-0 particles, such as the Higgs boson and the pion. Since Maxwell's equations are linear, the electromagnetic field provides another example of a free field. We will discuss this field

in Chapter 7, since gauge invariance complicates things when we quantize the theory. In Chapter 6 we discuss the Dirac field, another free field. It describes spin-1/2 particles such as the electron. But we shall begin with the Klein-Gordon Lagrangian, which provides the simplest introduction to the ideas of quantum field theory.

3.1   The Harmonic Oscillator

Before we tackle the quantization of the Klein-Gordon equation, let us review a simpler problem: the quantum harmonic oscillator. Classical, the Hamiltonian for a harmonic oscillator is

$$H = \frac{1}{2}p^2 + \frac{\omega}{2}q^2. \tag{3.5}$$

To quantize this theory, we imposed the canonical commutation relation:

$$[q, p] = i \tag{3.6}$$

where we have set $\hbar$ to 1. We now define the raising and lowering operators

$$a = \frac{1}{\sqrt{2\omega}}(\omega q + ip), \quad a^\dagger = \frac{1}{\sqrt{2\omega}}(\omega q - ip). \tag{3.7}$$

The position and momentum operators can be rewritten in terms of the raising and lower operators as

$$q = \frac{1}{\sqrt{2\omega}}(a + a^\dagger), \quad p = -i\sqrt{\frac{\omega}{2}}(a - a^\dagger). \tag{3.8}$$

Using the canonical commutation relation, we find that

$$[a, a^\dagger] = 1, \quad H = \omega\left(a^\dagger a + \frac{1}{2}\right). \tag{3.9}$$

We furthermore can calculate

$$[H, a^\dagger] = \omega a^\dagger, \quad [H, a] = -\omega a. \tag{3.10}$$

Armed with these commutation relations, we can now completely solve the theory. Let $|E\rangle$ be any energy eigenstate, then

$$Ha^\dagger |E\rangle = (E + \omega)a^\dagger |E\rangle, \quad Ha |E\rangle = (E - \omega)a |E\rangle. \tag{3.11}$$

By using the raising and lower operators on $|E\rangle$, we are hence able to find an infinite series of eigenvectors, with eigenvalues

$$..., E - 2\omega, E - \omega, E, E + \omega, E + 2\omega, ... \tag{3.12}$$

But we know that our theory has to have a ground-state $|0\rangle$. For this state to have the lowest energy, we need to require that $a |0\rangle = 0$; otherwise, $a |0\rangle$ would have lower energy! From this condition, we can calculate the ground-state energy:

$$H |0\rangle = \omega\left(a^\dagger a + \frac{1}{2}\right) |0\rangle = \frac{\omega}{2}. \tag{3.13}$$

Operatoring the raising operator on the vacuum state, we can calculate the excited states

$$|n\rangle = (a^\dagger)^n |0\rangle, \quad H |n\rangle = \omega(n + \frac{1}{2}) |n\rangle. \tag{3.14}$$

3.2   Quantizing the Klein-Gordon Field

The Klein-Gordon field can be thought of as an infinite collection of harmonic oscillators. To see how, let us Fourier transform our field:

$$\phi(\mathbf{p}, t) = \int \frac{d^3x}{(2pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}} \phi(\mathbf{x}, t). \tag{3.15}$$

The Klein-Gordon equation now becomes

$$\left(\frac{\partial^2}{\partial t^2} + \mathbf{p}^2 + m^2\right) \phi(\mathbf{p}, t) = 0. \tag{3.16}$$

This is the classical equation for a harmonic oscillator of frequency $\omega_\mathbf{p} = \sqrt{\mathbf{p}^2 + m^2}$! Notice that each Fourier mode decouples from all of the other modes. This is because the Klein-Gordon equation is linear.

Inspired by the similarities with the harmonic oscillator, let us rewrite

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

$$\pi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} - i\sqrt{\frac{\omega_\mathbf{p}}{2}} \left(a_\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{x}} - a_\mathbf{p}^\dagger e^{-i\mathbf{p}\cdot\mathbf{x}}\right). \tag{3.17}$$

This is complete analogous to (3.8). It is possible to invert the relations in order to write $a_\mathbf{p}$ in terms of ϕ and π, but we will not need to this.

Now we apply the canonical commutation relations. It is straightforward to show that the commutators

$$[\phi(\mathbf{x}), \phi(\mathbf{y})] = [\pi(\mathbf{x}), \pi(\mathbf{y})] = 0, \quad [\phi(\mathbf{x}), \pi(\mathbf{y})] = i\delta(\mathbf{x} - \mathbf{y}). \tag{3.18}$$

are equivalent to the commutators

$$[a_\mathbf{p}, a_\mathbf{q}] = [a_\mathbf{p}^\dagger, a_\mathbf{q}^\dagger] = 0, \quad [a_\mathbf{p}, a_\mathbf{q}^\dagger] = (2\pi)^3 \delta(\mathbf{p} - \mathbf{q}). \tag{3.19}$$

Next we compute the Hamiltonian in terms of the ladder operators:

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

$$= \int \frac{d^3p}{(2\pi)^3} \omega_\mathbf{p}(a_\mathbf{p} a_\mathbf{p}^\dagger + a_\mathbf{p}^\dagger a_\mathbf{p}) \tag{3.20}$$

$$= \int \frac{d^3p}{(2\pi)^3} \omega_\mathbf{p} a_\mathbf{p}^\dagger a_\mathbf{p} + \int d^3p \, \frac{\omega_\mathbf{p}}{2}\delta(\mathbf{0})$$

This looks suspiciously similar to the Hamiltonian for a harmonic oscillator. There is one slight catch: the constant $\frac{1}{2}\omega$ in the Hamiltonian for a harmonic oscillator has given way

to the term

$$\int d^3p \, \frac{\omega_{\mathbf{p}}}{2} \delta(\mathbf{0}), \tag{3.21}$$

an infinite integral of an infinite delta function. We will return to this term in a moment; but first, let us continue to follow our recipe for quantizing the harmonic oscillator.

Analogous to (3.11), we find that

$$[H, a_{\mathbf{p}}^\dagger] = \omega_{\mathbf{p}} a_{\mathbf{p}}^\dagger, \quad [H, a_{\mathbf{p}}] = -\omega_{\mathbf{p}} a_{\mathbf{p}}. \tag{3.22}$$

So if we act raising and lowering operators on any energy eigenstate $|E\rangle$, we find

$$H a_{\mathbf{p}}^\dagger |E\rangle = (E + \omega_{\mathbf{p}}) a_{\mathbf{p}}^\dagger |E\rangle, \quad H a_{\mathbf{p}} |E\rangle = (E - \omega_{\mathbf{p}}) a_{\mathbf{p}} |E\rangle. \tag{3.23}$$

Just like for the harmonic oscillator, this can be used to find an infinite number of eigenstates, with eigenvalues

$$..., E - 2\omega_{\mathbf{p}}, E - \omega_{\mathbf{p}}, E, E + \omega_{\mathbf{p}}, E + 2\omega_{\mathbf{p}}, ... \tag{3.24}$$

If we want a ground state to exist, then we must demand the existence of a ket $|0\rangle$ so that

$$a_{\mathbf{p}} |0\rangle = 0 \tag{3.25}$$

for every $\mathbf{p}$. All other energy eigenstates can derived by applying creation operators to $|0\rangle$, so the theory is solved!

3.3   The Vacuum State

Solving the theory is just the beginning. We now need to understand what our solution means, physically. Let's start with ground state $|0\rangle$. As we shall soon see, this state has a rather simple interpretation — it is the state with no particles, just empty spacetime. For this reason, the ground state of a quantum field theory is often called the *vacuum state.*

What is the energy of the vacuum? A quick computation gives us:

$$\begin{aligned}
\langle 0| H |0\rangle &= \langle 0| \left[\int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}} a_{\mathbf{p}}^\dagger a_{\mathbf{p}} + \int \frac{d^3p}{(2\pi)^3} \frac{\omega_{\mathbf{p}}}{2} (2\pi)^3 \delta(\mathbf{0}) \right] |0\rangle \\
&= \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}} \langle 0| a_{\mathbf{p}}^\dagger a_{\mathbf{p}} |0\rangle + \langle 0|0\rangle \int \frac{d^3p}{(2\pi)^3} \frac{\omega_{\mathbf{p}}}{2} (2\pi)^3 \delta(\mathbf{0}) \\
&= \int d^3p \, \frac{\omega_{\mathbf{p}}}{2} \delta(\mathbf{0}).
\end{aligned} \tag{3.26}$$

Apparently the strange, infinite quantity that bothered us is the energy of the vacuum!

From a pragmatic perspective, having an infinite energy ground state is not a problem. After all, we can never measure the energy of the ground state; all we can do is measure the energy difference between the ground state and the excited states. We could easily redefine our Hamiltonian:

$$H \to H - \langle 0| H |0\rangle = \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}} a_{\mathbf{p}}^\dagger a_{\mathbf{p}}, \tag{3.27}$$

and now our vacuum state has zero energy. This is what we shall do for the rest of the course. For now however, let us explore in more detail the vacuum energy.

First of all we should write (3.26) in a more useful form. Using the Fourier space representation of the delta function

$$\delta(\mathbf{q}) = \int \frac{d^3x}{(2\pi)^3} e^{-i\mathbf{q}\cdot\mathbf{x}} \tag{3.28}$$

we find that

$$\delta(0) = \lim_{q\to 0} \int \frac{d^3x}{(2\pi)^3} e^{-i\mathbf{q}\cdot\mathbf{x}} = \int \frac{d^3x}{(2\pi)^3} = V \tag{3.29}$$

where V is the volume of space we integrating over. This infinite now makes sense: we were calculating the total energy of our system, but what we really should have been calculating is the *energy density*:

$$\mathcal{H}_0 = \frac{\langle 0|H|0\rangle}{V} = \int \frac{d^3p}{(2\pi)^3} \frac{\omega_{\mathbf{p}}}{2}. \tag{3.30}$$

This type of divergence, which arise from the fact that space is infinite, is called an *infrared* divergence, because it occurs at large length scales.

Our formula for the energy density is still infinite; as $\mathbf{p}$ becomes larger, the integrand $\omega_{\mathbf{p}}$ does not go to zero, and so the integral diverges. Just as the ground state of a harmonic oscillator has energy $\frac{1}{2}\omega$, each mode of our Klein-Gordon field contributes $\frac{1}{2}\omega_{\mathbf{p}}$ to the energy density. Summing over these infinite modes results in an infinite energy density. This type of divergence is called an *ultraviolet* divergence, because it occurs at high momentum, as $|\mathbf{p}| \to \infty$.

3.3.1   The Casimir Effect

Imagine that we had two infinite, parallel, conducting plates a distance l apart. We will take the planes to be in the y-z axis. As you know from previous courses, the electric field in a conductor is zero. The solutions to Maxwell's equations must therefore be standing waves with momentum

$$\mathbf{k} = \left(\frac{\pi n}{l}, k_y, k_z\right) \tag{3.31}$$

so that electric field vanishes at the conducting plates. Casimir's insight was that this will disturb the vacuum, creating a measurable shift in the vacuum energy between the plates.

Why? Just like the scalar field, the electromagnetic field is quantized, and the vacuum energy becomes

$$\mathcal{H}_0^{EM} = \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}}. \tag{3.32}$$

Only difference between this and (3.30) is a factor of 2, which originates because the photon has two possible polarizations. The energy density per unit area of the plate is therefore

$$\frac{E}{A} = \sum_n \int \frac{dk_y dk_z}{(2\pi)^2} \sqrt{\frac{\pi^2 n^2}{l^2} + k_y^2 + k_z^2}. \tag{3.33}$$

Although this quantity is still formally infinite, it depends on l. And while we cannot measure the energy density, we can measure the change in energy density; this will give rise to a force between the two plates.

To actually calculate the Casimir effect requires a bit of work, so we will make three simplifications:
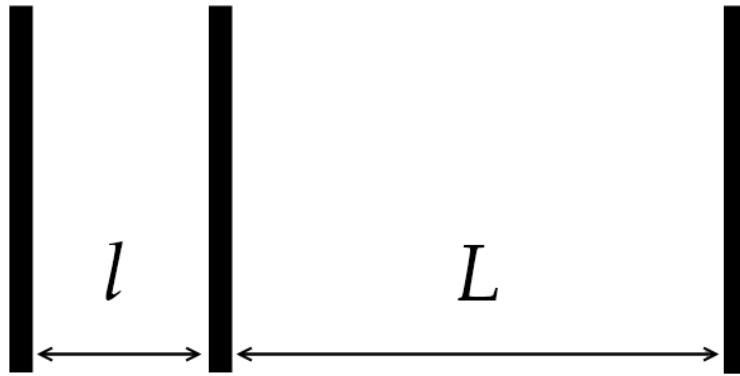
Figure 3.1: Calculating the Casimir effect with three plates.

1. The conducting plates modify the vacuum energy between the plates, but they also modify the vacuum energy outside the two plates. This second shift in energy is difficult to compute, because the volume of the outside world is infinite. To fix this problem, we can introduce a third plate, at a distance L from the first one. Moving only the middle plate, the energy density of the outside world is unchanged. After our calculation we can take $L \to \infty$ to recover the Casimir force. Our introduction of L here is known as an *infrared* regulator, as it removes the infrared difficulties we encountered in the previous section.

2. We will work with a massless scalar field rather than the electromagnetic field, so that we don't have to worry about polarization and the like.

3. We will work in one spatial dimension, so that there are no pesky integrals over k_y and k_z.

With these simplifications, the energy of our system is

$$E = E_P(l) + E_P(L - l) \tag{3.34}$$

where $E_P(l)$ is the energy between two plates at distance l apart:

$$E_P(l) = \frac{1}{2} \sum_{n=1}^{\infty} \omega_n = \frac{\pi}{2l} \sum_{n=1}^{\infty} n \tag{3.35}$$

Things aren't looking so good: that $\sum n$ is clearly infinite! What has gone wrong? Well, real metals do not perfect conduct. In fact, at sufficiently high frequencies, the light will be so energetic that the mirrors will be completely transparent to it — try reflecting gamma radiation with a mirror![1]

To incorporate this piece of physics, we introduce a factor of $e^{-a\omega_n/\pi}$ in $E_P(l)$. Here a is the characteristic frequency of the conducting plates: if $\omega_n \gg 1/a$ then the frequency does not contribute to our sum. This factor is called an *ultraviolet* regularization, as it

---

[1]Physically, conductors exist because free electrons in metals can move to balance out any electric fields. At high energies, the electric field of light will oscillate at high frequencies, and so the electrons will have to move faster and faster to counteract the field. There are physical limits to how fast the electrons can move, and so eventually the light wins.

removes the ultraviolet divergence we had previously. Our new expression for $E(l)$ is

$$E_P(l) = \frac{\pi}{2l} \sum_{n=1}^{\infty} n e^{-an/l} = -\frac{\pi}{2} \frac{d}{da} \left(\sum_{n=1}^{\infty} e^{-an/l} \right)$$
$$= -\frac{\pi}{2} \frac{d}{da} \left(\frac{1}{1 - e^{-a/l}} \right) = \frac{\pi}{2l} \frac{e^{a/l}}{(e^{a/l} - 1)^2}. \tag{3.36}$$

Since we are interested in the case where a is small (that is, when the distance between the plates is much greater than the characteristic scale of the conductor), we can expand

$$E_P(l) = \frac{l}{2\pi a^2} - \frac{\pi}{24l} + O\left(\frac{a^2}{l^3} \right) \tag{3.37}$$

As $a \to 0$, this expressions blows up. This is just as we would expect, because in this limit we end up back at (3.35). But the physical force between two plates should not be infinite:

$$F = -\frac{dE}{dl} = -E'_P(l) + E'_P(L - l)$$
$$= -\left(\frac{1}{2\pi a^2} - \frac{\pi}{24l^2} \right) + \left(\frac{1}{2\pi a^2} - \frac{\pi}{24(L - l)^2} \right) + O\left(\frac{a^2}{l^3} \right) \tag{3.38}$$

The troublesome a^{-2} terms cancel, and so we can now safely take both $a \to 0$ and $L \to \infty$:

$$F = \lim_{a \to 0} \lim_{L \to \infty} -\frac{\pi}{24} \left(\frac{1}{l^2} - \frac{1}{(L - l)^2} \right) + O\left(\frac{a^2}{l^3} \right) = -\frac{\pi}{24l^2}. \tag{3.39}$$

There you have it, an attractive, measurable force exists between two conducting plates!

If we were to repeat this calculation in 3 spatial dimensions and for photons, we find that the Casimir force is

$$\frac{1}{A} \frac{\partial E_P}{\partial l} = \frac{\pi^2}{240l^4}. \tag{3.40}$$

The Casimir force has been measured to within a few percent accuracy experimentally. In *Figure 3.2*, we can see the results of one such experiment.

Universality of the Casimir Effect

You may justifiably be bothered by the fact that we had to introduce some arbitrary factor $e^{-a\omega_n/\pi}$ in order to get a finite value. What we had used chosen a factor of $(1 + a\omega_n)^{-4}$, or just ignored all frequencies with $\omega_n > a^{-1}$?

As it turns out, we would still have found the same force in the limit where $a \to 0$! This means that the Casimir effect does not depend on the detailed properties of the conductors, but only on the properties of the quantum fields. On the other hand, if $a \neq 0$, we know that the Casimir force has corrections:

$$F = -\frac{\pi}{24l^2} + O(a). \tag{3.41}$$

Unlike the leading order term, these corrections do depend on the physics of the conductor at high frequencies. But so long as $l \gg a$, these corrections are not important, and so the Casimir effect is universal at large distances.
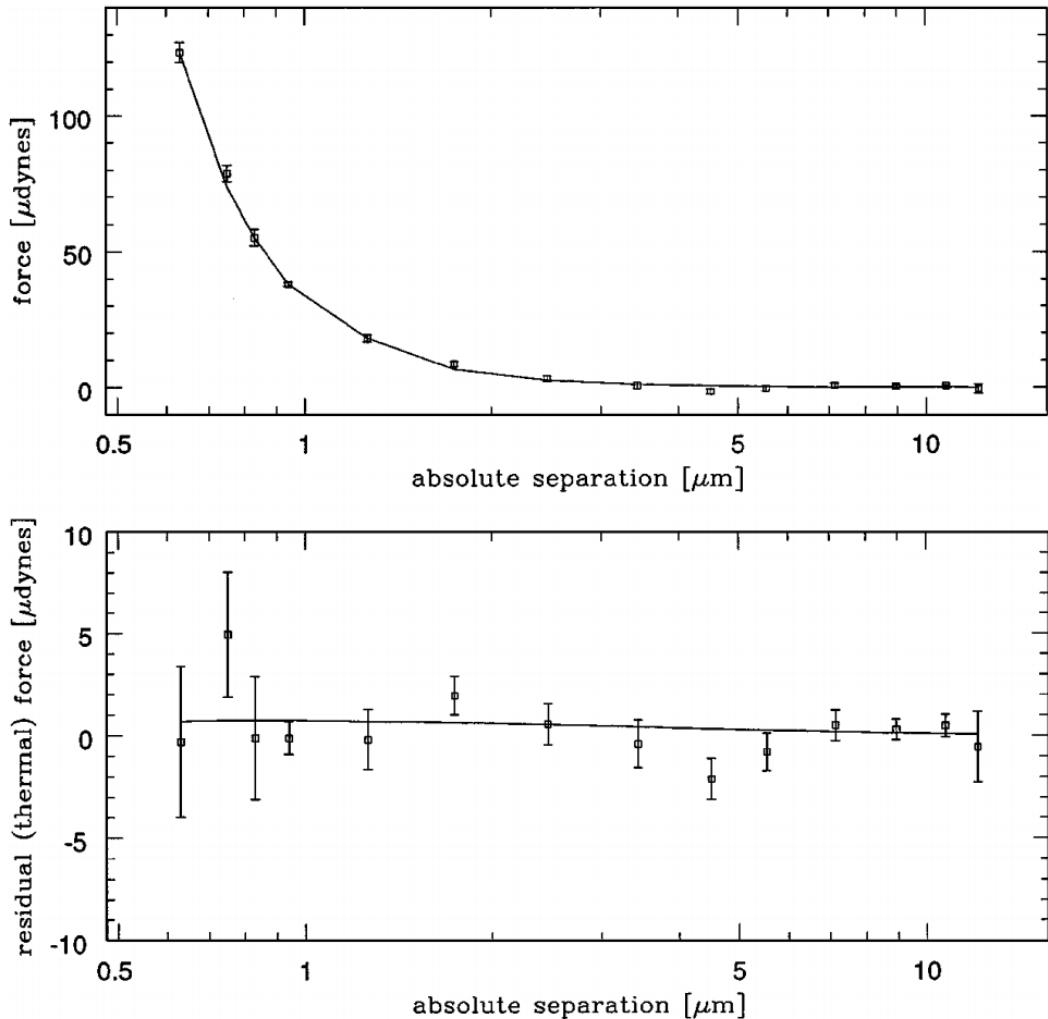
Figure 3.2: An experimental demonstration of the Casimir effect, measured by *Lamoreaux* in 1997. The top graph is the total force measured, the second graph is with the theoretically predicted Casimir effect subtracted off. This shows thermal residual forces. From *https://doi.org/10.1103/PhysRevLett.78.5.*

3.3.2   The Cosmological Constant

We said previously that our infinite vacuum energy has no physical effect. But this isn't quite true; gravity couples to the energy-momentum tensor, and so our vacuum energy should give rise to a *cosmological constant* in the Einstein field equations:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = -8\pi G T_{\mu\nu} + \mathcal{H}_0 g_{\mu\nu}. \tag{3.42}$$

We have good reason to believe that the Standard Model is valid to at least 1 TeV, and if we calculate $\mathcal{H}_0$ up to this frequency, we find

$$\mathcal{H}_0 \sim (1 \text{ TeV})^4. \tag{3.43}$$

Restoring units of $\hbar$ and c, and plugging our vacuum energy into the Einstein field equations, you would discover that the universe doubles in size roughly every

$$\frac{c^{5/2}\hbar^{3/2}}{\sqrt{G\mathcal{H}_0}} \sim 10^{-11} \text{ seconds.} \tag{3.44}$$

The entire universe would be destroyed in an instant! Since this clearly violates experimental observations, physicists unsuccessfully laboured for many decades to prove that $\mathcal{H}_0$ must be exactly 0. But while theorists tried to prove that $\mathcal{H}_0$, cosmologists[2] discovered that, indeed, a cosmological constant does exists. It just happens very small:

$$\Lambda_{\text{cosmos}} \sim 10^{-3} \text{ eV}^4 = 10^{-60}(1 \text{ TeV})^4. \tag{3.45}$$

This 10^{60} order of magnitude difference between theory and experiment has been dubbed 'the worst prediction in theoretical physics'. No one knows the origin of this cosmological constant, nor why the contribution from quantum fields delicately cancel to sixty decimal places.

3.4 Particles

3.4.1 Single Particle States

Having explored the vacuum, we will now turn to the excited states of the Klein-Gordon field. As we noted previously, any energy eigenstate can be obtained by acting creation operators on the vacuum. Let us start by considering the state

$$a_{\mathbf{p}}^\dagger |0\rangle . \tag{3.46}$$

This state has energy $\omega_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2}$:

$$H\left(a_{\mathbf{p}}^\dagger |0\rangle\right) = [H, a_{\mathbf{p}}^\dagger] |0\rangle + a_{\mathbf{p}}^\dagger H |0\rangle = \sqrt{\mathbf{p}^2 + m^2} \left(a_{\mathbf{p}}^\dagger |0\rangle\right) . \tag{3.47}$$

We can likewise calculate the momentum of the particle. To do so, we can use (2.45):

$$P_i = -\int d^3x \; \dot{\phi}\nabla_i\phi = \int \frac{d^3p}{(2\pi)^3} p_i a_{\mathbf{p}} a_{\mathbf{p}}^\dagger, \tag{3.48}$$

and hence find that

$$\mathbf{P}\left(a_{\mathbf{p}}^\dagger |0\rangle\right) = \mathbf{p}\left(a_{\mathbf{p}}^\dagger |0\rangle\right) . \tag{3.49}$$

As you will remember from special relativity, the energy of a particle with mass m and momentum $\mathbf{p}$ is given by $E = \sqrt{\mathbf{p}^2 + m^2}$. This leads to a very natural interpretation of $a_{\mathbf{p}}^\dagger |0\rangle$: it represents a single-particle state!

What about the spin of $|p\rangle$? The proper way to calculate the spin of a particle is to use Noether's theorem to calculate the angular momentum operators, and then to apply these operators to $a_{\mathbf{p}}^\dagger |0\rangle$:. Rather than go through this tedious process, we will instead note that, for $\mathbf{p} = 0$, there is only one state, $a_{\mathbf{0}}^\dagger |0\rangle$. If however our particle carried spin, we know that there would have to be multiple states with $\mathbf{p} = 0$, each labelled by a different value of s_z. So our particles have to be spin 0.

---

[2] Including the ANU's very own Brian Schmidt.

Relativistic Normalization

We now have to deal with the problem of normalization. We would like to define the single particle state

$$|\mathbf{p}\rangle = \mathcal{N}(\mathbf{p})a_{\mathbf{p}}^{\dagger}|0\rangle, \tag{3.50}$$

where $\mathcal{N}(\mathbf{p})$ is a normalization factor. You might think that we should just set $\mathcal{N}(\mathbf{p}) = 1$, but this choice turns out to not be relativistically invariant. It turns out that if we want our 1-particle states to be relativistically normalized, we should set

$$\mathcal{N}(\mathbf{p}) = \sqrt{2\omega_{\mathbf{p}}}. \tag{3.51}$$

(The factor of 2 is for later convenience). We then find

$$\langle\mathbf{p}|\mathbf{q}\rangle = \mathcal{N}(\mathbf{p})\mathcal{N}(\mathbf{q})\delta(\mathbf{p} - \mathbf{q}) = 2\omega_{\mathbf{p}}\delta(\mathbf{p} - \mathbf{q}) \tag{3.52}$$

To understand why this is the correct choice, consider the integral

$$\int d^4p \; \delta(p^2 - m^2)\theta(k^0)f(p^0, \mathbf{p}) \tag{3.53}$$

for some Lorentz invariant function f. Here θ is the Heaviside step function:

$$\theta(x) = 1 \text{ if } x > 0, \quad \theta(x) = 0 \text{ otherwise.} \tag{3.54}$$

Since no Lorentz transformation can change the sign of p^0, we conclude that the integrand in (3.53) is Lorentz invariant. A quick computation then shows that

$$\int d^4p \; \delta(p^2 - m^2)\theta(p^0)f(p^0, \mathbf{p}) = \int \frac{d^3p}{2\omega_{\mathbf{p}}}f(\omega_{\mathbf{p}}, \mathbf{p}), \tag{3.55}$$

implying that the measure

$$\int \frac{d^3p}{2\omega_{\mathbf{p}}} \tag{3.56}$$

is Lorentz invariant. This in turn implies that

$$2\omega_{\mathbf{p}}\delta(\mathbf{p} - \mathbf{q}) \tag{3.57}$$

is Lorentz invariant, just like we wanted.

Consider the state

$$\phi(\mathbf{x})|0\rangle = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \left(a_{\mathbf{p}}e^{i\mathbf{p}\cdot\mathbf{x}} + a_{\mathbf{p}}^{\dagger}e^{-i\mathbf{p}\cdot\mathbf{x}}\right)|0\rangle$$
$$= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}}e^{-i\mathbf{p}\cdot\mathbf{x}}|\mathbf{p}\rangle \tag{3.58}$$

which is just a linear superposition of momentum eigenstates. Other than the factor of $(2\omega_{\mathbf{p}})^{-1}$, this is the same expression for $|\mathbf{x}\rangle$ in terms of $|\mathbf{p}\rangle$ we use in non-relativistic quantum mechanics. Hence we can give an interpretation of $\phi(\mathbf{x})$ — when it acts on the vacuum, it creates a particle at position $\mathbf{x}$.

3.4.2   Multiparticle States

Now let us consider the state

$$\sqrt{4\omega_{\mathbf{p}}\omega_{\mathbf{q}}}a_{\mathbf{p}}^{\dagger}a_{\mathbf{q}}^{\dagger}\left|0\right\rangle=\left|\mathbf{p},\mathbf{q}\right\rangle. \tag{3.59}$$

Running through the same maths as before, we find that this state has energy

$$\omega_{\mathbf{p}}+\omega_{\mathbf{q}}=\sqrt{\mathbf{p}^{2}+m^{2}}+\sqrt{\mathbf{q}^{2}+m^{2}} \tag{3.60}$$

and momentum

$$\mathbf{p}+\mathbf{q}. \tag{3.61}$$

We therefore conclude that $\left|\mathbf{p},\mathbf{q}\right\rangle$ is a state with two particles, each of mass m, but which may have different momenta.

Because $a_{\mathbf{p}}^{\dagger}$ and $a_{\mathbf{q}}^{\dagger}$ commute, we find that

$$\left|\mathbf{p},\mathbf{q}\right\rangle=\sqrt{4\omega_{\mathbf{p}}\omega_{\mathbf{q}}}a_{\mathbf{p}}^{\dagger}a_{\mathbf{q}}^{\dagger}\left|0\right\rangle=\sqrt{4\omega_{\mathbf{p}}\omega_{\mathbf{q}}}a_{\mathbf{q}}^{\dagger}a_{\mathbf{p}}^{\dagger}\left|0\right\rangle=\left|\mathbf{q},\mathbf{p}\right\rangle. \tag{3.62}$$

Furthermore, if $\mathbf{p}=\mathbf{q}$ then we have two-particle state where the particles are both in the same state. We therefore conclude that our particles satisfy Bose-Einstein statistics!

This entire analysis carries through to any energy eigenstate

$$\sqrt{2\omega_{\mathbf{p}}2\omega_{\mathbf{q}}2\omega_{\mathbf{r}}...}(a_{\mathbf{p}}^{\dagger}a_{\mathbf{q}}^{\dagger}a_{\mathbf{r}}^{\dagger}...)\left|0\right\rangle=\left|\mathbf{p},\mathbf{q},\mathbf{r},...\right\rangle, \tag{3.63}$$

which we interpret as a multi-particle state. The entire setup leads to the number notation that you will be familiar with from many-body quantum mechanics.

Just like in many-body quantum mechanics, we can define the number operator

$$N=\int\frac{d^{3}p}{(2\pi)^{3}}a_{\mathbf{p}}^{\dagger}a_{\mathbf{p}}, \tag{3.64}$$

which measures the number of particles in our system:

$$N\left|p_{1},p_{2},...,p_{n}\right\rangle=n\left|p_{1},p_{2},...,p_{n}\right\rangle. \tag{3.65}$$

Because $[H,N]=0$, the number of particles is conserved. This, however, is a special feature of free theories, and in interacting theories we will generically find particle creation and destruction.

In this section, we have shown that combining quantum mechanics and classical fields gives rise to a theory containing particles! More specifically, we have shown that after quantizing a scalar field, we get the many-body theory of a spin-0 boson. Have you ever wondered why identical particles act so strangely in quantum mechanics? It seems very counter-intuitive to think that every photon and electron in the universe are identical, with the same, spin, mass and charge.

But, as we have just seen, this fact has a very straightforward origin in quantum field theory. Identical particles look the same because they are all quanta of the same field! Two photons look the same for the same reason that two waves of the electromagnetic field look the same.

3.5 Complex Fields

We now move onto studying the complex scalar field

$$\mathcal{L} = \partial_\mu \Phi^* \partial^\mu \Phi - m^2 \Phi^* \Phi. \tag{3.66}$$

The idea of a complex field may seem mysterious. But by writing

$$\Phi = \frac{1}{\sqrt{2}} \left(\phi_1 + i\phi_2 \right), \tag{3.67}$$

we find that our Lagrangian is just

$$\mathcal{L} = \frac{1}{2}(\partial \phi_1)^2 - \frac{1}{2}m^2 \phi_1^2 + \frac{1}{2}(\partial \phi_2)^2 - \frac{1}{2}m^2 \phi_2^2. \tag{3.68}$$

So our complex field is secretly two real Klein-Gordon theories, both of mass m. Nevertheless, it is often much more useful to work over one complex field rather than two real ones.

We quantize our theory in an identical manner to the real Klein-Gordon equation

$$\Phi = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \left(b_{\mathbf{p}} e^{ipx} + c_{\mathbf{p}}^\dagger e^{-ipx} \right)$$

$$\Phi^\dagger = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \left(b_{\mathbf{p}}^\dagger e^{-ipx} + c_{\mathbf{p}} e^{ipx} \right). \tag{3.69}$$

For the real Klein-Gordon equation, $\Phi^\dagger = \Phi$, and this force $b_{\mathbf{p}}^\dagger = c_{\mathbf{p}}$. Now our fields are not real, and so these operators are different.

The canonical momentum is

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\Phi}} = \dot{\Phi}^*, \tag{3.70}$$

and the canonical commutation relations are

$$[\Phi(\mathbf{x}), \pi(\mathbf{y})] = i\delta(\mathbf{x} - \mathbf{y}), \quad [\Phi(\mathbf{x}), \pi^\dagger(\mathbf{y})] = 0, \tag{3.71}$$

along with the usual

$$[\Phi(\mathbf{x}), \Phi(\mathbf{y})] = [\Phi^\dagger(\mathbf{x}), \Phi(\mathbf{y})] = [\pi(\mathbf{x}), \pi(\mathbf{y})] = [\pi^\dagger(\mathbf{x}), \pi(\mathbf{y})] = 0. \tag{3.72}$$

Using (3.69), it is straightforward to show that these are equivalent to

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

$$[c_{\mathbf{p}}, c_{\mathbf{q}}^\dagger] = (2\pi)^3 \delta(\mathbf{p} - \mathbf{q}) \tag{3.73}$$

along with

$$[b_{\mathbf{p}}, b_{\mathbf{q}}] = [b_{\mathbf{p}}, c_{\mathbf{q}}] = [b_{\mathbf{p}}, c_{\mathbf{q}}^\dagger] = [c_{\mathbf{p}}, c_{\mathbf{q}}] = 0. \tag{3.74}$$

So when we quantize the complex scalar field, we get two raising operators $b_{\mathbf{p}}^\dagger$ and $c_{\mathbf{p}}^\dagger$. These will give rise to two different types of particle. Repeating the steps for the real field, we can shows that these both have mass m, and that they are both spin zero. Akin to our interpretation of $\phi(\mathbf{x})$, we interpret $\Phi(\mathbf{x})$ as creating a particle c at point $\mathbf{x}$, and $\Phi^\dagger(\mathbf{y})$ creates a particle b at point $\mathbf{y}$.

When we discussed Noether's theorem in Chapter 2, we showed that classically, our theory has the conserved charge

$$Q = i \int d^3x \, (\dot{\Phi}^*\Phi - \Phi^*\dot{\Phi}) = i \int d^3x \, (\pi\Phi - \Phi^*\pi^*). \tag{3.75}$$

In the quantum theory, we can rewrite this operator as

$$Q = \int \frac{d^3p}{(2\pi)^3} \left(c_\mathbf{p}^\dagger c_\mathbf{p} - b_\mathbf{p}^\dagger b_\mathbf{p} \right) = N_c - N_b. \tag{3.76}$$

Hence $c_\mathbf{p}^\dagger$ creates a particle with charge $+1$, whereas $b_\mathbf{p}^\dagger$ creates a particle of charge -1. This is a theory of both particles and antiparticles. Of course, the particular choice of which is the particle versus antiparticle is a matter of convention.

It is not difficult to show that $[H, Q] = 0$, so this is also a conserved charge of the quantum theory. In a free field this is not so special, since N_b and N_c are separately conserved. If we move to interacting theories though, we find that both N_b and N_c will no longer be conserved. The charge Q will however remain conserved, even in the interacting theory, as long as the theory remains symmetric under

$$\Phi \to e^{i\alpha}\Phi. \tag{3.77}$$

3.6   Non-Relativistic Fields

While quantum field theory was developed for relativistic systems, many modern applications of the theory are to many-body quantum mechanics and to condensed matter physics. In fact, the many-body quantum mechanics you learnt last semester is secretly quantum field theory! We will begin calculating the non-relativistic limit of the Klein-Gordon field.

3.6.1   The Non-Relativistic Limit

Let's add the factors of c back into the complex Klein-Gordon equation:

$$\partial^2\psi + m^2\psi = \frac{1}{c^2}\partial_t^2\psi - \nabla^2\psi + m^2c^4\psi = 0. \tag{3.78}$$

Clearly we cannot just set $c \to \infty$, as the m^2 term will blow up. To fix this, note that if a mode has energy $E = mc^2 + \varepsilon$, then field will oscillate in time as $\psi \propto e^{-iEt} \approx e^{-imc^2t}$. We hence introduce the field

$$\Psi(\mathbf{x}, t) = e^{imc^2t}\psi(\mathbf{x}, t). \tag{3.79}$$

Plugging this back into the Klein-Gordon equation, we find

$$\frac{1}{c^2}\partial_t^2\psi - \nabla^2\psi + m^2c^4\psi = e^{-imc^2t} \left(\frac{1}{c^2}\partial_t^2\Psi - 2im\partial_t\Psi - \nabla^2\Psi \right) = 0. \tag{3.80}$$

Dividing through by e^{-imc^2t}, we can now safely take the $c \to \infty$. This results in a very familiar equation

$$i\frac{\partial}{\partial t}\Psi = -\frac{\nabla^2}{2m}\Psi. \tag{3.81}$$

It looks like the Schrödinger equation, but it isn't! We started with a classical Klein-Gordon field, and now have a classical field satisfying the Schrödinger equation. Quantum mechanics has not yet entered the picture: the field has no probabilistic interpretation.

We already met the Schrödinger field in Chapter 2, where we derived it from the Lagrangian

$$\mathcal{L} = \frac{i}{2}\left(\Psi^*\dot{\Psi} - \dot{\Psi}^*\Psi\right) - \frac{1}{2m}\nabla\Psi^* \cdot \nabla\Psi. \tag{3.82}$$

This form of the Lagrangian makes the symmetry between Ψ and Ψ^* manifest. Using integration by parts, we can rewrite

$$S = \int d^3x\, dt\, \frac{i}{2}\left(\Psi^*\dot{\Psi} - \dot{\Psi}^*\Psi\right) - \frac{1}{2m}\nabla\Psi^* \cdot \nabla\Psi = \int d^3x\, dt\, \Psi^*\left(i\dot{\Psi} + \frac{1}{2m}\nabla^2\Psi\right), \tag{3.83}$$

which is less symmetric, but often more useful. The Schrödinger equation then follows from apply the Euler-Lagrange equation to Ψ^*.

Because our theory has a complex field Ψ, it is easy to see that we have a phase symmetry $\Psi \to e^{i\alpha}\Psi$. Applying Noether's theorem, we find the conserved current

$$J_0 = |\Psi|^2, \quad \mathbf{J} = \frac{i}{2m}\left(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*\right). \tag{3.84}$$

These equations should be familiar from quantum mechanics, where they express conservation of probability. We again emphasise however that we are working in a classical field theory, so we do not have a probabilistic interpretation for Ψ.

3.6.2 Quantizing the Schrödinger Field

Having dealt with the classical Schrödinger field, let's quantize it. From the Hamiltonian formulation,

$$\pi = \frac{\partial\mathcal{L}}{\partial\dot{\Psi}} = i\Psi^*. \tag{3.85}$$

The conjugate momentum is dependent on Ψ! This might seem a little odd, but we can trace it back to the fact that the Schrödinger Lagrangian, and hence the Schrödinger equation, contains only first-order time derivatives. If we want to know how $\Psi(\mathbf{x}, t)$ evolves, we only need to specify $\Psi(\mathbf{x}, t_0)$ at the initial time $t = t_0$. We can now calculate the Hamiltonian

$$\mathcal{H} = \pi\Psi - \mathcal{L} = -\frac{1}{2m}\Psi^*\nabla^2\Psi. \tag{3.86}$$

To quantize our theory, we postulate the canonical commutator relations

$$[\Psi(\mathbf{x}), \pi(\mathbf{y})] = i\delta(\mathbf{x} - \mathbf{y}) \implies [\Psi(\mathbf{x}), \Psi^\dagger(\mathbf{y})] = \delta(\mathbf{x} - \mathbf{y}). \tag{3.87}$$

These commutators should be familiar from many-body quantum mechanics! The commutation relations they satisfy look very similar the those satisfied by raising and lower operators.

To further study the structure of our theory, we will now switch to Fourier space:

$$\Psi(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} a_\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{x}}. \tag{3.88}$$

Our commutator relations then require

$$[a_{\mathbf{p}}, a_{\mathbf{q}}^\dagger] = (2\pi)^3 \delta(\mathbf{p} - \mathbf{q}), \tag{3.89}$$

whilst the Hamiltonian becomes

$$H = \int d^3x \ -\frac{1}{2m}\Psi^*\nabla^2\Psi = \int \frac{d^3p}{(2\pi)^3} \ \frac{\mathbf{p}^2}{2m} a_{\mathbf{p}}^\dagger a_{\mathbf{p}}. \tag{3.90}$$

The vacuum will satisfy $a_{\mathbf{p}}|0\rangle = 0$, while the single-particle states are

$$|\mathbf{p}\rangle = a_{\mathbf{p}}|0\rangle, \quad H|\mathbf{p}\rangle = \frac{\mathbf{p}^2}{2m}|\mathbf{p}\rangle. \tag{3.91}$$

Many-particle states are of the form

$$a_{\mathbf{p}_1}^\dagger ... a_{\mathbf{p}_2}^\dagger |0\rangle, \tag{3.92}$$

with the particles satisfying Bose-Einstein statistics. From this we conclude that non-relativistic many-body quantum mechanics really is just a field theory. If our particles are in some potential well $U(\mathbf{x})$, then we can simply add an additional term to our Lagrangian

$$H = \int d^3x \ -\frac{1}{2m}\Psi^\dagger(\mathbf{x})\nabla^2\Psi(\mathbf{x}) + U(\mathbf{x})\Psi^\dagger(\mathbf{x})\Psi(\mathbf{x}). \tag{3.93}$$

Quantum Mechanics is Weird

Having quantized the Schrödinger field, let's dig a little deeper and ask, "Why does quantum mechanics look so different to quantum field theory?" Single-particle wavefunctions are much easier to deal with than field operators, so why can we use the former for non-relativistic systems but not relativistic ones?

The reason is that "non-relativistic" theories are not, in fact, non-relativistic. There are no absolute velocities in the non-relativistic world: two observers, travelling at different but constant velocities, will still experience the same laws of physics. This is called *Galilean relativity*, and was discovered by Galileo in 1632!

One consequence of Galilean symmetry is conservation of mass.[3] For the Schrödinger field, the mass is

$$M = \int d^3x \ m\Psi^\dagger(\mathbf{x})\Psi(\mathbf{x}), \tag{3.94}$$

which, as we showed previously, is conserved. In terms of the raising and lower operators,

$$M = \int \frac{d^3p}{(2\pi)^3} \ ma_{\mathbf{p}}^\dagger a_{\mathbf{p}} = mN, \tag{3.95}$$

where N is the number operator. So from this we conclude that the number of particles must be conserved. This holds even if we add interactions to our Lagrangian

$$\mathcal{L}_{\text{int}} = \lambda(|\Psi|^2)^2 + \mu(|\Psi|^2)^3 + ..., \tag{3.96}$$

---

[3]For something so basic in physics, the origin of mass conservation is obscure, and I do not know a good reference for the subject. Ultimately mass originates from the existence of a *projective representation of the Galilean group*. Explaining what exactly this means is beyond the scope of this footnote.

since mass is always conserved for a Galilean invariant system.

In particular, if we start in a state with one particle, then we always must remain in a state with one particle. We could describe this one-particle state using a wavefunction

$$|\psi\rangle = \int \frac{d^3p}{(2\pi)^3} \, \psi(p) a_{\mathbf{p}}^{\dagger} |0\rangle = \int d^3x \, \psi(x)\Psi(\mathbf{x}) |0\rangle \tag{3.97}$$

and then evolve the wavefunction according to the Schrödinger equation.

The point here is that it is not quantum field theory that is weird: non-relativistic quantum mechanics is very special. It only works because particle number is conserved, which itself is a consequence of Galilean invariance. As soon as we allow particle number to vary, we have to use quantum field theory. Special relativity is one example where this is the case, but many others exist. If we want to consider chemical reactions, or quantum optics, or quantum systems interacting with an environment, then there too we will need to use quantum field theory.

3.7 The Heisenberg Picture

So far we have been working in the Schrödinger picture. In this picture, the operators depend on their position in space, while the states evolve in time. In a Lorentz invariant theory, this division of space-time into space and time is not very natural. We should really try to put space and time on an equal footing. Fortunately, the Heisenberg picture allows us to do this.

In the Heisenberg picture, all time evolution is put into operators. Just as in quantum mechanics, we can make our operators evolve via

$$\phi(x) = \phi(\mathbf{x}, t) = e^{iHt}\phi(\mathbf{x})e^{-iHt}, \quad \pi(x) = \pi(\mathbf{x}, t) = e^{iHt}\pi(\mathbf{x})e^{-iHt}. \tag{3.98}$$

Our fields are now variables of a space-time x, rather than just a spatial coordinate $\mathbf{x}$.

For any given operator $M(x)$, the Heisenberg equation of motion is

$$\frac{\partial M(x)}{\partial t} = i[H, M(x)]. \tag{3.99}$$

For the Klein-Gordon field, we find

$$\frac{\partial \phi}{\partial t} = \pi, \quad \frac{\partial \pi}{\partial t} = \nabla^2\phi - m^2\phi \tag{3.100}$$

which are identical to the classical equations of motion for the field. Combining these two equations, we find that $\phi(x)$ satisfies the Klein-Gordon equation:

$$\partial^2\phi + m^2\phi = 0. \tag{3.101}$$

Because the ladder operators satisfy

$$[H, a_{\mathbf{p}}^{\dagger}] = \omega_{\mathbf{p}} a_{\mathbf{p}}^{\dagger}, \quad [H, a_{\mathbf{p}}] = \omega_{\mathbf{p}} a_{\mathbf{p}}, \tag{3.102}$$

they in a particularly simple fashion:

$$e^{iHt}a_{\mathbf{p}}e^{-iHt} = e^{-i\omega_{\mathbf{p}}t}a_{\mathbf{p}}, \quad e^{iHt}a_{\mathbf{p}}^{\dagger}e^{-iHt} = e^{+i\omega_{\mathbf{p}}t}a_{\mathbf{p}}^{\dagger}. \tag{3.103}$$

Using (3.17), we can rewrite

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \left(a_{\mathbf{p}} e^{-ip\cdot x} + a_{\mathbf{p}}^{\dagger} e^{ip\cdot x} \right) \tag{3.104}$$

where $x = (t, \mathbf{x})$ and $p = (\omega_{\mathbf{p}}, \mathbf{p})$. It is easy to check that this expression for $\phi(x)$ automatically satisfies the Klein-Gordon equation.

3.7.1   The Propagator

Let us consider the propagation of a particle from some point y to another point x. We can study this process by considering

$$
\begin{aligned}
\langle 0|\phi(x)\phi(y)|0\rangle &= \int \frac{d^3p\, d^3q}{(2\pi)^6} \frac{1}{4\omega_{\mathbf{p}}\omega_{\mathbf{p}'}} \langle 0|a_{\mathbf{p}} a_{\mathbf{q}}^{\dagger}|0\rangle \, e^{-ip\cdot x + iq\cdot y} \\
&= \int \frac{d^3p\, d^3q}{(2\pi)^6} \frac{1}{4\omega_{\mathbf{p}}\omega_{\mathbf{p}'}} \langle \mathbf{p}|\mathbf{q}\rangle e^{-ip\cdot x + iq\cdot y}.
\end{aligned}
\tag{3.105}
$$

The other combinations of the ladder operators, $a_{\mathbf{p}}^{\dagger} a_{\mathbf{q}}^{\dagger}$, $a_{\mathbf{p}}^{\dagger} a_{\mathbf{q}}$, and $a_{\mathbf{p}} a_{\mathbf{q}}$, do not contribute, as annihilation operators annihilate $|0\rangle$, and creation operators annihilate $\langle 0|$. Physically, this is because we have to create a particle before we can annihilate it.

The left-hand side of (3.105) is called a *correlation function*, because it measure the correlation of the field at different points. We can however that $\phi(x)|0\rangle$ creates a particle at position $\mathbf{x}$ at time x^0, and so we can interpret right-hand side of (3.105) as describing the propagation of a particle from x to y.

This is one aspect of quite a general theme in quantum field theory. Later in this course, we shall find that more complicated correlation functions of the form

$$\langle 0|\phi(x_1)\phi(x_2)...\phi(x_N)|0\rangle \tag{3.106}$$

can be related to scattering involving N particles. This allows us to study scattering amplitudes by calculating correlation functions. Particles and fields are tightly interwoven in quantum field theory.

Proceeding with our calculation, we find

$$
\begin{aligned}
\langle 0|\phi(x)\phi(y)|0\rangle &= \int \frac{d^3p\, d^3q}{(2\pi)^3} \frac{1}{\sqrt{4\omega_{\mathbf{p}}\omega_{\mathbf{p}'}}} \delta(\mathbf{p} - \mathbf{q}) e^{-ip\cdot x + iq\cdot y} \\
&= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}} e^{-ip\cdot(x-y)} = D(x - y).
\end{aligned}
\tag{3.107}
$$

The function $D(x - y)$ is called the *propagator*. We can use it to study many properties of our field theory.

First, a very logical question: is $D(x - y)$ Lorentz invariant? We better hope so, or else our theory would not be compatible with relativity. Fortunately, we recognize that if we insert $f(p) = e^{-ip\cdot(x-y)}$ into (3.55), then

$$\int d^4p \, \delta(p^2 - m^2)\theta(p^0) e^{-ip\cdot(x-y)} = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}} e^{-ip\cdot(x-y)} = D(x - y), \tag{3.108}$$

making the Lorentz invariance manifest.

Causality

Next we should check whether our propagator respects causality. When x and y are spacelike separated, $(x - y)^2 < 0$ and we can show that

$$D(x - y) \sim e^{-m|\mathbf{x}-\mathbf{y}|}. \tag{3.109}$$

This exponential vanishes outside the light cone, but isn't zero. Does this mean our theory breaks causality?

We should be careful not to confuse correlation with causation. All we have shown is that there can be non-local correlations in our theory. But this is possible even with Lorentz invariance. We could after all imagine a quantum fluctuation of the field at some point z in the past of both x and y. This fluctuation will propagate to both points, giving rise to a non-local correlation.

So this is the wrong kind of question to ask — to investigate causality, we should really ask whether a measurement performed at one point will non-causally effect other points. We can study this by calculating the commutator of our two fields: if this vanishes, then the order in which we measure $\phi(x)$ and $\phi(y)$ does not matter. Now we compute

$$\langle 0|[\phi(x), \phi(y)]|0\rangle = \langle 0|\phi(x)\phi(y)|0\rangle - \langle 0|\phi(y)\phi(x)|0\rangle = D(x - y) - D(y - x). \tag{3.110}$$

If $(x - y)^2 < 0$, there is a Lorentz transformation that maps $x - y \to y - x$. Lorentz invariance then dictates that $D(x - y) = D(y - x)$, and hence

$$\langle 0|[\phi(x), \phi(y)]|0\rangle = 0 \text{ for } (x - y)^2 < 0. \tag{3.111}$$

On the other hand, when $(x - y)^2 > 0$ no such Lorentz transformation exists, and in fact we do find that $D(x - y) \neq D(y - x)$ in this case.

To give a physical picture to the calculation, we can note that for $(x - y)^2 < 0$ the events cannot be ordered in a Lorentz invariant way. So given any physical process where a particle propagates from $x \to y$, we could also consider a process where the particles propagates from $y \to x$. In any measurable process, the amplitudes of these two events cancels.

Microcausality

So far all we have shown is that $\langle 0|[\phi(x), \phi(y)]|0\rangle = 0$ for spacelike intervals. It is not hard to show however, that:

$$[\phi(x), \phi(y)] = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)}}{2\omega_{\mathbf{p}}} = D(x - y) - D(y - x). \tag{3.112}$$

We hence find that the commutator is Lorentz invariant, and that it vanishes if x and y are spacelike separated. Moreover, the commutator of these two operators is a c-number. This last property is a peculiarity of free fields, but the previous two properties are more general.

We could consider more generally any *local operator* in our theory; these are operators such as

$$\phi(x)^2, \quad \partial_\mu\phi(x), \quad (\partial\phi(x))^2, \tag{3.113}$$

which depends on the value of the Klein-Gordon field and its derivatives evaluated at a

single point in spacetime. Because $\phi(x)$ commute when spacelike separated, we conclude that for any local operators $\mathcal{O}_1(x)$ and $\mathcal{O}_2(x)$,

$$[\mathcal{O}_1(x), \mathcal{O}_2(y)] = 0 \text{ for } (x-y)^2 < 0. \tag{3.114}$$

This condition is called *microcausality*, and it states that no matter what measurements you perform, if they are spacelike separated they cannot effect each other. As a statement of causality, it holds in all quantum field theories. In fact, it is usually given as one of the axioms of quantum field theory.

Functional Methods

With the tools of canonical quantization, we have quantized the Klein-Gordon field. We have discovered that this theory is just a generalization of the harmonic oscillator. Although a field theory, particles natural arise. We could now develop perturbation theory in a manner analogous to quantum mechanics, and from this derive the Feynman rules for interacting theories.

This approach is however quite tedious. Lorentz symmetry is not manifest, since we favour Hamiltonian rather than the Lagrangian. Because we work with operators which do not compute, we have to be very careful about the order of operators. Perturbation theory soon becomes a series of frustrating reorderings of ladder operations.

Feynman's path integral formulation of quantum mechanics provides a completely different approach to quantum field theory. It uses the Lagrangian rather than the Hamiltonian, and is manifestly relativistic. Novel calculation methods can be developed more easily in this framework: path-integral provide a completely different intuition to the canonical approach. In this chapter we will study the path-integral in some details. Our overall goal is to develop the tools we will need when deriving the Feynman rules in the next chapter.

Having mentioned the flaws of the canonical approach, we should also mention that the path integral approach has its share of deficits. Particles do not appear in the formalism: a big problem if you want to understanding how scattering works! The path-integral also obscures quantum mechanical aspects of quantum field theory, such as the operator structure of the theory. There even exists quantum field theory which have no Lagrangian formulation! Needless to say, if you want to master quantum field theory you will need a thorough understanding of both canonical and functional methods.

4.1 Path Integrals in Quantum Mechanics

Consider a one-dimensional particle, evolving according to the Hamiltonian

$$H = \frac{p^2}{2m} + V(q). \tag{4.1}$$

We wish to calculate the probability that a particle will travel from a position q_i at time t_i to q_f at time t_f. In the canonical approach, the amplitude for this process is

$$\mathcal{A} = \langle q_f | e^{-iHT} | q_i \rangle \tag{4.2}$$

where $T = t_f - t_i$.

The path-integral approach to calculating $\mathcal{A}$ is very different. We begin by considering

every possible path from q_i to q_f. For each given path we can calculate the action

$$S[q] = \int_{t_1}^{t_2} L(q, \dot{q}) \, dt. \tag{4.3}$$

The action is an example of what known as a *functional*: a function which maps functions to numbers. We then calculate $\mathcal{A}$ by summing over all possible paths

$$\mathcal{A} = \sum_{\text{all paths}} e^{iS[q]} = \int \mathcal{D}q \, e^{iS[q]}. \tag{4.4}$$

Here $\int \mathcal{D}q$ is called a *functional integral*: an integral which integrates over functionals rather than numbers. We can also differentiate functionals, using the *functional derivative*

$$\frac{\delta F[q]}{\delta q(t)}. \tag{4.5}$$

Formally defining these operations is tricky; we will derive more precise definitions in the next few sections.

Although the canonical and path-integral approach look very different, it is not too difficult to show that they are equivalent. Starting from the canonical approach, let us split the time interval into $N + 1$ intervals of duration $\delta t = t/(N + 1)$. Using the completeness relation, we can then write

$$\langle q_f | e^{-iHT} | q_i \rangle = \int \prod_{j=1}^{N} dq_k \langle q_f | e^{-iH\delta t} | q_N \rangle \langle q_N | e^{-iH\delta t} | q_{N-1} \rangle \ldots \langle q_1 | e^{-iH\delta t} | q_f \rangle. \tag{4.6}$$

where each integral is from $-\infty$ to ∞. Using the Baker-Campbell-Haussdorf formula,

$$e^{A+B} = e^A e^B e^{-\frac{1}{2}[A,B]+\cdots}, \tag{4.7}$$

we find that for small δt,

$$e^{-iH\delta t} = e^{-ip^2\delta t/2m} e^{-iV(q)\delta t} + O(\delta t^2). \tag{4.8}$$

Utilizing this, we calculate that in the limit as $\delta t \to 0$,

$$\langle q_{k+1} | e^{-iH\delta t} | q_k \rangle = \langle q_{k+1} | e^{-ip^2\delta t/2m} e^{-iV(q)\delta t} | q_k \rangle = e^{-iV(q_k)\delta t} \langle q_{k+1} | e^{-ip^2\delta t/2m} | q_k \rangle. \tag{4.9}$$

This last factor is just the propagator for a free particle, which is much easier to evaluate in momentum space:

$$\begin{aligned}
\langle q_{j+1} | e^{-ip^2\delta t/2m} | q_j \rangle &= \int \frac{dk}{2\pi} \langle q_{j+1} | e^{-ip^2\delta t/2m} | k \rangle \langle k | q_j \rangle \\
&= \int \frac{dk}{2\pi} e^{-ik^2\delta t/2m} e^{ik(q_{j+1}-q_j)} \\
&= \sqrt{\frac{-im}{2\pi\delta t}} \exp\left(\frac{im\delta t}{2} \left(\frac{q_{j+1} - q_j}{\delta t} \right)^2 \right)
\end{aligned} \tag{4.10}$$

Substituting our computations in (4.6), we find that

$$\mathcal{A} = \left(\frac{-im}{2\pi\delta t}\right)^{N/2} \int \prod_{j=1}^{N} dq_k \exp\left(\sum_{j=0}^{N} i\delta t \left[\frac{m}{2}\left(\frac{q_{j+1}-q_j}{\delta t}\right)^2 - iV(q)\right]\right) \tag{4.11}$$

where we have taken $q_0 = q_i$ and $q_{N+1} = q_f$.

We can now take the continuum limit where $N \to \infty$ and $\delta t \to 0$:

$$\frac{q_{j+1}-q_j}{\delta t} \to \dot{q}$$
$$\sum_{j=0}^{N} i\delta t \to \int_0^T dt \tag{4.12}$$
$$\left(\frac{-im}{2\pi\delta t}\right)^{N/2} \int \prod_{j=1}^{N} dq_k \to \int_{q_1 \to q_2} \mathcal{D}q.$$

This procedure gives us a definition of the functional integral. First we discretize our function, and then integrating over the different possible values at each discrete point. The functional integral is then the limit of this process as $\delta t \to 0$. We can now write our amplitude as:

$$\mathcal{A} = \int_{q_1 \to q_2} \mathcal{D}q \ \exp\left(i\int_0^T dt \left[\frac{m}{2}\dot{q}^2 - V(q)\right]\right) = \int_{q_1 \to q_2} \mathcal{D}q \ \exp\left(i\int_0^T dt \ L(q,\dot{q})\right), \tag{4.13}$$

which is precisely the formula given by the path-integral.

4.1.1   Correlation Functions

In many applications of quantum field theory, we are interested in correlation functions such as

$$\langle\Omega|\phi(x_1)\phi(x_2)...\phi(x_N)|\Omega\rangle, \tag{4.14}$$

where $|\Omega\rangle$ is the ground state of our theory. One strength of the path-integral is that it allows us to efficiently compute these correlation functions.

Consider for instance the expression

$$I = \int_{q_1 \to q_2} \mathcal{D}q \ q(t_1)q(t_2) \ \exp\left(i\int_{-T}^T dt \ L(q,\dot{q})\right). \tag{4.15}$$

We here are assuming that $t_1, t_2 \in [-T, T]$. There are two possible cases, depending on whether t_1 or t_2 is larger.

First assume that $t_1 < t_2$. We can our integral into three parts: the integral from q_1 at $-T$ to $q(t_1)$:

$$I_1[q] = \int_{q_1 \to q} \mathcal{D}q \exp\left(i\int_{-T}^{t_1} dt \ L(q,\dot{q})\right) = \langle q|e^{-iH(t_1+T)}|q_1\rangle, \tag{4.16}$$

the integral from t_1 to t_2

$$I_2[q, q'] = \int_{q \to q'} \mathcal{D}q \exp\left(i\int_{t_1}^{t_2} dt \ L(q,\dot{q})\right) = \langle q'|e^{-iH(t_2-t_1)}|q\rangle, \tag{4.17}$$

and finally the integral from t_2 to T:

$$I_3[q'] = \int_{q' \to q_2} \mathcal{D}q \exp\left(i \int_{t_2}^{T} dt\ L(q, \dot{q}) \right) = \langle q_2 | e^{-iH(T-t_2)} | q' \rangle, \tag{4.18}$$

In terms of these, I can be rewritten as:

$$\begin{aligned}
I &= \int dq\ dq' I_3[q'] q' I_2[q', q] q I_1[q] \\
&= \int dq\ dq'\ \langle q_2 | e^{-iH(T-t_2)} | q' \rangle\ q' \langle q' | e^{-iH(t_2-t_1)} | q \rangle\ q \langle q | e^{-iH(t_1+T)} | q_1 \rangle
\end{aligned} \tag{4.19}$$

where the q and q' integrals are the integrals over all possible values of $q(t_1)$ and $q(t_2)$. We can recognize the Schrödinger picture operators

$$q_s = \int dq q\, |q\rangle \langle q|, \tag{4.20}$$

allowing use to rewrite

$$I = \langle q_2 | e^{-iH(T-t_2)} q_S e^{-iH(t_2-t_1)} q_S e^{-iH(t_1+T)} | q_1 \rangle \tag{4.21}$$

Converting these operators to the Heisenberg picture, we then find that

$$I = \langle q_2 | e^{-iHT} q(t_2) q(t_1) e^{-iHT} | q_1 \rangle. \tag{4.22}$$

Our analysis has so far assumed that $t_1 < t_2$; for the opposite case of $t_2 < t_1$, we would instead find that

$$I = \langle q_2 | e^{-iHT} q(t_1) q(t_2) e^{-iHT} | q_1 \rangle. \tag{4.23}$$

From this, we conclude that

$$\int_{q_1 \to q_2} \mathcal{D}q\ q(t_1) q(t_2)\ e^{i \int_{-T}^{T} dt\ L(q, \dot{q})} = \langle q_2 | e^{-iHT} \mathcal{T}\{q(t_1) q(t_2)\} e^{-iHT} | q_1 \rangle \tag{4.24}$$

where $\mathcal{T}$ is the time-order operation:

$$\mathcal{T}\{\phi(t_1)\phi(t_2)\} = \begin{cases} \phi(t_1)\phi(t_2) & \text{if } t_1 > t_2 \\ \phi(t_2)\phi(t_1) & \text{if } t_2 > t_1. \end{cases} \tag{4.25}$$

This formula holds for arbitrary states, but we are most interested in the ground state $|\Omega\rangle$. To single this state out, we can take the limit $T \to \infty(1 - i\varepsilon)$ for some small $\varepsilon > 0$. Introducing a complete set of energy eigenstates $|E_n\rangle$, we find

$$\begin{aligned}
e^{-iHT(1-i\varepsilon)} | q_1 \rangle &= \sum_n e^{-iHT(1-i\varepsilon)} | E_n \rangle \langle E_n | q_1 \rangle \\
&= \sum_n e^{-iE_n T(1-i\varepsilon)} | E_n \rangle \langle E_n | q_1 \rangle \\
&= e^{-E_0 \varepsilon T} \sum_n e^{-(E_n - E_0)\varepsilon T} e^{-iE_n T} | E_n \rangle \langle E_n | q_1 \rangle.
\end{aligned} \tag{4.26}$$

In the limit where $T \to \infty$, the contribution from excited states will be exponential

suppressed compared to those in the ground state, and hence:

$$\lim_{T \to \infty} e^{-iHT(1-i\varepsilon)} |q_1\rangle = \lim_{T \to \infty} e^{-E_0 \varepsilon T} e^{-iE_0 T} |E_0\rangle \langle E_0|q_1\rangle \tag{4.27}$$

We can now rewrite:

$$\lim_{T \to (1-i\varepsilon)\infty} \int_{q_1 \to q_2} \mathcal{D}q \; q(t_1)q(t_2) \; e^{i \int_{-T}^{T} dt L(q,\dot{q})}$$
$$= \lim_{T \to (1-i\varepsilon)\infty} \langle q_2|\Omega\rangle\langle\Omega|q_1\rangle e^{-i2E_0 T(1-i\varepsilon)} \langle\Omega|\mathcal{T}\{q(t_1)q(t_2)\}|\Omega\rangle. \tag{4.28}$$

We still have some annoying phase factors and overlap factors (we are also implicitly assuming here that $\langle\Omega|q_1\rangle \neq 0$). But we can get the same factors by computing

$$\lim_{T \to (1-i\varepsilon)\infty} \int_{q_1 \to q_2} \mathcal{D}q \; e^{i \int_{-T}^{T} dt L(q,\dot{q})} = \lim_{T \to (1-i\varepsilon)\infty} \langle q_2|\Omega\rangle\langle\Omega|q_1\rangle e^{-i2E_0 T(1-i\varepsilon)} \langle\Omega|\Omega\rangle, \tag{4.29}$$

and so can conclude that

$$\langle\Omega|\mathcal{T}\{q(t_1)q(t_2)\}|\Omega\rangle = \lim_{T \to (1-i\varepsilon)\infty} \frac{\int \mathcal{D}q \; q(t_1)q(t_2) \; \exp\left[i \int_{-T}^{T} dt \; L(q,\dot{q})\right]}{\int \mathcal{D}q \; \exp\left[i \int_{-T}^{T} dt \; L(q,\dot{q})\right]}. \tag{4.30}$$

Generalizing this formula to higher correlation function is easy: you just add additional factors of $q(t)$ onto either side of the above expression.

4.1.2 The Generating Functional

There is an even more efficient method for calculating correlation functions. To describe it requires a definition of the *functional derivative*, which we promised earlier. The basic axiom of this derivative is the equation

$$\frac{\delta}{\delta J(t)} J(\tau) = \delta(t - \tau). \tag{4.31}$$

For linear functionals the derivative is particularly simple

$$L[J] = \int d\tau \; J(\tau)q(\tau) \implies \frac{\delta}{\delta J(t)} L[J] = q(t). \tag{4.32}$$

These rules generalize the usual rules of calculus, which for discrete vectors give

$$\frac{\partial}{\partial x_i} x_j = \delta_{ij}, \quad \frac{\partial}{\partial x_i} \sum_j x_j p_j = p_i. \tag{4.33}$$

To calculate more complicated functionals, we simply use the chain rule:

$$\frac{\delta}{\delta J(t)} f\left[\int d\tau \; J(\tau)q(\tau)\right] = q(t)f'\left[\int d\tau \; J(\tau)q(\tau)\right]. \tag{4.34}$$

Slightly trickier to calculate are cases involving derivatives:

$$
\begin{aligned}
\frac{\delta}{\delta J(t)} J'(\tau) &= \frac{\delta}{\delta J(t)} \int d\sigma \ \delta(\tau - \sigma) J'(\sigma) \\
&= -\frac{\delta}{\delta J(t)} \int d\sigma \ \delta'(\tau - \sigma) J(\sigma) = \delta'(t - \tau).
\end{aligned}
\tag{4.35}
$$

This procedure generalizes to higher order derivatives in the obvious way.

Functional derivatives allow us to very quickly study calculus of variations. For instance, minimizing the action by requiring

$$
\frac{\delta S[q]}{\delta q(t)} = 0
\tag{4.36}
$$

provides an efficient derivation of the Euler-Lagrange equations.

With this formal development under our belt, let us define the *generating functional*

$$
Z[J] = \int \mathcal{D}q \ \exp\left[i \int dt \ [\mathcal{L}[q] + J(t)q(t)]\right].
\tag{4.37}
$$

Here we have added to our Lagrangian a *source term J*. You can think of it as an external device which is used to probe the theory. If we now differentiate this functional, we find

$$
\begin{aligned}
\frac{\delta Z}{\delta J(t_1)} &= \int \mathcal{D}q \ \frac{\delta}{\delta J(t_1)} \exp\left[i \int dt \ [\mathcal{L}[q] + J(t)q(t)]\right] \\
&= \int \mathcal{D}q \ i\phi(t_1) \exp\left[i \int dt \ [\mathcal{L}[q] + J(t)q(t)]\right].
\end{aligned}
\tag{4.38}
$$

Rearranging this and setting $J = 0$, we then find that

$$
\langle \Omega | \phi(t_1) | \Omega \rangle = \frac{-i}{Z[0]} \frac{\delta}{\delta J(t_1)} Z[J] \bigg|_{J=0}.
\tag{4.39}
$$

Likewise, the two-point function is simply

$$
\langle \Omega | \mathcal{T}\phi(t_1)\phi(t_2) | \Omega \rangle = \frac{1}{Z[0]} \left(-i\frac{\delta}{\delta J(t_1)}\right) \left(-i\frac{\delta}{\delta J(t_2)}\right) Z[J] \bigg|_{J=0},
\tag{4.40}
$$

and we can generalize for n-point functions in the obvious way.

The generating functional encodes the full information of the quantum theory. If we compute $Z[J]$, then we can quickly calculate any n-point function. All we have to do is the path-integral (4.37)! The difficulty, unfortunately, is that for an interacting theory $Z[J]$ is nigh impossible to calculate.

4.1.3   The Harmonic Oscillator

We will now apply the path-integral approach to our favourite quantum system: the harmonic oscillator! The Lagrangian for the harmonic oscillator is

$$
L = \frac{1}{2}\dot{q}^2 - \frac{1}{2}\omega^2 q^2,
\tag{4.41}
$$

where, as in the previous chapter, we have set $m = 1$. Our aim is to compute

$$Z[J] = \int Dq \, \exp[iS[q, J]] = \int Dq \, \exp\left[i \int dt \, \left[\frac{1}{2}\dot{q}^2 - \frac{1}{2}\omega^2 q^2 + Jq\right]\right] \tag{4.42}$$

Using integration by parts, we can rewrite the integral as

$$Z[J] = \int Dq \, \exp\left[i \int dt \, \left[-\frac{1}{2}q\left(\frac{d^2}{dt^2} + \omega^2\right)q + Jq\right]\right] \tag{4.43}$$

In this form, our integral looks suspiciously like a Gaussian integral.

We will now Fourier transform our variables:

$$q(E) = \int dt \, e^{iEt} q(t), \quad J(E) = \int dt \, e^{iEt} J(t), \tag{4.44}$$

where as per usual we abuse notation, using the same symbol for a variable and its Fourier transform. Using the inverse relations

$$q(t) = \int \frac{dE}{(2\pi)} e^{iEt} q(E), \quad J(t) = \int \frac{dE}{(2\pi)} e^{iEt} J(E), \tag{4.45}$$

we can rewrite the action as

$$S[q, J] = \int dt \int \frac{dE \, dE'}{(2\pi)^2} e^{i(E+E')t} \left[-\frac{1}{2}q(E')(-E^2 + \omega^2)q(E) + J(E')q(E)\right]. \tag{4.46}$$

Performing the integral over t now gives a delta function $2\pi\delta(E + E')$, which then allows us tio integrate over E'. Our expression hence simplifies to

$$S[q, J] = \int \frac{dE}{2\pi} \left[-\frac{1}{2}q(-E)(-E^2 + \omega^2)q(E) + J(-E)q(E)\right]. \tag{4.47}$$

Our next trick is to perform a change of variables, defining

$$x(E) = q(E) + \frac{J(E)}{E^2 - \omega^2}. \tag{4.48}$$

We then get:

$$S[x, J] = \int \frac{dE}{2\pi} \left[-\frac{1}{2}x(-E)(-E^2 + \omega^2)x(E) - \frac{1}{2}\frac{J(-E)J(E)}{E^2 - \omega^2}\right]. \tag{4.49}$$

Furthermore, the integration measure is unchanged by a constant shift, so $\mathcal{D}x = \mathcal{D}q$. Substituting this all back into our generating functional, we have

$$Z[J] = \exp\left[-\frac{i}{2}\int \frac{dE}{2\pi}\frac{J(-E)J(E)}{E^2 - \omega^2}\right] \int \mathcal{D}x \, \exp\left[\frac{i}{2}\int \frac{dE}{2\pi}x(-E)(-E^2 + \omega^2)x(E)\right]$$
$$= Z[0]\exp\left[-\frac{i}{2}\int \frac{dE}{2\pi}\frac{J(-E)J(E)}{E^2 - \omega^2}\right]. \tag{4.50}$$

Since in any calculation of correlation functions the $Z[0]$'s cancel, we will not bother to evaluate $Z[0]$.

We have almost calculated the path-integral — but there has been a light hiccup. If

you examine the integral

$$\int \frac{dE}{2\pi} \frac{J(-E)J(E)}{E^2 - \omega^2}, \tag{4.51}$$

you notice that there is are poles at $E = \pm\omega$. This means that our integral is not well defined. The way to fix this is to add an infinitesimal complex component to these poles:

$$\int \frac{dE}{2\pi} \frac{J(-E)J(E)}{E^2 - \omega^2 + i\varepsilon}. \tag{4.52}$$

This can be traced back to the ε we used in (4.30) to define correlation factors.[1]

We can now switch back to the time-domain:

$$Z[J] = Z[0] \exp\left[\frac{i}{2} \int dt \, dt' \, J(t)G(t - t')J(t')\right] \tag{4.53}$$

where

$$G(t - t') = \int \frac{dE}{2\pi} \frac{-e^{iE(t-t')}}{E^2 - \omega^2 + i\varepsilon} = \frac{i}{2\omega} e^{-i\omega|t-t'|}. \tag{4.54}$$

Using (4.40), we can now compute the two-point function

$$\begin{aligned}
\langle\Omega|\mathcal{T}\phi(t_1)\phi(t_2)|\Omega\rangle &= -\frac{\delta}{\delta J(t_1)} \frac{\delta}{\delta J(t_2)} \exp\left[\frac{i}{2} \int dt \, dt' \, J(t)G(t - t')J(t')\right]\Bigg|_{J=0} \\
&= -i\frac{\delta}{\delta J(t_1)} \left(\int dt \, G(t_2 - t)J(t)\right) Z[J]\Bigg|_{J=0} \\
&= -iG(t_2 - t_1) + \text{ terms containing } J\,|_{J=0} \\
&= -iG(t_2 - t_1).
\end{aligned} \tag{4.55}$$

The same process works for any other correlation functions. When we have an odd number of q operators, there will be left-over factors of J, and so the result is always zero. For an even number of q operators, we can pair up each of the functional derivatives to get a factor of $G(t_2 - t_1)$. For example

$$\begin{aligned}
\langle\Omega|\mathcal{T}q(t_1)q(t_2)q(t_3)q(t_4)|\Omega\rangle = -[&G(t_1 - t_2)G(t_3 - t_4) \\
&+ G(t_1 - t_3)G(t_2 - t_4) \\
&+ G(t_1 - t_4)G(t_2 - t_3)],
\end{aligned} \tag{4.56}$$

and higher correlators can be found in a similar fashion.

4.2 Path Integrals for Free Fields

4.2.1 Path Integrals for Quantum Fields

Our results from the last section generalize to quantum field theory. Given a Lagrangian $\mathcal{L}(\phi)$, we can compute the amplitude

$$\langle\phi_a(\mathbf{x})|e^{-iHT}|\phi_b(\mathbf{x})\rangle = \int_{\phi_a(\mathbf{x},0)\to\phi_b(\mathbf{x},T)} D\phi \, \exp\left[\int_0^T \int d^3x \, \mathcal{L}[\phi]\right]. \tag{4.57}$$

---

[1]See, for instance, *Srednicki*, in which careful care is taken to retain this ε.

In this context, a path from ϕ_a to ϕ_b is a field configuration $\phi(\mathbf{x}, t)$ so that

$$\phi(\mathbf{x}, 0) = \phi_a(\mathbf{x}), \quad \phi(\mathbf{x}, T) = \phi_b(\mathbf{x}). \tag{4.58}$$

Although in principle (4.57) allows us to calculate any quantity of interest, it is rarely useful. In any application, the quantum state will be a superposition of $|\phi(\mathbf{x})\rangle$'s; that is, a functional with a number for every possible $\phi(\mathbf{x})$. For this reason, wavefunctions are useless in quantum field theory.

Generating functionals, on the other hand, are extremely useful in quantum field theory. Consider for instance

$$Z[J] = \int D\phi \, \exp\left[\int d^4x \, \mathcal{L}[\phi] - J(x)\phi(x)\right]. \tag{4.59}$$

The only difference from quantum mechanics is that our source term is now a function of both space and time. Functional derivatives generalize in the obvious way:

$$\frac{\delta}{\delta J(y)} J(x) = \delta^{(4)}(x - y). \tag{4.60}$$

The manipulation we used to derive (4.40) generalize to quantum field theory:

$$\langle\Omega|\mathcal{T}\phi(x)\phi(y)|\Omega\rangle = \frac{1}{Z[0]} \left(-i\frac{\delta}{\delta J(x)}\right)\left(-i\frac{\delta}{\delta J(y)}\right) Z[J]\bigg|_{J=0}, \tag{4.61}$$

and similar formulae hold for higher correlation functions.

Microcausality and Time Ordering

You may be wondering how time ordering can be defined in relativity. For any two spatially separated events x and y, observers will disagree as to event occurred first. Say that in your reference frame, $x^0 > y^0$, so that x occurs after y. You then use the (4.61) to conclude

$$\frac{1}{Z[0]} \left(-i\frac{\delta}{\delta J(x)}\right)\left(-i\frac{\delta}{\delta J(y)}\right) Z[J]\bigg|_{J=0} = \langle\Omega|\phi(x)\phi(y)|\Omega\rangle. \tag{4.62}$$

But now assume that I am in a reference frame where $y^0 > x^0$. Since the left-hand side of the above expression is Lorentz invariant, I will compute the same quantity that you did. But since our time orderings are different, I conclude

$$\frac{1}{Z[0]} \left(-i\frac{\delta}{\delta J(x)}\right)\left(-i\frac{\delta}{\delta J(y)}\right) Z[J]\bigg|_{J=0} = \langle\Omega|\phi(y)\phi(x)|\Omega\rangle. \tag{4.63}$$

The only way this is possible is if

$$\langle\Omega|\phi(x)\phi(y)|\Omega\rangle = \langle\Omega|\phi(y)\phi(x)|\Omega\rangle, \tag{4.64}$$

which we can rewrite as

$$\langle\Omega|[\phi(x), \phi(y)]|\Omega\rangle = 0 \text{ for } (x - y)^2 < 0. \tag{4.65}$$

We have so far considered the field $\phi(x)$, but the argument we gave should work for

any two local operators $\mathcal{O}_1(x)$ and $\mathcal{O}_2(y)$. We therefore conclude that

$$\langle \Omega | [\mathcal{O}_1(x), \mathcal{O}_2(y)] | \Omega \rangle \text{ for } (x-y)^2 < 0. \tag{4.66}$$

This condition, which is required if our time ordering is to make sense, is simply the microcausality condition applied to the ground state. That this condition falls out of both canonical and functional approaches to field theory speaks to the fundamental importance of causality in our theories.

4.2.2   The Klein-Gordon Field

We now turn to calculating the generating functional for the Klein-Gordon field. Our derivation will be almost identical to the case of the harmonic oscillator. We begin with the generating functional

$$Z[J] = \int \mathcal{D}\phi \ \exp\left(iS[q,J]\right) = \int \mathcal{D}\phi \exp\left[i \int d^4x \ \frac{1}{2}(\partial\phi)^2 - \frac{m^2}{2}\phi^2 + J\phi\right]$$
$$= \int \mathcal{D}\phi \exp\left[i \int d^4x \ \frac{1}{2}\phi(-\partial^2 - m^2)\phi + J\phi\right] \tag{4.67}$$

where in the last equality we made use of integration by parts. Now we Fourier transform

$$\phi(x) = \int \frac{d^4k}{(2\pi)^4} \ e^{ikx}\phi(k), \quad J(x) = \int \frac{d^4k}{(2\pi)^4} \ e^{ikx}J(k). \tag{4.68}$$

Substituting these expressions into the action, we can rewrite the action as

$$S[\phi, J] = \int \frac{d^4k}{(2\pi)^4} \left[\frac{1}{2}\phi(-k)(k^2 - m^2)\phi(k) + J(-k)\phi(k)\right]. \tag{4.69}$$

Now we perform the change of variable

$$\varphi(k) = \phi(k) - \frac{J(k)}{k^2 - m^2}. \tag{4.70}$$

This is a mere constant shift, so $\mathcal{D}\varphi = \mathcal{D}\phi$. In terms of our new variables, the action is now

$$S[\varphi, J] = \int \frac{d^4k}{(2\pi)^4} \left[\frac{1}{2}\varphi(-k)(k^2 - m^2)\varphi(k) - \frac{1}{2}\frac{J(-k)J(k)}{k^2 - m^2}\right]. \tag{4.71}$$

We now substitute this back into $Z[J]$:

$$Z[J] = \exp\left[-\frac{i}{2}\int \frac{d^4k}{(2\pi)^4}\frac{J(-k)J(k)}{k^2 - m^2}\right]\int \mathcal{D}\varphi \exp\left[i\int\frac{d^4k}{(2\pi)^4}\left[\frac{1}{2}\varphi(-k)(k^2 - m^2)\varphi(k)\right]\right]$$
$$= Z[0]\exp\left[-\frac{i}{2}\int\frac{d^4k}{(2\pi)^4}\frac{J(-k)J(k)}{k^2 - m^2}\right]. \tag{4.72}$$

Just as for the harmonic oscillator, the integral over φ simply produces a constant Z_0 which cancels in any calculation.

Going back into real space, we can write

$$Z[J] = Z[0]\exp\left[-\frac{1}{2}\int d^4x \ d^4y \ J(x)D_F(x-y)J(y)\right] \tag{4.73}$$

where we define the *Feynman* propagator:

$$D_F(x-y) = \int \frac{d^4k}{(2\pi)^4} \frac{ie^{-ik(x-y)}}{k^2 - m^2 + i\varepsilon}. \tag{4.74}$$

Note that we have reintroduced the infinitesimal $i\varepsilon$ which is required for the integral to be well-defined.

4.2.3   The Feynman Propagator

The Feynman propagator is of fundamental importance for quantum field theory. It is, for instance, a core element of Feynman diagrams. We can use (4.61) to relate it to a physical observable:

$$\langle \Omega | \mathcal{T} \phi(x)\phi(y) | \Omega \rangle$$
$$= \frac{1}{Z[0]} \left(-i\frac{\delta}{\delta J(x)} \right) \left(-i\frac{\delta}{\delta J(y)} \right) \left(Z[0] \exp\left[-\frac{1}{2} \int d^4w \; d^4z \; J(w) D_F(w-z) J(z) \right] \right)\Bigg|_{J=0}$$
$$= -\frac{\delta}{\delta J(x)} \left(-\left[\int dw^4 J(w) D_F(w-y) \right] \exp\left[-\frac{1}{2} \int d^4x \; d^4y \; J(x) D_F(x-y) J(y) \right] \right)\Bigg|_{J=0}$$
$$= D_F(x-y). \tag{4.75}$$

But we already know that if $x^0 > y^0$,

$$\langle \Omega | \phi(x)\phi(y) | \Omega \rangle = D(x-y) = \int \frac{d^3p}{(2\pi)^3} \frac{e^{ip(x-y)}}{2\omega_\mathbf{p}}, \tag{4.76}$$

which we showed in chapter 2. We can hence write

$$D_F(x-y) = \begin{cases} D(x-y) & x^0 > y^0 \\ D(y-x) & y^0 > x^0 \end{cases} \tag{4.77}$$

Alternatively, we can evaluate the k^0 integral in our definition of $D_F(x-y)$ by using a contour integral. When we do this, we find that

$$D_F(x-y) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_\mathbf{p}} \left(e^{ip(x-y)}\theta(x^0 - y^0) + e^{ip(y-x)}\theta(y^0 - x^0) \right), \tag{4.78}$$

where θ is the Heaviside step-function. It is not difficult to see that (4.77) and (4.78) are equal.

4.2.4   Higher Correlation Functions

We have seen how to compute the two-point function. What about higher correlation functions? Let's start with the four point function. To simply notation, we write:

$$J(x_i) \to J_i$$
$$D_F(x-y) \to D_{xy} \tag{4.79}$$
$$D_F(x_i - x_j) \to D_{ij}.$$

Furthermore, we will integrate over repeated variables.

Using this notation, the four-point function can computed as

$$
\begin{aligned}
\langle\Omega|\mathcal{T}\phi_1\phi_2\phi_3\phi_4|\Omega\rangle &= (-i)^4\frac{\delta}{\delta J_1}\frac{\delta}{\delta J_2}\frac{\delta}{\delta J_3}\frac{\delta}{\delta J_4}e^{-\frac{1}{2}J_xD_{xy}J_y}\bigg|_{J=0}\\
&= \frac{\delta}{\delta J_1}\frac{\delta}{\delta J_2}\left[-D_{34}+J_xD_{x4}J_yD_{y3}\right]e^{-\frac{1}{2}J_xD_{xy}J_y}\bigg|_{J=0}\\
&= \frac{\delta}{\delta J_1}\left[D_{34}J_xD_{x2}+D_{24}J_yD_{y3}+J_xD_{x4}D_{23}\right]e^{-\frac{1}{2}J_xD_{xy}J_y}\bigg|_{J=0}\\
&= D_{12}D_{34}+D_{13}D_{24}+D_{14}D_{23}.
\end{aligned}
\tag{4.80}
$$

This procedure generalizes to any n-point function. To derive the general form of n-point function, we first expand $Z[J]$ is a power series:

$$
\exp\left[-\frac{1}{2}J_xD_{xy}J_y\right] = \sum_{k=0}^{\infty}\frac{1}{k!}\left(-\frac{1}{2}J_xD_{xy}J_y\right)^k.
$$

Each time we differentiate by J_i, we remove a power of J. So if we differentiate the nth term in the power series $2n$ times, then this will gives us a constant term. Hence:

$$
\begin{aligned}
\langle\Omega|\mathcal{T}\phi_1...\phi_{2n}|\Omega\rangle &= i^{2n}\frac{\delta}{\delta J_1}...\frac{\delta}{\delta J_{2n}}\sum_{k=0}^{\infty}\frac{1}{k!}\left(-\frac{1}{2}J_xD_{xy}J_y\right)^k\bigg|_{J=0}\\
&= \left((-1)^n\frac{\delta}{\delta J_1}...\frac{\delta}{\delta J_{2n}}\left[\frac{1}{n!}\left(-\frac{1}{2}J_xD_{xy}J_y\right)^n\right]+O(J^2)\right)\bigg|_{J=0}\\
&= \frac{1}{2^nn!}\frac{\delta}{\delta J_1}...\frac{\delta}{\delta J_{2n}}(J_xD_{xy}J_y)^n.
\end{aligned}
\tag{4.81}
$$

Notice that if we instead had an odd number of fields, our first term would have been $O(J)$. This proves that all odd correlations functions are zero.

Turning back to the even case, we can use the product rule repeatedly to calculate

$$
\frac{1}{2^nn!}\frac{\delta}{\delta J_1}...\frac{\delta}{\delta J_{2n}}(J_xD_{xy}J_y)^n = \frac{1}{n!2^n}\sum_{\text{all permutations}}D_{i_1i_2}D_{i_3i_4}...D_{i_ni_{2n}}.
\tag{4.82}
$$

To simplify this further we need to use some combinatorics. If we swap the indices i_1 and i_2, we would then get a new permutation. But since $D_{xy}=D_{yx}$, this new permutation will give the same contribution as the old one. There are 2^n ways to swap each of the terms in the propagator, and each of these gives an equal contribution to the sum. Similarly, we could swap i_1 and i_2 with i_3 and i_4. This new permutation will again gives the same contribution to our sum. Rather than sum over all permutations of the numbers 1 through to $2n$, we should instead consider all unique ways of pairing these numbers. We can then calculate the correlation function as

$$
\langle\Omega|\mathcal{T}\phi_1...\phi_{2n}|\Omega\rangle = \sum_{\text{all pairings}}D_{i_1i_2}D_{i_3i_4}...D_{i_{2n-1}i_{2n}}.
\tag{4.83}
$$

Feynman Diagrams

There is a very nice way to write our computations diagrammatically, using *Feynman diagrams*. We could for instance write the two-point function as

$$\langle \Omega | \mathcal{T} \phi_1 \phi_2 | \Omega \rangle = D_{12} = 1 \bullet\!\!-\!\!\!-\!\!\!-\!\!\bullet \, 2 \tag{4.84}$$

and the four point function as

$$\langle \Omega | \mathcal{T} \phi_1 \phi_2 \phi_3 \phi_4 | \Omega \rangle = D_{12} D_{34} \qquad + D_{13} D_{24} \qquad + D_{14} D_{23}$$

$$\tag{4.85}$$

To calculate the $2n$-point function, we draw a point for each of the $2n$ indices, connecting pairs via a line. To each connect pair we associate the propagator $D_{i_1 i_2}$, producting together all of the propagators. We then sum over all of the different possible diagrams.

These are our first examples of Feynman diagrams. Since we are working in a free theory, the rules are kind of trivial. But as we shall see, they provide the foundations for the Feynman rules of interacting theories.

4.2.5   The Inverse-Square Law

So far our description of $Z[J]$ has been very abstract. To make it more concrete, we shall try to connect it to the energy of our system. If our source J are time-dependent, we know that the Hamiltonian for our system,

$$H[J] = \int d^3x \, \Pi \frac{\partial \mathcal{L}}{\partial \dot{\phi}} - \mathcal{L} = \int d^3x \, H_0 - \phi J. \tag{4.86}$$

will be time-independent. Using the definition of the $Z[J]$ as a path-integral,

$$Z[J] = \lim_{T \to (1-i\varepsilon)\infty} \langle \Omega | e^{-iH[J]T} | \Omega \rangle = \lim_{T \to (1-i\varepsilon)\infty} e^{-iW[J]T} \tag{4.87}$$

where $W[J]$ is the ground-state energy for the source configuration J. Taking the logarithm of both sides,

$$i \log(Z[J]) = \lim_{T \to \infty} T W[J]. \tag{4.88}$$

These expression are both formally infinite. If however, we carefully divide through by T and then take $T \to \infty$, we can extract the energy of our state, $E[J]$.

Specializing now to the Klein-Gordon theory:

$$i \log(Z[J]) = i \log \left(Z[0] \exp \left[-\frac{1}{2} \int d^4x d^4y J(x) D_F(x - y) J(y) \right] \right)$$
$$= i \log(Z[0]) - \frac{i}{2} \int d^4x d^4y J(x) D_F(x - y) J(y). \tag{4.89}$$

The first term is the vacuum energy for the Klein-Gordon field. We calculated this quantity in Section 3.3, and know it is infinite. But as in Section 3.3, we know that the vacuum

energy is not observable and does not depend on J. We can simply ignore it, and take

$$\lim_{T\to\infty} TE[J] = \lim_{T\to\infty} T(W[J] - W[0]) = -\frac{i}{2}\int d^4x\; d^4y\; J(x)D_F(x-y)J(y). \quad (4.90)$$

Using this formula, we can calculate the energy of any configuration of sources.

To study this quantity, specialize even further to the case of static point sources:

$$J(\mathbf{x}) = J_1(\mathbf{x}_1) + J_2(\mathbf{x}_2) = \delta^{(3)}(\mathbf{x} - \mathbf{x}_1) + \delta^{(3)}(\mathbf{x} - \mathbf{x}_2) \quad (4.91)$$

This can be thought of as the charge distribution given by two particles at $\mathbf{x}_1$ and $\mathbf{x}_2$, a trick which should be familiar from electrostatics. Our energy is now

$$\lim_{T\to\infty} TE[J]$$
$$= -\frac{i}{2}\int d^4x d^4y\, (J_1(x)D_F(x-y)J_1(y) + 2J_1(x)D_F(x-y)J_2(y) + J_2(x)D_F(x-y)J_2(y))\,. $$
$$(4.92)$$

We are interested in the interaction of our points, given by the $J_1 J_2$ term in the above expression. The $J_1 J_1$ term on the other hand will be the same regardless of where $\mathbf{x}_2$ is; we can interpret it as the self-interaction of the point. Likewise we do not care about $J_2 J_2$. So we compute

$$\lim_{T\to\infty} TE_{12} = -i\int d^4x\; d^4y\; J_1(x)D_F(x-y)J_2(y). \quad (4.93)$$

This will tell us how the energy between to two charges varies with distance. Inserting our formula for the Feynman propagator:

$$\begin{aligned}
\lim_{T\to\infty} TE_{12} &= \int d^4x\; d^4y\; J_1(x)J_2(y)\int \frac{d^4k}{(2\pi)^4}\frac{e^{ik(x-y)}}{k^2-m^2-i\varepsilon} \\
&= \int \frac{d^4k}{(2\pi)^4}d^4x\; d^4y\; \delta(\mathbf{x}-\mathbf{x}_1)\delta(\mathbf{y}-\mathbf{y}_1)\frac{e^{ik(x-y)}}{k^2-m^2-i\varepsilon} \\
&= \int \frac{d^4k}{(2\pi)^4}dx^0\; dy^0\; \frac{e^{ik^0(x^0-y^0)-i\mathbf{k}\cdot(\mathbf{x}_2-\mathbf{x}_1)}}{(k^0)^2-\mathbf{k}^2-m^2-i\varepsilon}
\end{aligned} \quad (4.94)$$

Performing the integral over y^0 gives us a delta function $2\pi\delta(k^0)$. We can then integrate k^0 to get

$$\lim_{T\to\infty} TE_{12} = \int \frac{d^3k}{(2\pi)^3}d^0x\; \frac{e^{-i\mathbf{k}\cdot(\mathbf{x}_2-\mathbf{x}_1)}}{-\mathbf{k}^2-m^2-i\varepsilon} = \left(\int dx^0\right)\int \frac{d^3k}{(2\pi)^3}\frac{e^{-i\mathbf{k}\cdot(\mathbf{x}_2-\mathbf{x}_1)}}{-\mathbf{k}^2-m^2-i\varepsilon}. \quad (4.95)$$

Our integral over x^0 just gives us a factor of T, cancelling the factor of T on the left-hand side. All we have to do now is evaluate

$$E_{12} = -\int \frac{d^3k}{(2\pi)^3}\frac{e^{-i\mathbf{k}\cdot(\mathbf{x}_2-\mathbf{x}_1)}}{\mathbf{k}^2+m^2}. \quad (4.96)$$

The factor of $i\varepsilon$ could be removed because the denominator can never be zero. Solving the integral is a little tricky, and yields

$$E_{12} = -\frac{1}{4\pi r}e^{-mr}, \quad (4.97)$$

where $r = |\mathbf{x}_1 - \mathbf{x}_2|$ is the distance between the particles. Since this is negative and monotonically increases with r, we have found that the force between the two points is positive! This type of potential is known as a Yukawa potential.

Yukawa discovered his potential while studying atomic nuclei. He reasoned from the observed range of the nuclear force $R \sim 1$ fm that there exists a new particle of mass $\sim 1/R$. A particle with the correct mass, the pion, was discovered soon after.

Quantum field theory gives us a new perspective on forces. They are not a fundamental concept, but arise from the interaction of particles with a field. Because our potential was derived from the Feynman propagator, we often say that forces are generated by the virtual exchange of a particle; in this case a ϕ particle. We can associate the photon to the electromagnetic force, and the graviton to the gravitational force.

You may have wondered why both Newton's law and Coulomb's law use a inverse-square force. Taking $m \to 0$, we find that for a massless field the potential

$$E_{12} = -\frac{1}{4\pi r}. \tag{4.98}$$

We conclude that massless particles result in inverse-square forces! Of course, the photon and graviton are not scalar fields, and this causes some subtleties. We will study the photon in Chapter 7, showing that the familiar properties of electromagnetism can be reproduced.

4.3 Statistical Mechanics and Finite Temperatures

4.3.1 The Euclidean Path-Integral

In our discussion of the path-integral, we have never stopped to ask is whether the path-integral is actually well-defined. After all, if we consider

$$Z = \int \mathcal{D}\phi \, \exp\left[i \int d^3x \, dt \, \frac{1}{2}(\partial\phi)^2 - V(\phi)\right], \tag{4.99}$$

then the integrand oscillates more quickly as the action varies, but always has norm one. We have no guarantee that the integral converges! This problem is not unique to path-integrals; consider, for instance, the *Fresnel integral*

$$\int_{-\infty}^{\infty} dx \exp[ix^2] = \int_{-\infty}^{\infty} dx \, [\cos(x^2) + i\sin(x^2)]. \tag{4.100}$$

This integral rapidly oscillates as $x \to \infty$, and so it is not immediately clear how one should define this integral.[2]

Wick rotation is a clever trick to make these integrals better defined. The idea is simple: make time imaginary. By setting $x_4 = it$, we can define the *Euclidean path-integral*

$$Z = \int \mathcal{D}\phi \, \exp\left[-\int d^4x \, \frac{1}{2}(\nabla\phi)^2 + V(\phi)\right] \tag{4.101}$$

---

[2]There exist various definitions will allow us to assign meaning to otherwise divergent integrals, such as the *Cauchy Principal Value*. But care is necessary for these cases, because manipulations which are valid for ordinary integrals may fail when applied to these pathological cases.

where, $dx_4 = idt$ and ∇ is the Euclidean derivative

$$(\nabla \phi)^2 = \left(\frac{\partial \phi}{\partial x_1} \right)^2 + \left(\frac{\partial \phi}{\partial x_2} \right)^2 + \left(\frac{\partial \phi}{\partial x_3} \right)^2 + \left(\frac{\partial \phi}{\partial x_4} \right)^2. \tag{4.102}$$

It is now time for some useful language. We say a field theory has dimension $d + 1$ if it has d spatial dimensions and 1 time dimension. For most of these notes we have focussed on theories with dimension $3 + 1$, since our universe appears to have three spatial dimensions. Quantum mechanics is just quantum field theory in $0 + 1$ dimensions, since we don't have any spatial dimensions. Quantum field theory is even easier in $0 + 0$ dimensions — in that case it is just plain old integration!

In Euclidean field theory there is no time direction, so we say that we have a $d + 0$ dimensional theory. Our Wick rotation has related a $d + 1$ QFT to a $(d + 1) + 0$ dimensional QFT.

To link this back to statistical mechanics, recall that the partition function for a statistical system is

$$Z = \sum_{\text{states}} \exp\left(-\beta E_s \right) \tag{4.103}$$

where E_s is the energy of a state s, $\beta = 1/T$, and $e^{-\beta E}$ is the Boltzmann factor. If for instance, we had an N-particle system

$$E(p_i, q_i) = \sum_i \frac{1}{2m} p_i^2 + V(q_1, ..., q_N), \tag{4.104}$$

then the partition function is

$$\begin{aligned}
Z &= \prod_i \int dp_i \ dq_i \ \exp\left(-\beta E(p, q) \right) \\
&= \left[\prod_i \int dp_i \ \exp\left(-\beta \sum_i \frac{1}{2m} p_i^2 \right) \right] \left[\prod_i \int dq_i \ \exp\left(-\beta V(q_1, ..., q_N) \right) \right].
\end{aligned} \tag{4.105}$$

In statistical field theories, we are interested not in particles, but in statistical fields. Such fields are very common in physics. We could, for instance, consider the density of a fluid, or the magnetisation of a lump of iron. At large scales, we can average over the atoms and simply treat these systems as fields. Statistical field theory describes the fluctuations in these fields. Although usually these fluctuations are small, around second order phase transitions they become very important.

In the relativistic world, cosmologists need statistic field theory to describe the earliest stages of our universe, where all matter was a hot soup of particles. The cosmic microwave background, a relic of these era, is a classic example of a statistical field. The tools we develop here can even be applied to the stock market, which is really just a $1 + 0$ dimensional field theory!

For our purposes, it is easiest to consider a scalar field in d dimensions, with Hamiltonian

$$H[\pi, \phi] = \int d^d x \ \frac{1}{2} \pi^2 + \frac{1}{2} (\nabla \phi)^2 + V(\phi). \tag{4.106}$$

This is similar to the Hamiltonian we derived in Section 2.3, except we have allowed a more general potential $V(\phi)$. Instead of integrating over coordinates such as p and q, we should integrate over the allowed field configurations: a path-integral! We can now write
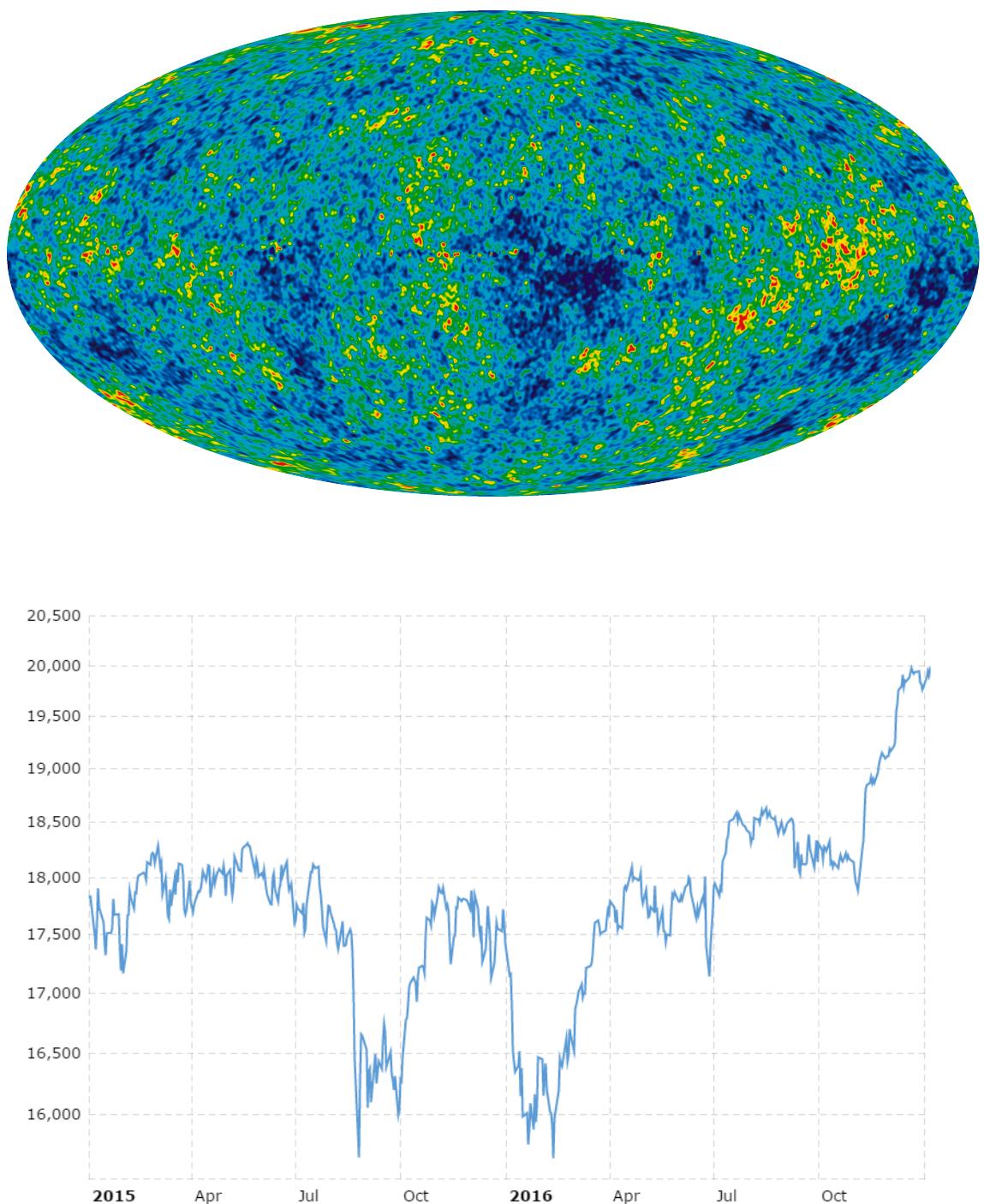
Figure 4.1: Some examples of statistical fields. The top image is of the cosmic microwave background, the bottom of the Dow-Jones index.

the partition function for our field

$$
\begin{aligned}
Z &= \int \mathcal{D}\pi \, \mathcal{D}\phi \, \exp\left(-\beta H[\pi, \phi]\right) \\
&= \left[\int \mathcal{D}\pi \, \exp\left(-\beta \int d^d x \, \frac{1}{2}\pi^2\right)\right]\left[\int \mathcal{D}\phi \, \exp\left(-\beta \int d^d x \, \frac{1}{2}(\nabla\phi)^2 + V(\phi)\right)\right].
\end{aligned}
$$
(4.107)

The right-hand factor governs the probability distribution for the field ϕ; if this is all we care about, we can just consider the *reduced partition function*:

$$
Z_{\text{red}} = \int \mathcal{D}\phi \, \exp\left(-\beta \int d^d x \, \frac{1}{2}(\nabla\phi)^2 + V(\phi)\right).
$$
(4.108)

Here comes the punchline: apart from the factor of β, this is just the Euclidean path-integral (4.101)! We conclude that Euclidean $d+1$ quantum field theories are equivalent to classical statistical field theory in $d+1$ spatial dimensions.

4.3.2 Quantum Statistical Field Theory

Having shown that classical statistical field theory is really just quantum field theory, let's now move on to quantum statistical mechanics. Given a Hamiltonian H with eigenstates $|E_i\rangle$, the density matrix for the system will be

$$
\rho = \sum_i e^{-\beta E_i} |E_i\rangle \langle E_i| = e^{-\beta H},
$$
(4.109)

and so the partition function is

$$
Z = \sum_i e^{-\beta E_i} = \text{tr}(\rho) = \text{tr}(e^{-\beta H}) = \sum_n \langle n| e^{-\beta H} |n\rangle
$$
(4.110)

for any orthonormal basis $|n\rangle$.

We already know from the path-integral that

$$
\langle F| e^{-iHT} |I\rangle = \int_{I \to F} \mathcal{D}q \, \exp\left(i \int_0^T dt \, \mathcal{L}[q]\right).
$$
(4.111)

Now we simply Wick rotate $T \to -i\beta$ and set $|I\rangle = |F\rangle = |n\rangle$, to find

$$
Z = \text{tr}(e^{-\beta H}) = \int_{\text{PBC}} \mathcal{D}q \, \exp\left(-\int_0^\beta \mathcal{L}_{\text{E}}[q]\right).
$$
(4.112)

The path-integral is now performed over all periodic boundary conditions, where $q(0) = q(\beta)$.

The extension to fields is obvious. If H is the Hamiltonian of our quantum field theory, we calculate

$$
Z = \text{tr}\left(e^{-\beta H}\right) = \int_{PBC} \mathcal{D}\phi \, \exp\left(-\int_0^\beta d\tau \int d^d x \, \mathcal{L}[\phi]\right),
$$
(4.113)

where periodic boundary conditions require that $\phi(\mathbf{x}, 0) = \phi(\mathbf{x}, \beta)$. Quantum statistical field theory in d-dimensions is just Euclidean field theory in $(d+1)$-dimensions with

periodic boundary conditions! I think that this is one of the most amazing results in all of physics. At heart, it comes from the similarity between the time-evolution operator e^{-iHT} and the Boltzmann factor $e^{-\beta H}$.

Like all great results in physics, this connection is a deeply practical one. In this course we will develop a number of methods to compute things in quantum field theory. With minor modifications, all of these methods carry over to statistical mechanics.

Interacting Fields

So far we have only discussed free fields. By solving the Klein-Gordon theory, we have seen that it is just a collection of non-interacting relativistic particles. This means that nothing interesting ever happens: the particles never scatter off each other, and nor will they ever interact with a measuring device.

In this chapter, we will begin to study more complicated theories. We could for instance examine the interacting scalar field:

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda_3}{3!}\phi^3 - \frac{\lambda_4}{4!}\phi^4 - ... \tag{5.1}$$

where the coefficients λ_n are known as *coupling constants*. Writing the ϕ^3 and ϕ^4 operators using raising and lowering operators, we now expect our Hamiltonian to include terms such as

$$a_{\mathbf{p}}^\dagger a_{\mathbf{p}}^\dagger a_{\mathbf{p}}^\dagger a_{\mathbf{p}}^\dagger \text{ and } a_{\mathbf{p}}^\dagger a_{\mathbf{p}}^\dagger a_{\mathbf{p}}^\dagger a_{\mathbf{p}}, \tag{5.2}$$

creating and destroying particles. This hints that our theory no longer preserves particle number; we could confirm this by calculating $[H, N] \neq 0$.

In the previous chapter, we learnt that to compute correlation functions, we simply need the generating functional

$$Z[J] = \int \mathcal{D}\phi \, \exp\left[i \int d^4x \, \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda_3}{3!}\phi^3 - \frac{\lambda_4}{4!}\phi^4 + ... + J\phi\right]. \tag{5.3}$$

The bad news is that this integral is impossible to solve. The good news is that if the coupling constants are small, we can use perturbation theory to systematically calculate $Z[J]$, using the tools of the previous chapter. We can then calculate any n-point function to arbitrary accuracy.

Correlation functions are all well and good, but how does this lead to experimentally observable quantities, such as scattering amplitudes and decays? Linking the two will require us to use the canonical framework. Once we have developed this connection, we will be able to systematically calculate physical observables in our theory.

5.1 Perturbation Theory and Unit Analysis

What does it mean for the coupling constant λ_n to be small? At first you might expect us to require $\lambda_n \ll 1$, but with a little though you'll see that this doesn't work. The problem is that λ_n will not necessarily be unitless.

Because we work in natural units where $\hbar = 1$, the action has to unitless: $[S] = 0$. Using

the convention that energy has units $+1$, the integration measure has units $[d^4x] = -4$, and so

$$[\mathcal{L}] = [S] - [d^4x] = +4. \tag{5.4}$$

What are the units of the scalar field? Since $[\partial_\mu] = +1$, we can deduce that

$$[\phi] = 1, \quad [m] = 1, \quad [\lambda_n] = 4 - n. \tag{5.5}$$

Herein lies the problem; to say that λ_n is small, we need to know what energy scale Λ we are working at. The coupling constants divide into three categories:

1. *Relevant couplings*: $[\lambda_3] = 1$. In a process of energy Λ, the dimensionless quantity is $\lambda_3 \Lambda^{-1}$. This means that at high energies with $\Lambda \gg \lambda_3$, the $\lambda_3 \phi^3$ term will be a small perturbation, but at low energies will be a large one. This type of coupling is called *relevant* because it will be most important at low energies. For a relativistic theory, the energy Λ is always greater than the mass m of the particle. So if $m \gg \lambda_3$, then λ_3 will always be a small perturbation.

2. *Marginal couplings*: $[\lambda_4] = 0$. Since this is unitless, we simple require $\lambda_4 \ll 1$ for perturbation theory to work.

3. *Irrelevant couplings*: $[\lambda_n] < 0$ for $n > 4$. The dimensionless quantity is $\lambda_n \Lambda^{4-n}$, which is small at low energies but large at high energies. We call these perturbations *irrelevant*, as they can be ignored at low energies (which are the easiest energies to probe experimentally).

In quantum field theory, we almost always consider only relevant and marginal couplings. Such theories are called *renormalizable theories*. If on the other hand a theory contains an irrelevant coupling, it is *non-renormalizable*. This causes major problems in perturbation theory; infinities plague calculations and they cannot be removed. To fully understand the reasons for this require a understanding of *renormalisation*, which is unfortunately beyond the scope of this course. Nevertheless, there is an immediate pay-off to these considerations: we can focus on theories involving ϕ^3 and ϕ^4 couplings only, whilst ignoring an infinite number of more complicated ϕ^n couplings.

Why is QFT simple?

Considering only renormalisable theories simplifies quantum field theory immensely. If we were studying quantum mechanics for instance, we would consider the Schrödinger equation with any possible potential $V(x)$. We can have harmonic oscillators, infinite square wells, sine waves, tanh potentials, or any other arbitrary shape.

Quantum field theory is however, much more discerning. If we have a scalar field theory, then we only need to worry about ϕ^3 and ϕ^4 terms. All we have to do is measure m, λ_3 and λ_4, and then we are done! Of course, we could have multiple scalar particles, and perhaps more interesting things like fermions or gauge bosons. But even then, all we have to do is write down the renormalisable theory, and then measure the masses and coupling constants.

The Standard Model, for instance, is the only renormalisable theory compatible with the particles we have observed. This doesn't necessarily mean that the Standard Model is correct; what it does mean is that, if the Standard Model is wrong, there are new particles out there waiting to be discovered.

We know that the Standard Model cannot be the final word in particle physics. For a start, gravity should become important at energies around $M_{\text{pl}} = 10^{19}$ GeV. But the energy probed by the LHC is only 10^3 GeV, a 10^{-16} difference. So the real reason QFT is simple is that there is a large difference in energy scales, and so non-renormalisable terms are extremely difficult to observe.

Extremely difficult is, however, not the same as impossible. There are three important, non-renormalisable effects which we can (potentially) observe:

1. Gravity. The Einstein-Hilbert action is famously non-renormalisable, and this is why gravity has proved so hard to quantize. The reason we experience gravity at all is because is long-ranged and universally attractive — but it takes a whole planet-load of particles before gravity becomes really noticeable!

2. Neutrino masses come from non-renormalisable terms of dimension -1. This gives them tiny, but non-zero, masses through the *seesaw mechanism.*

3. Proton decay. This comes from terms of dimension -2 in the Standard Model. If protons decay, they will do so *extremely* slowly: current experimental evidence show that the half-life must be greater than 10^{34} years! Nevertheless, the absence of proton decay gives us important constraints on possible high energy physics.

5.1.1 $\lambda\phi^4$ Theory

Putting aside both unit analysis and philosophy, we can now introduce the hero of this chapter: $\lambda\phi^4$ theory. This is the interacting field theory with Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda^4}{4!}\phi^4. \tag{5.6}$$

We will work with this theory because of its simplicity. (You might think that an ϕ^3 term would be simpler again, but such a theory has no ground state; fixing this would require an extra ϕ^4 term.) Since λ is unitless, we can use perturbation theory when $\lambda \ll 1$.

For the moment we are using $\lambda\phi^4$ theory for purely pedagogical reasons. Nevertheless, it is a very important theory in the real world. The Higgs self-interaction is described by a ϕ^4 interaction, and in statistical mechanics ϕ^4 theory is used to describe phase transitions.

5.2   Feynman Diagrams

Our aim is to calculate correlation functions in $\lambda\phi^4$ theory when $\lambda \ll 1$. We will start by considering the two-point function. Using (4.61),

$$\langle\Omega|\mathcal{T}\phi(x)\phi(y)|\Omega\rangle = \frac{\int \mathcal{D}\phi \ \phi(x)\phi(y) \exp\left[i\int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4\right]}{\int \mathcal{D}\phi \ \exp\left[i\int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4\right]}. \tag{5.7}$$

To calculate the two-point function we need to evaluate the path-integrals in both the denominator and the numerator. We will start with the numerator

$$\begin{aligned}
G_\lambda(x,y) &\equiv \int \mathcal{D}\phi \ \phi(x)\phi(y) \exp\left[\int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4\right] \\
&= \int \mathcal{D}\phi \ \phi(x)\phi(y) \exp\left[-i\int d^4z\frac{\lambda}{4!}\phi^4\right] \exp\left[i\int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2\right].
\end{aligned} \tag{5.8}$$

This manipulation suggests that we should expand the first exponential as a power series:

$$\exp\left[-i\int d^4z \frac{\lambda}{4!}\phi^4\right] = 1 - i\int d^4z \frac{\lambda}{4!}\phi^4 - \frac{1}{2}\left(\int d^4z \frac{\lambda}{4!}\phi^4\right)^2 + \dots$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{-i\lambda}{4!}\right)^k \left(\int d^4z \ \phi(z)^4\right)^k. \tag{5.9}$$

Inserting this back into (5.8), we get

$$G_\lambda = \int \mathcal{D}\phi \ \sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{-i\lambda}{4!}\right)^k \left(\int d^4z \ \phi(z)^4\right)^k \phi(x)\phi(y)\exp\left[i\int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2\right]$$

$$= \sum_{k=0}^{\infty} \left(\frac{-i\lambda}{4!}\right)^k \frac{1}{k!}\int \mathcal{D}\phi \ \left(\int d^4z \ \phi(z)^4\right)^k \phi(x)\phi(y)\exp\left[i\int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2\right]$$

$$= G^0 + \lambda G^{(1)} + \lambda^2 G^{(2)} + \dots, \tag{5.10}$$

allowing us to systematically calculate G_λ as a power series.

The first term is given by

$$G^{(0)} = \int \mathcal{D}\phi \ \phi(x)\phi(y)\exp\left[i\int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2\right]. \tag{5.11}$$

We now recognise this path-integral as a correlation function in the free scalar theory:

$$\langle\Omega|\phi(x)\phi(y)|\Omega\rangle_F = \frac{1}{Z_0}\int \mathcal{D}\phi \ \phi(x)\phi(y)\exp\left[i\int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2\right] \tag{5.12}$$

where the subscript F will be used to denote correlators in the free field theory, and where

$$Z_0 = \int \mathcal{D}\phi \ \exp\left[i\int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2\right]. \tag{5.13}$$

In section 4.2.4 we showed that

$$\langle\Omega|\phi(x)\phi(y)|\Omega\rangle_F = D_F(x-y), \tag{5.14}$$

and so we conclude that

$$G^{(0)} = Z_0 D_F(x-y) = Z_0 \ \ x \bullet\!\!-\!\!\!-\!\!\bullet y \quad . \tag{5.15}$$

Here, as in Section 4.2.4, we have introduced a diagrammatic way to write this quantity. The dots x and y represent points, and the line between them represents the factor of $D_F(x-y)$.

Now we move onto the first-order correction

$$G^{(1)} = \int \mathcal{D}\phi \ \phi(x)\phi(y)\left(-i\int d^4z \frac{\phi^4}{4!}\right)\exp\left[i\int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2\right]$$

$$= \frac{-i}{4!}\int d^4z \int \mathcal{D}\phi \ \phi(x)\phi(y)\phi(z)^4 \exp\left[i\int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2\right]. \tag{5.16}$$

Once again we have reduce the problem to the computation of a correlation function in

the free theory — this time, we need the 6-point function:

$$G^{(1)} = \frac{-i}{Z_0 4!} \int d^4z \, \langle\Omega|\mathcal{T}\phi(x)\phi(y)\phi(z)^4|\Omega\rangle. \tag{5.17}$$

In light of this, it is worth recalling the formula we derived in Section 4.2.4:

$$\langle\Omega|\mathcal{T}\phi(x_1)...\phi(x_{2n})|\Omega\rangle = \sum_{\text{all pairings}} D_F(x_{i_1} - x_{i_2})...D_F(x_{i_{2n-1}} - x_{i_{2n}}) \tag{5.18}$$

For (5.17) there are 15 different pairings, but many of them equal. If we pair $\phi(x)$ and $\phi(y)$, there are three different ways to pair the four $\phi(z)$'s. If instead we pair x and y to a z each, there are $4 \cdot 3 = 12$ ways to do this. Hence

$$\langle\Omega|\mathcal{T}\phi(x)\phi(y)\phi(z)^4|\Omega\rangle = 3D_F(x-y)D(z-z)^2 + 12D(x-z)D(y-z)D(z-z). \tag{5.19}$$

Again we will try to represent $G^{(1)}$ diagrammatically, with each point as a dot, and each D_F as a line. We now however have to deal both with not only 'external' points x and y, but also the point z, which we associate with a factor $(-i\lambda) \int d^4z$. For the moment we will not worry about the constant factors, and simply write

$$G^{(1)} = \left(x\, \bullet\!\!-\!\!\bullet\, y \quad \infty \right) + \left(x\, \bullet\!\!-\!\!\bullet\!\!-\!\!\bullet\, y \right) \tag{5.20}$$

The lines in the diagrams are called *propagators*, since they give factors of D_F, and the internal points are called *vertices*. Since the propagator gives the amplitude for a free particle to propagate from x to y, it is tempting to interpret these diagrams as processes, where particle are created, live and die.

We can extend this process to any order in λ. In general

$$\begin{aligned}
G^{(n)} &= \frac{1}{n!} \int \mathcal{D}\phi \, \phi(x)\phi(y) \left(-i\int d^4z \, \frac{\phi^4}{4!}\right)^n \exp\left[i\int d^4z \, \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2\right] \\
&= \frac{(-i)^n}{n!(4!)^n} \int d^4z_1...d^4z_n \int \mathcal{D}\phi \, \phi(x)\phi(y)\phi(z_1)^4...\phi(z_n)^4 \exp\left[i\int d^4z \, \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2\right] \\
&= \frac{Z_0(-i)^n}{n!(4!)^n} \int d^4z_1...d^4z_n \langle\Omega|\mathcal{T}\phi(x)\phi(y)\phi(z_1)^4...\phi(z_n)^4|\Omega\rangle_F.
\end{aligned}$$

$$\tag{5.21}$$

The correlation function can then be evaluated as the product of Feynman propagators, summing over all the different ways to pair the vertices and external points. Each possible pairing can be associated to a Feynman diagram, made from propagators, vertices, and external points. To evaluate the diagram we use the *Feynman rules*:

1. For each propagator, $z_i\, \bullet\!\!-\!\!-\!\!-\!\!\bullet\, z_j = D_F(z_i - z_j)$

2. For each vertex, $z = (-i\lambda)\int d^4z$ (5.22)
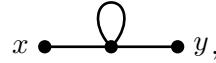
3. For each external point, $x\, \bullet\!\!-\!\!-\!\!-\! = 1$

4. Divide by the symmetry factor.

The only thing we now have to work out is the symmetry factor; this arises from all the 'different' ways to pair vertices that give the same Feynman diagram. To work this out, first we note that we can always interchange the n internal vertices. Since there are $n!$ ways do this, this exactly cancels the factor of $1/n!$ in (5.21). Each vertex has four lines attaches, and so the various placement of these lines generates a factor of 4! for each vertex. This cancels the factor of $1/(4!)^n$ in (5.21).

This counting scheme unfortunately over-counts certain diagrams. If a line starts and ends on the same vertex for instance, our scheme counts these as two different pairings — but we know that these are actually the same. This means that for the diagram

$$x \bullet \! \bullet y,$$

the symmetry factor is 2. Likewise, the symmetry factor for

is 8, since there are two different pairs of vertices which can be swapped, plus we can also swap the two propagators. In general, the symmetry factor is the number of possible ways to interchange components of a diagram without changing the diagram itself. Fortunately, it is rare to have to deal with symmetry factors larger than 2, so there is no need panic! If you are ever in doubt about the symmetry factor, it is easiest to work out all of the equivalent pairings which give the same Feynman diagram, and then divide by the factor of $n!(4!)^n$.

So far have only discussed the 2-point function. But our arguments can easily be generalized; for the n-point function we simply consider all Feynman diagrams with n external lines:

$$\int \mathcal{D}\phi \ \phi(x_1)...\phi(x_n) \exp\left[i \int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4\right]$$

$$= Z_0 \times (\text{sum of Feynman diagrams with } n \text{ external points}). \qquad (5.23)$$

The Feynman Rules in Momentum Space

The Feynman rules we just defined are often called the *position-space* Feynman rules, since the external points and vertices are associated with points in spacetime. For most calculations however, it is much more convenient to work in momentum space, where the propagator is

$$D_F(x - y) = \int \frac{d^4p}{(2\pi)^4} \frac{ie^{ip(x-y)}}{p^2 - m^2 + i\varepsilon}. \qquad (5.24)$$

Now to each propagator we associate a 4-momentum p. We add to each propagator an arrow, to represent the direction of the momentum. Flipping the arrow is the same as flipping the sign of the momentum.

What happens at a vertex? When four lines meet at a point z, our z-dependent factors

are

$$\to -i\lambda \int d^4z \ e^{-ip_1 z} e^{-ip_1 z} e^{ip_3 z} e^{ip_4 z}$$

$$= (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4).$$

(5.25)

This tells us that momentum is conserved at each of the vertices. The delta functions allow us to integrate over many of the propagator momenta. After doing this, we get the following *momentum-space Feynman rules:*

1. For each propagator, $= \dfrac{i}{p^2 - m^2 + i\varepsilon}$

2. For each vertex, $= -i\lambda$

3. For each external point, x $= e^{-ipx}$

4. Impose conservation of momentum at each vertex.

5. Integrate over each undetermined momentum: $\displaystyle\int \frac{d^4p}{(2\pi)^4}$

6. Divide by the symmetry factor.

(5.26)

It is again very tempting to give a particle interpretation to these diagrams, describing the creation, scattering, and destruction of particles of a give momentum. By integrating over the possible momenta, we are simply summing the amplitudes for the different ways the process can occur. In the next section we will make this intuition exact. But for now, let us continue our calculation of the n-point correlation function.

Connected Diagrams

So much for the numerator of (5.7). We now simply need to deal with the denominator:

$$\int \mathcal{D}\phi \ \exp\left[i \int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4\right].$$

(5.27)

Notice however, that this is really the '0-point' function, and so we can use (5.23) to compute is:

$$\int \mathcal{D}\phi \ \exp\left[i \int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4\right]$$

$$= Z_0 \times (\text{sum of Feynman diagrams with no external points}).$$

(5.28)

(Note that this sum includes the 'empty' Feynman diagram with no lines and no vertices, to which we assign the value 1.) What does it mean to divide through by these Feynman diagrams?

For any Feynman diagram, we can break it up into the *connected part* — those vertices in the diagram which are connected to external vertices — and the *disconnected part*, which

are not. For instance in the diagram

$$x \bullet\!\!-\!\!\bullet y \qquad \qquad (5.29)$$

the left-hand part is connected, and the right-hand part is not. Given any connected contribution C to the n-point function, we can form various Feynman diagrams

$$= \sum_{\text{disconnect diagrams}} C \times (\text{disconnected diagrams})$$

$$= C \times (\text{sum of Feynman diagrams with no external points}). \qquad (5.30)$$

The contribution to the n-function factorize, into the product of the connected and the disconnected diagrams:

$$\int \mathcal{D}\phi \ \phi(x_1)...\phi(x_n) \exp\left[i \int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4\right]$$

$$= Z_0 \times \begin{pmatrix}\text{sum of connected diagrams}\\\text{with } n \text{ external points}\end{pmatrix} \times \begin{pmatrix}\text{sum of connected diagrams}\\\text{with no external points}\end{pmatrix}. \qquad (5.31)$$

Inserting this into our formula for the n-point function:

$$\langle\Omega|\mathcal{T}\phi(x_1)...\phi(x_n)|\Omega\rangle = \frac{\int \mathcal{D}\phi \ \phi(x_1)...\phi(x_n) \exp\left[i \int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4\right]}{\int \mathcal{D}\phi \ \exp\left[i \int d^4z \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4\right]}.$$

$$= \frac{Z_0 \times \begin{pmatrix}\text{sum of connected diagrams}\\\text{with } n \text{ external points}\end{pmatrix} \times \begin{pmatrix}\text{sum of connected diagrams}\\\text{with no external points}\end{pmatrix}}{Z_0 \times \begin{pmatrix}\text{sum of connected diagrams}\\\text{with no external points}\end{pmatrix}}$$

and so

$$\langle\Omega|\mathcal{T}\phi(x_1)...\phi(x_n)|\Omega\rangle = \text{sum of connected diagrams with } n \text{ external points.} \qquad (5.32)$$

Odd Correlation Functions

You may have noticed some very interesting about odd correlation functions. Try, for instance, to draw a Feynman diagram that will contribute to the ϕ field *vacuum expectation value* (vev):

$$\langle\Omega|\phi(x)|\Omega\rangle. \qquad (5.33)$$

You should quickly be able to convince yourself that no diagram can exist. From this we conclude that

$$\langle\Omega|\phi(x)|\Omega\rangle = 0. \qquad (5.34)$$

This is an exact result, true to any order of λ. Similarly, we cannot draw any Feynman diagrams containing any odd number of fields, so we conclude

$$\langle\Omega|\phi(x_1)...\phi(x_{2n+1})|\Omega\rangle = 0. \tag{5.35}$$

There is a deep reason that odd correlation function disappear in ϕ^4-theory. Under the transformation

$$Z^{-1}\phi(x)Z = -\phi(x) \tag{5.36}$$

the Lagrangian

$$\mathcal{L} = \int d^4x \, \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4 \tag{5.37}$$

is left unchanged. We therefore say that ϕ^4-theory has a $\mathbb{Z}_2$ symmetry — $\mathbb{Z}_2$ because $Z^2 = (-1)^2 = 1$, so the group structure is the same as modulo 2 arithmetic.

Since the Lagrangian is unchanged by Z, we deduce that $[H, Z] = 0$ for the Hamiltonian H. If our theory has a unique ground state, then as

$$HZ|\Omega\rangle = ZH|\Omega\rangle = E_0 Z|\Omega\rangle, \tag{5.38}$$

we require that

$$Z|\Omega\rangle = e^{-i\alpha}|\Omega\rangle \tag{5.39}$$

for some phase factor α.

Here comes the clincher: as our theory is invariant under Z:

$$\begin{aligned}
\langle\Omega|\phi(x_1)...\phi(x_n)|\Omega\rangle &= \langle\Omega|(ZZ^{-1})\phi(x_1)(ZZ^{-1})\phi(x_2)...\phi(x_n)(ZZ^{-1})|\Omega\rangle \\
&= \langle\Omega|Z(Z^{-1}\phi(x_1)Z)(Z^{-1}\phi(x_2)Z)...(Z^{-1}\phi(x_n)Z)Z^{-1}|\Omega\rangle \\
&= (-1)^n\langle\Omega|\phi(x_1)...\phi(x_n)|\Omega\rangle.
\end{aligned} \tag{5.40}$$

When $n = 2k + 1$ is odd,

$$\langle\Omega|\phi(x_1)...\phi(x_{2k+1})|\Omega\rangle = -\langle\Omega|\phi(x_1)...\phi(x_{2k+1})|\Omega\rangle, \tag{5.41}$$

and thus any odd correlation function must vanish.

5.3   The S-Matrix

We achieved the first goal of this chapter: developing an elegant formalism, Feynman diagrams, for an abstract quantity, the n-point function. In this section, we will try to link these to physical observables like particle scattering and decays.

The classic particle experiment is to have two beams of particles with momentum $\mathbf{k}_1$ and $\mathbf{k}_2$. These collide and briefly interact, and produce a shower of particles with momentum $\mathbf{p}_1$, $\mathbf{p}_2$, ... $\mathbf{p}_n$, which quickly spread apart and are detected by detectors. Writing the initial and final states as

$$|i\rangle = |\mathbf{k}_1, \mathbf{k}_2\rangle, \quad |f\rangle = |\mathbf{p}_1, ..., \mathbf{p}_n\rangle, \tag{5.42}$$

the amplitude for our scattering process is

$$\lim_{T\to\infty} \langle f|e^{-i2TH}|i\rangle \equiv \langle f|S|i\rangle. \tag{5.43}$$

The matrix S is know, rather imaginatively, as the *S-matrix*. In any scattering process, there is a chance that nothing happens. The particles may just miss each other. To extract the interesting part of the S matrix, we write

$$S = 1 + iT. \tag{5.44}$$

Since momentum is always conserved, we can always write

$$\langle \mathbf{p}_1, ..., \mathbf{p}_n | T | \mathbf{k}_1, \mathbf{k}_2 \rangle = (2\pi)^4 \delta \left(k_1 + k_2 - \sum p_i \right) \mathcal{M}(k_1 k_2 \rightarrow p_1...p_n). \tag{5.45}$$

The only non-trivial part of the S-matrix left for us to calculate is the *invariant matrix element* $\mathcal{M}$.

We have achieved a cute division of labour. The matrix element $\mathcal{M}$ tells us how particle interact in our theory — it gives us the "dynamics". But of course, if we want to compare to experiments, we will need to relate $\mathcal{M}$ to cross sections — the "kinematics" of the theory.

5.3.1 From Correlation Functions to the S-Matrix

The general method for turning correlation functions into S-matrix amplitudes is called the *LSZ reduction formula*, first derived by Lehmann, Symanzik, and Zimmerman. The argument is unfortunately rather technical, and we will not present it here.[1] Instead we will simply try to motivate it.

Imagine the case of $2\phi \rightarrow 2\phi$ scattering. We could consider the correlation function:

$$\begin{aligned} G(\mathbf{y}_1, \mathbf{y}_2; \mathbf{x}_1, \mathbf{x}_2) &= \langle \Omega | \phi(\mathbf{y}_1, T)\phi(\mathbf{y}_2, T)\phi(\mathbf{x}_1, -T)\phi(\mathbf{x}_2, -T) | \Omega \rangle \\ &= \langle \Omega | \phi(\mathbf{y}_1)\phi(\mathbf{y}_2)e^{-2iHT}\phi(\mathbf{x}_1)\phi(\mathbf{x}_2) | \Omega \rangle . \end{aligned} \tag{5.46}$$

In free field theory, we interpreted $\phi(\mathbf{x}) | \Omega \rangle$ as the state of a single particle at $\mathbf{x}$. So if we were working in free field theory, we could interpret the above quantity as the S-matrix for a process where two particles start at $\mathbf{x}_1$ and $\mathbf{x}_2$, and after time T, scatter to $\mathbf{y}_1$ and $\mathbf{y}_2$. Unfortunately, we are not working in a free field theory, and so we will have to work a little harder to extract a scattering amplitude.

Just like in quantum mechanics, in a quantum field theory the completeness relation

$$I = \sum_\lambda | \lambda \rangle \langle \lambda | \tag{5.47}$$

holds, where the sum is all states in our theory. In a quantum field theory, the set of all states is extremely massive — it includes one-particle states, two-particle states, three-particles states, and so on — but nevertheless, the completeness relation still holds. Inserting it into the correlation function G, we find that

$$G(\mathbf{y}_1, \mathbf{y}_2; \mathbf{x}_1, \mathbf{x}_2) = \sum_{\lambda, \lambda'} \langle \Omega | \phi(\mathbf{y}_1)\phi(\mathbf{y}_2) | \lambda \rangle \langle \lambda | e^{-2iHT} | \lambda \rangle \langle \lambda | \phi(\mathbf{x}_1)\phi(\mathbf{x}_2) | \Omega \rangle . \tag{5.48}$$

This middle factor, $\langle \lambda | e^{-2iHT} | \lambda' \rangle$, is just the general form of an S-matrix element. In

---

[1] A good discussion can be found in Chapter 7.1 and 7.2 of *Peskin & Schroeder*.

particular, if we are interested in two-body scattering, we could consider the states

$$|\lambda\rangle = |\mathbf{k}_1, \mathbf{k}_2\rangle \text{ and } |\lambda'\rangle = |\mathbf{p}_1, \mathbf{p}_2\rangle. \tag{5.49}$$

These states describe two distantly separated particles with moment $\mathbf{k}_1$ and $\mathbf{k}_2$ scattering to two distantly separated particles with moment $\mathbf{p}_1$ and $\mathbf{p}_2$. We can now rewrite G as

$$
\begin{aligned}
&G(\mathbf{y}_1, \mathbf{y}_2; \mathbf{x}_1, \mathbf{x}_2) \\
&\quad = \langle \Omega | \phi(\mathbf{y}_1)\phi(\mathbf{y}_2) | \mathbf{k}_1, \mathbf{k}_2\rangle \langle \mathbf{k}_1, \mathbf{k}_2 | e^{-2iHT} | \mathbf{p}_1, \mathbf{p}_2\rangle \langle \mathbf{p}_1, \mathbf{p}_2 | \phi(\mathbf{x}_1)\phi(\mathbf{x}_2) | \Omega\rangle + \text{other terms.}
\end{aligned}
\tag{5.50}
$$

The S-matrix element we want is contained in the G; we just need to know how to extract it! This is the job of the LSZ formula.

Stating the LSZ formula is easiest using Feynman diagram. First we use the momentum space rules to write $G(\mathbf{y}_1, \mathbf{y}_2; \mathbf{x}_1, \mathbf{x}_2)$ in Fourier space:

$$G(q_1, q_2; q_1', q_2') = \quad \times \quad + \quad \bowtie \quad + \quad \text{\large O} \quad + \quad \times \quad + \ldots \tag{5.51}$$

In these diagrams, time goes up the page. The two bottom propagators have incoming momentum q_1' and q_2', whereas the top two propagators have outgoing momentum q_1 and q_2.

Our next step is to put the propagators on *mass shell*; that is, to consider only incoming and outgoing momentum where $q^2 = m^2$. Any physical particle must be on mass shell. In our specific example, this means that we set

$$
\begin{aligned}
(q_1')^\mu &= (\omega_{\mathbf{p}_1}, \mathbf{p}_1), & (q_2')^\mu &= (\omega_{\mathbf{p}_2}, \mathbf{p}_2) \\
q_1^\mu &= (\omega_{\mathbf{k}_1}, \mathbf{k}_1), & q_2^\mu &= (\omega_{\mathbf{k}_2}, \mathbf{k}_2).
\end{aligned}
\tag{5.52}
$$

This causes some problems; since each leg comes with a factor of $(q^2 - m^2)^{-1}$. Putting our momentum on mass shell causes everything to blow up![2]

To fix this problem, we have to *amputate* the diagrams. For the first three diagrams in (5.51) the meaning of this is clear: we simply ignore the external propagators. The fourth diagram is a little more tricky:

$$
\tag{5.53}
$$

The problem is that the propagator between the vertices gives a factor of $(q^2 - m^2)^{-1}$. But conservation of momentum requires that $q = k_1$ and hence that $q^2 = m^2$. Our diagram blows up.

To make sense of this, we should not only exclude external propagators, but also

---

[2]This observation is actually at the heart of the LSZ formula. The basic idea is to look for the places where the correlation functions is blows up (has a 'pole'), and the use complex analysis to extract the 'residue', which is in fact the S-matrix element.
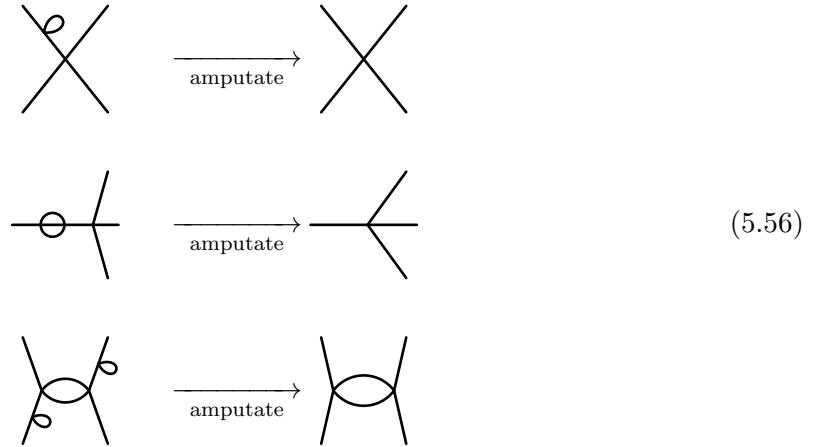
any loops on the external propagators. In the LSZ formalism, the amputated diagrams contribute to the S-matrix

$$\langle \mathbf{k}_1, \mathbf{k}_2 | e^{-2iHT} | \mathbf{p}_1, \mathbf{p}_2 \rangle \tag{5.54}$$

whereas loops connected to a single leg of a Feynman diagram contribute to the terms on either side:

$$\langle \Omega | \phi(\mathbf{y}_1) \phi(\mathbf{y}_2) | \mathbf{k}_1, \mathbf{k}_2 \rangle \ \text{and} \ \langle \mathbf{p}_1, \mathbf{p}_2 | \phi(\mathbf{x}_1) \phi(\mathbf{x}_2) | \Omega \rangle. \tag{5.55}$$

To define amputation, start at each external leg. Find the last point at which we can disconnect the diagram by removing a single propagator, such that this separates the external leg from the rest of the diagram. Cut here. For instance,

$$\tag{5.56}$$

So far, we have only talked about $2\phi \to 2\phi$ scattering. The generalization to more complicated scattering is, however, straightforward. To study $m \to n$ scattering process, we use the $m + n$-correlation function. We then draw the Feynman diagrams which contribute to this function, putting the external legs on-shell, and then amputating. Our general procedure can be summarized as[3]

$$i\mathcal{M}(2\pi)^4 \delta \left(\sum k_i - \sum p_j \right)$$
$$= \quad \begin{array}{l} \text{Sum of connected, amputated diagrams, with incom-} \\ \text{ing momentum } k_i \text{ and outgoing momentum } p_j. \end{array} \tag{5.57}$$

It is usually easiest to evaluate the Feynman diagrams in momentum space. Our final set

---

[3]This formula isn't quite correct; slight modifications arise when we *renormalise* the theory. Renormalisation is unfortunately beyond the scope of this course. The interested reader can find a full explanation of this important topic in Chapters 7 and 10 of *Peskin & Schroeder*.

of Feynman rules for $\lambda \phi^4$ is now:

    1. For each propagator, $\dfrac{i}{p^2 - m^2 + i\varepsilon}$

    2. For each vertex, $= -i\lambda$

    3. For each external leg, $= 1$

    4. Impose conservation of momentum at each vertex.

    5. Integrate over each undetermined momentum: $\displaystyle\int \frac{d^4p}{(2\pi)^4}$

    6. Divide by the symmetry factor.

$$\tag{5.58}$$

5.3.2   Scattering Amplitudes in $\lambda\phi^4$ Theory

Having spent the last few chapters developing formal machinery, we can now calculate something! Let's use our tools to study $2\phi \to 2\phi$ scattering in $\lambda\phi^4$ theory. We will label the incoming momentum k_1^μ and k_2^μ, and the outgoing momentum is p_1^μ and p_2^μ.

Zero Order Contributions

According to our Feynman rules, the lowest order (ie $O(\lambda)$) contributions we can consider are

$$= \delta(p_1 - k_1)\delta(p_2 - k_2) + \delta(p_2 - k_1)\delta(p_1 - k_2). \tag{5.59}$$

The interpretation of these diagrams are simple; they represent two particles that do not interact. As ϕ particles are bosons, we cannot distinguish between the two cases $k_1 \to p_1$ and $k_1 \to p_2$, and so we have to include two different diagrams. Calling these diagrams "scattering processes" is a little rich. These diagrams contribute to the S-matrix, but they do not contribute to the T-matrix — which as we recall is the interesting part of the S-matrix.
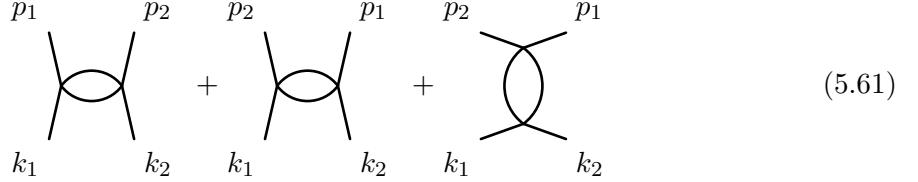
First Order Contributions

Having calculated the rather uninteresting case of no scattering, let's consider the first contribution to the T-matrix. This is the $O(\lambda)$ term

$$= -i\lambda(2\pi)^4\delta(p_1 + p_2 - k_1 - k_2) \tag{5.60}$$

and so $\mathcal{M} = -\lambda$. This shows us that, to lowest order, the probability of particles scattering off each other is completely independent of their momenta, as long as the momentum is conserved. In the next section we shall learn how to convert our matrix element into a cross-section, which we can then measure in the lab.

Higher Order Contributions

Flushed with success, we can move on to the $O(\lambda^2)$ contributions. Following our Feynman rules, there three diagrams to consider:

$$\tag{5.61}$$

The first and second diagrams are related by Bose symmetry. For the moment we will focus on the first diagram; the second and third diagrams are almost identical. Including the internal momenta, our diagram is

$$= -(2\pi)^4 \delta(p_1 + p_2 - k_1 - k_2)\lambda^2 \int \frac{d^4q}{(2\pi)^4} \frac{1}{(q^2 - m^2)(q'^2 - m^2)} \tag{5.62}$$

where conservation of momentum requires that $k_1 + k_2 = q + q'$, and so $q' = q - k_1 - k_2$. We hence find that

$$\mathcal{M} = i\lambda^2 \int \frac{d^4q}{(2\pi)^4} \frac{1}{((q^2 - m^2)\left((q - k_1 - k_2)^2 - m^2\right)} \tag{5.63}$$

The good news is that all we have to do know is evaluate a single integral! The bad news is that it is a really hard integral to evaluate. The even worse news is that the integral doesn't converge.

Why? Well, in the regime where $q^2 >> m^2, (k_1 + k_2)^2$, our integral becomes

$$\mathcal{M} \to i\lambda^2 \int \frac{d^4q}{(2\pi)^4} \frac{1}{(q^2)^2} \propto \int_0^\infty \frac{r^3}{r^4} dr = \infty. \tag{5.64}$$
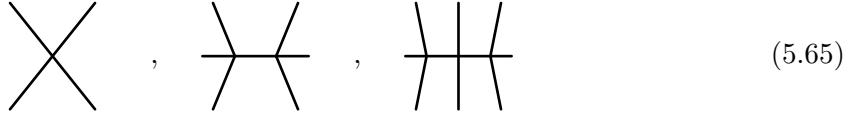
Here we have used the change of variables $r^2 = q^2$ to simplify the integral.

We have a serious problem. When we turn this amplitude into a cross section, we will find that the cross section for the process must also be infinite. But the cross section cannot be infinite, it is a physical quantity we can go out and measure! Something has gone very wrong.

5.3.3   Tree and Loop Diagrams

To discuss the problem further, we should introduce some terminology. A *tree-level diagram* is a diagram which does not contain any loops in it. In $\lambda\phi^4$ theory these include

diagrams such as

$$\text{(5.65)}$$

In any tree-level diagram, we can determine the momentum going through each propagator just from the external momentum. We won't have to evaluate any integrals if we stick to tree-level diagrams.

Loop diagrams, on the other hand, contain closed loops, and around each closed loop we will have an undetermined momentum to integrate over. Most of the time, these integrals diverge. This is not just a problem in our $\lambda\phi^4$ toy model, but in every quantum field theory. Similar problems occur in quantum electrodynamics, the quantum theory describing light. If you naively calculate the mass and charge of an electron, you will find that both are infinite.

In spite of these divergences, it is possible to push ahead and get a finite predictions out of quantum field theory. The process by which this is done is known as *renormalisation*. We will not cover the subject in this course, but you can read about it in a quantum field theory text.[4]

To further illuminate the difference between tree and loop diagrams, restore $\hbar$ in our path-integral:

$$Z = \int \mathcal{D}\phi \ \exp[S[\phi]/\hbar] = \int \mathcal{D}\phi \ \exp\left[\frac{i}{\hbar}\int d^4x \ \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4!}\lambda\right]. \qquad \text{(5.66)}$$

Each propagator now gets an additional factor of $\hbar$, and each vertex a factor of $\hbar^{-1}$. It is not hard to convince yourself that each loop gives an extra power of $\hbar$. For instance

$$\text{(5.67)}$$

If we organise our Feynman diagrams by the number of loops, then this is equivalent to an expansion in Planck's constant, and tree-level diagrams are the leading terms. This is called the *semi-classical expansion*. In some sense, the tree-level diagrams are really the classical result, and the loop diagrams are the quantum correction.

5.4 Scattering and Decay

5.4.1 Scattering

Imagine that we have two beams of particles, of momentum $\mathbf{k}_1$ and $\mathbf{k}_2$. In any real experiment the particles of each beam will be far away, and we can consider just the collisions between one particle from each beam. This is a very practical reason why the most interesting scattering problems are $2 \to N$ — the chance of three particles scattering

---

[4]See for instance *Peskin & Schroeder*, Chapter 6, 7, and 10. If you really want to know how to calculate the integral in (5.63), the answer can be found in Section 10.2.

together is for all intents and purposes zero in any experiment. We define the *cross section* as

$$\sigma_f = \frac{N_f}{T\Phi}.$$ (5.68)

In this formula, N_f is the number of final products f produced, which we divide both the time T the beams were run for and the flux F of the beams. Different products will have different cross-section: in an experiment, we may consider the cross section to produce e^+e^-, or $\mu^+\mu^-$, or $e^+e^-\gamma$ and so on. Many experiments are even more discerning then this, measuring not only the final products, but also their momenta. We could then consider the *differential cross section*

$$d\sigma_f(\mathbf{p}_1, \mathbf{p}_2, ...\mathbf{p}_n) = \frac{N_f(\mathbf{p}_1, \mathbf{p}_2, ...\mathbf{p}_n)}{T\Phi}.$$ (5.69)

Integrating over all allowed momenta will then give us the total cross section.

Now imagine there are only two particles, one in each beam, in our entire volume V over time T. Our expression for the cross section is now

$$d\sigma_f = \frac{dP_f}{T\Phi},$$ (5.70)

where P_f is the probability of a collision. The flux is now easy to calculate, because there are only two particles. If one of the particles was at rest, the flux would just be the speed of the incoming particle divided by the volume

$$\Phi = \frac{|\mathbf{v}|}{V}.$$ (5.71)

More generally, both particles will be moving towards each other, and so

$$\Phi = \frac{|\mathbf{v}_1 - \mathbf{v}_2|}{V}.$$ (5.72)

Next we deal with the probability P_f. We know from quantum mechanics that to calculate probabilities, we take the amplitude and mod-square it. So, generalizing to field theory, we expect that

$$dP_f = \frac{\langle f| S |i\rangle}{\langle f|f\rangle\langle i|i\rangle} d\Pi$$ (5.73)

where $d\Pi$ is the volume element for the final state of momenta we are interested in. Since this must be proportional to the differential momentum of each final state and integrate to one, we find

$$d\Pi = \prod_j \frac{V}{(2\pi)^3} d^3p_j.$$ (5.74)

The reason we have $\langle i|i\rangle$ and $\langle f|f\rangle$ in the denominator is that these are not normalized to 1, but instead are relativistically normalized. For a single particle state $|\mathbf{p}\rangle$, we have

$$\langle \mathbf{p}|\mathbf{p}\rangle = 2\omega_{\mathbf{p}}(2\pi)^3 \delta(\mathbf{p} - \mathbf{p}).$$ (5.75)

We can interpret this delta function as

$$(2\pi)^3 \delta(\mathbf{p} - \mathbf{p}) = \int d^3x \; e^{i(\mathbf{p}-\mathbf{p})x} = V$$ (5.76)

and so, if we work over finite volume,

$$\langle \mathbf{p} | \mathbf{p} \rangle \rightarrow 2\omega_{\mathbf{p}} V. \tag{5.77}$$

For our multi-particle states, we hence have

$$\langle i | i \rangle = 4\omega_{\mathbf{k}_1}\omega_{\mathbf{k}_2} V^2, \quad \langle f | f \rangle = \prod_j 2\omega_{\mathbf{p}_j} V. \tag{5.78}$$

Turning now to the S-matrix element,

$$\langle f | S | i \rangle = \langle f | (1 + iT) | i \rangle = \delta^4 \left(\sum p \right) [1 + (2\pi)^4 i \langle f | \mathcal{M} | i \rangle]. \tag{5.79}$$

Since all we care about is the case where the particles collide

$$\langle f | S | i \rangle^2 \rightarrow \delta \left(\sum p \right)^2 (2\pi)^8 |\mathcal{M}_{if}|^2 = TV\delta \left(\sum p \right) (2\pi)^4 |\mathcal{M}_{if}|^2 \tag{5.80}$$

where we have interpret the extra delta function as a spacetime volume. Combining this together,

$$dP_f = \frac{(2\pi)^4 \delta \left(\sum p \right) TV}{(2E_1 V)(2E_1 V)} |\mathcal{M}_{if}|^2 \prod_{\text{final states } j} \frac{d^3 p_j}{(2\pi)^3 2\omega_{\mathbf{p}_j}} \tag{5.81}$$

The volume factors from the $d\Pi$ cancel the factors from $\langle f | f \rangle$. As we showed in chapter 3, the right-hand factor is relativistically invariant — it is known as the *relativistically invariant n-body phase space*

$$d\Pi_{\text{RI}} = (2\pi)^4 \delta \left(\sum p \right) \prod_{\text{final states } j} \frac{d^3 p_j}{(2\pi)^3 2\omega_{\mathbf{p}_j}}. \tag{5.82}$$

We can now finally calculate the cross section

$$\begin{aligned} d\sigma_f = \frac{dP_f}{T\Phi} &= \frac{V}{T|\mathbf{v}_1 - \mathbf{v}_2|} \frac{TV}{(2E_1 V)(2E_1 V)} |\mathcal{M}_{if}|^2 d\Pi_{\text{RI}} \\ &= \frac{1}{4E_1 E_2 |\mathbf{v}_1 - \mathbf{v}_2|} |\mathcal{M}_{if}|^2 d\Pi_{\text{RI}}. \end{aligned} \tag{5.83}$$

All of the V and T factors have cancelled out, leaving us with a finite expression for the differential cross section. To calculate the total cross section we simply need to integrate both sides.

A Special Case

The simplest type of scattering we could consider is $2 \rightarrow 2$ scattering in the centre of mass (COM) frame. In this special case, it is possible to simplify (5.83) further. Consider the process

$$\mathbf{k}_1 + \mathbf{k}_2 \rightarrow \mathbf{p}_3 + \mathbf{p}_4. \tag{5.84}$$

In the centre of mass frame, $\mathbf{p}_4 = -\mathbf{p}_3$, so we can now integrate $\mathbf{p}_4$ over the delta function in $d\Pi_{\text{RI}}$,

$$\int d\Pi_{\text{RI}} = \int \frac{d^3 p_3}{(2\pi)^3 4E_3 E_4} (2\pi)\delta(E_{\text{cm}} - E_3 - E_4). \tag{5.85}$$

Here E_{cm} is the centre of mass energy, and

$$E_3 = \sqrt{m_1^2 + \mathbf{p}_3{}^2}, \quad E_4 = \sqrt{m_2^2 + \mathbf{p}_4{}^2} = \sqrt{m_2^2 + \mathbf{p}_3{}^2}. \tag{5.86}$$

Integrating over the magnitude and direction of $\mathbf{p}_3$ separately,

$$\begin{aligned}
\int d\Pi_{\text{RI}} &= \int \frac{p_3^2 dp_3 d\Omega}{16\pi^2 E_3 E_4}(2\pi)\delta(E_{\text{cm}} - E_3 - E_4) \\
&= \int d\Omega \frac{p_3^2}{16\pi^2 E_1 E_2}\left(\frac{p_3}{E_1} + \frac{p_3}{E_2}\right) \\
&= \int d\Omega \frac{1}{16\pi^2}\frac{|\mathbf{p}_3|}{E_{\text{cm}}}.
\end{aligned} \tag{5.87}$$

Here the final $d\Omega$ integral integrates over the 2-sphere; that is, the different possible orientations of $\mathbf{p}_1/|\mathbf{p}_1|$.

Applying this formula to (5.83), we have found

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{CM}} = \frac{1}{64\pi^2 E_1 E_2 |\mathbf{v}_1 - \mathbf{v}_2|}\frac{|\mathbf{p}_3|}{E_{\text{cm}}}|\mathcal{M}_{if}|^2. \tag{5.88}$$

If all four particles have equal mass, then this further reduces to

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{CM}} = \frac{|\mathcal{M}_{if}|^2}{64\pi^2 E_{\text{cm}}^2}. \tag{5.89}$$

We can apply the above formula to the $2\phi \to 2\phi$ scattering considered in the previous section. Recall that for this process,

$$\mathcal{M} = -\lambda, \tag{5.90}$$

at leading order, and so

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{CM}} = \frac{\lambda^2}{64\pi^2 E_{\text{cm}}^2}. \tag{5.91}$$

The total cross section can then be calculated by integrating over $d\Omega$, and then dividing by 2 since we have two identical particles in the final state. Our final answer is then

$$\sigma_{total} = \frac{\lambda^2}{32\pi E_{\text{cm}}^2}. \tag{5.92}$$

5.4.2   Decay

Consider an unstable particle with momentum $\mathbf{k}_1$, decaying into some final products f.

$$\mathbf{k}_1 \to \mathbf{p}_1 + \mathbf{p}_2 + \dots \tag{5.93}$$

The decay rate is simply the probability that this decay occurs in an arbitrarily short time period

$$\Gamma_f = \lim_{t \to 0}\frac{P_f(t)}{t}. \tag{5.94}$$

We can also consider the differential decay rate $d\Gamma_f(\mathbf{p}_1, bop_2, \dots)$, which depends not only on the final products but also their momenta.

We can calculate the decay rate in a manner completely analogous to the cross section

$$d\Gamma_f = \lim_{t \to 0} \frac{dP_f}{t}.$$ (5.95)

The only difference is that we no longer have a flux factor to worry about. It is not hard to guess that

$$d\Gamma_f = \frac{1}{2E_1}|\mathcal{M}_{if}|^2 d\Pi_{\mathrm{RI}},$$ (5.96)

which can be confirmed by a calculation completely analogous to the case of scattering. The factor of $(2E_1)^{-1}$ gives us the usual time dilation factor. Normally we work in the rest frame of the particle, and so

$$d\Gamma_f = \frac{1}{2m_1}|\mathcal{M}_{if}|^2 d\Pi_{\mathrm{RI}}.$$ (5.97)

You may be concerned that this formula is nonsense. After all, we calculate $\mathcal{M}$ by assuming there is a particle at $-\infty$. But clearly an unstable particle cannot be sent into the distant past — it would decay! Although we cannot justify it here, it does turn out that (5.97) is the correct formula for decay. In a more sophisticated treatment we can derive the formula as a consequence of the *optical theorem*.

5.5   Cubic Vertices

So far we have only discussed the Feynman rules for $\lambda\phi^4$. But the tools we have developed can be generalized to any QFT. This is what makes them so powerful. Any time you wish to study an interacting theory, usually your first steps will be to find the Feynman rules for the theory.

Consider the Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 + \frac{1}{2}(\partial\Phi)^2 - \frac{1}{2}M^2\Phi^2 - \frac{1}{2}g\phi^2\Phi$$ (5.98)

with $g \ll M, m$. We now have two fields, ϕ and Φ, and therefore two particles, which we shall call 'mesons'.[5] This theory has a slight defect — the potential has a stable local minimum for $\phi = \Phi = 0$, but for sufficiently large Φ the potential become unbound from below. We should be careful not to push the theory too far.

To derive the Feynman rules for our theory, we should consider

$$Z = \int \mathcal{D}\phi \, \mathcal{D}\Phi \, \exp\left[i \int d^4x \, \mathcal{L}_{\mathrm{free}}[\phi, \Phi] - \frac{1}{2}g\phi^2\Phi\right]$$ (5.99)

where

$$\mathcal{L}_{\mathrm{free}}[\phi, \Phi] = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 + \frac{1}{2}(\partial\Phi)^2.$$ (5.100)

Now we just need to repeat the analysis of the last 16 pages for this new case.

Just kidding! Like most quantum field theorists, we won't properly re-derive the Feynman rules for each new theory we come across. Most of the time, it is actually pretty easy to guess the rules straight from the Lagrangian. There are a few traps where we need

---

[5]Strictly speaking, a *meson* is a quark-antiquark bound state, although it is sometimes used (especially in older literature) to describe any boson.

to be cautious[6], and in all these cases we can rerun our previous analysis to make sure we get everything correct. This is not one of those occasions.

Since we are primarily interested in scattering and decay processes, we will work in the momentum basis. Start with the propagators. When $g = 0$, our theory simply describes two non-interacting scalar theories. It is not hard to guess that the ϕ propagator is

$$\overline{}_{p} = \frac{1}{p^2 - m^2},\tag{5.101}$$

and the Φ propagator is
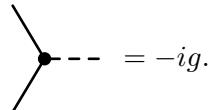
$$\text{-----}_{p} = \frac{1}{p^2 - M^2}.\tag{5.102}$$

We will used dashed lines for the Φ propagator so that we can distinguish between the two particles.

Next we need to consider the interaction term. To determine what this should look like diagrammatically, let's expand

$$\begin{aligned}
Z &= \int \mathcal{D}\phi \, \mathcal{D}\Phi \, \exp\left[i \int d^4x \, \mathcal{L}_{\text{free}}[\phi, \Phi] - \frac{1}{2}g\phi^2\Phi\right] \\
&= \int \mathcal{D}\phi \, \mathcal{D}\Phi \, \exp\left[i \int d^4y \, -\frac{1}{2}g\phi^2\Phi\right] \exp\left[i \int d^4x \, \mathcal{L}_{\text{free}}[\phi, \Phi]\right] \\
&= \int \mathcal{D}\phi \, \mathcal{D}\Phi \, \left(1 - \frac{ig}{2}\left[\int d^4y \, \phi^2\Phi\right] - \frac{g^2}{8}\left[\int d^4y \, \phi^2\Phi\right]^2 + ...\right) \exp\left[i \int d^4x \, \mathcal{L}_{\text{free}}[\phi, \Phi]\right].
\end{aligned}\tag{5.103}$$

Each term in this exponential can be written schematically as $(\phi^2\Phi)^n$. Compare this to $(\phi^4)^n$ expansion we obtained for $\lambda\phi^4$ theory. We now introduce the vertex:

$$= -ig.\tag{5.104}$$
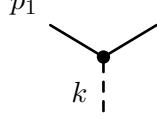
You can think of this as an ϕ-meson travelling along, emitting a Φ-meson, if we take time running up the page. Alternatively, we could think of this as the collision of two ϕ-mesons to produce an Φ-meson, with time running horizontally. If we reverse time we see an Φ meson decay into two ϕ mesons. Our vertex represents all of these processes.

The rest of our Feynman rules are the same as for the $\lambda\phi^4$ theory. We have to conserve momentum at each vertex, and integrate over all the undetermined momenta in loop diagrams. There is also the annoying issue of symmetry factors, but fortunately we rarely have to worry about these in a computation.

---

[6]Derivative couplings, symmetry factors, massless particles, fermionic minus signs, Yang-Mills 'ghosts', to name a few. You will encounter all of these as you continue to study QFT.

5.5.1   Meson Decay

Let's start by examining the simplest possible process, Φ-meson decay. As we noted above, this is described at leading order by the Feynman diagram

$$
= -ig(2\pi)^4 \delta(p_1 + p_2 - k_1) \tag{5.105}
$$

where an initial meson with momentum k decays into two ϕ-mesons, with momentum p_1 and p_2 respectively.

Before calculating further, we should consider the kinematics of this process. In the rest frame of the Φ, $k = (M, \mathbf{0})$, and so conservation of momentum dictates that $\mathbf{p}_1 = -\mathbf{p}_2$. This means that

$$
E_1 = \sqrt{m^2 + \mathbf{p}_1^2} = \sqrt{m^2 + \mathbf{p}_2^2} = E_2, \tag{5.106}
$$

so the two decay products have equal energy.

Furthermore, conservation of energy implies that

$$
E_1 + E_2 = k^0 = M. \tag{5.107}
$$

This means that

$$
E_1 = \sqrt{m^2 + \mathbf{p}_1^2} = \frac{M}{2} \tag{5.108}
$$

Since $E_1 \geq m$, this is only possible if $M > 2m$. Otherwise, the decay process cannot occur, and Φ is stable.

With this in mind, we can now calculate the decay rate of the Φ-meson:

$$
d\Gamma = \frac{1}{2M} |\mathcal{M}|^2 (2\pi)^4 \delta(p_1 + p_2 - k) \frac{d^3\mathbf{p}_1}{(2\pi)^3 2E_1} \frac{d^3\mathbf{p}_2}{(2\pi)^3 2E_2}. \tag{5.109}
$$
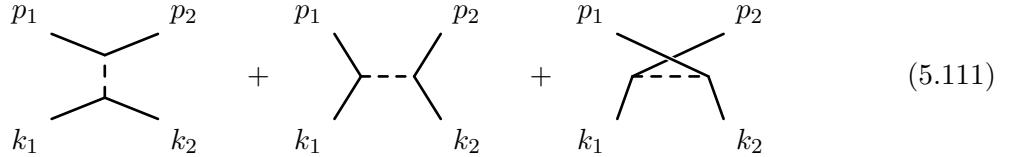
Calculating in the rest frame of the Φ meson, $E_1 = E_2 = \frac{1}{2}M$, and so

$$
\begin{aligned}
\Gamma &= \int d\Gamma \\
&= \frac{g^2}{2M^3(2\pi)^2} \int d\mathbf{p}_1 \, d\mathbf{p}_2 \, \delta\left(M - \sqrt{m^2 + \mathbf{p}_1^2} - \sqrt{m^2 + \mathbf{p}_2^2} \right) \delta(\mathbf{p}_1 - \mathbf{p}_2) \\
&= \frac{g^2}{2M^3(2\pi)^2} \int d\mathbf{p}_1 \, \delta\left(M - 2\sqrt{m^2 + \mathbf{p}_1^2} \right) \\
&= \frac{g^2}{2M^3(2\pi)^2} \int dr \, 4\pi r^2 \delta\left(M - 2\sqrt{m^2 + r^2} \right) \\
&= \frac{g^2}{16\pi M^2} \sqrt{M^2 - 4m^2}.
\end{aligned} \tag{5.110}
$$

In the second last step we used the spherical symmetry of the integral to make the substitution $r = |\mathbf{p}_1|$. So we have the decay rate of the Φ-meson at leading order in g. Higher order contributions include both loop-corrections to the process $\Phi \to 2\phi$, and also more complicated decays such as $\Phi \to 4\phi$.

5.5.2   Meson-Meson Scattering

Another process we could consider the $2\phi \to 2\phi$ scattering. We will label the incoming momentum k_1 and k_2, and the outgoing momentum p_1 and p_2. There are three diagrams which will contribute to this process

$$\tag{5.111}$$

As for our decay, it is fruitful to consider the kinematics of our process. When considering $2\phi \to 2\phi$ scattering, it is often useful to define three quantities, know as *Mandelstam variables*:

$$
\begin{aligned}
s &= (k_1 + k_2)^2 &&= (p_1 + p_2)^2 \\
t &= (p_1 - k_1)^2 &&= (p_2 - k_2)^2 \\
u &= (p_1 - k_2)^2 &&= (p_2 - k_1)^2
\end{aligned}
\tag{5.112}
$$

These three parameters give a Lorentz invariant description of the scattering process. The definition of t and u are interchangeable upon swapping p_1 and p_2; by convention we take t to be the squared difference of the momentum of the two most similar particle.

We can now evaluate the three diagrams as:

$$= \frac{-g^2}{(p_1 + p_2)^2 - M^2} \quad = \frac{-g^2}{s^2 - M^2}$$

$$= \frac{-g^2}{(p_1 - k_1)^2 - M^2} \quad = \frac{-g^2}{t^2 - M^2} \tag{5.113}$$

$$= \frac{-g^2}{(p_2 - k_1)^2 - M^2} \quad = \frac{-g^2}{u^2 - M^2}.$$

Here we have stripped the global factor of $i(2\pi)^4 \delta(k_1 + k_2 - p_1 - p_2)$. We shall do this systematically from now on, equating a diagram with its associated matrix element $\mathcal{M}$. We often say that the first process is in the s-channel, the second in the t-channel and the third in the u-channel. Each leads to a characteristic momentum dependence. Notice that whenever s, t or u are equal to M^2, the scattering amplitude blows up. This is how we discover new particles in a particle accelerator — we look for sharp peaks in scattering cross-sections. For this reason, unstable particles are also know as *resonances*, since their primary experimental signature is the observation of a sharp peak in scattering data.

In *Figure 5.1*, we can see an example of a resonance. When an electron and positron collide, they can produce a Z-boson. But the Z-boson is unstable, and decays back to other particles. These include *hadrons* — bound states of quarks such as the proton, neutron and pion. By plotting the cross section for $e^+ e^- \to$ hadrons, we can see a peak when the centre of mass energy is equal to the mass of the Z-boson.
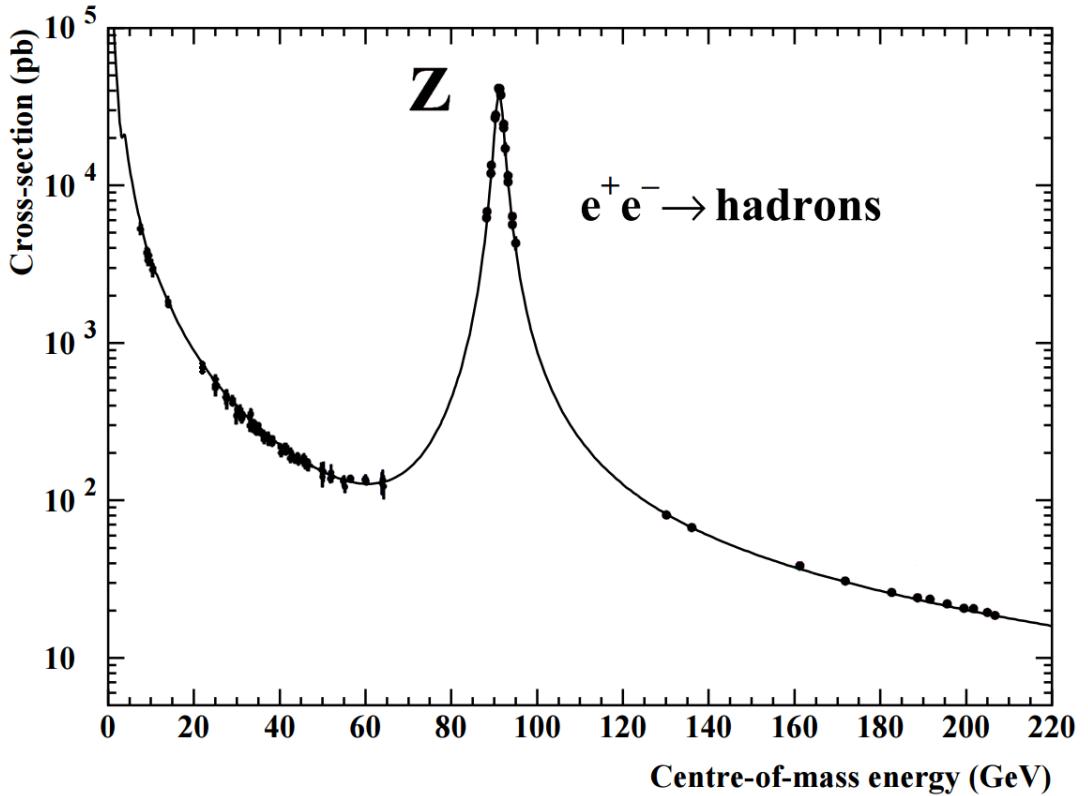
Figure 5.1: Cross section for an electron-positron collision to produces hadrons. The black dots are experimental results, and the solid line is the Standard Model prediction. The peak is caused by the existence of the Z-boson, which has a mass of 91 GeV. The image has been taken from *arXiv:hep-ex/0509008v3*.

More on Mandelstam Variables

To understand better then kinematics of 2-body scattering, we can evaluate s, u and t in the centre of mass frame. We can always choose a reference frame so that initially, the incoming particles are both travelling in the x direction. Furthermore, we can rotate so that the outgoing particles are travelling in the xy-plane. Assuming all our particles have mass m, we find

$$k_1 = (E, k, 0, 0), \qquad\qquad k_2 = (E, -k, 0, 0)$$
$$p_1 = (E, k\cos\theta, k\sin\theta, 0), \qquad p_2 = (E, -k\cos\theta, -k\sin\theta, 0), \tag{5.114}$$

where $E = \sqrt{k^2 + m^2}$. So we see that our collision can be described by two variables, k and θ. In terms of these variables, our Mandelstam variables are

$$s = (k_1 + k_2)^2 = (2E)^2 = E_{\text{CM}}^2$$
$$t = (k_1 - p_1)^2 = -k^2(1 - \cos\theta)^2 - k^2\sin^2(\theta) = -2k^2(1 - \cos\theta) \tag{5.115}$$
$$u = (k_1 - p_2)^2 = -k^2(1 + \cos\theta)^2 - k^2\sin^2(\theta) = -2k^2(1 + \cos\theta).$$

When $\theta = 0$, we find $t = 0$, whereas when $\theta = \pi$, then $u = 0$. Summing the Mandelstam variables, we discover

$$s + t + u = 4m^2. \tag{5.116}$$

This formula generalizes to

$$s + t + u = \sum_i m_i^2 \tag{5.117}$$
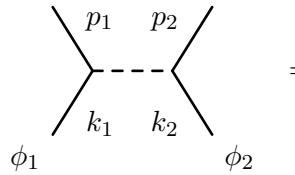
for the case where the particles have different masses.

The Yukawa Potential

Now let's generalize to the case where we have two species of ϕ-mesons, ϕ_1 and ϕ_2, both of mass m. Couple these to the Φ-meson via the interaction

$$\mathcal{L}_{\text{int}} = -\frac{g}{2} \left(\phi_1^2 + \phi_2^2 \right) \Phi. \tag{5.118}$$

We can now consider the scattering process $\phi_1 \phi_2 \rightarrow \phi_1 \phi_2$. Only one diagram contributes to this process, the t-channel process

$$= \frac{-g^2}{(k_1 - p_1)^2 - M^2} = \frac{-g^2}{t^2 - M^2}. \tag{5.119}$$

We will work in the centre of mass frame. . In this reference frame, $k \equiv k_1 = -k_2$, and $p \equiv p_1 = -p_2$. All particles have equal energy, and so $(k - p)^2 = -(\mathbf{k} - \mathbf{p})^2$.

We will now compare this to non-relativistic scattering. According to the Born approximation, the amplitude for non-relativistic scattering from a potential $U(\mathbf{x})$ is given by

$$\mathcal{M}_{\text{NR}} = \langle \mathbf{p} | \left(-\int d^3 x \, U(\mathbf{x}) \right) | \mathbf{k} \rangle = -\int d^3 x \, U(\mathbf{x}) e^{-i(\mathbf{p} - \mathbf{k}) \cdot \mathbf{x}}. \tag{5.120}$$

Equating this to our relativistic amplitude, we can calculate the scattering potential $U(\mathbf{x})$ in the non-relativistic limit:

$$\mathcal{M} = \frac{-g^2}{t^2 - M^2} = \frac{g^2}{(\mathbf{p} - \mathbf{k})^2 + M^2} = \int d^3 x \, U(\mathbf{x}) e^{-i(\mathbf{p} - \mathbf{k}) \cdot \mathbf{x}}. \tag{5.121}$$

The potential can now be calculated using an inverse-Fourier transform with respect to $\mathbf{q} = \mathbf{k} - \mathbf{p}$:

$$U(\mathbf{x}) = -g^2 \int \frac{d^3 p}{(2\pi)^3} \frac{e^{i\mathbf{q} \cdot \mathbf{x}}}{\mathbf{q}^2 + M^2}. \tag{5.122}$$

We already evaluated this integral in Section 4.2.5, where we found

$$U(\mathbf{x}) = -\frac{g^2}{4\pi r} e^{-Mr}. \tag{5.123}$$

In the non-relativistic limit, we recover an attractive Yukawa potential, in agreement with the results of 4.2.5. Scalar fields really do generate attractive forces between particles.

Fermions

So far we have only discussed scalar theories. In particle physics however, scalar particles are quite rare. There is only one scalar particle in the Standard Model: the Higgs boson. Even the existence of the Higgs boson is very troubling, leading to one of the most troubling problems in modern physics — the hierarchy problem.[1]

Instead, most particles are either fermions with spin-1/2 (for instance, the electron, the neutrino, and the proton), or gauge bosons with spin-1 (such as the photon, the Z^0 and the gluon). In this chapter we shall study fermions, and in the next chapter we shall study the simplest gauge boson, the photon.

6.1 Lorentz Invariance

In order to understand the origin of spin, we need to understand how Lorentz transformations act on fields. Recall that a Lorentz transformation is a linear map

$$x^\mu \to \Lambda^\mu_\nu x^\nu \tag{6.1}$$

under which the interval

$$x^\mu x_\mu = g_{\mu\nu} x^\mu x^\nu \tag{6.2}$$

is preserved. This condition implies that

$$g_{\mu\nu} \Lambda^\mu_\rho \Lambda^\nu_\sigma = g_{\rho\sigma} \tag{6.3}$$

for any Lorentz transformation Λ. Given any two Lorentz transformation Λ_1 and Λ_2, their product $\Lambda_1 \Lambda_2$ will also be a Lorentz transformation. This means that the set of Lorentz transformations form a *group*.

How does scalar field $\phi(x)$ change under a Lorentz transformation? We require that if $x \to x'$, then

$$\phi'(x') = \phi(x) \tag{6.4}$$

so that both observers agree on the value of ϕ at x. For this condition to hold we need

$$\phi'(x) = \phi(\Lambda^{-1} x). \tag{6.5}$$

The Lagrangian density itself should be a scalar under a Lorentz transformation. If this is the case, then the equations of motion will also be Lorentz invariant. Since the integration measure is always Lorentz invariant, this means that the path-integral will also

---

[1]To properly understand the hierarchy problem requires an understanding of renormalisation. Essential the question comes down to, why is the Higgs boson so light compared to the Planck scale?

be Lorentz invariant. Since any physical quantity can be calculated from the path-integral, this is enough to prove that the theory is relativistically invariant.

With this in mind, we can now check that our scalar field theories are consistent with special relativity. Under a Lorentz transformation

$$
\begin{aligned}
\mathcal{L}(x) &= \frac{1}{2}(\partial\phi(x))^2 - V(\phi(x)) \\
&\to \frac{1}{2}g^{\mu\nu}\partial_\mu\phi'(x)\partial_\nu\phi'(x) - V(\phi'(x)) \\
&= \frac{1}{2}g^{\mu\nu}\partial_\mu\phi(\Lambda^{-1}x)\partial_\nu\phi(\Lambda^{-1}x) - V(\phi(\Lambda^{-1}x)) \\
&= \frac{1}{2}g^{\mu\nu}[(\Lambda^{-1})^\rho_\mu\partial_\rho\phi(\Lambda^{-1}x)][(\Lambda^{-1})^\sigma_\nu\partial_\sigma\phi(\Lambda^{-1}x)] - V(\phi(\Lambda^{-1}x)) \\
&= \frac{1}{2}g^{\rho\sigma}\partial_\rho\phi(\Lambda^{-1}x)\partial_\sigma\phi(\Lambda^{-1}x) - V(\phi(\Lambda^{-1}x)) = \mathcal{L}(\Lambda^{-1}x),
\end{aligned}
\tag{6.6}
$$

where we have used the Lorentz invariance of the metric:

$$
g^{\rho\sigma} = g^{\mu\nu}(\Lambda^{-1})^\rho_\mu(\Lambda^{-1})^\sigma_\nu.
\tag{6.7}
$$

Since the Lagrangian transforms as a Lorentz scalar, we conclude that the theory is indeed Lorentz invariant.

6.1.1   Vectors and Tensors

The transformation law for ϕ is the simplest option, and the only one with a single component. But there are many options if we have a multicomponent field.

Vector fields, such as the electromagnetic potential $A^\mu(x)$, transform in a more interesting way under a Lorentz transformation:

$$
A^\mu(x) \to \Lambda^\mu_\nu A^\nu(\Lambda^{-1}x).
\tag{6.8}
$$

Not only do we transform spacetime, but we also need to transform the field A^μ itself. Building Lorentz invariant Lagrangians from these fields is a little more interesting then for scalars. Terms like A^0 or A^1 are not Lorentz invariant, but the terms such as

$$
(A)^2 = A^\mu A_\mu \quad \text{and} \quad \partial^\mu A_\mu
\tag{6.9}
$$

are all good. Vector fields are not just limited to vector theories; if ϕ is a scalar, then the derivative

$$
\partial_\mu\phi(x)
\tag{6.10}
$$

will transform as a Lorentz vector.

Using vectors, we can produce more tensors. For instance, a 2-rank tensor such as electromagnetic field strength $F^{\mu\nu}$ will transform as:

$$
F^{\mu\nu}(x) \to \Lambda^\mu_\rho\Lambda^\nu_\sigma F^{\rho\sigma}(\Lambda^{-1}x).
\tag{6.11}
$$

6.1.2   Group Representations

It turns out the vectors and tensors do not exhaust the possible ways a field can transform. In general we can imagine a field Φ_a as a vector with n-components. Under a Lorentz

boost, the field will transform as

$$\Phi_a(x) \to M_a^b(\Lambda)\Phi_b(\Lambda^{-1}x) \tag{6.12}$$

where $M(\Lambda)$ is some matrix. What are the allowed matrices $M(\Lambda)$?

The basic restriction on M is that if we apply two Lorentz transformations Λ_1 and Λ_2, then this should be the same as the Lorentz transformation $\Lambda_2\Lambda_1$

$$M(\Lambda_2)M(\Lambda_1)\Phi(\Lambda_1^{-1}\Lambda_2^{-1}x) = M(\Lambda_2\Lambda_1)\Phi\left((\Lambda_2\Lambda_1)^{-1}x\right). \tag{6.13}$$

This leads us to the requirement

$$M(\Lambda_2\Lambda_1) = M(\Lambda_2)M(\Lambda_1). \tag{6.14}$$

The mathematical jargon for this restriction is that M must form an n-dimensional *representation* of the Lorentz group. We just need to find the allowed representations.

Rather than confront the problem head-on, we will instead begin by considering a simpler example: rotations in three-dimensions. This group is known as $SO(3)$. As you learnt in quantum mechanics, to describe any rotation we simply need to know the angular momentum operators J^i, $n \times n$ matrices which satisfy the commutation relations

$$[J^i, J^j] = i\varepsilon^{ijk}J^k. \tag{6.15}$$

Any general rotation then has the form

$$R(\theta_i) = \exp(-iJ^i\theta_i). \tag{6.16}$$

The operators J^i can be thought of as infinitesimal rotations, since

$$R(\theta_i) = \exp(-iJ^i\theta_i) = 1 - iJ^i\theta_i + ... \tag{6.17}$$

For every n there is an n-dimensional representation of $SO(3)$, representing particles of spin $(n-1)/2$. For instance, when $n=1$ we have the representation

$$J^1 = J^2 = J^3 = 0. \tag{6.18}$$

This representation is called the *trivial representation*, because it isn't very interesting. We use it to describe spin-0 particles.

Much more interesting is the spin-1/2 representation,

$$J^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad J^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad J^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1. \end{pmatrix} \tag{6.19}$$

These matrices are the three Pauli matrices, σ^1, σ^2 and σ^3 respectively.

As a final example, recall the spin-1 representation:

$$iJ^1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad iJ^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad iJ^3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{6.20}$$

This is also called the vector representation, since it described how vectors transform when

rotated.

6.1.3 Lorentz Representations

Now let's consider how the 4-vector transform under boosts. If we boost space-time in the x^1 direction, then

$$
\begin{aligned}
x^0 &\to x^0 \cosh(\xi) + x^1 \sinh(\xi) = x^0 + \xi x^1 + O(\xi^2) \\
x^0 &\to x^0 \sinh(\xi) + x^1 \cosh(\xi) = x^1 + \xi x^0 + O(\xi^2).
\end{aligned}
\tag{6.21}
$$

We can hence read off the generator of boosts in the x^1 direction

$$
iK_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
\tag{6.22}
$$

Similar calculations reveal that

$$
iK_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad
iK_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.
\tag{6.23}
$$

Combined with the generators of rotation

$$
iJ_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad
iJ_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad
iJ_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad
\tag{6.24}
$$

we have a representation of the Lorentz group.

Now we compute the commutators

$$
\begin{aligned}
[J_i, J_j] &= i\varepsilon_{ijk} J_k \\
[J_i, K_j] &= i\varepsilon_{ijk} K_k \\
[K_i, K_j] &= -i\varepsilon_{ijk} J_k.
\end{aligned}
\tag{6.25}
$$

The first of these commutators tells us that the J_i's generate momentum, and the second relation tells us that boosts transform as a vector. The meaning of the third commutator is a little more obscure, but nevertheless is important in the physics of Thomas procession.

What does it mean to find a representation of the Lorentz group? All we need to do is find $n \times n$ matrices J^i and K^i so that the above commutator relations hold.

6.1.4 The Spinor

To find another representation of the Lorentz group, assume for the moment that we have $n \times n$ matrices γ^μ satisfying

$$
\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu} \times I,
\tag{6.26}
$$

where I is the identity matrix. This relation is called the *Clifford algebra*, and is very important not just in 4-dimensional Minkowski space, but in both Euclidean and Minkowski space of arbitrary dimension. We could then consider the matrices

$$J^k = \frac{i}{2}\varepsilon^{ijk}[\gamma^i, \gamma^j], \quad K^k = \frac{i}{2}[\gamma^0, \gamma^k]. \tag{6.27}$$

Using (6.26) many times, we can easily show that these matrices satisfy the commutator relations (6.25)!

All we need to do is find matrices satisfying (6.26). It turns out that we need the gamma matrices to be four dimensional. One such set of matrices is

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \tag{6.28}$$

where σ^i is the ith Pauli matrix. Many other possibilities exists: if you choose some matrix V, then if γ^μ satisfies the Clifford algebra so does $V^{-1}\gamma^\mu V$. It turns out however that all possible Clifford algebras[2] are four dimensional, and they are all equivalent to the above matrices modulo V. Our choice (6.28) is known as the *Weyl* or *chiral representation*.

What are the Lorentz transformations in this representation? Using (6.27), we can compute

$$J^i = \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}, \quad K^i = -i\begin{pmatrix} \sigma^i & 0 \\ 0 & -\sigma^i \end{pmatrix}. \tag{6.29}$$

A four-component field ψ which transforms under this representation is called a *Dirac spinor*. In general, the components in the spinor will be complex numbers. Notice that the rotation operators merely consists of two spin-1/2 rotation operators stacked.

6.2 The Dirac Equation

Now that we have defined the spinor ψ, what equations of motion can it follow that are Lorentz invariant? Naively, we could try using the Klein-Gordon equation

$$(\partial^2 + m^2)\psi = 0. \tag{6.30}$$

It turns out however that this equation doesn't work.[3] Instead we could try to use the gamma matrices γ^μ. It is not hard to show that under Lorentz transformations,

$$\Lambda_{\frac{1}{2}}^{-1}\gamma^\mu\Lambda_{\frac{1}{2}} = \Lambda_\nu^\mu\gamma^\nu. \tag{6.31}$$

What this equation means is that, if we rotate by θ_i and boost by β_j, then transforming the gamma matrices using the spinor transformation

$$\Lambda_{\frac{1}{2}} = \exp\left(-iJ^i\theta_i + iK^j\beta_j\right) \tag{6.32}$$

---

[2]Technically, all *irreducible* representations are four dimensional — you could always stack representations to get an representation that has dimension 8, 12, 16 or so on. Look up irreducibility in a representation theory text!

[3]The problem essential boils down to the fact that the equation is independently invariant under both Lorentz transformations, and also under transformation of just the spinor without changing spaectime.

gives the same result as transforming them as a vector. This allows us to take the Lorentz index μ on the gammas seriously. So we could try the *Dirac equation*

$$(i\gamma^\mu \partial_\mu - m)\psi = 0, \tag{6.33}$$

contracting the gamma index with the derivative index. To check that this equation is indeed Lorentz invariant, let us compute

$$
\begin{aligned}
(i\gamma^\mu \partial_\mu - m)\psi(x) &\to (i\gamma^\mu \left(\Lambda^\nu_\mu \partial_\nu\right) - m)\Lambda_{\frac{1}{2}}\psi(\Lambda^{-1}x) \\
&= \left[i\left(\gamma^\mu \Lambda_{\frac{1}{2}}\right)\left(\Lambda^\nu_\mu \partial_\nu\right) - \Lambda_{\frac{1}{2}}m\right]\psi(\Lambda^{-1}x) \\
&= \left[i\left(\Lambda_{\frac{1}{2}}\Lambda_{\frac{1}{2}}^{-1}\gamma^\mu \Lambda_{\frac{1}{2}}\right)\left(\Lambda^\nu_\mu \partial_\nu\right) - \Lambda_{\frac{1}{2}}m\right]\psi(\Lambda^{-1}x) \\
&= \Lambda_{\frac{1}{2}}\left[i\left(\Lambda^\mu_\rho \gamma^\rho\right)\left(\Lambda^\nu_\mu \partial_\nu\right) - m\right]\psi(\Lambda^{-1}x) \\
&= \Lambda_{\frac{1}{2}}(i\gamma^\mu \partial_\mu - m)\psi(\Lambda^{-1}x).
\end{aligned}
\tag{6.34}
$$

The gamma matrices mix the components of the Dirac spinor. Nevertheless, each component of the Dirac spinor also evolves according to the Klein-Gordon equation! To see this, compute

$$0 = (-i\gamma^\mu \partial_\mu - m)(i\gamma^\nu \partial_\nu - m)\psi = \left(\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu + m^2\right)\psi. \tag{6.35}$$

But since $\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu = \gamma^\mu \gamma^\nu \partial_\nu \partial_\mu = \gamma^\nu \gamma^\mu \partial_\mu \partial_\nu$, we can write

$$\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu = \frac{1}{2}\left(\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu + \gamma^\nu \gamma^\mu \partial_\mu \partial_\nu\right) = \frac{1}{2}\{\gamma^\mu, \gamma^\nu\}\partial_\mu \partial_\nu = g^{\mu\nu}\partial_\mu \partial_\nu = \partial^2, \tag{6.36}$$

and so we conclude that

$$(\partial^2 + m^2)\psi = 0. \tag{6.37}$$

It is very common to consider combinations such as $\gamma^\mu \partial_\mu$ and $\gamma^\mu p_\mu$. To simplify these expressions, Feynman invented the slash notation:

$$\slashed{\partial} \equiv \gamma^\mu \partial_\mu, \quad \slashed{p} \equiv \gamma^\mu p_\mu.$$

We will make liberal use of this notation, particularly in the next chapter. The Dirac equation can now be written as

$$(i\slashed{\partial} - m)\psi = 0. \tag{6.38}$$

6.2.1 The Dirac Lagrangian

What Lagrangian gives rise to the Dirac equation? To write down a Lagrangian, we need to know how to combine two Dirac spinors to get a scalar. The obvious guess,

$$\psi^\dagger \psi \tag{6.39}$$

doesn't work. Under a Lorentz transformation,

$$\psi^\dagger \psi \to \psi^\dagger \Lambda_{\frac{1}{2}}^\dagger \Lambda_{\frac{1}{2}}\psi, \tag{6.40}$$

and if our transformations were unitary $\Lambda_{\frac{1}{2}}^\dagger = \Lambda_{\frac{1}{2}}^{-1}$, everything would be fine. The problem is that the boost generators K^i are not hermitian — using (6.28), we in fact discover that they are antihermitian — and so under a boost

$$\psi^\dagger \psi \to \psi^\dagger \left(e^{-i\beta_i K^i} \right)^\dagger e^{-i\beta_i K^i} \psi = \psi^\dagger e^{-i2\beta_i K^i} \psi \neq \psi^\dagger \psi. \tag{6.41}$$

To fix this, we define

$$\overline{\psi} = \psi^\dagger \gamma^0. \tag{6.42}$$

Because (and you should check this!)

$$[J^i, \gamma^0] = 0, \quad [K^i, \gamma^0] = 2K^i, \tag{6.43}$$

we now find that under a boost,

$$\overline{\psi}\psi \to \psi^\dagger \left(e^{-i\beta_i K^i} \right)^\dagger \gamma^0 e^{-i\beta_i K^i} \psi = \psi^\dagger e^{-i\beta_i K^i} \psi^0 e^{-i\beta_i K^i} \psi = \psi^\dagger \psi^0 e^{i\beta_i K^i} e^{-i\beta_i K^i} = \overline{\psi}\psi, \tag{6.44}$$

and similarly under a rotation

$$\overline{\psi}\psi \to \psi^\dagger \left(e^{-i\beta_i J^i} \right)^\dagger \gamma^0 e^{-i\beta_i J^i} \psi = \psi^\dagger e^{i\beta_i J^i} \psi^0 e^{-i\beta_i J^i} = \psi^\dagger \psi^0 e^{i\beta_i J^i} e^{-i\beta_i J^i} = \overline{\psi}\psi. \tag{6.45}$$

We conclude that $\overline{\psi}\psi$ truly is a Lorentz scalar. A similar computation should convince you that $\overline{\psi}\gamma^\mu\psi$ is a Lorentz vector. The correct Lagrangian for the Dirac Equation is simply

$$\mathcal{L} = \overline{\psi}\left(i\gamma^\mu \partial_\mu - m \right)\psi. \tag{6.46}$$

Applying the Euler-Lagrange equation to $\overline{\psi}$ immediately yields the Dirac equation! If we instead apply the Euler-Lagrange equation to ψ, we get the Hermitian conjugate of the Dirac equation,

$$-i(\partial_\mu \overline{\psi})\gamma^\mu - m\overline{\psi} = 0. \tag{6.47}$$

Degrees of Freedom

How many degrees of freedom does a field have? The answer is obviously infinite, because we can vary the field at any point in spacetime to get a new field configuration. The number of degrees of freedom at a specific point however, is finite. For the scalar field we know that at any point we can specify $\phi(x)$ and $\pi(x)$; when we quantized the theory this gave rise to a single particle. We can hence say that the scalar field has one degree of freedom.

For the complex scalar field, the field and field momentum are complex, so there are four numbers to specify at any point. This gave rise to a theory with two particles, a particle and an anti-particle. The complex scalar field hence as two degrees of freedom.

What about the Dirac field? Naively, you may think there are eight degrees of freedom, corresponding to the four complex components of the spinor. Before jumping to conclusions, let's calculate the conjugate momentum

$$\pi_\psi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\overline{\psi}\psi^0 = i\psi^\dagger. \tag{6.48}$$

Notice that, unlike the scalar field, this is not independent of the field ψ! The reason

for this is that the Dirac equation is first order, not second order like the Klein-Gordon equation. So at any point we only have eight numbers to specify — the components of ψ. This means that the Dirac fields only has four degrees of freedom, and hence contains four particles. We will soon see that these particles can be interpreted as the spin-up and spin-down states of a particle, and an additional spin-up and spin-down for its anti-particle.

6.3   Solutions of the Dirac Equation

Before we quantize the Dirac field, we should try to solve the Dirac equation

$$(i\gamma^\mu \partial_\mu - m)\psi = 0. \tag{6.49}$$

Since any solution to the Dirac equation automatically solves the Klein-Gordon equation, we know that any plane-wave solution will have the form:

$$\psi(x) = e^{-ipx}u(p) + e^{ipx}v(p). \tag{6.50}$$

where $p^2 = m^2$ and where u and v are 4-spinors. Inserting this into the Dirac equation gives us conditions on u and v:

$$\begin{aligned}(\gamma^\mu p_\mu - m)u(p) &= 0 \\ (\gamma^\mu p_\mu + m)v(p) &= 0.\end{aligned} \tag{6.51}$$

We will first solve these equations in the rest frame, where $\mathbf{p} = 0$ and $p^0 = m$:

$$\begin{aligned}(\gamma^0 p_0 - m)u(\mathbf{0}) &= m \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} u(\mathbf{0}) = 0 \\ (\gamma^0 p_0 + m)v(\mathbf{0}) &= m \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} v(\mathbf{0}) = 0.\end{aligned} \tag{6.52}$$

In general, the solutions to these equations are

$$u(\mathbf{0}) = a \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad v(\mathbf{0}) = a \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} \tag{6.53}$$

where a and b are arbitrary coefficients.

How can we interpret these four independent solutions? Recall from (6.27) that

$$J^i = \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}, \tag{6.54}$$

so under a rotation

$$u_+(\mathbf{0}) = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \rightarrow \exp\left(-iJ^i\theta_i\right) u_+(\mathbf{0}) = \begin{pmatrix} e^{-iJ^i\theta_i}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ e^{-iJ^i\theta_i}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix}$$

$$u_-(\mathbf{0}) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \rightarrow \exp\left(-iJ^i\theta_i\right) u_-(\mathbf{0}) = \begin{pmatrix} e^{-iJ^i\theta_i}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ e^{-iJ^i\theta_i}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix}. \tag{6.55}$$

So the two spinors $u_\pm(\mathbf{0})$ transform like a spin-1/2 particle. Although it is premature to talk about particles and spin, is should not be surprising that, upon quantization, the two $u_\pm(\mathbf{0})$ spinors gives rise to the two spin states of a spin-1/2 particle. We similarly find that $v_\pm(\mathbf{0})$ gives rise to another spin-1/2 particle.

Now that we have the general form of $u(p)$ and $v(p)$ in the rest frame, we can apply boosts to calculate them in general. Given the boost $\Lambda_{\frac{1}{2}}(p)$ which takes $(m, \mathbf{0}) \rightarrow (\sqrt{m^2 + \mathbf{p}^2}, \mathbf{p})$, we define

$$u_{+1}(p) = \Lambda_{\frac{1}{2}}(p)\sqrt{m}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad u_{-1}(p) = \Lambda_{\frac{1}{2}}(p)\sqrt{m}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

$$v_{+1}(p) = \Lambda_{\frac{1}{2}}(p)\sqrt{m}\begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad v_{-1}(p) = \Lambda_{\frac{1}{2}}(p)\sqrt{m}\begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}. \tag{6.56}$$

The factor of $\sqrt{m}$ turns out to be the most useful normalization in calculations. Any solution of the Dirac equation with momentum p hence has the form

$$\psi(x) = \sum_{s=\pm 1} a_s e^{-ipx} u_s(p) + b_s e^{ipx} v_s(p) \tag{6.57}$$

where $a_{\pm 1}$ and $b_{\pm 1}$ are arbitrary complex numbers.

It is straightforward, albeit a little tedious, to express $u_s(p)$ and $v_s(p)$ in closed form. In most computations however, we can get away with a few useful formulas. The first is to note that, since $\overline{\psi}_1\psi_2$ is a Lorentz scalar,

$$\overline{u}_s(p)u_{s'}(p) = \overline{u}_s(\mathbf{0})u_{s'}(\mathbf{0}) = 2m\delta_{ss'}, \quad \overline{v}_s(p)v_{s'}(p) = -2m\delta_{ss'}, \quad \overline{u}_s(p)v_{s'}(p) = 0. \tag{6.58}$$

Note the minus sign for the v spinors!

Another useful formula allows us to sum over spins, which we need to compute unpolarized cross-sections. In the rest-frame we can calculate

$$\sum_s u_s(\mathbf{0})\overline{u}_s(\mathbf{0}) = m\gamma^0 + m, \quad \sum_s v_s(\mathbf{0})\overline{v}_s(\mathbf{0}) = m\gamma^0 - m, \tag{6.59}$$

so applying a Lorentz boost we find that

$$\sum_s u_s(p)\overline{u}_s(p) = \not{p} + m, \quad \sum_s v_s(p)\overline{v}_s(p) = \not{p} - m. \tag{6.60}$$

As a final example, we can consider

$$\begin{aligned}
\overline{u}_{s'}(\mathbf{0})\gamma^\mu u_s(\mathbf{0}) = 2m\delta_{s's} &\implies \overline{u}_{s'}(p)\gamma^\mu u_s(p) = 2p^\mu \delta_{s's} \\
\overline{v}_{s'}(\mathbf{0})\gamma^\mu v_s(\mathbf{0}) = 2m\delta_{s's} &\implies \overline{v}_{s'}(p)\gamma^\mu v_s(p) = 2p^\mu \delta_{s's}.
\end{aligned} \tag{6.61}$$

In quantum electrodynamics these formulae govern the coupling of the electron to the photon.

6.3.1   The Dirac Hamiltonian

So far we have been studying the classical Dirac equation. But what we haven't mentioned is that the Dirac equation harbours a deep, dark secret — it has no ground state! To see this, let us compute the Hamiltonian density

$$\mathcal{H} = \Pi_\psi \frac{\partial \mathcal{L}}{\partial \dot{\psi}} - \mathcal{L} = i\psi^\dagger \dot{\psi} - \overline{\psi}(i\not{\partial} - m)\psi = \overline{\psi}\left(-i\gamma^k \nabla_k + m\right)\psi. \tag{6.62}$$

If we calculate the energy density of a wave $e^{-ipx}u_s(p)$, then

$$\begin{aligned}
\mathcal{H}_+ &= e^{ipx}\overline{u}_s(p)\left(-i\gamma^k \nabla_k + m\right)(e^{-ipx}u_s(p)) \\
&= \overline{u}_s(p)\gamma^k u_s(p)p_k + m\overline{u}_s(p)u_s(p) \\
&= 2(\mathbf{p}^2 + m^2)
\end{aligned} \tag{6.63}$$

where we have made use of both (6.58) and (6.61). This energy is positive, and apart from the factor of two, looks like the energy of a single particle.

Now calculate the energy density of a wave $e^{ipx}v_s(p)$

$$\begin{aligned}
\mathcal{H}_- &= e^{-ipx}\overline{v}_s(p)\left(-i\gamma^k \nabla_k + m\right)(e^{ipx}v_s(p)) \\
&= -\overline{v}_s(p)\gamma^k v_s(p)p_k + m\overline{v}_s(p)v_s(p) \\
&= -2(\mathbf{p}^2 + m^2).
\end{aligned} \tag{6.64}$$

The energy is negative! This is a serious problem, both classically and in quantum mechanics. As soon as we add interactions, the vacuum will decay into a sea of high energy waves $u(\mathbf{p})$ and $v(\mathbf{p})$. We cannot quantize our theory in the same way we quantized the Klein-Gordon equation.

6.4   Quantizing the Dirac Equation

We will first quantize the Dirac equation using canonical quantization. Just like in Chapter 3, we can expand our fields in Fourier space as

$$\psi(\mathbf{x}) = \int \frac{dp}{(2\pi)^3}\frac{1}{\sqrt{2\omega_\mathbf{p}}}\sum_s a_\mathbf{p}^s u_s(p)e^{-i\mathbf{p}\cdot\mathbf{x}} + b_\mathbf{p}^{s\,\dagger}v_s(p)e^{i\mathbf{p}\cdot\mathbf{x}}, \tag{6.65}$$

where $\omega_{\mathbf{p}} = \sqrt{m^2 + \mathbf{p}^2}$. This is just the quantized version of (6.57), where the coefficients have been promoted to operators.

Following the steps for the Klein-Gordon field, we could apply the canonical quantization relations

$$[\psi_a(\mathbf{x}), \Pi_b(\mathbf{y})] = [\psi_a(\mathbf{x}), i\psi_b^\dagger(\mathbf{y})] = i\delta_{ab}\delta(\mathbf{x} - \mathbf{y}). \tag{6.66}$$

But we already know this is doomed to failure. The $v_s(p)e^{ipx}$ wave has negative energy, and so $b_{\mathbf{p}}^{s\,\dagger}$ will create particles with negative mass. You can find more details in *Peskin & Schroeder*, but the conclusion is simple: no matter how hard you try, (6.66) will not give a sensible quantum theory.

6.4.1 Fermionic Statistics

In Chapter 3, we discovered that the canonical commutators gave rise to the ladder operators

$$[a, a^\dagger] = 1, \tag{6.67}$$

which in turn describe a particle with bosonic statistics.

Electrons, however, are not bosons but fermions. It is impossible to have two electrons in the same state. Let $b_1^\dagger$ and $b_2^\dagger$ act on the vacuum $|0\rangle$ to create an electron in state 1 and state 2 respectively. The state

$$|1, 2\rangle = b_2^\dagger b_1^\dagger |0\rangle \tag{6.68}$$

then represents the state with both an electron in state 1 and in state 2. Fermi statistics then requires that

$$|2, 1\rangle = -|1, 2\rangle \implies \{b_1^\dagger, b_2^\dagger\} = b_1^\dagger b_2^\dagger + b_2^\dagger b_1^\dagger = 0. \tag{6.69}$$

Anticommutators, not commutators, hold the key to the Dirac equation.

As you will have learnt in many-body quantum mechanics, the fermionic ladder operators satisfy the anticommutation rules

$$\{b_i, b_j^\dagger\} = \delta_{ij}, \quad \{b_i, b_j\} = \{b_i^\dagger, b_j^\dagger\} = 0. \tag{6.70}$$

The generalization to fields is easy to guess

$$\begin{aligned} \{\psi_a(\mathbf{x}), \Pi_b(\mathbf{y})\} = \{\psi_a(\mathbf{x}), i\psi_b^\dagger(\mathbf{y})\} &= i\delta_{ab}\delta(\mathbf{x} - \mathbf{y}), \\ \{\psi_a(\mathbf{x}), \psi_b(\mathbf{y})\} = \{\psi_a^\dagger(\mathbf{x}), \psi_b^\dagger(\mathbf{y})\} &= 0. \end{aligned} \tag{6.71}$$

Writing the field operators in terms of the ladder operators

$$\begin{aligned} \psi(x) &= \int \frac{dp}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \sum_s a_{\mathbf{p}}^s u_s(p)e^{-ipx} + b_{\mathbf{p}}^{s\,\dagger} v_s(p)e^{ipx} \\ \overline{\psi}(x) &= \int \frac{dp}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \sum_s a_{\mathbf{p}}^{s\,\dagger} \overline{u}_s(p)e^{ipx} + b_{\mathbf{p}}^s \overline{v}_s(p)e^{-ipx}, \end{aligned} \tag{6.72}$$

we find that the ladder operators satisfy anticommutation relations

$$\begin{aligned} \{a_{\mathbf{p}}^s, a_{\mathbf{p}'}^{s\,\dagger}\} &= (2\pi)^3 \delta^{ss'} \delta(\mathbf{p} - \mathbf{p}') \\ \{b_{\mathbf{p}}^s, b_{\mathbf{p}'}^{s\,\dagger}\} &= (2\pi)^3 \delta^{ss'} \delta(\mathbf{p} - \mathbf{p}'), \end{aligned} \tag{6.73}$$

with all other commutators being zero. The Hamiltonian is

$$
\begin{aligned}
H &= \int d^3x \, \overline{\psi} \left(-i\gamma^k \nabla_k + m \right) \psi \\
&= \int \frac{d^3p}{(2\pi)^3} \, \omega_{\mathbf{p}} \left(\sum_s a_{\mathbf{p}}^{s\,\dagger} a_{\mathbf{p}}^s - b_{\mathbf{p}}^s b_{\mathbf{p}}^{s\,\dagger} \right) \\
&= \int \frac{d^3p}{(2\pi)^3} \, \omega_{\mathbf{p}} \left(\sum_s a_{\mathbf{p}}^{s\,\dagger} a_{\mathbf{p}}^s + b_{\mathbf{p}}^{s\,\dagger} b_{\mathbf{p}}^s \right) - \int d^3p \, 2\omega_{\mathbf{p}} \delta^{(3)}(\mathbf{0}).
\end{aligned}
\tag{6.74}
$$

The Vacuum Energy

Just like for the Klein-Gordon equation, we have found that the Hamiltonian is infinite! And just like the vacuum energy of the Klein-Gordon field, we can ignore the vacuum energy of the Dirac field in almost all calculations. Nevertheless, it is interesting to compare the two expressions. Recalling from Chapter 3 that

$$
H_0^{\text{KG}} = \int d^3p \, \frac{\omega_{\mathbf{p}}}{2} \delta(\mathbf{0}),
\tag{6.75}
$$

we find that

$$
H_0^{\text{Dirac}} = -4 H_0^{\text{KG}}.
\tag{6.76}
$$

The factor of 4 is easy to understand — the Dirac equations describes four different particles, whereas the Klein-Gordon equation has only one. Much more interesting is the difference in sign. Whereas the Klein-Gordon vacuum has positive infinite energy, the Dirac vacuum energy is negative infinity!

Could this solve the cosmological constant problem discussed in Chapter 3? In certain types of theories, those possessing *supersymmetry*, the bosonic and fermionic vacuum energies are equal and opposite, and so the resulting ground state has exactly zero energy. But alas, our universe does not appear to be supersymmetric, at least at the energies we can currently observe.

6.4.2   Particles

Having discussed the vacuum energy, we will now drop this term from our Hamiltonian:

$$
H = \int \frac{d^3p}{(2\pi)^3} \, \omega_{\mathbf{p}} \left(\sum_s a_{\mathbf{p}}^{s\,\dagger} a_{\mathbf{p}}^s + b_{\mathbf{p}}^{s\,\dagger} b_{\mathbf{p}}^s \right).
\tag{6.77}
$$

Notice that our Hamiltonian treats the $a_{\mathbf{p}}^s$ and $b_{\mathbf{p}}^s$ symmetrically. Although we have been using anti-commutators, the Hamiltonian satisfies the following commutator relations

$$
\begin{aligned}
{}[H, a_{\mathbf{p}}^s] &= -\omega_{\mathbf{p}} a_{\mathbf{p}}^s, & [H, a_{\mathbf{p}}^{s\,\dagger}] &= -\omega_{\mathbf{p}} a_{\mathbf{p}}^{s\,\dagger} \\
[H, b_{\mathbf{p}}^s] &= -\omega_{\mathbf{p}} b_{\mathbf{p}}^s, & [H, b_{\mathbf{p}}^{s\,\dagger}] &= -\omega_{\mathbf{p}} b_{\mathbf{p}}^{s\,\dagger}.
\end{aligned}
\tag{6.78}
$$

So just like the Klein-Gordon equation, we can create a tower of energy eigenstates by acting the creation and annihilation operators on a given energy eigenstate. In particular, the vacuum state must be annihilated by the annihilation operators

$$
a_{\mathbf{p}}^s \ket{0} = b_{\mathbf{p}}^s \ket{0} = 0.
\tag{6.79}
$$

As you may guess, the state

$$a_{\mathbf{p}}^{s\,\dagger}|0\rangle \tag{6.80}$$

is a single-particle state, with energy $\omega_{\mathbf{p}}$. A quick calculation of the momentum operator

$$\mathbf{P} = \int d^3x\ \psi^\dagger(-i\nabla)\psi = \int \frac{d^3p}{(2\pi)^3}\ \sum_s \mathbf{p}\left(a_{\mathbf{p}}^{s\,\dagger}a_{\mathbf{p}}^s + b_{\mathbf{p}}^{s\,\dagger}b_{\mathbf{p}}^s\right) \tag{6.81}$$

confirms that this state has momentum $\mathbf{p}$. By the same logic, $b_{\mathbf{p}}^{s\,\dagger}|0\rangle$ is also a single-particle state. By convention, we refer to the particles created by $a_{\mathbf{p}}^{s\,\dagger}$ as *fermions* and those created by $b_{\mathbf{p}}^{s\,\dagger}$ as *antifermions*. To relativistically normalize the one-particle states, we define

$$|\mathbf{p}, s\rangle = \sqrt{2\omega_{\mathbf{p}}}a_{\mathbf{p}}^{s\,\dagger}|0\rangle\,. \tag{6.82}$$

As we have discussed previously, the two fermions $|\mathbf{p}, \pm\rangle$ refer to the two spin states of the fermion, each carrying momentum $\hbar/2$. Demonstrating this properly is a little tedious — see page 60 of *Peskin & Schroeder* for details.

Spin and Statistics

Multiparticle states can be produced by acting multiple raising operators on the vacuum. Because our raising operators now anti-commute, our particles now satisfy Fermi-Dirac statistics

$$\left|\mathbf{p}, s; \mathbf{q}, s'\right\rangle = a_{\mathbf{p}}^{s\,\dagger}a_{\mathbf{p}'}^{s'\,\dagger}|0\rangle = -a_{\mathbf{p}'}^{s'\,\dagger}a_{\mathbf{p}}^{s\,\dagger}|0\rangle = -\left|\mathbf{q}, s'; \mathbf{p}, s\right\rangle\,. \tag{6.83}$$

In particular, the fermion satisfy the Pauli Exclusion principle $|\mathbf{p}, s; \mathbf{p}, s\rangle = 0$.

We have hence found that spin-0 particles when quantized give rise to bosons, and spin-1/2 particles when quantized give rise to fermions. If we try to quantize the Dirac equation using commutators, the ground state would not be stable. Similarly, if you try to quantize a spin-0 particle using anticommutators you will find that it does not lead to a consistent quantum theory.[4]

The *spin-statistics theorem* states that any particle with an integer spin must be a boson, and that any particle with half-integer spin is a fermion. A properly rigorous proof is not easy to give — quantum field theory is notoriously difficult to define — but the theorem ranks amongst the most important in all of physics.

Charge

Like the complex scalar field, the Dirac Lagrangian

$$\mathcal{L} = \overline{\psi}(i\slashed{\partial} - m)\psi \tag{6.84}$$

is symmetric under phase rotation

$$\psi \rightarrow e^{i\alpha}\psi. \tag{6.85}$$

This symmetry is important because it will continue to hold when we add interactions; as long as our theory is built from Dirac bilinears such as $\overline{\psi}\psi$ and $\overline{\psi}\gamma^\mu\psi$, phase invariance will continue to hold. Applying Noether's theorem, we discover that the conserved current associated to this symmetry is

$$j^\mu = \overline{\psi}\gamma^\mu\psi. \tag{6.86}$$

---

[4]See II.4 of *Zee* or Chapter 4 of *Srednicki* for details. Or work it out for yourself using Chapter 3!

The conserved quantity is hence

$$Q = \int d^3x \, \overline{\psi}\gamma^0\psi = \int d^3x \, \psi^\dagger\psi = \int \frac{d^3p}{(2\pi)^3} \sum_s \left(a_{\mathbf{p}}^{s\,\dagger} a_{\mathbf{p}}^s - b_{\mathbf{p}}^{s\,\dagger} b_{\mathbf{p}}^s \right). \tag{6.87}$$

From this we conclude that $a_{\mathbf{p}}^{s\,\dagger}$ creates a fermion of charge $+1$, and that $b_{\mathbf{p}}^{s\,\dagger}$ creates a fermion of charge -1.

When we discuss quantum electrodynamics, we will discover that Q is proportional to charge of the particle. For instance, in QED $a_{\mathbf{p}}^{s\,\dagger}$ creates an electron of charge $-e$, and $b_{\mathbf{p}}^{s\,\dagger}$ creates a positron of charge $+e$.

6.4.3   The Dirac Propagator

Have so far worked in the Schrödinger pictures, let us now switch to the Heisenberg picture to calculate the Feynman propagator

$$(S_F)_{ab}(x-y) = \langle 0|\mathcal{T}\psi_a(x)\overline{\psi}_b(y)\,|0\rangle. \tag{6.88}$$

We have included indices to make obvious that $S_F(x-y)$ is a 4×4 matrix. Because of the anti-commuting nature of ψ, time-ordering now comes with an extra minus sign:

$$\mathcal{T}\psi_a(x)\overline{\psi}_b(y) = \theta(x^0-y^0)\psi_a(x)\overline{\psi}_b(y) - \theta(y^0-x^0)\overline{\psi}_b(y)\psi_a(x). \tag{6.89}$$

Computing the propagator is a straightforward task using (6.65):

$$\begin{aligned}
\langle 0|\psi_a(x)\overline{\psi}_b(y)\,|0\rangle &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{p}}} \sum_s u_a^s(p)\overline{u}_b^s(p)e^{-ip(x-y)} \\
&= \int \frac{d^3p}{(2\pi)^3}\frac{\slashed{p}+m}{2\omega_{\mathbf{p}}}e^{-ip(x-y)} \\
\langle 0|\overline{\psi}_b(y)\psi_a(x)\,|0\rangle &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{p}}} \sum_s v_a^s(p)\overline{v}_b^s(p)e^{-ip(y-x)} \\
&= \int \frac{d^3p}{(2\pi)^3}\frac{\slashed{p}-m}{2\omega_{\mathbf{p}}}e^{-ip(y-x)}.
\end{aligned} \tag{6.90}$$

Combining these, the Feynman propagator is

$$S_F(x-y) = \int \frac{d^3p}{(2\pi)^3}\frac{i}{2\omega_{\mathbf{p}}} \left(\theta(x^0-y^0)(\slashed{p}+m)e^{-ip(x-y)} - \theta(y^0-x^0)(\slashed{p}-m)e^{ip(x-y)} \right). \tag{6.91}$$

Like the Klein-Gordon propagator, this can be elegantly written as an integral over four momentum:

$$S_F(x-y) = \int \frac{d^4p}{(2\pi)^4}\frac{i(\slashed{p}+m)}{p^2-m^2+i\varepsilon}e^{-ip(x-y)} \tag{6.92}$$

where ε is an infinitesimal positive number. When we draw Feynman diagrams, this is the quantity we associate to the fermion lines.

6.5   The Fermionic Path-Integral

The Dirac field satisfies anticommutation relations. This means that we cannot just apply the path-integral naively. As we saw previously this results gives commutation relations, such as the microcausality condition discussed in Section 4.2.1, which are not compatible with fermionic statistics. The solution is to invent a new type of number, known as *Grassmann numbers*. At first these may seem a little *ad hoc*, just as imaginary numbers once seemed strange. Nevertheless, Grassmann numbers are the correct way to study classical fields which anticommute.

6.5.1   Grassmann Numbers

The defining feature of Grassmann numbers is that they *anticommute*; that is, for any two numbers θ and η,

$$\theta\eta = -\eta\theta. \tag{6.93}$$

As a consequence of this relation, Grassmann numbers square to zero: $\theta^2 = 0$. Grassmann numbers can be added together and multiplied by ordinary numbers, just like in a vector space. So for instance

$$(a\theta + b\eta)(c\theta + d\eta) = ad\theta\eta + bc\eta\theta = (ad - bc)\theta\eta, \tag{6.94}$$

where a, b, c and d are ordinary (that is, real or complex) numbers. Another interesting property is that the product of two Grassmann numbers commutes with any Grassman number

$$(\theta\eta)\xi = \theta\eta\xi = -\theta\xi\eta = \xi\theta\eta = \xi(\theta\eta). \tag{6.95}$$

We define any function of a Grassmann variable by a power series:

$$f(\theta) = a + b\theta, \tag{6.96}$$

for two ordinary numbers a and b. Because $\theta^2 = 0$, this is the most general form a function can have on a single variable θ.

Integration

The main reason Grassmann variables exist is so that they can be integrated over. Our eventual aim, after all, is to describe the fermionic path-integral. We define the integral to have two properties:

1. Integration must be linear

2. Integration must be invariant under a change of variables $\theta \to \theta + \eta$

This last property is analogous to the invariance of the integral

$$\int_{-\infty}^{\infty} f(x) \, dx = \int_{-\infty}^{\infty} f(y - a) \, dy; \tag{6.97}$$

our Grassmann integral is the analogy of these integrals. Using linearity, we find that

$$\int d\theta \, f(\theta) = \int d\theta(a + b\theta) = a \int d\theta + b \int d\theta \, \theta. \tag{6.98}$$

Now applying the second property of Grassmann integration

$$\int d\theta \ f(\theta) = \int d\theta \ f(\theta - \eta) \implies (a + b\eta) \int d\theta + b \int d\theta \ \theta = a \int d\theta + b \int d\theta \ \theta. \quad (6.99)$$

This is only possible if $\int d\theta = 0$. We now define the Grassmann integral to be

$$\int d\theta(a + b\theta) = b. \quad (6.100)$$

For the case of multiple integrals there is a further sign ambiguity; the convention adopted is

$$\int d\theta \int d\eta \ \eta\theta = +1 \quad (6.101)$$

performing the inside integral first.

Differentiation

We can also differentiate Grassmann variables, defining

$$\frac{d}{d\theta}f(\theta) = \frac{d}{d\theta}(a + b\theta) = b, \quad (6.102)$$

just like for regular derivatives. Note that this gives the same result as integration! Just like integration, there is a sign ambiguity when performing multiple derivatives, which we fix by defining

$$\frac{d}{d\eta}(\eta\theta) = +\theta. \quad (6.103)$$

6.5.2   Evaluating the Path-Integral

To define the Dirac path-integral, we consider the Dirac field classically as a *Grassmann field*; that is, a field which anticommutes with itself. We then define the path-integral in the same way as for scalar fields

$$Z_0 = \int \mathcal{D}\overline{\psi} \ \mathcal{D}\psi \exp\left[i \int d^4x \ \overline{\psi}(i\displaystyle{\not}\partial - m)\psi\right] \quad (6.104)$$

except here our path-integral is over Grassmann variables. The results of Chapter 4 hold with only cosmetic changes; for instance,

$$\langle\Omega|\mathcal{T}\psi(x_1)\overline{\psi}(x_2)|\Omega\rangle = \frac{\int \mathcal{D}\overline{\psi} \ \mathcal{D}\psi \ \psi(x_1)\overline{\psi}(x_2) \ \exp\left[i \int d^4x \ \overline{\psi}(i\displaystyle{\not}\partial - m)\psi\right]}{\int \mathcal{D}\overline{\psi} \ \mathcal{D}\psi \exp\left[i \int d^4x \ \overline{\psi}(i\displaystyle{\not}\partial - m)\psi\right]} \quad (6.105)$$

The Generating Functional

In practice, the simplest way to calculate correlation functions is to define the generating functional

$$Z[\eta, \overline{\eta}] = \int \mathcal{D}\overline{\psi} \ \mathcal{D}\psi \exp\left[i \int d^4x \ \overline{\psi}(i\displaystyle{\not}\partial - m)\psi + \overline{\eta}\psi + \overline{\psi}\eta\right], \quad (6.106)$$

where η is a Grassmann-valued spinor. To evaluate this integral, we will complete the square, just like we did in Section 4.2.2. Going into Fourier space, we find that

$$Z[\eta, \overline{\eta}] = \int \mathcal{D}\overline{\psi} \, \mathcal{D}\psi \exp\left[i \int d^4k \, \overline{\psi}(-k)(\not{k} - m)\psi(k) + \overline{\eta}(-k)\psi(k) + \overline{\psi}(-k)\eta(k)\right]. \tag{6.107}$$

We now define the change of variables

$$\Psi(k) = \psi(k) - \frac{\not{k} + m}{k^2 - m^2}\eta(k) \tag{6.108}$$

$$\overline{\Psi}(k) = \Psi^*(k)\gamma^0 = \overline{\psi}(k) - \overline{\eta}(k)\frac{\not{k} + m}{k^2 - m^2}. \tag{6.109}$$

This is simply a shift of variables and hence does not change the integration measure. Using the identity

$$(\not{k} + m)(\not{k} - m) = (\not{k} - m)(\not{k} + m) = k^2 - m^2, \tag{6.110}$$

our path-integral becomes

$$\begin{aligned}Z[\eta, \overline{\eta}] &= \int \mathcal{D}\overline{\Psi} \, \mathcal{D}\Psi \exp\left[i \int d^4k \, \overline{\Psi}(-k)(\not{k} - m)\Psi(k) + \overline{\eta}(-k)\frac{\not{k} + m}{k^2 - m^2}\eta(k)\right] \\ &= Z_0 \exp\left[i \int d^4k \, \overline{\eta}(-k)\frac{\not{k} + m}{k^2 - m^2}\eta(k)\right]. \end{aligned} \tag{6.111}$$

Using the Feynman propagator

$$S_F(x - y) = \int \frac{d^4k}{(2\pi)^4} \frac{i(\not{k} + m)}{k^2 - m^2 + i\varepsilon} e^{-ik(x-y)}, \tag{6.112}$$

we can now write the path-integral in position space as

$$Z[\eta, \overline{\eta}] = Z_0 \exp\left[-\int d^4x \, d^4y \, \overline{\eta}(x)S_F(x - y)\eta(y)\right], \tag{6.113}$$

where, as for the scalar field, the infinitesimal $i\varepsilon$ guarantees that our integrals are well-defined.

6.5.3   Correlation Functions for the Dirac Field

It is now straightforward to calculate correlation functions; we just need to be careful to keep track of signs when differentiating. For instance, the two-point function can be calculated as

$$\begin{aligned}\langle\Omega|\mathcal{T}\psi(x)\overline{\psi}(y)|\Omega\rangle &= \frac{1}{Z_0}\left(-i\frac{\delta}{\delta\overline{\eta}(x)}\right)\left(+i\frac{\delta}{\delta\eta(y)}\right) Z[\overline{\eta}, \eta]\Big|_{\eta=0} \\ &= S_F(x - y). \end{aligned} \tag{6.114}$$

Higher correlation function can be calculated in a similar way. A large numbers of these functions are actually zero:

$$\langle\Omega|\mathcal{T}\psi(x_1)...\psi(x_n)\overline{\psi}(y_1)...\overline{\psi}(y_m)|\Omega\rangle = 0 \text{ if } m \neq n, \tag{6.115}$$

since otherwise there will always be unpaired η or $\overline{\eta}$ after we differentiate.

To understand more deeply why these correlators vanish, recall that the Dirac equation is symmetric under the transformation $\psi \to e^{i\alpha}\psi$. Since the correlation function must be invariant under this symmetry, we deduce that

$$\langle\Omega|\mathcal{T}\psi(x_1)...\psi(x_n)\overline{\psi}(y_1)...\overline{\psi}(y_m)|\Omega\rangle = e^{i\alpha(m-n)}\langle\Omega|\mathcal{T}\psi(x_1)...\psi(x_n)\overline{\psi}(y_1)...\overline{\psi}(y_m)|\Omega\rangle$$
(6.116)

as long as the ground state is invariant under the transformation. For $m \neq n$, this is possible only if the correlation function vanishes. This argument is almost identical to the one given in Section 5.1, where we showed that odd correlation functions were zero in $\lambda\phi^4$ theory. The only difference is that here we have a continuous rather than a discrete symmetry.

The only non-trivial correlators left to calculate are of the form:

$$\langle\Omega|\mathcal{T}\psi(x_1)...\psi(x_n)\overline{\psi}(y_1)...\overline{\psi}(y_n)|\Omega\rangle .$$
(6.117)

These can be evaluated diagrammatically, in much the same way as for the scalar field. Although it is straightforward to derive these from (6.113), we will simply state the results. First we draw a point for each x_i and y_j. We then pair each x_i with a single y_j, associating to this a factor of $S_F(x_i - y_j)$.

Now comes the tricky part: if our permutation is an odd permutation, we associate a minus sign to the diagram. This comes from the fermionic nature of the field. We then sum over all possible diagrams. As an example, the four-point function is

$$\langle\Omega|\mathcal{T}\psi(x_1)\psi(x_2)\overline{\psi}(y_1)\overline{\psi}(y_2)|\Omega\rangle =$$ (6.118)

$$= S_F(x_1 - y_1)S_F(x_2 - y_2) - S_F(x_1 - y_2)S_F(x_2 - y_1)$$

6.6 Yukawa Theory

We have spent a long time studying the free Dirac field. Having developed the path-integral for the Dirac field, we are now in a position to derive the Feynman rules for interacting theories involving fermions. The simplest such theory is Yukawa theory

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 + \overline{\psi}(i\not{\partial} - M)\psi - g\phi\overline{\psi}\psi,$$
(6.119)

which describes the interaction of a scalar ϕ with a fermion ψ. Couplings of the form $g\phi\overline{\psi}\psi$ are called *Yukawa couplings*, and appear in the Standard Model between fermions and the Higgs boson.[5]

What is the dimension of g? To calculate this, we note that the kinetic term

$$\int d^4x\, \overline{\psi}(i\not{\partial})\psi$$
(6.120)

---

[5]In fact, of the 18 parameters in the Standard model, 13 come from Yukawa couplings.

must be unitless. This then requires that $[\psi] = 3/2$. The dimesnion of $[\phi\overline{\psi}\psi]$ is 4, and so g is unitless.

Yukawa theory was first proposed as a crude model of the nuclear force. The nucleon field is represented by ψ, and the meson field by ϕ. This is of course not a very accurate model of nuclear force, not the least because there are two types of nucleons — the proton and the neutron. Nevertheless, the theory is a good toy model in which to study interacting fermions. In the next chapter, we will apply what we have learnt to the more physically relevant theory of QED.

6.6.1   Feynman Rules

To calculate the Feynman rules, we evaluate the path-integral

$$Z = \int \mathcal{D}\phi \,\mathcal{D}\psi \,\mathcal{D}\overline{\psi} \,\exp\left[i\int d^4x \,\frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 + \overline{\psi}(i\slashed{\partial} - M)\psi - g\phi\overline{\psi}\psi\right] \quad (6.121)$$

as a power series in g, just like we did for scalar fields.
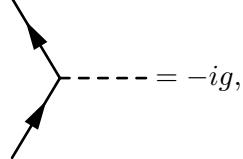
It is not very hard to guess the Feynman rules for our theory. To the meson field we associate the propagator

$$\text{-------}_p \quad = \frac{i}{p^2 - m^2} \quad (6.122)$$

while to the nucleon we associate the propagator

$$\xrightarrow{\quad\quad}_p \quad = \frac{i(\slashed{p} + m)}{p^2 - M^2}. \quad (6.123)$$

Next we introduce the vertex

$$= -ig, \quad (6.124)$$

which represents the $g\overline{\psi}\psi$ interaction in the Lagrangian. The regular rules for momentum apply — momentum is to be conserved at every vertex, and all undetermined momenta are to be integrated over.

External Lines

External lines will be amputated when we are calculating scattering amplitudes. For the mesons we do this by setting each external leg to 1. For nucleons it is a little tricker, because we have to keep track of the nucleon spin. To find the correct factors, recall that, in the Schrödinger picture,

$$\psi(\mathbf{x}) = \int \frac{dp}{(2\pi)^3} \frac{1}{\sqrt{2\omega_\mathbf{p}}} \sum_s a_\mathbf{p}^s u_s(p)e^{-i\mathbf{p}\cdot\mathbf{x}} + b_\mathbf{p}^{s\,\dagger} v_s(p)e^{i\mathbf{p}\cdot\mathbf{x}}$$

$$\overline{\psi}(\mathbf{x}) = \int \frac{dp}{(2\pi)^3} \frac{1}{\sqrt{2\omega_\mathbf{p}}} \sum_s a_\mathbf{p}^s \overline{u}_s(p)e^{i\mathbf{p}\cdot\mathbf{x}} + b_\mathbf{p}^{s\,\dagger} \overline{v}_s(p)e^{-i\mathbf{p}\cdot\mathbf{x}}. \quad (6.125)$$

From this we calculate

$$
\begin{aligned}
\langle\Omega|\psi(\mathbf{x})\,|\mathbf{p},s\rangle_f &= u_s(p)e^{-i\mathbf{p}\cdot\mathbf{x}}, & {}_a\langle\mathbf{p},s|\psi(\mathbf{x})\,|0\rangle &= v_s(p)e^{i\mathbf{p}\cdot\mathbf{x}}, \\
\langle\Omega|\overline{\psi}(\mathbf{x})\,|\mathbf{p},s\rangle_a &= \overline{v}_s(p)e^{-i\mathbf{p}\cdot\mathbf{x}}, & {}_f\langle\mathbf{p},s|\psi(\mathbf{x})\,|0\rangle &= \overline{u}_s(p)e^{i\mathbf{p}\cdot\mathbf{x}},
\end{aligned}
\tag{6.126}
$$

where the f and a subscripts denote fermion and antifermion states respectively. This gives us the factors[6] we need for external fermions:

$$
\begin{aligned}
&\text{Incoming fermion} &&= u_s(\mathbf{p}), &&\text{Outgoing fermion} &&= \overline{u}_s(\mathbf{p}) \\
&\text{Incoming antifermion} &&= \overline{v}_s(\mathbf{p}), &&\text{Outgoing antifermion} &&= v_s(\mathbf{p}).
\end{aligned}
\tag{6.127}
$$

Symmetry Factors and Minus Signs

We are almost done with the Yukawa Feynman rules — we just have to deal with some pesky factors. The good news is that all symmetry factors in Yukawa theory are one (hooray!). This is because the vertex $\phi\overline{\psi}\psi$ has no symmetries: we can always distinguish between the scalar, the fermion, and the antifermion. So we never have to worry about symmetry factors.
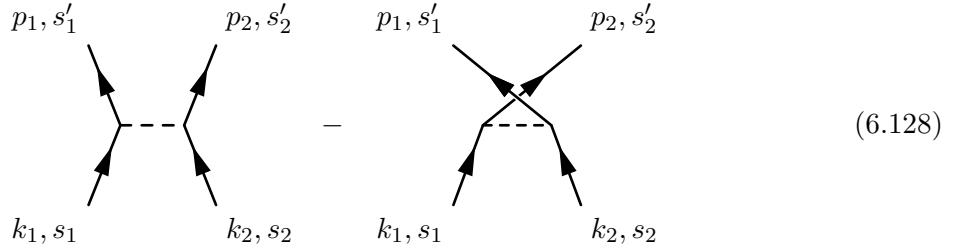
The bad news is that, because ψ is a fermion, there is a new crop of annoying factors to deal with. This is because fermions anticommute, and so when we swap to fermions we get an additional minus sign. We saw this already in Section 6.5.3 when we calculated the four-point function for a free Dirac field. Although there do exist general methods to determine the sign of a diagram, in practise we only need to remember two things

1. Every closed fermionic loop has an additional factor of (-1).

2. When combining the effects of two diagrams, we should consider the statistics of the initial and final states.

We will not need to worry too much about the first rule, since it is only relevant for loop diagrams. The second rule is however quite important, as we shall see in the next section.

6.6.2   Scattering of Indistinguishable Particles

Consider the process of two fermions scattering off each other. Label the incoming momenta and spin k_1, s_1 and k_2, s_2, and label the outgoing momenta and spin p_1, s_1' and p_2, s_2'. At tree level, are two diagrams which contribute to this process:

$$\tag{6.128}$$

Note the crucial minus sign between these two diagrams. This comes from Fermi statistics: the fermions in the final state have been swapped in the second diagram, and we hence need to include a minus sign to reflect this.

---

[6]A more rigorous proof of this would, like for the scalar case, require us to use the LSZ reduction formula.

It is now straightforward to write down the scattering amplitude

$$\mathcal{M} = -g^2 \left[\overline{u}_{s_1'}(p_1)u_{s_1}(k_1)\frac{1}{(p_1-k_1)^2-m^2}\overline{u}_{s_2'}(p_2)u_{s_2}(k_2) \right.$$
$$\left. -\overline{u}_{s_2'}(p_2)u_{s_1}(k_1)\frac{1}{(p_2-k_1)^2-m^2}\overline{u}_{s_1'}(p_1)u_{s_2}(k_2) \right]. \tag{6.129}$$

Converting this to a cross-section requires a bit more work; we postpone this for scattering in QED, which is more interesting than the less physically relevant Yukawa theory.

The Yukawa Potential

Another chapter, another discussion of the Yukawa potential. This time we can ask the question: what force do fermions feel? We will concentrate on the case of distinguishable fermions, where we have to worry only about the first, not the second diagram above. Next enter in the centre of mass frame, where as we recall from the previous chapter

$$k_1 = (E, \mathbf{k}), \qquad k_2 = (E, -\mathbf{k})$$
$$p_1 = (E, \mathbf{p}), \qquad p_2 = (E, -\mathbf{p}). \tag{6.130}$$

Therefore, $(p_1-k_1)^2 = -(\mathbf{p}-\mathbf{k})^2$.

Consider the non-relativistic limit, where $\mathbf{p}^2 = \mathbf{k}^2 \ll m^2$. In this limit,
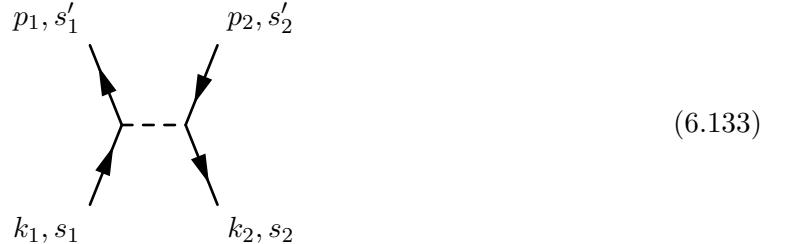
$$\overline{u}_{s_1'}(p_1)u_{s_1}(k_1) \to \overline{u}_{s_1'}(\mathbf{0})u_{s_1}(\mathbf{0}) = 2M\delta_{s_1's_1}$$
$$\overline{u}_{s_2'}(p_2)u_{s_2}(k_2) \to \overline{u}_{s_2'}(\mathbf{0})u_{s_2}(\mathbf{0}) = 2M\delta_{s_2's_2}. \tag{6.131}$$

Our amplitude now becomes

$$\mathcal{M} = \frac{g^2}{(\mathbf{p}-\mathbf{k})^2+m^2}4M^2\delta_{s_1's_1}\delta_{s_2's_2}. \tag{6.132}$$

The factor of $4M^2$ comes from our relativistic normalization, and will cancel when calculating the cross-section. Other than this factor our amplitude is identical to (5.121). We hence conclude that fermions interact according to an attractive Yukawa potential.

What about a fermion scattering from an antifermion? We would then calculate the diagram

$$\tag{6.133}$$

From this we calculate the amplitude

$$
\begin{aligned}
\mathcal{M} &= (-1)\frac{g^2}{(\mathbf{p}-\mathbf{k})^2+m^2}4M^2\overline{u}_{s_1'}(p_1)u_{s_1}(k_1)\overline{v}_{s_1'}(p_1)v_{s_1}(k_1)\\
&\to (-1)\frac{g^2}{(\mathbf{p}-\mathbf{k})^2+m^2}(-1)4M^2\delta_{s_1's_1}\delta_{s_2's_2} \qquad\qquad (6.134)\\
&= \frac{g^2}{(\mathbf{p}-\mathbf{k})^2+m^2}4M^2\delta_{s_1's_1}\delta_{s_2's_2}.
\end{aligned}
$$

since $\overline{v}_{s'}(\mathbf{0})v_s(\mathbf{0}) = -2m\delta_{ss'}$. Notice there is also a global minus, which is the factor of minus one out front. One way to understand this minus sign is to notice that if we connected the incoming fermion-antifermion, and the outgoing fermion-antifermion, we would have a loop diagram, which we associated with an additional minus sign. This minus sign cancels out the minus sign from the contraction of v with $\overline{v}$. We conclude that a fermion and antifermion are also attracted to each other by a Yukawa potential.

All that is left is the case of two antifermions. Unsurprisingly, this also results in an attractive Yukawa potential. We conclude that scalar fields are universally attractive.

Quantum Electrodynamics

In this chapter, we will quantize the electromagnetic field. We will then couple the field to fermions, we will get *quantum electrodynamics*, the quantum theory of light and matter.

From previous courses on electromagnetism, you should remember the Maxwell Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \tag{7.1}$$

where the field strength is defined by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{7.2}$$

The Euler-Lagrange equations now give us the equations of motion for our theory:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu A_\nu)} \right) = -\partial_\mu F^{\mu\nu} = 0. \tag{7.3}$$

The electromagnetic tensor also satisfies the *Bianchi Identity*

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\mu F_{\lambda\nu} = 0, \tag{7.4}$$

which can be derived straight from the definition of the F.

You are probably used to thinking in terms of the electric and magnetics fields, $\mathbf{E}$ and $\mathbf{B}$ respectively. These are embedded in the electromagnetic tensor as

$$F^{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{7.5}$$

The equation of motion (7.3) is equivalent to the equations

$$\nabla \cdot \mathbf{E} = 0, \quad \nabla \times \mathbf{B} = \frac{\partial \mathbf{E}}{\partial t}, \tag{7.6}$$

whilst the Bianchi identity (7.4) gives other two Maxwell equations

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \cdot \mathbf{B} = 0. \tag{7.7}$$

The electromagnetic fields interacts with the charge current $J^\mu = (\rho, \mathbf{J})$. To include

this interaction, we modify our Lagrangian to

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + A^\mu J_\mu. \tag{7.8}$$

The equations of motion then become

$$\partial_\mu F^{\mu\nu} = J^\nu, \tag{7.9}$$

which then implies that

$$\partial_\nu J^\nu = \partial_\nu \partial_\mu F^{\mu\nu} = -\partial_\nu \partial_\mu F^{\mu\nu} \implies \partial_\nu J^\nu = 0. \tag{7.10}$$

So we can only couple the electromagnetic field to conserved currents; otherwise our theory will be inconsistent.

7.1   Gauge Invariance

You may be thinking that, because A^μ has four components, when we quantize we shall get four different particles. In reality however, the photon only has two degrees of freedom, corresponding to the two different polarizations of light. What happens to the other degrees of freedom?

One of the degrees of freedom is easy to account for. If you look carefully at the Maxwell Lagrangian, you will notice that $\dot{A}_0$ does not appear. Using the equations of motion,

$$\partial_\mu F^{\mu\nu} = 0 \implies 0 = \partial_i F^{i0} = \nabla^2 A_0 + \nabla \cdot \frac{\partial \mathbf{A}}{\partial t} = 0. \tag{7.11}$$

From this equation, we can solve to find

$$A_0(\mathbf{x}) = \int d^3y \, \frac{1}{4\pi|\mathbf{x} - \mathbf{y}|} \nabla \cdot \frac{\partial \mathbf{A}(\mathbf{y})}{\partial t}. \tag{7.12}$$

This means that A_0 is fixed by $\mathbf{A}$; it isn't something we get to choose ourselves. This eliminates one degree of freedom.

The other fake degree of freedom comes from *gauge symmetry*. As you have learnt in previous courses, under the *gauge transformation*

$$A^\mu(x) \to A^\mu(x) + \partial^\mu \alpha(x), \tag{7.13}$$

the electromagnetic tensor, and hence the action, does not change:

$$F_{\mu\nu} \to \partial_\mu(A_\nu + \partial_\nu \alpha) - \partial_\nu(A_\mu + \partial_\mu \alpha) = F_{\mu\nu} + \partial_\mu \partial_\nu \alpha - \partial_\nu \partial_\mu \alpha = F_{\mu\nu}. \tag{7.14}$$

This leads to an infinite number of symmetries, one for each field configuration $\alpha(x)$ you care to dream up.

What do we do about these gauge symmetries? In electrodynamics, we know that the observable quantities are the electric and magnetic fields, not A^μ. If you calculate using A^μ, and I calculate using $A^\mu + \partial^\mu \alpha$, then we should still make the same physical predictions. The gauge symmetries are not true symmetries of our system; rather, they represent a redundancy in our description.

7.1.1   Why Gauge Invariance?

Given this redundancy, why do we bother with A^μ? Many people have attempted to rewrite electromagnetism with reference to only **E** and **B**. These attempts run in to two difficulties:

1. To couple a current J^μ to the electromagnetic field, we require a term $A^\mu J_\mu$ in the Lagrangian. This doesn't violate gauge invariance because, under a gauge transformation

$$\int d^4x \; A^\mu J_\mu \rightarrow \int d^4x \; (A^\mu + \partial^\mu \alpha)\, J_\mu = \int d^4x \; A^\mu J_\mu - \alpha \partial^\mu J_\mu = \int d^4x \; A^\mu J_\mu.$$
(7.15)

   Both integration by parts and the conservation of J^μ play an essential role in this calculation.

2. There exist subtle quantum effects which seem to require A^μ, such as the experimentally observed Aharanov-Bohm effect. These rely on the idea that, if we integrate over a closed curve C, then under a gauge transformation

$$\int_C A^\mu \cdot dx_\mu \rightarrow \int_C (A^\mu + \partial^\mu \alpha) \cdot dx_\mu = \int_C A^\mu \cdot dx_\mu$$
(7.16)

   since $\partial^\mu \alpha$ is a conservative field. We do not have the time to explore this fascinating topic, which is deeply connected to ideas in differential geometry, and has applications in both particle and condensed matter physics.

It looks like we are stuck with the highly redundant A^μ. But so far we discovered gauge invariance by staring at the Maxwell Lagrangian. Is there a deeper way to understand gauge invariance?

Naively, we could study vector fields with Lagrangians similar to the Klein-Gordon field

$$\mathcal{L} = -\frac{1}{2}(\partial^\mu A^\nu)(\partial_\mu A_\nu) + \frac{1}{2}m^2 A^2.$$
(7.17)

Let us do something a little odd, and split A into components: $A^\mu = (A^0, A^i)$. The Lagrangian then becomes

$$\mathcal{L} = \sum_{i=1}^{3} \left(\frac{1}{2}(\partial^\mu A^i)(\partial_\mu A_i) - \frac{1}{2}m^2 A^i A_i \right) - \frac{1}{2}(\partial^\mu A^0)(\partial_\mu A_0) + \frac{1}{2}m^2 A^0 A_0.$$
(7.18)

This looks like the Lagrangian of four non-interacting Klein-Gordon fields. But notice that the A^0 field enters the Lagrangian with the wrong sign! When we compute the Hamiltonian, this will mean that the A^0 field will have negative energy, which would be a disaster.

The problem is due to the fact that in $A^\mu A_\mu$, the time component of A^μ enters with opposite sign to the spatial components. We have to think very carefully about the interactions of vector fields if we want a stable ground state. Although it is beyond our current scope to show it, a consistent theory require us to introduce gauge symmetries to get rid of the pesky A^0 term.[1]

---

[1] There exists a partial exemption for an $m^2 A^2$ term, which can nevertheless be understood in terms of the *Higgs mechanism*, discussed in the next chapter.

The Maxwell field is the simplest possible gauge theory. *Yang-Mills theories* are built on more complicated gauge symmetries, and in Standard Model are used to describe both the strong and the weak force. *General relativity* is also a gauge theory; this time, the gravitational field is not a vector A^μ but the metric tensor $g^{\mu\nu}$. This redundancy explains why, whilst there are 10 independent components in the metric tensor (the metric tensor must be symmetric), gravitational waves have only two polarizations.

So all four of the fundamental forces are built on gauge symmetries. It is hard not to conclude that Nature really loves gauge theories.

7.1.2 Path-Integrals for the Electromagnetic Field

It is possible to canonically quantize the electromagnetic field, just like we quantized the Klein-Gordon and Dirac equations. It is however a tedious subject, because of difficulties introduce by gauge invariance.[2]

We will instead use the path-integral to study the electromagnetic field. Consider the path-integral

$$
\begin{aligned}
Z[J] &= \int \mathcal{D}A \ \exp\left(i \int d^4x \ -\frac{1}{4} F_{\mu\nu}F^{\mu\nu} + J^\mu A_\mu \right) \\
&= \int \mathcal{D}A \ \exp\left(i \int d^4k \ \frac{1}{2} A_\mu(-k) \left(-k^2 g^{\mu\nu} + k^\mu k^\nu \right) A_\nu(k) + A_\mu(-k) J^\mu(k) \right)
\end{aligned}
\tag{7.19}
$$

where in the second line we have done our usual trick of going into Fourier space. The action of the field will vanish for any field configuration $A^\mu(k) = k^\mu \alpha(k)$. Going back to position space, we recognize that this problem is due to gauge invariance: given a field configuration $A^\mu(x)$, the configuration $A^\mu(x) + \partial^\mu \alpha(x)$ is an equally good representation of the field.

Say we managed to choose one gauge configuration A^μ_{phys} for each set of gauge equivalent fields $A^\mu_{\text{phys}} + \partial^\mu \alpha$. We could then rewrite our path-integral as

$$
Z[J] = \int \mathcal{D}\alpha \ \mathcal{D}A_{\text{phys}} \ \exp\left(S[A + \partial\alpha] \right) = \left(\int \mathcal{D}\alpha \right) \left(\int \mathcal{D}A_{\text{phys}} \exp\left(S[A_{\text{phys}}] \right) \right), \tag{7.20}
$$

where we have separated out the gauge transformation from the physically distinct field configurations. The problem, then, comes from the integral over α, which gives an infinite factor we need to somehow get rid of.

To do this we gauge fix — we choose some condition that we want A_{phys} to satisfy, but that $A_{\text{phys}} + \partial\alpha$ will not. You have probably met a few different gauge conditions in electromagnetism:

1. The *Coulomb gauge*, $\nabla \cdot \mathbf{A} = 0$. Using the formula for A^0 we found in terms of $\mathbf{A}$, this condition gives $A^0 = 0$.

2. The *Axial gauge*, $A^3 = 0$. This is very useful for certain calculations in QED, but is much more useful when studying Yang-Mills theory.

3. The *Light-cone gauge*, $A^0 - A^1 = 0$. If you want to study string theory, this is a gauge you will one day become very familiar with.

---

[2]The electromagnetic field is canonical quantized (in two different ways!) in Chapter 6 of Tong's lecture notes.

4. The *Lorenz gauge*, $\partial^\mu A_\mu = 0$. Unlike the previous three gauges, this gauge is Lorentz invariant. For this reason it is usually the most convenient in calculations.

Whatever gauge we use, we can write the gauge condition in the form $G(A) = 0$ for some function G. We then only want to integrate over configurations A_{phys}, where $G(A_{\text{phys}}) = 0$. To do this, we insert a delta function in our path integral:[3]

$$\int \mathcal{D}A_{\text{phys}} \, \exp\left(S[A]\right) = \int \mathcal{D}A \, \delta(G(A)) \exp\left(S[A]\right). \tag{7.21}$$

We now work in a modified version of the Lorenz gauge, with $G(A) = \partial A - \omega$ for some scalar field $\omega(x)$. This gives us

$$Z[J] = \mathcal{N} \int \mathcal{D}A \, \delta(\partial A - \omega) \exp\left(S[A]\right), \tag{7.22}$$

where $\mathcal{N} = \int \mathcal{D}\alpha$ is just a normalization constant. But since the field ω is arbitrary, we can integrate over all possible ω:

$$\begin{aligned}
Z[J] &= \mathcal{N} \int \mathcal{D}\omega \left[\exp\left(-\frac{i}{2\xi}\int d^4x \, \omega^2\right) \int \mathcal{D}A \, \delta(\partial A - \omega) \exp\left(S[A]\right)\right] \\
&= \mathcal{N} \int \mathcal{D}A \left[\exp\left(S[A]\right) \int \mathcal{D}\omega \exp\left(-\frac{i}{2\xi}\int d^4x \, \omega^2\right) \delta(\partial A - \omega)\right] \\
&= \mathcal{N} \int \mathcal{D}A \left[\exp\left(S[A]\right) \exp\left(-\frac{i}{2\xi}\int d^4x \, (\partial^\mu A_\mu)^2\right)\right]. \\
&= \mathcal{N} \int \mathcal{D}A \exp\left(i\int d^4x \, -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}(\partial^\mu A_\mu)^2 + J^\mu A_\mu\right).
\end{aligned} \tag{7.23}$$

Effectively, all we have done is added a new $(\partial A)^2$ term to the Lagrangian — this new Lagrangian is called the *gauge-fixed* Lagrangian. We have the freedom to choose ξ to be whatever we like; in any physical calculation is should cancel out, giving us a check on our calculations.

We still haven't dealt with the normalization factor, but fortunately, we don't have to. Using the results of Chapter 4, we calculate physical observables using

$$\begin{aligned}
\langle\Omega|\mathcal{O}(A)|\Omega\rangle &= \frac{\int \mathcal{D}A \, \mathcal{O}(A) \exp\left(i\int d^4x \, -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}\right)}{\int \mathcal{D}A \, \exp\left(i\int d^4x \, -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}\right)} \\
&= \frac{\int \mathcal{D}A \, \mathcal{O}(A) \exp\left(i\int d^4x \, -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}(\partial^\mu A_\mu)^2\right)}{\int \mathcal{D}A \, \exp\left(i\int d^4x \, -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}(\partial^\mu A_\mu)^2\right)};
\end{aligned} \tag{7.24}$$

the constant $\mathcal{N}$ cancels.

---

[3]We have been even more cavalier with rigour than usual here. Our arguments only work if G is linear in A, but fortunately this is true for all gauges we consider here. More generally, one has to take into account how the delta function changes under a gauge transformation, which is known as *Fadeev-Popov* quantization. You can find the details in both *Peskin & Schroeder* 9.4, and in *Zee* 3.2. While somewhat overkill for electromagnetism, the method is critical when quantizing Yang-Mills and gravity.

Evaluating the Path-Integral

Now that we have gauge fixed, we can turn to the problem of solving the path-integral

$$Z[J] = \int \mathcal{D}A \, \exp\left(i \int d^4x \, -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2\xi} (\partial^\mu A_\mu)^2 + J^\mu A_\mu \right)$$
$$= \int \mathcal{D}A \, \exp\left(i \int d^4k \, \frac{1}{2} A_\mu(-k) \left(-k^2 g^{\mu\nu} + \left(1 - \frac{1}{\xi}\right) k^\mu k^\nu \right) A_\nu(k) + A_\mu(-k) J^\mu(k) \right).$$
$$(7.25)$$

We then make the change of variables

$$B_\mu(k) = A_\mu(k) + \frac{J^\nu(k)}{k^2} \left(g_{\mu\nu} - (1 - \xi) \frac{k_\mu k_\nu}{k^2} \right) \tag{7.26}$$

Since this is just a constant shift of the integration variable, the integration measure does not change. Upon changing the variable, we now find that

$$Z[J] = \int \mathcal{D}A \, \exp\left(i \int d^4k \, \frac{1}{2} A_\mu(-k) \left(-k^2 g^{\mu\nu} + \left(1 - \frac{1}{\xi}\right) k^\mu k^\nu \right) A_\nu(k) \right)$$
$$\times \exp\left(\frac{i}{2} \int d^4k \, \frac{J^\mu(-k) J^\nu(k)}{k^2} \left(g_{\mu\nu} - (1 - \xi) \frac{k_\mu k_\nu}{k^2} \right) \right) \tag{7.27}$$
$$= Z[0] \exp\left(\frac{i}{2} \int d^4k \, \frac{J^\mu(-k) J^\nu(k)}{k^2} \left(g_{\mu\nu} - (1 - \xi) \frac{k_\mu k_\nu}{k^2} \right) \right).$$

Just like for the Klein-Gordon equation, the path integral gives a factor $Z[0]$ which is irrelevant in an physical computation. In position space, we can write

$$Z[J] = Z[0] \exp\left(-\frac{1}{2} \int d^4x \, d^4y \, J_\mu(x) D_F^{\mu\nu}(x - y) J_\nu(y) \right) \tag{7.28}$$

where we define

$$D_F^{\mu\nu}(x - y) = \int d^4k \frac{-i e^{-ik(x-y)}}{k^2 + i\varepsilon} \left(g^{\mu\nu} - (1 - \xi) \frac{k^\mu k^\nu}{k^2} \right). \tag{7.29}$$

This is the photon propagator — like the Klein-Gordon propagator, it will play a central role in the Feynman rules. We have introduced an infinitesimally positive $i\varepsilon$ in the denominator to make the integral well-defined.

As we noted before, ξ is an arbitrary number, and should cancel in any computation of a physical quantity. The most common gauges are the *Landau gauge*, where $\xi = 0$, and the *Feynman gauge*, where $\xi = 1$. Other gauges do occasionally prove useful; for instance, the *Yennie gauge* $\xi = 3$ has found applications in some bound state computations.

Photon Correlation Functions

Correlation functions work the same way for electrodynamics as they do for the Klein-Gordon field; we simply differentiate $Z[J]$ with respect to J. We then find that

$$\langle \Omega | A^\mu(x) A^\nu(y) | \Omega \rangle = \frac{1}{Z[0]} \left(-i \frac{\delta}{\delta J_\mu(x)} \right) \left(-i \frac{\delta}{\delta J_\nu(y)} \right) Z[J] = D_F^{\mu\nu}(x - y). \tag{7.30}$$

Because our expression for the propagator depends on ξ, we immediately recognize that the right-hand side is not gauge invariant; this shouldn't be too surprising since the left-hand side isn't either.

Higher correlation functions follow all the same rules as before — we now just need to use the photon propagator, keeping track of the indices of A^μ.

7.2 Coupling Matter to Light

7.2.1 Fermions

Now that we have quantized light, we need to couple it to matter. The simplest such theory is *quantum electrodynamics* (QED), describing the interaction of photons with charged fermions. The Lagrangian for the theory is

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \overline{\psi}(i\slashed{\partial} - m)\psi - eA_\mu\overline{\psi}\gamma^\mu\psi, \tag{7.31}$$

where e is a dimensionless coupling constant. Most commonly, this theory is used to describe the interaction of electrons and positrons. If we wanted to include heavier particles — such as muons, quarks, or protons — we can simply add in extra fermionic fields, and substitute $e \to Qe$ where Q is the charge of the particle. This simple Lagrangian can account for the behaviour of electrons, light, electric and magnetic fields, chemistry, and all of the complex systems that are built on these.

Remember that the Dirac field is symmetric under phase rotation $\psi \to e^{i\alpha}\psi$, and this gives rise to the conserved current

$$J^\mu = \overline{\psi}\gamma^\mu\psi. \tag{7.32}$$

Our photon field hence couples to fermions by coupling to the conserved current J^μ. We therefore equate the conserved charge with the electric charge of the fermion (up to a constant factor e):

$$Q = \int d^3x \; eJ^0 = \int \frac{d^3p}{(2\pi)^3} \sum_s e\left(a_\mathbf{p}^{s\,\dagger}a_\mathbf{p}^s - b_\mathbf{p}^{s\,\dagger}b_\mathbf{p}^s\right). \tag{7.33}$$

Fermions will have charge $+e$, and antifermions have charge $-e$. For this reason, a theory of electrons and light must also contain positrons.

Gauge invariance was a crucial feature of electrodynamics, and for the theory to be consistent, gauge invariance must hold even in interacting theories. Is QED still gauge invariant? The answer is yes, under the gauge transformations

$$A_\mu \to A_\mu + \partial_\mu\lambda, \quad \psi \to e^{-ie\lambda}\psi. \tag{7.34}$$

To see this, we note that the F^2 term and the $m\overline{\psi}\psi$ term are both manifestly gauge invariant. We only need to worry about the two terms

$$\begin{aligned}
\overline{\psi}i\slashed{\partial}\psi - eA_\mu\overline{\psi}\gamma^\mu\psi &\to \overline{\psi}e^{ie\lambda}\gamma^\mu\left(i\partial_\mu - eA_\mu - e\partial_\mu\lambda\right)\left(e^{-ie\lambda}\psi\right) \\
&= \overline{\psi}\gamma^\mu\left(i\partial_\mu + e\partial_\mu\lambda - eA_\mu - e\partial_\mu\lambda\right)\psi \\
&= \overline{\psi}\gamma^\mu\left(i\partial_\mu - eA_\mu\right)\psi
\end{aligned} \tag{7.35}$$

So while neither of the above terms is by itself gauge invariant, the combination

$$\overline{\psi}(i\slashed{\partial} - e\slashed{A})\psi \tag{7.36}$$

is. Because of this, we define the *gauge covariant derivative*

$$D_\mu = \partial_\mu + ieA_\mu, \tag{7.37}$$

so that $D\psi$ transforms like the Dirac spinor under a gauge transformation:

$$D_\mu(e^{-ie\lambda}\psi) \to e^{-ie\lambda}D_\mu\psi. \tag{7.38}$$

With this piece of notation we rewrite the QED Lagrangian:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \overline{\psi}(i\slashed{D} - m)\psi. \tag{7.39}$$

Using the Euler-Lagrange equations, we find that the equations of motion for QED are

$$(i\slashed{D} - m)\psi = 0, \quad \partial_\mu F^{\mu\nu} = e\overline{\psi}\gamma^\nu\psi. \tag{7.40}$$

7.2.2   Scalars

The electomagnetic field must couple to a conserved current. For the real scalar field, no such current exists and so real fields cannot carry electric charge. For the complex field Φ, we have a phase rotation symmetry $\Phi \to e^{i\alpha}\Phi$ akin to the symmetry of the Dirac field. Mimicking the fermionic theory, postulate that under a local gauge transformation,

$$\Phi \to e^{-ie\lambda}\Phi. \tag{7.41}$$

The normal derivative is no longer gauge covariant, but the covariant derivative is. Hence the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + D^\mu\Phi^*D_\mu\Phi - m^2\Phi^*\Phi \tag{7.42}$$

is gauge invariant. This theory is known as *scalar QED*, and describes the interaction of light with the particles such as the Higgs boson and the pion.
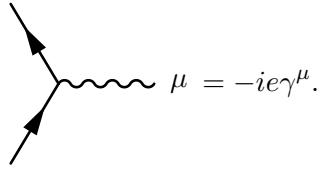
7.3   Feynman Rules for QED

It is not hard to guess the Feynman rules for QED; they are almost identical similar to the rules for Yukawa theory. To the electron field we associated the propagator

$$\xrightarrow[p]{} \quad = \frac{i(\slashed{p} + m)}{p^2 - m^2}. \tag{7.43}$$

and to the photon

$$\mu \,\sim\!\!\sim\!\!\sim\!\!\sim\, \nu = \frac{-i}{p^2}\left(g^{\mu\nu} - (1-\xi)\frac{p^\mu p^\nu}{p^2}\right) \tag{7.44}$$

The $eA_\mu\overline{\psi}\gamma^\mu\psi$ interaction term gives rise to the vertex
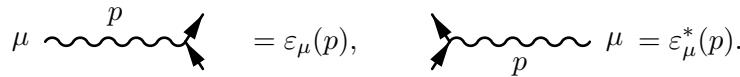

$$\mu \; = -ie\gamma^\mu. \tag{7.45}$$

As usual, momentum is conserved at each vertex, and all undetermined momenta are integrated over. Just like Yukawa theory, we will not have any symmetry factors, because the vertex itself has no symmetries, but we do have fermionic minus signs to deal with.

7.3.1   Polarization of Light

For the Dirac field, we have to label the spin state of an external leg, associated with an additional factor of u or v. Likewise, photons have two polarization states, so on external legs we associated polarization vectors:

$$\mu \underset{p}{\wwwww}\; = \varepsilon_\mu(p), \qquad \underset{p}{\wwwww}\; \mu \; = \varepsilon_\mu^*(p). \tag{7.46}$$

You have probably come across polarization before, when discussing light in a classical electrodynamics course. As you will recall, we describe a monochromatic, polarized wave as

$$\mathbf{E}(\mathbf{x},t) = e^{i(|\mathbf{k}|t - \mathbf{k}\cdot\mathbf{x})}\mathbf{n}, \quad \mathbf{B}(\mathbf{x},t) = e^{i(|\mathbf{k}|t - \mathbf{k}\cdot\mathbf{x})}\mathbf{k} \times \mathbf{n} \tag{7.47}$$

where $\mathbf{k}$ is the 3-momentum of the laser, and $\mathbf{n}$ is the polarization vector. In order to satisfy Maxwell's equation, $\mathbf{n}\cdot\mathbf{k} = 0$, and so there are two possible polarizations of a laser.

In a relativistically covariant description, we describe the momentum by the 4-vector k^μ and the polarization vector by n^μ. We should also work with the electromagnetic potential

$$A^\mu(x) = e^{ikx}n^\mu, \tag{7.48}$$

rather $\mathbf{E}$ and $\mathbf{B}$. Our condition on the polarization vector becomes $k_\mu n^\mu = 0$. There is a problem though: there are three linearly independent vectors which satisfy $k_\mu n^\mu = 0$. But we know there are only two polarizations, so how can we get rid of one?

Gauge invariance comes to the rescue! Under the gauge transformation $\lambda(x) = ce^{ikx}$, the electromagnetic potential transforms as

$$A^\mu(x) = e^{ikx}n^\mu \to A^\mu(x) + c\partial^\mu\lambda(x) = e^{ikx}(n^\mu + ck^\mu). \tag{7.49}$$

Two polarizations are hence physically equivalent if they differ by a multiple of k^μ. For this reason, there are only two physical polarizations.

If you think about it, this puts a strong condition on Feynman diagrams in QED. Any process involving a photon with polarization $\varepsilon^\mu(k)$ will have amplitude $\mathcal{M}(k) = \varepsilon_\mu(k)\mathcal{M}^\mu(k)$. But under a gauge transformation

$$\varepsilon_\mu(k) \to \varepsilon_\mu(k) + ck_\mu \implies \mathcal{M}(k) \to \varepsilon_\mu(k)\mathcal{M}^\mu(k) + ck_\mu\mathcal{M}^\mu(k). \tag{7.50}$$

But any physical process must be gauge invariant, and so we deduce that

$$k_\mu \mathcal{M}^\mu(k) = 0. \tag{7.51}$$

This equation is known as the *Ward-Takahashi identity.* It can be proved either diagrammatically, by proving that it holds for all Feynman diagrams of a given order, or formally from the path-integral, using symmetry arguments. We will not delve further into this topic here; the interested reader should look it up in their favourite QFT text.

7.3.2   Scalar QED

In order to deduce the Feynman rules for scalar QED, let's expand

$$
\begin{aligned}
D^\mu \Phi^* D_\mu \Phi &= [(\partial^\mu - ieA^\mu)\Phi^*]\,[(\partial_\mu + ieA_\mu)\Phi] \\
&= \partial^\mu \Phi^* \partial_\mu \Phi + ieA_\mu(\Phi^* \partial^\mu \Phi - \Phi \partial^\mu \Phi^*) + e^2 A^2 |\Phi|^2.
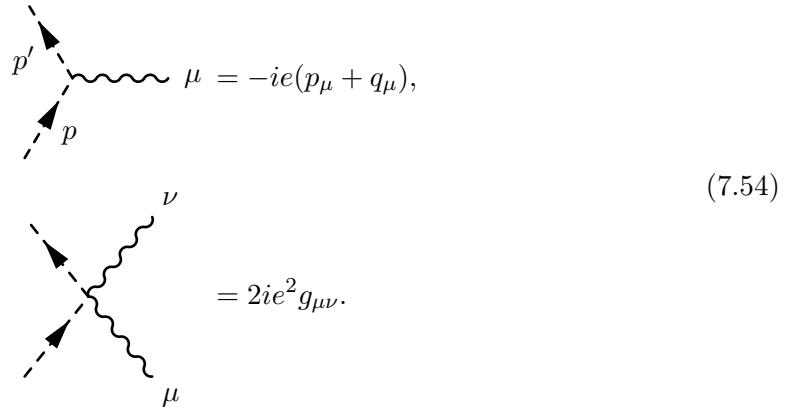\end{aligned}
\tag{7.52}
$$

We now have two interaction terms, each of which have an associated vertex. Furthermore, one of these interactions includes a derivative term.

It is not hard to derive the propagator for a complex field using the results of Section 3.5. The result,

$$
\begin{array}{c}
\text{-- -- -- }\blacktriangleright\text{-- --}\\[-4pt]
\hphantom{xxxx}p
\end{array}
\quad = \frac{i}{p^2 - m^2}, \tag{7.53}
$$

is in fact identical to the propagator for a real field — which is less surprising when we remember that a complex field is secretly just two real fields $\Phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$.

We can now write the two vertices for scalar QED:

$$
\begin{aligned}
&\mu \;= -ie(p_\mu + q_\mu), \\[40pt]
&\quad\;\; = 2ie^2 g_{\mu\nu}.
\end{aligned}
\tag{7.54}
$$

The first vertex depends on the momentum of the scalar particle — our derivative term results in a momentum dependent coupling! Since the second diagram has two identical photons in it, we need to include an additional factor of 2.

7.4   Electrons and Muons

Having spent the last few chapters developing formalism, we can finally apply our tools to realistic calculations. The first processes we shall study involve the interaction of electrons and muons. Muons are just heavier versions of the electron, with a mass
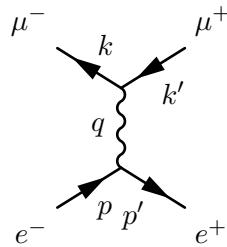
$$m_\mu = 105.6 \text{ MeV} \approx 200 m_e. \tag{7.55}$$

The muon is the second lightest charge particle, and so at low energies is the second most important fermion, after the electron. With only minor modifications, we can also describe processes involving the tau[4], quarks, or protons.

7.4.1   Electron-Positron Annihilation to Muons

The process $e^+e^- \to \mu^+\mu^-$ is amongst both the simplest and also most important process in high-energy physics. Electron-positron colliders are common, and this is partially because e^+e^- collisions are theoretically easy to understand.[5] The $e^+e^- \to \mu^-\mu^-$ process is in fact used to calibrate these machines.

Only one diagram contributes to $e^+e^- \to \mu^+\mu^-$ at tree-level, from which we can immediately write down the amplitude:

$$= \overline{v}_{s'}(p')(-ie\gamma^\mu)u_s(p)\left(\frac{-ig_{\mu\nu}}{q^2}\right)\overline{u}_r(k)\,(-ie\gamma^\nu)\,v_{r'}(k'). \qquad (7.56)$$

Here we work in the Feynman gauge $\xi = 1$. Leaving the spin subscripts implicit and simplifying slightly, we have a relatively simple expression for the amplitude:

$$\mathcal{M} = \frac{ie^2}{q^2}\left(\overline{v}(p')\gamma^\mu u(p)\right)\left(\overline{u}(k')\gamma_\mu v(k)\right). \qquad (7.57)$$

To turn this into a cross-section, we need to calculate the conjugate of $\mathcal{M}$. This is easy to do using the formula

$$(\overline{v}\gamma^\mu u)^* = u^\dagger(\gamma^\mu)^\dagger(\gamma^0)^\dagger v = u^\dagger\gamma^0\gamma^\mu v = \overline{u}\gamma^\mu v. \qquad (7.58)$$

Our squared matrix element is therefore

$$|\mathcal{M}|^2 = \frac{e^4}{q^4}\left[\overline{v}(p')\gamma^\mu u(p)\right]\left[\overline{u}(k')\gamma_\mu v(k)\right]\left[\overline{v}(k)\gamma_\nu u(k')\right]\left[\overline{u}(p)\gamma^\nu v(p')\right]. \qquad (7.59)$$

The bracketing is to emphasise that each of the terms in brackets is just a number for each μ or ν; we can rearrange these in any order. We can now input any particular spinors and momenta to calculate the scattering amplitude. For our purposes however, the simplest and most interesting quantity is the unpolarized cross-section. Quite commonly, experiments are blind to polarization, so this is an important quantity to calculate. Polarized beams of electron and positrons are hard to produce, and muon sensors are often blind to polarization.

---

[4]An even heavier version of the electron, with a mass of 1.78 GeV.

[5]The most powerful particle accelerator ever built is the LHC, which collides protons together. There are plans to build an electron-positron collider of similar power, but which will allow mucg better precision measurements.

To calculate the unpolarized cross-section, we simply sum over the different spin states:

$$|\mathcal{M}_{\text{UP}}|^2 = \frac{1}{4} \sum_{s,s',p,p'} |\mathcal{M}(s,s',p,p')|^2. \tag{7.60}$$

This can be evaluated using the spin-sums from Section 6.3.1:

$$\sum_s u_s(p)\overline{u}_s(p) = \not{p} + m, \quad \sum_s v_s(p)\overline{v}_s(p) = \not{p} - m. \tag{7.61}$$

If we apply the first formula to the expression

$$\sum_{ss'} \left[\overline{v}_{s'}(p')\gamma^\mu u_s(p)\right] \left[\overline{u}_s(p)\gamma^\nu v_{s'}(p')\right] = \sum_{s'} \overline{v}_{s'}(p')\gamma^\mu(\not{p} + m_e)\gamma^\nu v_{s'}(p')$$
$$= \text{tr}[(\not{p}' - m_e)\gamma^\mu(\not{p} + m_e)\gamma^\nu]. \tag{7.62}$$

A similar trick works for the the muons terms, and so we find

$$|\mathcal{M}_{\text{UP}}|^2 = \frac{e^4}{4q^4}\text{tr}[(\not{p}' - m_e)\gamma^\mu(\not{p} + m_e)\gamma^\nu]\text{tr}[(\not{k}' - m_\mu)\gamma^\mu(\not{k} + m_\mu)\gamma^\nu]. \tag{7.63}$$

We have successfully eliminated the spinors from our cross-section, and now just need to evaluate the two traces.

Evaluating Traces

Traces of gamma matrices can be easily evaluated by appealing to their algebraic properties. When there are no gamma matrices, then

$$\text{tr}(\mathbf{1}) = 4, \tag{7.64}$$

since we just sum the four 1's on the diagonal. The case of a single gamma matrix is not much harder:

$$\text{tr}(\gamma^0) = \text{tr}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = 0, \quad \text{tr}(\gamma^i) = \text{tr}\begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} = 0. \tag{7.65}$$

Next we evaluate the product of two gamma matrices:

$$\text{tr}(\gamma^\mu\gamma^\nu) = \text{tr}(2g^{\mu\nu} - \gamma^\nu\gamma^\mu) = 8g^{\mu\nu} - \text{tr}(\gamma^\mu\gamma^\nu) \tag{7.66}$$

and so we conclude

$$\text{tr}(\gamma^\mu\gamma^\nu) = 4g^{\mu\nu}. \tag{7.67}$$

Higher products of gamma matrices can be evaluated in a similar manner, repeatedly using the anticommutation relation and the cyclic nature of the trace. In general, the trace of an odd number of gamma matrices always vanishes. For even matrices we get an increasingly complicated series of metric tensors. For instance,

$$\text{tr}(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma) = 4\left(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma} + g^{\mu\sigma}g^{\nu\rho}\right). \tag{7.68}$$

The Unpolarized Cross-Section

We can now finish the job of evaluating the cross-section. Using our trace formulae, we find

$$\text{tr}[(\not{p}' - m_e)\gamma^\mu(\not{p} + m_e)\gamma^\nu] = 4[p'^\mu p^\nu + p'^\nu p^\mu - g^{\mu\nu}(p \cdot p' + m_e^2)]$$
$$\text{tr}[(\not{k} + m_e)\gamma_\mu(\not{k}' - m_\mu)\gamma_\nu] = 4[k'_\mu k_\nu + k'_\nu k_\mu - g_{\mu\nu}(k \cdot k' + m_\mu^2)].$$

(7.69)

From now on we shall neglect the electron mass, setting $m_e \to 0$. The error caused by this approximation is of order $m_e/m_\mu \approx 1/200$, which is smaller than the loop contributions to this process.

Under this approximation, we now find that

$$|\mathcal{M}_{\text{UP}}|^2 = \frac{8e^4}{q^4} \left[(p \cdot k)(p' \cdot k') + (p \cdot k')(p' \cdot k) + m_\mu^2(p \cdot p') \right].$$

(7.70)

To understand the implications of this formula further, let's go into the centre of mass frame. We will take the electron-positron beam to be in the x-direction, and will assume that the muons scatter in the x-y plane. Our momenta our now given by

$$p = (E, E, 0, 0), \qquad\qquad p' = (E, E, 0, 0)$$
$$k = (E, |\mathbf{k}| \cos\theta, |\mathbf{k}| \sin\theta, 0), \qquad k = (E, -|\mathbf{k}| \cos\theta, -|\mathbf{k}| \sin\theta, 0)$$

(7.71)

where $|\mathbf{k}| = \sqrt{E^2 - m_\mu^2}$, and where $q^2 = (p + p')^2 = 4E^2$. In this reference frame, our amplitude becomes

$$|\mathcal{M}_{\text{UP}}|^2 = \frac{8e^4}{16E^4} \left[E^2(E - |\mathbf{k}| \cos\theta)^2 + E^2(E + |\mathbf{k}| \cos\theta)^2 + m_\mu^2 E^2 \right]$$
$$= e^4 \left[\left(1 + \frac{m_\mu^2}{E^2} \right) + \left(1 - \frac{m_\mu^2}{E^2} \right) \cos^2\theta \right].$$

(7.72)

From here, calculating the cross-section is just an application of a few formulas from Section 5.3. Using these, the differential cross-section can be found to be

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{16E^2} \sqrt{1 - \frac{m_\mu^2}{E^2}} \left[\left(1 + \frac{m_\mu^2}{E^2} \right) + \left(1 - \frac{m_\mu^2}{E^2} \right) \cos^2\theta \right],$$

(7.73)

where α is the *fine-structure constant*:

$$\alpha = \frac{e^2}{4\pi} \approx \frac{1}{137.04}.$$

(7.74)

We then just integrate over $d\Omega$ to calculate the total cross-section

$$\sigma = \frac{\pi\alpha^2}{3E^2} \sqrt{1 - \frac{m_\mu^2}{E^2}} \left(1 + \frac{1}{2} \frac{m_\mu^2}{E^2} \right).$$

(7.75)

Often we are interested in the limit where both the electrons and muons are moving relativistically. If, for instance, our machine is to produce Higgs bosons, we will need the centre of mass energy to be at least $m_H \approx 125$ GeV, which is over a thousand times
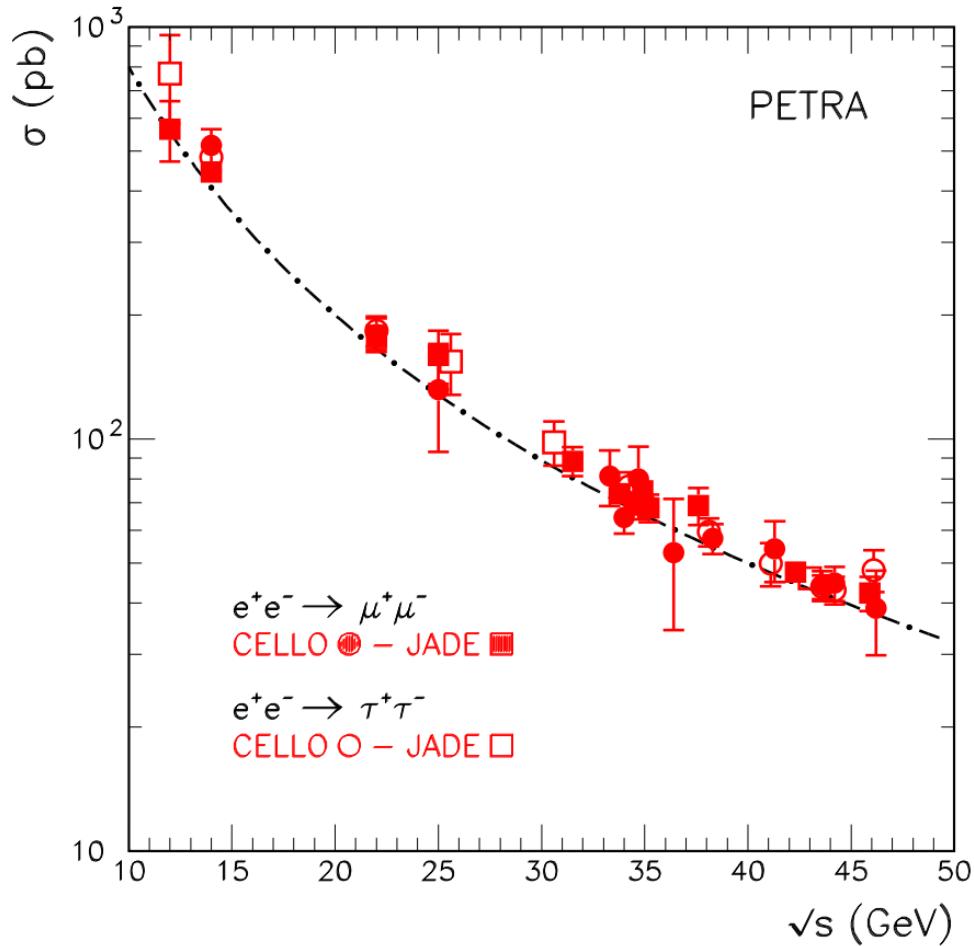
Figure 7.1: Cross section for e^+e^- to produce $\mu^+\mu^-$ and $\tau^+\tau^-$. The experimental results are from PETRA. The x-axis is the centre of mass energy $E_{\mathrm{cm}} = 2E$, and the dashed curve is the theoretical prediction given by (7.76). The source of the image is *arXiv:0906.1271v4*.

heavier than the muon! Taking $E \gg m_\mu$, our expressions simplify to

$$\frac{d\sigma}{d\Omega} \xrightarrow[E \gg m_\mu]{} \frac{\alpha^2}{16E^2}(1 + \cos\theta), \quad \sigma \xrightarrow[E \gg m_\mu]{} \frac{\pi\alpha^2}{3E^2}. \tag{7.76}$$

In *Figure 7.1*, the experimental cross section for $e^+e^- \to \mu^+\mu^-$ is plotted against the centre-of-mass energy $E_{\mathrm{cm}} = 2E$. Our theoretical cross section matches the data very well.

Tau-Antitau Pairs

Having calculated the cross-section of $e^+e^- \to \mu^+\mu^-$, let's apply it to other particles. The simplest case is the τ, which behaves just like a muon, but is heavier. All we have to do is replace the muon mass with the tau mass. In the ultrarelativistic limit, we discovered that the cross-section was independent of the muon mass. From this, we conclude that

$$\frac{\sigma(e^+e^- \to \tau^+\tau^-)}{\sigma(e^+e^- \to \mu^+\mu^-)} \xrightarrow[E \gg m_\tau]{} 1. \tag{7.77}$$

At really high energies, QED cannot tell the difference between the muon and the tau!

In *Figure 7.1*, cross sections for both $e^+e^- \to \mu^+\mu^-$ and $e^+e^- \to \tau^+\tau^-$ are plotted for energies much greater than m_τ. From this graph, it is clear that at high energies, the $\mu^+\mu^-$ and $\tau^+\tau^-$ pairs are equally likely to be produced.

Quark-Antiquark Pairs

Next we can try applying our formula to quark-antiquark pairs, $e^+e^- \to \bar{q}q$. The only difference between this case and that of the tau is that quarks also interact with each other via the strong interaction. This interaction is described by a theory called *Quantum Chromodynamics* (QCD), where each quark comes in three different 'colours'. We just need to modify our cross-section by a few factors. First, we need to replace the muon charge with the quark charge. For the u, c and t quarks the charge is $+2/3$, while for the d, s and b quarks have charge $-1/3$. Next we have to include a factor of three, to take into account the three different colours of each quark. And finally, we need to take into account the effect of the strong force.

As you may have heard, the strong force is strong. So strong in fact, that we never see quarks and antiquarks by themselves, but only in bound states called *hadrons* — such as the proton, the neutron and the pion. You might then think, rather reasonably, that calculating the cross-section of $e^+e^- \to \bar{q}q$ is a complete joke; we will never ever see the quarks. This is true, but it turns out that in the limit of high energy, QCD actually becomes less and less strong, and that its effects can be ignored at these energies. We will, of course, still not see any quarks. But what we will see are hadrons being produced, and the cross-section for producing these hadrons is just the cross-section for producing quarks. Hence,

$$R = \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} \xrightarrow[E\to\infty]{} 3\sum_i Q_i^2, \tag{7.78}$$

where we sum over all of the quarks with mass below E (any quark with mass greater than E is to heavy to be produced in the collision).

Quantum chromodynamical effects modify the value of R. These effects are most important when $E \approx m_q$ for some quark. Bound states of quarks, such as the J/ψ, will result in sharp spikes in the cross-section.

The experimental value of R is plotted in *Figure 7.2*, for centre of mass energies between 2.5 and 40 GeV. There are three distinct regions, depending on which quarks contribute to the sum in (7.78). At low energies, the up, down and strange quarks are the only quarks light enough to be produced, and so

$$R(E_{\text{cm}} < 2m_c) = 3\left(Q_u^2 + Q_d^2 + Q_s^2\right) = 3\left(\left(\frac{2}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 + \left(-\frac{1}{3}\right)^2\right) = 2. \tag{7.79}$$

At intermediate energies, the charm quark can also be produced, and so

$$R(2m_c < E_{\text{cm}} < 2m_b) = R(E_{\text{cm}} < 2m_c) + 3Q_c^2 = 2 + 3\left(\frac{2}{3}\right)^2 = \frac{10}{3}. \tag{7.80}$$

At high energies we can also produce the bottom quark, and so

$$R(2m_c < E_{\text{cm}} < 2m_b) = R(2m_c < E_{\text{cm}} < 2m_b) + 3Q_b^2 = \frac{10}{3} + 3\left(-\frac{1}{3}\right)^2 = \frac{11}{3}. \tag{7.81}$$
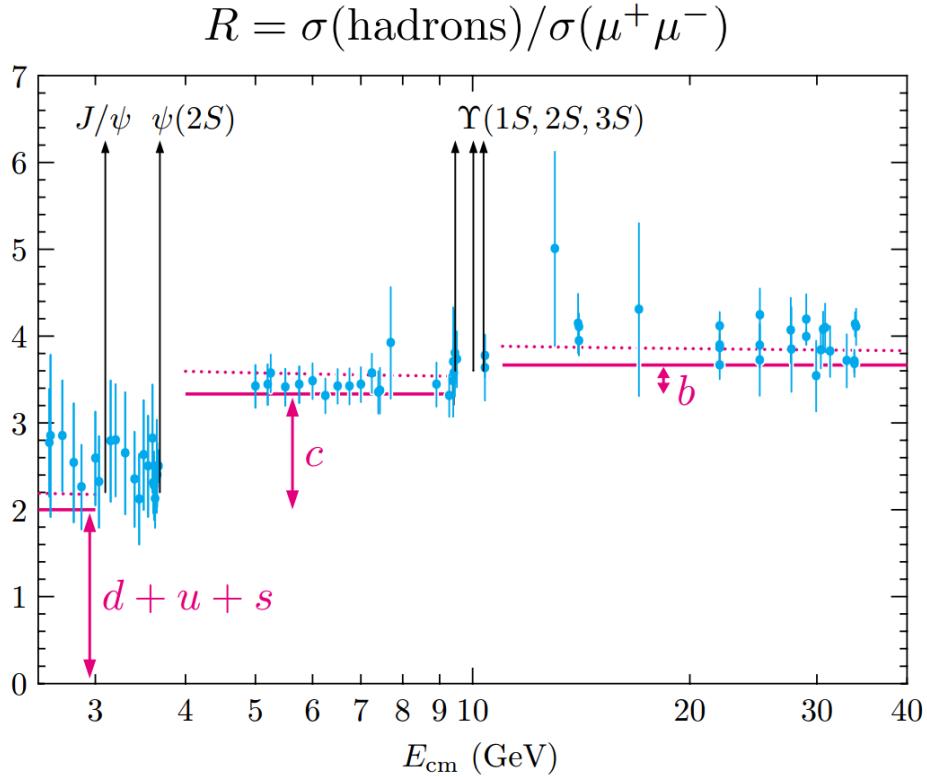
Figure 7.2: The ratio of cross-sections for electron-positron annihilation into hadrons and electron-positron annihilation into muons. The solid line is prediction (7.78), whereas the dotted line is a more sophisticated prediction which includes QCD corrections to R.

7.4.2   Electron-Muon Scattering

Now consider the scattering of an electron and a muon: $e^-\mu^- \to e^-\mu^-$. The Feynman diagram for this process is

$$= \frac{ie^2}{q^2}\overline{u}(p_1')\gamma^\mu u(p_1)\overline{u}(p_2')\gamma_\mu u(p_2). \tag{7.82}$$

This Feynman diagram looks familiar — it is in fact just the diagram for $e^+e^- \to \mu^+\mu^-$ put on its side! Computing the unpolarised amplitude for this process, we find

$$\mathcal{M}_{\mathrm{UP}} = \frac{e^4}{4q^4}\mathrm{tr}\left[(\not{p}_1' + m_e)\gamma^\mu \not{p}_1 + m_e)\gamma^\nu\right]\left[(\not{p}_2' + m_\mu)\gamma^\mu \not{p}_2 + m_\mu)\gamma^\nu\right]. \tag{7.83}$$

This amplitude is not just a suspiciously similar to the $e^+e^- \to \mu^+\mu^-$ result (7.69) — it is identical, under the substitutions

$$p \to p_1, \quad p' \to -p_1', \quad k \to p_2', \quad k' \to -p_2. \tag{7.84}$$

We can hence use our previous results to calculate

$$\mathcal{M}_{\text{UP}} = \frac{8e^4}{q^4} \left[(p_1 \cdot p_2')(p_1' \cdot p_2) + (p_1 \cdot p_2)(p_1' \cdot p_2' - m_\mu^2 p_1 \cdot p_1') \right] \tag{7.85}$$

where we have once again set $m_e = 0$. As before, we now move into the centre of mass frame to evaluate this expression:

$$\begin{aligned} p_1 &= (k, k, 0, 0), & p_2 &= (E, -k, 0, 0) \\ p_1' &= (k, k\cos\theta, k\sin\theta, 0), & p_2' &= (E, -k\cos\theta, -k\sin\theta, 0) \end{aligned} \tag{7.86}$$

where $E^2 = k^2 + m_\mu^2$. In terms of E, k and θ our amplidue becomes

$$\mathcal{M}_{\text{UP}} = \frac{2e^4}{k^2(1 - \cos\theta)^2} \left[(E + k)^2 + (E + k\cos\theta)^2 - m_\mu^2(1 - \cos\theta) \right]. \tag{7.87}$$

From this we can evaluate the cross-section and then compare to previous results. The answer is a little messy; in the high-energy limit where $E \gg m_\mu$, the differential cross-section simplifies to

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{8E^2(1 - \cos\theta)^2} \left(4 + (1 + \cos\theta)^2 \right). \tag{7.88}$$

The Inverse-Square Law

For the fourth and final time, let's discuss the inverse square law. We are now in a position to show that Coulomb's law can be recovered from QED.

Let's consider the non-relativistic limit of electron-muon scattering. We will work in the centre of mass frame, and assume that the energy of each particle is much less than the electron mass. In this limit,

$$\begin{aligned} \mathcal{M} &= \frac{ie^2}{q^2} \overline{u}(p_1')\gamma^\mu u(p_1)\overline{u}(p_2')\gamma_\mu \mu(p_2) \\ &\to \frac{ie^2}{q^2} \overline{u}_s(\mathbf{0})\gamma^\mu u_{s'}(\mathbf{0})\overline{u}_r(\mathbf{0})\gamma_\mu u_{r'}(\mathbf{0}) \\ &= \frac{ie^2}{q^2}(+2m_e\delta_{ss'})(+2m_\mu\delta_{rr'}) \end{aligned} \tag{7.89}$$

where we have used (6.61) to simplify the last line. In the non-relativistic limit,

$$q^2 = (p_1' - p_1)^2 = -(\mathbf{p}_1' - \mathbf{p}_1)^2 \tag{7.90}$$

and so our amplitude becomes

$$\mathcal{M} = \frac{-ie^2}{(\mathbf{p}_1' - \mathbf{p}_1)^2} 4m_e m_\mu \delta_{ss'}\delta_{rr'}. \tag{7.91}$$
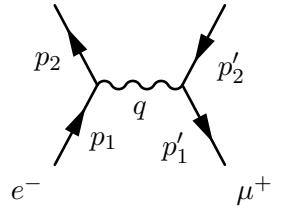
This is almost identical to (6.132) — the only differences are that the photon is massless, and that there is an additional minus sign. Using the results of Section 5.4.2, we conclude

that our scattering is equivalent to the potential

$$U(\mathbf{x}) = + \left.\frac{e^2}{4\pi|\mathbf{x}|}e^{-M|\mathbf{x}|}\right|_{M=0} = \frac{\alpha}{|\mathbf{x}|}. \tag{7.92}$$

This gives a repulsive inverse square force, just as Coulomb's law predicts.

What if we scattered an electron off of an antimuon, $e^-\mu^+ \to e^-\mu^+$? The amplitude for this process is

$$= -\frac{ie^2}{q^2}\overline{u}(p_1')\gamma^\mu u(p_1)\overline{v}(p_2')\gamma_\mu v(p_2). \tag{7.93}$$

Notice that we have to include a global minus sign for this amplitude, just like we had for the Yukawa theory. Taking the non-relativistic limit of this process

$$\begin{aligned}
\mathcal{M} &\to -\frac{ie^2}{q^2}\overline{u}(\mathbf{0})\gamma^\mu u(\mathbf{0})\overline{v}(\mathbf{0})\gamma_\mu v(\mathbf{0}) \\
&= -\frac{ie^2}{q^2}(+2m_e\delta_{ss'})(+2m_\mu\delta_{rr'}).
\end{aligned} \tag{7.94}$$

where we have used (6.61) to simplify our expression. Note the all critical plus sign in the term $\overline{v}\gamma^\mu v$. From this we conclude

$$\mathcal{M}(e^-\mu^+ \to e^-\mu^+) = -\mathcal{M}(e^-\mu^- \to e^-\mu^-) \tag{7.95}$$

and so we have the potential

$$U(\mathbf{x}) = -\frac{\alpha}{|\mathbf{x}|}. \tag{7.96}$$
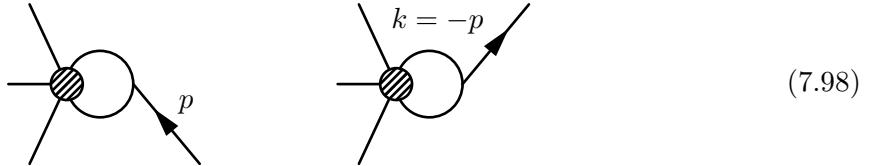
Opposite charges attract, just as you would expect.

7.4.3 Crossing Symmetry

We have noticed a close relationship between electron-muon scattering and the process $e^+e^- \to \mu^+\mu^-$. Using the amplitude of the latter process, we quickly calculate the amplitude in the former process. This trick is known as *crossing symmetry*, and is a general property of all quantum field theory. It states that the amplitude of a process involving an incoming particles ϕ of momentum p is equal to the amplitude of the same process, except where there is an outgoing antparticle $\overline{\phi}$ of momentum $k = -p$:

$$\mathcal{M}(\phi(p) + ... \to ...) = \mathcal{M}(... \to ... + \overline{\phi}(k)). \tag{7.97}$$

This relation can be derived directly from Feynman diagrams. Compare the pair of diagrams

$$(7.98)$$

If ϕ is a scalar, these Feynman diagrams will be equal, and so the crossing symmetry is manifest. We hence see there is a natural correspondence between diagrams with an incoming $\phi(p)$ and an outgoing $\overline{\phi}(k)$.

If instead we consider a fermion, the analysis is a little more complicated because we have to include the external leg contributions u and v. Nevertheless, crossing symmetry continues to hold in this case.

Scattering in QED

Crossing symmetry greatly simplifies the number of amplitudes we need to calculate in QED. Considering only $2 \to 2$ scattering process involving photons, electrons, positrons, muons and antimuons, you may naively think there are going to be hundreds of process to consider.

But QED has many symmetries. It is not hard to check that not only do we have charge conservation, but that the electron number N_e (defined as the number of electron minus the number of positrons) and the muon number N_μ are independently conserved. For this reason, the process $e^- e^- \to \mu^- \mu^-$ is forbidden, even though it does not violate charge conservation.

Crossing symmetry simplifies the number of amplitudes we need even further. All $2 \to 2$ scattering process in QED fall into one of four classes:

1. Process involving two distinct fermions, such as muons and electrons. These were the process we considered in this section. They are

    - $e^+ e^- \to \mu^+ \mu^-$: Electron-positron annihilation into a muon-antimuon pair.
    - $\mu^+ \mu^- \to e^+ e^-$: Muon-antimuon annihilation into an electron-positron pair.
    - $e^- \mu^- \to e^- \mu^-$: Electron-muon scattering. This is just Rutherford scattering, except with a muon rather than a proton.
    - $e^- \mu^+ \to e^- \mu^+$: Electron-antimuon scattering
    - $e^+ \mu^- \to e^+ \mu^-$: Positron-muon scattering
    - $e^+ \mu^+ \to e^+ \mu^+$: Positron-antimuon scattering

2. Process involving two identical fermions, such as process involving only electron and positrons:

    - $e^+ e^- \to e^+ e^-$: This process is known as Bhabha scattering, calculated by Homi Bhabha in 1936.
    - $e^- e^- \to e^- e^-$: Møller scattering, calculated in 1931. Considered to be a rather boring process at the time, someone only bothered to measure it in 1950.
    - $e^+ e^+ \to e^+ e^+$: Positron-positron scattering.

3. Processes involving fermions and photons, such as electron-light interaction:

- $e^-e^+ \to \gamma\gamma$: Electron-positron annihilation into photons.

- $e^-\gamma \to e^-\gamma$: Compton scattering, first observed by Arthur Compton in 1923. The cross-section was calculated theoretically by Klein and Nishina in 1929.

- $e^+\gamma \to e^+\gamma$: Compton scattering for positrons.

4. Final we have light-light interactions, $\gamma\gamma \to \gamma\gamma$. This process is known as Halpern scattering, and was calculated in 1933 by Otto Halpern. There are no tree-level contributions, and so the process is much rarer than the others listed here. It has yet to be observed directly in an experiment.

Having studied the first class of amplitudes, we will quickly say a few words about the second class, before moving on to a more detailed study of electron-photon interactions.

When studying the process $e^-e^- \to e^-e^-$, we have contributions from two diagrams:

$$\mathcal{M} = \qquad\qquad - \qquad\qquad \tag{7.99}$$

These diagrams can be evaluated in much the same way as for muon-electron scattering. The only difference is that we now get interesting (and tedious to calculate) interference effects between the two diagrams.

Under crossing symmetry, these two diagrams can be used to describe Bahba scattering $e^+e^- \to e^+e^-$:

$$\mathcal{M} = \qquad\qquad - \qquad\qquad \tag{7.100}$$

7.5 Electrons and Photons

In this section we will study process in QED involving both photons and electrons. Again for simplicity we will only calculate unpolarized cross sections. This will require us to sum over the electron and the photon polarization states. For the electron we can use the methods we developed in the previous section. A similar trick works for photons.

To derive it, consider an process in which a photon of momentum k and polarization $\varepsilon_\mu(k)$ is produced. The amplitude for this process is

$$\mathcal{M} = \qquad\qquad = -\mathcal{M} = i\mathcal{M}^\mu(k)\varepsilon_\mu(k)^*. \tag{7.101}$$

Let's now choose a reference frame where $k^\mu = (k, 0, 0, k)$. The two physical polarizations of the photon correspond to the vectors

$$\varepsilon_1^\mu = (0, 1, 0, 0), \quad \varepsilon_1^\mu = (0, 0, 1, 0) \tag{7.102}$$

and so

$$\sum_{\text{polarizations}} |\mathcal{M}|^2 = \sum_{\text{polarizations}} \mathcal{M}^\mu(k)\mathcal{M}^{\nu*}(k)\varepsilon_\mu(k)^* \varepsilon_\nu(k)^* = |\mathcal{M}^1|^2 + |\mathcal{M}^2|^2. \tag{7.103}$$

We now need to use Ward-Takahashi identity, which we introduced in Section 7.3.1:

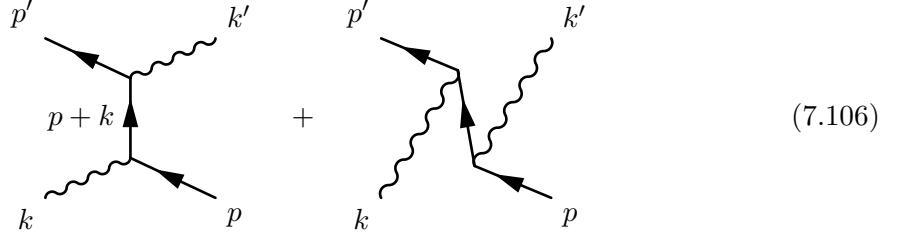$$k_\mu \mathcal{M}^\mu(k) = 0 \implies k\mathcal{M}^0 - k\mathcal{M}^3 = 0. \tag{7.104}$$

This implies $\mathcal{M}^0 = \mathcal{M}^3$, and so

$$\sum_{\text{polarizations}} \mathcal{M}^\mu \mathcal{M}^{\nu*} \varepsilon_\mu^* \varepsilon_\nu^* = |\mathcal{M}^1|^2 + |\mathcal{M}^2|^2$$
$$= |\mathcal{M}^1|^2 + |\mathcal{M}^2|^2 + |\mathcal{M}^3|^2 - |\mathcal{M}^0|^2 = -g_{\mu\nu}\mathcal{M}^\mu \mathcal{M}^{\nu*}. \tag{7.105}$$

So to calculate an unpolarized photon cross section, we simply replace $\varepsilon_\mu \varepsilon_\nu^*$ with $-g_{\mu\nu}$.

7.5.1   Compton Scattering

There are two diagrams which contribute to Compton scattering

$$\tag{7.106}$$

From these diagrams, we can immediately write done the amplitude for this process

$$\begin{aligned}
i\mathcal{M} = &\overline{u}(p')(-ie\gamma^\mu)\varepsilon_\mu^*(k')\frac{i(\not{p} + \not{k} + m)}{(p+k)^2 - m^2}(-e\gamma^\nu)\varepsilon_\nu(k)u(p) \\
&+ \overline{u}(p')(-ie\gamma^\nu)ek_\nu^*(k)\frac{i(\not{p} - \not{k}' + m)}{(p-k')^2 - m^2}(-e\gamma^\nu)\varepsilon_\nu(k')u(p) \\
&= -ie^2\varepsilon_\mu^*(k')\varepsilon_\nu(k)\overline{u}(p')\left[\frac{\gamma^\mu(\not{p} + \not{k} + m)\gamma^\nu}{(p+k)^2 - m^2} + \frac{\gamma^\nu(\not{p} - \not{k}' + m)\gamma^\mu}{(p-k')^2 - m^2}\right]u(p).
\end{aligned} \tag{7.107}$$

There are a few simplification we can make to this expression. Using $p^2 = m^2$ and $k^2 = 0$, we find

$$(p + k)^2 - m^2 = 2p \cdot k, \quad (p - k')^2 - m^2 = -2p \cdot k'. \tag{7.108}$$

A little Dirac algebra simplifies the numerator:

$$\begin{aligned}
(\not{p} + m)\gamma^\nu u(p) &= (\{\gamma^\mu, \gamma^\nu\}p_\mu - \gamma^\nu \not{p} + \gamma^\nu m)u(p) \\
&= 2p^\mu u(p) - \gamma^\nu(\not{p} - m)u(p) = 2p^\mu u(p),
\end{aligned} \tag{7.109}$$

where we have used (6.51) to cancel $(\not{p} - m)u(p) = 0$. This trick works on both numerators, and so

$$\mathcal{M} = -e^2 \varepsilon^*_\mu(k') \varepsilon_\nu(k) \overline{u}(p') \left[\frac{\gamma^\mu \not{k} \gamma^\nu + 2\gamma^\mu p^\nu}{2p \cdot k} + \frac{\gamma^\nu \not{k'} \gamma^\mu - 2\gamma^\nu p^\mu}{2p \cdot k'} \right] u(p). \tag{7.110}$$

Now we take the amplitude squared and sum over the polarizations:

$$\begin{aligned}
\frac{1}{4} \sum_{\text{spin}} |\mathcal{M}|^2 =& \frac{e^4}{4} g_{\mu\rho} g_{\nu\sigma} \text{tr} \left((\not{p}' + m) \left[\frac{\gamma^\mu \not{k} \gamma^\nu + 2\gamma^\mu p^\nu}{2p \cdot k} + \frac{\gamma^\nu \not{k'} \gamma^\mu - 2\gamma^\nu p^\mu}{2p \cdot k'} \right] \right. \\
& \left. (\not{p} + m) \left[\frac{\gamma^\sigma \not{k} \gamma^\rho + 2\gamma^\sigma p^\rho}{2p \cdot k} + \frac{\gamma^\rho \not{k'} \gamma^\sigma - 2\gamma^\rho p^\sigma}{2p \cdot k'} \right] \right).
\end{aligned} \tag{7.111}$$

There are only 16 terms in the trace for us to evaluate, using the trace identities of the previous section. Rather than evaluate them in gruesome detail, here is the result

$$\frac{1}{4} \sum_{\text{spin}} |\mathcal{M}|^2 = 2e^4 \left[\frac{p \cdot k'}{p \cdot k} + \frac{p \cdot k}{p \cdot k'} + 2m^2 \left(\frac{1}{p \cdot k} - \frac{1}{p \cdot k'} \right) + m^4 \left(\frac{1}{p \cdot k} - \frac{1}{p \cdot k'} \right)^2 \right]. \tag{7.112}$$

Our next job is to evaluate this in a reference frame. For a change, we won't work in the centre of mass frame, but instead in the rest frame of the initial electron. We will take the incoming photon to be in the x-direction, and the outgoing photon in the x-y direction:

$$\begin{aligned}
p &= (m, 0, 0, 0), & p' &= (E', \mathbf{p}') \\
k &= (\omega, \omega, 0, 0), & k' &= (\omega', \omega' \cos\theta, \omega' \sin\theta, 0).
\end{aligned} \tag{7.113}$$

The amplitude in this reference frame is particularly simple:

$$\frac{1}{4} \sum_{\text{spin}} |\mathcal{M}|^2 = 2e^4 \left[\frac{\omega'}{\omega} + \frac{\omega}{\omega'} + 2m \left(\frac{1}{\omega} - \frac{1}{\omega'} \right) + m^2 \left(\frac{1}{\omega} - \frac{1}{\omega'} \right)^2 \right] \tag{7.114}$$

With a bit more effort, we can turn this into a cross-section

$$\frac{d\sigma}{d\cos\theta} = \frac{\pi\alpha^2}{m^2} \left(\frac{\omega'}{\omega} \right)^2 \left[\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \sin^2\theta \right], \tag{7.115}$$
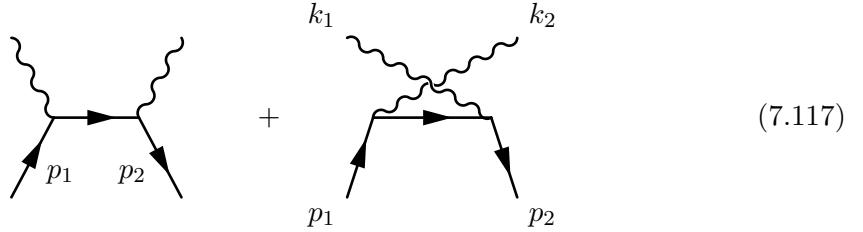
where ω' is given by the Compton formula

$$\omega' = \frac{\omega}{1 + \frac{\omega}{m}(1 - \cos\theta)}. \tag{7.116}$$

7.5.2   Electron-Positron Annihilation to Photons

We have time for one last process in QED: the annihilation of an electron-positron pair into photons. The simplest such process is $e^+ e^- \rightarrow 2\gamma$; there are two diagrams which

contribute to this at tree-level:

$$p \to p_1, \quad p' \to -p_2, \quad k \to -k_1, \quad k' \to k_2 \tag{7.118}$$

These are just the Compton diagrams on their sides; by setting

$$p \to p_1, \quad p' \to -p_2, \quad k \to -k_1, \quad k' \to k_2 \tag{7.118}$$

we can use the Compton amplitude to calculate:

$$\frac{1}{4}\sum_{\text{spin}}|\mathcal{M}|^2 = -2e^4\left[\frac{p_1 \cdot k_2}{p_1 \cdot k_1} + \frac{p_1 \cdot k_1}{p_1 \cdot k_2} + 2m^2\left(\frac{1}{p_1 \cdot k_1} + \frac{1}{p_1 \cdot k_2}\right) - m^4\left(\frac{1}{p_1 \cdot k_1} + \frac{1}{p_1 \cdot k_2}\right)^2\right].$$
$$\tag{7.119}$$

As per usual, we specialize to the centre of mass frame:

$$\begin{aligned}
p_1 &= (E, p, 0, 0), & p_2 &= (E, -p, 0, 0)\\
k_1 &= (E, E\cos\theta, E\sin\theta, 0), & k_2 &= (E, -E\cos\theta, -E\sin\theta, 0),
\end{aligned}\tag{7.120}$$

where $E^2 = p^2 + m^2$. With a little computation, the cross-section becomes

$$\frac{d\sigma}{d\cos\theta} = \frac{\pi\alpha^2}{2Ep}\left[\frac{E^2 + p^2\cos^2\theta}{m^2 + p^2\sin^2\theta} + \frac{2m^2}{m^2 + p^2\sin^2\theta} - \frac{2m^4}{(m^2 + p^2\sin^2\theta)^2}\right]\tag{7.121}$$

Although this expression is somewhat complicated, it simplifies considerably at high energies:

$$\frac{d\sigma}{d\cos\theta} \xrightarrow[E\gg m]{} \frac{\pi\alpha^2}{2E}\left(\frac{1 + \cos^2\theta}{\sin^2\theta}\right).\tag{7.122}$$

This approximation breaks down however for small angles, where $\sin\theta \sim m/p$ or smaller.

Spontaneous Symmetry Breaking

8.1   Linear Sigma Models

8.1.1   Discrete Symmetry Breaking

Spontaneous symmetry breaking is one of the most important ideas in modern physics. Like many great ideas, the core concept is rather simple, and can be used to explain the properties of many otherwise disparate systems. We'll begin our study of the phenomenon with some toy models.

Imagine that we took our $\lambda\phi^4$ theory, except that instead of having a negative mass term $-\frac{1}{2}m^2\phi^2$, we replace it with a positive term:

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 + \frac{1}{2}\mu^2\phi^2 - \frac{\lambda^4}{4!}\phi^4. \tag{8.1}$$

This theory has a discrete symmetry, under the operation $\phi \to -\phi$. How does our theory behave? If we tried to blindly apply our previous results to this theory, we would conclude that the ϕ field creates a particle of mass $\sqrt{-\mu^2} = i\mu$. This cannot be correct.

To understand what went wrong, let us calculate the Hamiltonian for our system:

$$H = \int d^3x \; \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 - \frac{1}{2}\mu^2\phi^2 + \frac{\lambda^4}{4!}\phi^4. \tag{8.2}$$

Classical the energy density will be minimized by a constant field $\phi(\mathbf{x}) = \phi_0$, in terms of which

$$\mathcal{H} = V(\phi_0) = -\frac{1}{2}\mu^2\phi_0^2 + \frac{\lambda}{4!}\phi_0^4. \tag{8.3}$$

If we plot $V(\phi_0)$, we can see immediately that something strange is happening. At $\phi_0 = 0$ we have not a minimum but a maximum of the potential. This is unstable. The imaginary 'mass' of the ϕ particle is one symptom of the instability; for historic reasons it is sometimes known as a *tachyonic instability*.

It is hard to find the two true minima of $V(\phi)$. They are given by

$$\phi_0 = \pm v = \pm\sqrt{\frac{6}{\lambda}}\mu, \tag{8.4}$$

where v is known as the *vacuum expectation value* (or *VEV*) of ϕ.

At low energies, we expect $\phi(x)$ to be close to either v or $-v$. Assuming the former, we will define a new field

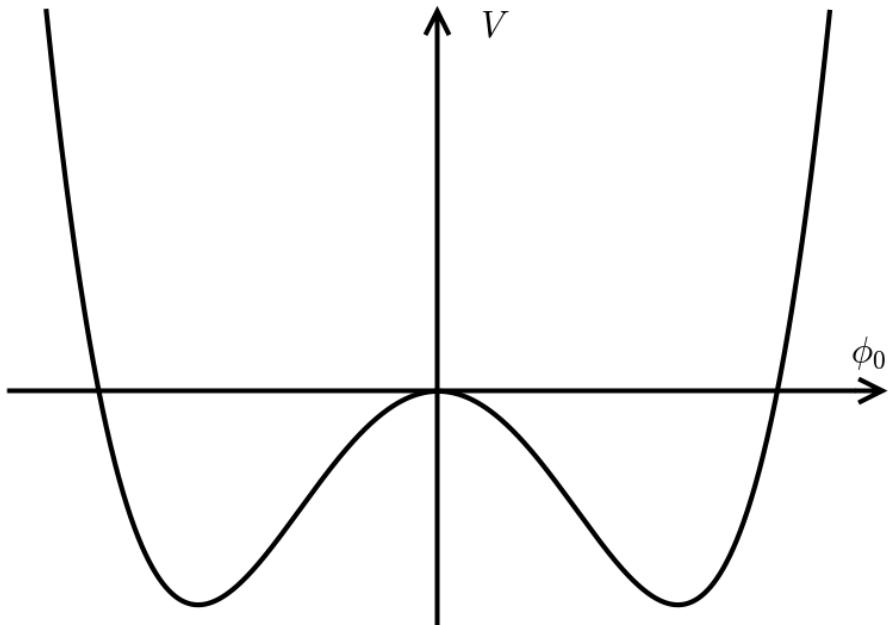$$\sigma(x) = \phi(x) - v. \tag{8.5}$$

Figure 8.1: Potential for discrete symmetry breaking.

Since $\phi(x) \sim v$, $\sigma(x)$ will be small at low energies. In terms of our new field, the Lagrangian is

$$\mathcal{L} = \frac{1}{2}(\partial\sigma)^2 - \frac{1}{2}(2\mu^2)\sigma^2 - \sqrt{\frac{\lambda}{6}}\mu\sigma^3 - \frac{\lambda}{4!}\sigma^4, \tag{8.6}$$

where we have dropped the physically irrelevant constant term. This theory now looks just like any other scalar theory. It contains a σ particle of mass $\sqrt{2}\mu$. Using Feynman diagrams, we could study the scattering of the σ particle, which interacts with itself via both cubic and quartic terms.

The symmetry of our system is no longer manifest; if I were to hand you (8.6) without any context, it would not be obvious at all that there was a hidden symmetry. Only by studying the interrelationships between the σ^2, σ^3 and σ^4 terms could you notice that our theory has a symmetry $\phi \to -\phi$. For this reason we say that the symmetry has been *spontaneously broken.*

8.1.2   Complex Fields

More interesting behaviour occurs when we break a continuous symmetry. Take the Lagrangian

$$\mathcal{L} = (\partial^\mu\Phi^*)(\partial_\mu\Phi) + \mu^2(\Phi^*\Phi) - \frac{\lambda}{2}(\Phi^*\Phi)^2, \tag{8.7}$$

where Φ is a complex field. This theory is invariant under the $U(1)$ symmetry $\Phi \to e^{i\alpha}\Phi$. We will now reparametrize the field as

$$\Phi(x) = \rho(x)e^{i\theta(x)}, \tag{8.8}$$

in terms of which the Lagrangian becomes

$$\mathcal{L} = \rho^2(\partial\theta)^2 + (\partial\rho)^2 + \mu^2\rho^2 - \lambda\rho^4. \tag{8.9}$$

The $U(1)$ symmetry of our theory is spontaneously broken, because the potential is minimised when

$$\rho = |\Phi| = \sqrt{\frac{\mu^2}{\lambda}} = v. \tag{8.10}$$

Introducing a new field $\chi = \rho - v$, our final form of the Lagrangian is

$$\mathcal{L} = (\partial\theta)\left[v^2 + v\chi + \chi^2\right] + \left[(\partial\chi)^2 - 2\mu^2\chi - 4\chi^3\sqrt{\frac{\mu^2\lambda}{2}} - \lambda\chi^4\right] \tag{8.11}$$

where we have again dropped the constant term. This theory has two particles: a massless ϕ field and a massive χ field. There are a number of interaction terms, but these ultimately are all dependent on the two parameters μ and λ.

8.1.3   The Linear Sigma Model

We can generalize our results to a whole class of theories, known as *linear sigma models*. They involve a set of N real scalar fields ϕ^a, interacting under the Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi^a)^2 + \frac{1}{2}\mu^2(\phi^a)^2 - \frac{\lambda}{4}[(\phi^a)^2]^2. \tag{8.12}$$

We implicitly sum over a in each factor of $(\phi^a)^2$. Our theory generalizes the previous two; when $N = 1$ we recover $\lambda\phi^4$, while when $N = 2$, we rewrite $\Phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$ to recover our complex field theory.

Our Lagrangian is invariant under the symmetry

$$\phi^a \to M^{ab}\phi^b \tag{8.13}$$

for any $N \times N$ orthogonal matrix R. This group of transformations is simple the rotation group in N dimensions, alias the N-dimensional *orthogonal group*, or $O(N)$ for short. The lowest-energy configuration is just a constant field ϕ_0^a, which must minimize

$$V(\phi^a) = -\frac{1}{2}\mu^2(\phi^a)^2 + \frac{\lambda}{4}[(\phi^a)^2]^2. \tag{8.14}$$

This potential is minimized by any field satisfying

$$(\phi_0^a)^2 = \frac{\mu^2}{\lambda} = v^2, \tag{8.15}$$

but the direction the field points in can be arbitrary. For convenience we will assume that

$$\phi_0^a = (0, 0, ..., 0, v), \tag{8.16}$$

and define the new set of fields

$$\phi^a(x) = (\pi^b(x), v + \sigma(x)) \qquad \text{for } b = 1, ..., N - 1. \tag{8.17}$$

In terms of these new fields, the Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}(\partial\pi^b)^2 + \frac{1}{2}(\partial\sigma)^2 - \frac{1}{2}(2\mu^2)\sigma^2 - \sqrt{\lambda}\mu\sigma^3 - \frac{\lambda}{4}\sigma^4 - \sqrt{\lambda}\mu(\pi^b)^2\sigma - \frac{\lambda}{2}(\pi^b)^2\sigma^2 - \frac{\lambda}{4}[(\pi^b)^2]^2. \tag{8.18}$$

We have one massive σ field, along with $N-1$ massless π^b fields. The $O(N)$ symmetry has been broken, but there is still an $O(N-1)$ symmetry which acts only on the π^b fields. We say that the $O(N)$ symmetry has been broken down to an $O(N-1)$ subgroup.

Feynman Rules

Since the linear sigma model is just a theory of interacting scalar fields, it is not hard to derive the Feynman rules for the theory. The main interesting feature is that there are $N-1$ identical π^a fields. The propagator for the π^a field is hence

$$a \;\; \underline{\qquad\qquad}_{\;p} \;\; b = \frac{i\delta^{ab}}{p^2}. \tag{8.19}$$

By using one propagator to represent the π^a fields, we have made good use of the $O(N-1)$ symmetry of our theory. The σ propagator is just the regular propagator for a particle of mass $m = \sqrt{2}\mu$:

$$- - - - - -_{\;p} \;\; = \frac{i}{q^2 - 2\mu^2}. \tag{8.20}$$

There are five interaction vertices, which can be read straight off the Lagrangian:

$$
\begin{aligned}
& = -6i\frac{\mu}{\sqrt{\lambda}}, && = -6i\delta^{ab}\frac{\mu}{\sqrt{\lambda}} \\[2em]
& = -6i\lambda, && = -2i\delta^{ab}\lambda \\[2em]
& = -2i\lambda\left[\delta^{ab}\delta^{cd} + \delta^{ac}\delta^{bd} + \delta^{ad}\delta^{bc}\right]
\end{aligned}
\tag{8.21}
$$

Notice that all of these interactions depend only on two parameters, λ and μ. Even though the $O(N)$ symmetry is spontaneously broken, it can be found hidden in the interrelations of the interaction terms.

8.2   Nambu-Goldstone Bosons

We have seen that in the linear sigma model, spontaneous symmetry breaking results in massless particles. This is a very general results, known as *Goldstone's theorem*. The resultant particles are called *Nambu-Goldstone bosons*.

To prove the theorem, recall that for every continuous symmetry there exists a conserved charge Q, which commutes with the Hamiltonian:

$$[H, Q] = 0. \tag{8.22}$$

Denote the ground state of the theory by $|\Omega\rangle$, and assume (without loss of generality) that $H|\Omega\rangle = 0$. This implies that

$$H(Q|\Omega\rangle) = HQ|\Omega\rangle = QH|\Omega\rangle = 0, \tag{8.23}$$

so $Q|\Omega\rangle$ is also a ground state of the theory. There are now two options; either $Q|\Omega\rangle$ is linearly dependent on $|\Omega\rangle$, or there exists a second ground state $|\Omega'\rangle$ with

$$Q|\Omega\rangle = c_1|\Omega\rangle + c_2|\Omega'\rangle. \tag{8.24}$$

Are arguments so far have been very general. QFT now enters the picture by demanding that we have a local current

$$Q = \int d^D x \; J^0(\mathbf{x}, t) \tag{8.25}$$

where we integrate over space, but not time. We next define the state

$$|\mathbf{k}\rangle = \int d^D x \; e^{-i\mathbf{k}\cdot\mathbf{x}} J^0(\mathbf{x}, t)|\Omega\rangle. \tag{8.26}$$

The momentum of this state is $\mathbf{k}$:

$$P_i|\mathbf{k}\rangle = \int d^D x \; e^{-i\mathbf{k}\cdot\mathbf{x}}[P_i, J^0(\mathbf{x}, t)]|\Omega\rangle = \int d^D x \; e^{-i\mathbf{k}\cdot\mathbf{x}}\left[-i\nabla_i J^0(\mathbf{x}, t)\right]|\Omega\rangle = k_i|\mathbf{k}\rangle, \tag{8.27}$$

where in the first equality we used the fact that vacuum has zero momentum. As $\mathbf{k} \to \mathbf{0}$, $|\mathbf{k}\rangle \to Q|\Omega\rangle$, which we have already established is a ground state of the theory. So as the momentum of the state goes to zero, the energy of the state also goes to zero. As we know from relativity, this implies that $|\mathbf{k}\rangle$ describes a massless particle.

The Goldstone theorem is extraordinarily general: it applies to any theory with a spontaneously broken symmetry, regardless of any other details.

If there is more than one broken symmetry, we have a series of conserved charges Q^a. By the above argument, each of these will generate a Nambu-Goldstone boson. In the non-linear sigma model for instance, we broke $O(N) \to O(N-1)$, and this resulted in $N-1$ massless bosons. More generally, if we have a group G which is broken to a subgroup H, the number of Nambu-Goldstone bosons is

$$\dim(G) - \dim(H), \tag{8.28}$$

where $\dim(G)$ is the dimension of the group. If you want to get really fancy, you can say that Nambu-Goldstone bosons live in the *coset space* G/H.

8.2.1   Examples of Spontaneous Symmetry Breaking

The Ferromagnet

Spontaneous symmetry breaking was first discovered not by particle physicists but by condensed matter physicists. Consider a lump of iron. We can consider the material to be made of lots of tiny magnetic dipoles. These magnetic dipoles want to align with their neighbours, and in the ground state the entire metal will comprise of dipoles pointing in the same direction. This breaks the rotational symmetry of the system. In real life the

situation is slightly more complicated, as long-ranged interactions tend to split the metal up into regions of uniform magnetisation. But with a little coaxing, we can get these different domains to align, giving us a magnetic field on a macroscopic scale.

The Nambu-Goldstone bosons of a magnetic are called *magnons*, and they determine the low-energy behaviour of a magnetic material.

Crystals and Phonons

The crystal structure of many solids breaks the translational invariance of physics to a discrete subgroup. Phonons — long-ranged vibrations of the crystal lattice — are the Nambu-Goldstone boson associated with this symmetry breaking.

Sound Waves

In a fluid, sound waves propagate via the wave equation

$$
\left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right] \phi(\mathbf{x}, t) = 0. \tag{8.29}
$$

This is just the Klein-Gordon equation for a massless field. In particular, this means that sound waves are *gapless*: as the wavelength becomes large, the frequency of the wave goes to zero. The classical version of Goldstone's theorem states that in a spontaneously broken symmetry results in the existence of a gapless excitation.

This leads to the logical question: can we interpret sound waves as the Nambu-Goldstone modes of a broken symmetry? The answer is yes. The "ground state" of a fluid breaks the Galilean invariance of hydrodynamics, because we have to go into the reference frame where the fluid is at rest. Sound waves then emerge as the Nambu-Goldstone modes of this symmetry.[1]

Superfluids and Superconductivity

These phenomena both results from spontaneous symmetry breaking. The mathematical analysis is a little more involved, so we will return to this topic once we have discussed the Higgs mechanism.

Pions and QCD

The pion is the lightest meson. It has a mass of ~ 139 MeV, which is only about 15% that of the proton. Although not massless, we can understand the lightness of the pion using spontaneous symmetry breaking. In the limit where the up and down quark are massless, QCD has a symmetry known as *chiral symmetry*. It is this symmetry that is spontaneously broken, with the pion as the Nambu-Goldstone boson. In reality, the quark masses explicitly break the symmetry, giving a small mass to the pion. Nevertheless, the approximate spontaneous symmetry breaking allows us to develop a rich and systematic theory of the strong interaction, at least at low energies.

---

[1]More details can be found in Chapter 1 of *Boulevard of Broken Symmetries* by Schakel.

8.3   The Abelian Higgs Mechanism

What happens when we spontaneously break a gauge symmetry? This might seem like an academic question, but it turns out to be the key to understanding the weak force. The Standard Model has a gauge symmetry $SU(2) \times U(1)$, which is broken down to a $U(1)$ gauge symmetry by spontaneous symmetry breaking. This is known as the *Higgs mechanism*, and is responsible for many features of the Standard Model:

1. Gauge symmetry forces gauge bosons to be massless. On the other hand, the W and Z bosons which mediate the weak force are massive. This caused by the Higgs mechanism, which gives mass to these gauge bosons.

2. In the Standard Model all fermions must be massless due to gauge symmetry. The Higgs mechanism also gives mass to fermions in the Standard Model.

3. The Higgs mechanism predicts the existence of a new boson, known as the *Higgs boson*. In one of the great triumphs of modern physics, the Higgs boson was discovered by the LHC in 2012. With it comes the end of an era for particle physics: every particle predicted by the Standard Model has now been observed experimentally.

To understand the electroweak theory, we would unfortunately need to understand the structure of non-abelian gauge theories. This is beyond the scopeof the course, so we will have to make do by studying a simpler example — spontaneous symmetry breaking in QED. This is called the *abelian Higgs mechanism* and although not directly applicable to particle physics, plays a critical role in the behaviour of superconductors.

We will start with the scalar QED Lagrangian:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + D^{\mu}\Phi D_{\mu}\Phi + \mu^2 \Phi^*\Phi - \frac{\lambda}{2}(\Phi^*\Phi)^2. \tag{8.30}$$

Going into polar coordinates

$$\Psi(x) = \rho(x)e^{i\theta(x)}, \tag{8.31}$$

the Lagrangian becomes

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \rho^2(\partial_{\mu}\theta - eA_{\mu})^2 + (\partial\rho)^2 + \mu^2\rho - \lambda\rho^4. \tag{8.32}$$

Next define the field

$$B^{\mu} = A^{\mu} - \frac{1}{e}\partial^{\mu}\theta. \tag{8.33}$$

This new field is actually gauge invariant, since under a gauge transformation:

$$A^{\mu} \to A^{\mu} + \partial^{\mu}\alpha, \quad \theta \to \theta + e\alpha, \tag{8.34}$$

and these two transformation cancel out. Furthermore, it is not hard to show that

$$F^{\mu\nu} = \partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu} = \partial^{\mu}B^{\nu} - \partial^{\nu}B^{\mu}. \tag{8.35}$$

With this we can now remove $A^{\mu}(x)$ and $\theta(x)$ from the Lagrangian:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + e^2\rho^2 B_{\mu}^2 + (\partial\rho)^2 + \mu^2\rho - \lambda\rho^4. \tag{8.36}$$

Other than the additional gauge terms, our theory is a carbon-copy of that in Section 8.1.2, and once again it is minimised when

$$\rho = |\Phi| = v = \sqrt{\frac{\mu^2}{\lambda}}. \tag{8.37}$$

This leads us to reparameterize

$$\rho(x) = \chi(x) + v, \tag{8.38}$$

in terms of which the Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}(ev)^2 B_\mu^2 + \frac{1}{2}(\partial\chi)^2 - \mu^2\chi^2 + e^2 v\chi B_\mu^2 + \frac{1}{2}e^2\chi^2 B_\mu^2 - \sqrt{\lambda}\mu\chi^3 - \frac{\lambda}{4}\chi^4. \tag{8.39}$$

From this Lagrangian, we see that there are two particles in our theory. We have a vector boson B^μ of mass ev, and a real scalar field χ of mass $\sqrt{2}\mu$.

What happened to the Nambu-Goldstone boson? When we studied the complex field in Section 8.1.2, the Nambu-Goldstone boson arose from the phase field θ. But we got rid of this field by combining it with the massless gauge field A^μ. THis gave us a massive vector field B^μ. A massless gauge field has two degrees of freedom, whilst a massive vector field has three degrees of freedom, corresponding to the three states of a spin-1 particle. This extra degree of freedom came from the θ field!

The appearance of a massive gauge boson was discovered by many physicists independently, and is known as the Higgs mechanism. The leftover χ field is called the Higgs boson.

8.3.1   Propagator for a Massive Vector Boson

We can now read the Feynman rules for our theory straight from the Lagrangian. The only question we need to ask is: what is the propagator for B^μ? After all, we have only ever derived the propagator of a massless vector field.

The relevant terms in the Lagrangian are

$$-\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}M^2 B_\mu^2 \tag{8.40}$$

where $M = ev$ is the mass of the vector boson. We know how to calculate the propagator when $M = 0$; in the Feynman gauge it is given by:

$$\mu \,\,\wwwww\,\, \nu = \frac{-ig^{\mu\nu}}{p^2}. \tag{8.41}$$

We can then treat M as a perturbation on this propagator. Because the mass term has only two powers of B^μ, we can immediate write down the Feynman vertex as

$$\mu \,\,\wwwww\bullet\wwwww\,\, \nu = ig_{\mu\nu}M^2. \tag{8.42}$$

What we want to calculate is the propagator, which is really just the Fourier transform of

2-point function:

$$D_F^{\mu\nu}(p) = \int d^4x \ e^{-ipx} \langle B^\mu(x) B^\nu(0) \rangle. \tag{8.43}$$

We can calculate this quantity by summing Feynman diagrams:

$$
\begin{aligned}
D_F^{\mu\nu}(p) = \quad &= \quad + \quad + \quad + \quad \dots \\
&= -\frac{i}{p^2}g^{\mu\nu} - \frac{iM^2}{p^4}g^{\mu\sigma}g_{\sigma\rho}g^{\rho\nu} - \frac{iM^4}{p^6}g^{\mu\sigma}g_{\sigma\rho}g^{\rho\tau}g_{\tau\lambda}g^{\lambda\nu} - \dots \\
&= -\frac{i}{p^2}g^{\mu\nu} - \frac{iM^2}{p^4}g^{\mu\nu} - \frac{iM^4}{p^6}g^{\mu\nu} + \dots
\end{aligned}
\tag{8.44}
$$

Notice that we do not amputate the legs of these diagrams, as we are calculating a correlation function, not a scattering amplitude. It is not hard to see the general pattern:

$$D_F^{\mu\nu}(p) = \sum_{k=0}^{\infty} \frac{-ig^{\mu\nu}}{p^2}\left(\frac{M^2}{p^2}\right)^k = -\frac{ig^{\mu\nu}}{p^2}\frac{1}{1-M^2/p^2} = -\frac{ig^{\mu\nu}}{p^2-M^2}. \tag{8.45}$$

To sum the series in the second equality we used the formula

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots \tag{8.46}$$

So the propagator for a massive vector field is similar to that of a scalar field — we just have to add a factor of $g^{\mu\nu}$.

8.3.2 Generating Fermion Masses

Using the abelian Higg mechanism, we have illustrated how gauge bosons can gain mass. Another aspect of the electroweak theory is that the spontaneous symmetry breaking gives rise to fermionic masses. A full explanation of why this is the case — why fermion masses are otherwise forbidden in the Standard Model — will have to wait for a different course. Here we simply illustrate the general idea using Yukawa theory.

Take the theory

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 + \frac{1}{2}\mu^2\phi^2 - \frac{\lambda}{4!}\phi^4 + \overline{\psi}(i\slashed{\partial} - M)\psi - g\phi\overline{\psi}\psi, \tag{8.47}$$

where we have coupled a scalar field from 8.2.1 to a fermion via a Yukawa interaction. As we showed early, the minimum energy configuration is

$$\phi_0 = \pm v = \pm\frac{6\mu}{\lambda} \tag{8.48}$$

Defining the new field $\sigma = \phi - v$, our Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}(\partial\sigma)^2 - 2\mu^2\sigma^2 - \sqrt{\frac{\lambda}{6}}\mu\sigma^3 - \frac{\lambda}{4!}\sigma^4 + \overline{\psi}(i\slashed{\partial} - M - gv)\psi - g\sigma\overline{\psi}\psi. \tag{8.49}$$

After spontaneous symmetry breaking, the mass of the fermion has changed:

$$M \to M + gv = M + \frac{6g\mu}{\lambda}. \tag{8.50}$$

In particular, if our fermion had originally been massless, it now acquires a mass.

8.4   Superfluids

In this section and the next, we will explore apply spontaneous symmetry breaking to non-relativistic systems. Recall from Chapter 3 the Schrödinger Lagrangian:

$$\mathcal{L} = \Psi^* \left[i \frac{\partial}{\partial t} + \frac{1}{2m} \nabla^2 \right] \Psi, \tag{8.51}$$

which when quantized gives rise to a non-relativistic theory of bosons. In the context of superfluids, these bosons are not fundamental particles but instead are atoms, usually at very cold temperatures. These atoms interact with each other in a potentially quite complicated fashion. For simplicity we will however assume that the only interaction is a two-body interaction, adding an additional term to our Lagrangian

$$\mathcal{L} = \Psi^* \left[i \frac{\partial}{\partial t} + \frac{1}{2m} \nabla^2 \right] \Psi - \frac{\lambda}{4} (\Psi^* \Psi)^2. \tag{8.52}$$

Using the Euler-Lagrange equations, the equation of motion for the field is

$$-i \frac{\partial}{\partial t} \Psi = \frac{1}{2m} \nabla^2 \Psi + \frac{1}{2} \lambda |\Psi|^2 \Psi. \tag{8.53}$$

This equation is called the *Non-linear Schrödinger equation*, and can be used to describe Bose-Einstein condensates.

Finite Density

So far in this course we have concerned ourselves with the vacuum, or with a small handful of particles interacting in the void. In condensed matter physics we deal with very different kinds of systems. A Bose-Einstein condensate may comprise of millions of atoms, and a lump of metal $\sim 10^{23}$ electrons. How can we study theories at finite density?

Let us modify our interaction term slightly in (8.52):

$$\mathcal{L} = \Psi^* \left[i \frac{\partial}{\partial t} + \frac{1}{2m} \nabla^2 \right] \Psi - \frac{\lambda}{4} (\Psi^* \Psi - \rho_0)^2. \tag{8.54}$$

The Hamiltonian for this theory is

$$H = \int d^3 x \, \frac{1}{2m} (\nabla \Psi^*) \cdot (\nabla \Psi) + \frac{\lambda}{4} (\Psi^* \Psi - \rho_0)^2, \tag{8.55}$$

and so the energy is minimized when

$$|\Psi(\mathbf{x})|^2 = \rho_0. \tag{8.56}$$

Recalling that the number operator for a Schrödinger field is

$$N = \int d^3x \ |\Psi(\mathbf{x})|^2, \tag{8.57}$$

we see that by varying ρ_0, we can change the density of our field. What changing ρ_0 corresponds to is changing the chemical potential of our system.

As per usual, we will work in polar coordinates

$$\Psi(x) = \sqrt{\rho(x)}e^{i\theta(x)}, \tag{8.58}$$

our Lagrangian becomes

$$\mathcal{L} = -\rho\frac{\partial\theta}{\partial t} - \frac{1}{2m}\left[\frac{1}{4\rho}(\nabla\rho)^2 + \rho(\nabla\theta)^2\right] - \frac{\lambda}{4}(\rho - \rho_0)^2, \tag{8.59}$$

where the $\partial_t\rho$ has been dropped because it is a total derivative. Next we expand $\sqrt{\rho(x)} = \sqrt{\rho_0} + \chi(x)$, making the assumption that $\chi(x) \ll \rho_0$. We can then expand our Lagrangian

$$\mathcal{L} = -2\sqrt{\rho_0}\chi\frac{\partial\theta}{\partial t} - \frac{\rho_0}{2m}(\nabla\theta)^2 - \frac{1}{2m}(\nabla\chi)^2 - \lambda\rho_0\chi^2 + O(\chi^3) \tag{8.60}$$

where once again we dropped total derivatives (this time the $\rho_0\dot{\theta}$ term). From now on we will work only with the truncated Lagrangian. This is an approximation, but is a good one so long as $\chi(x) \ll \rho_0$.

The Effective Action

Let's now calculate the path-integral for our theory:

$$
\begin{aligned}
Z &= \int \mathcal{D}\theta \ \mathcal{D}\chi \exp\left[i\int d^4x \ -2\sqrt{\rho_0}\chi\frac{\partial\theta}{\partial t} - \frac{\rho_0}{2m}(\nabla\theta)^2 - \frac{1}{2m}(\nabla\chi)^2 - \lambda\rho_0\chi^2\right] \\
&= \int \mathcal{D}\theta \ \mathcal{D}\chi \exp\left[i\int d^4x \ -\frac{\rho_0}{2m}(\nabla\theta)^2 + \chi\left(\frac{1}{2m}\nabla^2 - \lambda\rho_0\right)\chi - \left(2\sqrt{\rho_0}\frac{\partial\theta}{\partial t}\right)\chi\right].
\end{aligned} \tag{8.61}
$$

If we can evaluate this, then we can calculate all other physical quantities in our theory. As per usual, the main catch is that path-integrals are hard to calculate.

What we will try to do is partially evaluate the integral — integrating over χ but not over θ. The χ integral is really just a Gaussian:

$$Z_\chi[J] = \int \mathcal{D}\chi \exp\left[i\int d^4x \ \chi\left(\frac{1}{2m}\nabla^2 - \lambda\rho_0\right)\chi - \chi J\right] \tag{8.62}$$

where the source is defined as

$$J(\mathbf{x}, t) = 2\sqrt{\rho_0}\dot{\theta}(\mathbf{x}, t). \tag{8.63}$$

We can evaluate the integral via the usual method — Fourier transforming and then a change of variable. For a bit of a change, we will only Fourier transform the spatial

coordinates; our path-integral now becomes:

$$Z_\chi[J] = \int \mathcal{D}\chi \exp\left[i \int d^3k \, dt \, \chi(-\mathbf{k},t)\left(\frac{1}{2m}\mathbf{k}^2 - \lambda\rho_0\right)\chi(\mathbf{k},t) - \chi(\mathbf{k},t)J(-\mathbf{k},t)\right] \quad (8.64)$$

Next we define the field

$$\chi(\mathbf{k},t) = h(\mathbf{k},t) + \frac{1}{2}\frac{J(\mathbf{k},t)}{\frac{1}{2m}\mathbf{k}^2 - \lambda\rho_0} \quad (8.65)$$

(This is the last time we will introduce a new field, I promise!) Inserting this into the path-integral (8.62), we find that

$$\begin{aligned}
Z_\chi[J] &= \int \mathcal{D}h \exp\left[i \int d^3k \, dt \, h(-\mathbf{k},t)\left(\frac{1}{2m}\mathbf{k}^2 - \lambda\rho_0\right)h(\mathbf{k},t) - \frac{1}{4}\frac{J(-\mathbf{k},t)J(\mathbf{k},t)}{\frac{1}{2m}\mathbf{k}^2 - \lambda\rho_0}\right] \\
&= \exp\left[i \int d^3k \, dt - \frac{1}{4}\frac{J(-\mathbf{k},t)J(\mathbf{k},t)}{\frac{1}{2m}\mathbf{k}^2 - \lambda\rho_0}\right] \\
&\quad \times \int \mathcal{D}h \exp\left[i \int d^3k \, dt \, h(-\mathbf{k},t)\left(\frac{1}{2m}\mathbf{k}^2 - \lambda\rho_0\right)h(\mathbf{k},t)\right].
\end{aligned}$$

$$(8.66)$$

The second factor is just a normalization constant, and is of no interest to us. The first term, however, will modify the behaviour of the θ field. Using the above expression to simplify (8.61), we find

$$Z = \int \mathcal{D}\theta \, \exp\left[i \int d^4x \, -\frac{\rho_0}{2m}(\nabla\theta)^2\right]\exp\left[i \int d^3k \, dt - \frac{1}{4}\frac{J(-\mathbf{k},t)J(\mathbf{k},t)}{\frac{1}{2m}\mathbf{k}^2 - \lambda\rho_0}\right] \quad (8.67)$$

All reference to the χ field has disappeared. The behaviour of θ can be computed in terms of the *effective action*

$$\begin{aligned}
S_{\text{eff}}[\theta] &= -\int d^4x \frac{\rho_0}{2m}(\nabla\theta)^2 + \int d^3k \, dt \, -\frac{1}{4}\frac{J(-\mathbf{k},t)J(\mathbf{k},t)}{\frac{1}{2m}\mathbf{k}^2 - \lambda\rho_0} \\
&= -\int d^4x \frac{\rho_0}{2m}(\nabla\theta)^2 + \frac{1}{\lambda}\int d^3k \, dt \, \frac{\dot\theta(-\mathbf{k},t)\dot\theta(\mathbf{k},t)}{1 - \Lambda^2\mathbf{k}^2}
\end{aligned}$$

$$(8.68)$$

where $\Lambda = (2m\rho_0\lambda)^{-1/2}$. Our path-integral (8.61) now has a particularly simple form:

$$Z = \mathcal{N}\int \mathcal{D}\theta \, \exp\left[iS_{\text{eff}}[\theta]\right]. \quad (8.69)$$

Using the effective action, we can calculate the behaviour of the θ field whilst ignoring the χ field.

Low Energy Expansion

Though we have achieved our goal of integrating out χ, we still have a very ugly $S_{\text{eff}}[\theta]$ to to work with. For a start, this action is explicitly non-local; with a bit of clever integration we can in fact evaluate in position space

$$\frac{1}{\lambda}\int d^3k \, dt \, \frac{\dot\theta(-\mathbf{k},t)\dot\theta(\mathbf{k},t)}{1 - \Lambda^2\mathbf{k}^2} = \frac{1}{4\pi\lambda}\int d^3x \, d^3y \, dt \, \dot\theta(\mathbf{x},t)\dot\theta(\mathbf{y},t)\frac{e^{-|\mathbf{x}-\mathbf{y}|/\Lambda}}{|\mathbf{x}-\mathbf{y}|}. \quad (8.70)$$

The θ fields interact via a Yukawa potential!

We know that a Yukawa potential decays exponentially over distances greater than Λ^{-1}. In condensed matter physics we are usually interested in the macroscopic behaviour of a system — which in natural units corresponds to the low energy limit. Using the power series

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + ... \tag{8.71}$$

we can expand in powers of Λ^{-1}:

$$\begin{aligned}
\int d^3k \ dt \ \frac{\dot{\theta}(-\mathbf{k},t)\dot{\theta}(\mathbf{k},t)}{1-\Lambda\mathbf{k}^2} &= \int d^3k \ dt \ \dot{\theta}(-\mathbf{k},t)\dot{\theta}(\mathbf{k},t)\left(1+\Lambda^2\mathbf{k}^2+\Lambda^4(\mathbf{k}^2)^2+...\right) \\
&= \int d^3x \ dt \ \dot{\theta}(x)^2 - \Lambda^2\dot{\theta}(x)\nabla^2\dot{\theta}(x) + \Lambda^4\dot{\theta}(x)\nabla^4\dot{\theta}(x) - ...
\end{aligned} \tag{8.72}$$

So the non-local effective action can be expanded as an infinite series of local terms. Having done all this work, we finally can expand

$$\begin{aligned}
S_{\text{eff}} &= \int d^3t \ dt \ -\frac{\rho_0}{2m}(\nabla\theta)^2 + \frac{1}{\lambda}\dot{\theta}(x)^2 - \frac{\Lambda^2}{\lambda}\dot{\theta}(x)\nabla^2\dot{\theta}(x) + ... \\
&\approx \int d^4x \ \frac{1}{\lambda}\dot{\theta}(x)^2 - \frac{\rho_0}{2m}(\nabla\theta)^2.
\end{aligned} \tag{8.73}$$

After the dust has settled, we have discovered an approximate form of the θ Lagrangian which is accurate at length scales much greater than Λ^{-1}. Using the Euler-Lagrange equations, we find that θ evolves according to the wave equation:

$$\left[\frac{\partial^2}{\partial t^2} - \frac{\lambda\rho_0}{2m}\frac{\partial^2}{\partial x^2}\right]\theta = 0. \tag{8.74}$$

The θ field is the Nambu-Goldstone mode for the theory, and governs the low-energy behaviour of our system. It is also known as a *linear dispersion mode*, because the dispersion relation is

$$\omega(\mathbf{k}) = \sqrt{\frac{\lambda\rho_0}{2m}}|\mathbf{k}|. \tag{8.75}$$

The higher order derivative terms will give corrections to this relation of order $\Lambda^2\mathbf{k}^2$, but these are only relevant at short wavelengths.

Our Lagrangian is really just the Lagrangian for a massless scalar field. When we quantize the theory, the θ field will give rise to quantized excitations with momentum $\hbar\mathbf{k}$ and energy $\hbar\omega(\mathbf{k})$.

In *Figure 8.2*, the relationship between wavelength and energy is plotted for superfluid helium–4. For wavelengths above about 0.2 nm, the energy depends linearly on the momentum. At shorter wavelengths, other excitations, known as maxons and rotons, become important. Their effects can be qualitatively described in more sophisticated models.

8.4.1   Philosophy of Effective Field Theory

Having just completed a long computation, we should take a few seconds to catch our breath and evaluate the steps we just took:

1. First we chose a good parametrization of our field. Originally our Lagrangian was written in terms of a complex field Ψ. But because our system exhibited spontaneous
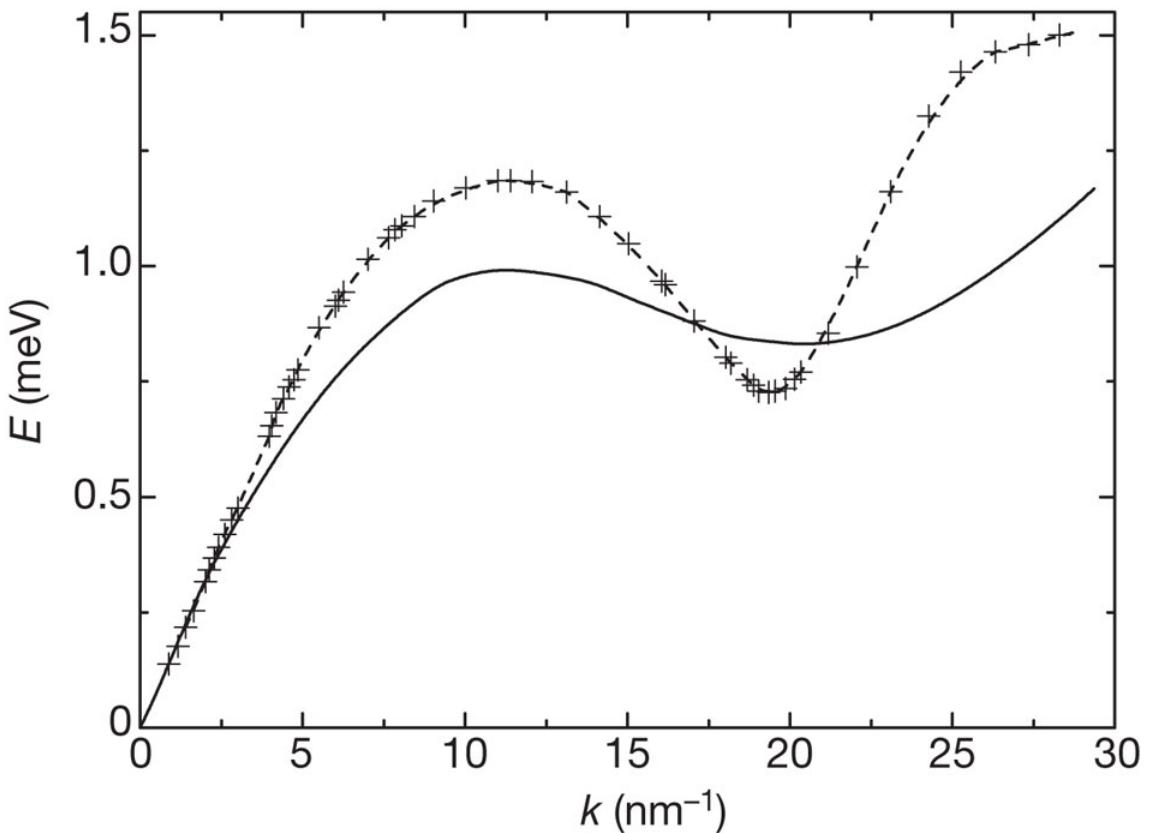
Figure 8.2: Elementary excitations in superfluid helium–4. The solid line is the dispersion relation predicted by Landau, whilst the crosses correspond to experimental data determined through neutron scattering. Notice that there is a linear dispersion relation at low energies, just as we predicted. The figure has been adapted from *doi:10.1038/nature10919.*

symmetry breaking, we realised that a more physically meaningful parametrization was

$$\Psi(\mathbf{x}, t) = (\sqrt{\rho_0} + \chi(\mathbf{x}, t))e^{i\theta(\mathbf{x}, t)}. \tag{8.76}$$

2. Next we isolated the fields of interest. In our case, we knew that at low energies there would be a Nambu-Goldstone mode associated with the broken phase symmetry. It was not hard to guess that this field should be the phase field θ. To study just the θ field, we integrated the other fields out of the path-integral.

3. The process of integrating out χ gave us a non-local effective action for θ. Since however we are only interested in the low energy behaviour of the system, we could expand the action as an infinite series of local terms. Truncating this series allows us to approximate the dynamics of θ at long wavelengths.

These steps illustrate the technique known as *effective field theory*. At heart the idea is simple — at low energies, things should look simple. All of the heavy particles will be too heavy to produce, leaving only a handful of light particles to consider. For the case of spontaneous symmetry breaking, the low-energy physics is dominated by the Nambu-Goldstone bosons. In QED, the low energy physics is dominated by the massless photons. This is the reason you could study classical electrodynamics without knowing about the electron, muon, or any other charged particle.

Effective field theory is central to our modern understanding of quantum field theory, connecting such disparate topics as neutrino masses, gravitational wave scattering, the quantum hall effect and phase transitions. It holds the key to understanding why loop diagram diverge, and how to fix this problem via renormalisation. Sadly we do not have the time to explore this crucial topic — you will need to do another course on quantum field theory.

8.4.2   Linear Dispersion Modes

After all this work, we should get to the punchline. A linear dispersion mode implies superfluidity. Why? Imagine that we had a mass M of fluid flowing through a pipe with velocity $\mathbf{v}$. This could slow down by emitting an excitation of the θ field. Conservation of momentum dictates that

$$M\mathbf{v} = M\mathbf{u} + \hbar\mathbf{k} \tag{8.77}$$

where $\mathbf{u}$ is the new velocity of the fluid. But this is only possible if there is sufficient energy:

$$\begin{aligned}
\frac{1}{2}M\mathbf{v}^2 &\geq \frac{1}{2}M\mathbf{u}^2 + \hbar\omega(\mathbf{k}) \\
&= \frac{1}{2}M\left(\mathbf{v} - \frac{\hbar\mathbf{k}}{M}\right)^2 + \hbar\omega(\mathbf{k})
\end{aligned} \tag{8.78}$$

where we have used conservation of momentum to eliminate $\mathbf{u}$. Rearranging this inequality gives us

$$\hbar\mathbf{v}\cdot\mathbf{k} \geq \hbar\omega(\mathbf{k}) + \frac{1}{2M}\hbar^2\mathbf{k}^2 \tag{8.79}$$

In the limit where M is macroscopic the second term can be neglected. Using the dispersion relation (8.75), we find that

$$\mathbf{v}\cdot\mathbf{k} \geq \sqrt{\frac{\lambda\rho_0}{2m}}|\mathbf{k}| \tag{8.80}$$

and hence

$$|\mathbf{v}| \geq v_c = \sqrt{\frac{\lambda\rho_0}{2m}}. \tag{8.81}$$

If the fluid is moving with a velocity less than v_c, it will be unable to slow down. Because of this, superfluids have zero viscosity, and will flow indefinitely.

What is special about a superfluid?

To understand why superfluids are special, consider instead a gas of free bosons. In this gas, our mass of fluid can give momentum $\hbar\mathbf{k}$ to any boson at the cost of $\frac{\hbar^2\mathbf{k}^2}{2m}$ energy. Repeating our previous argument, our fluid flow can loose momentum if it can satisfy the inequality

$$\hbar\mathbf{v}\cdot\mathbf{k} \geq \hbar\omega(\mathbf{k}) + \frac{1}{2M}\hbar^2\mathbf{k}^2 = \left(\frac{1}{2M} + \frac{1}{2m}\right)\hbar^2\mathbf{k}^2 \tag{8.82}$$

The left-hand side is linear in momentum, whereas the right-hand side is quadratic. By setting $|\mathbf{k}|$ to be sufficiently small, we can satisfy this inequality for any velocity. Superfluidity cannot occur in a regular fluid!

The special property of a superfluid is the $\lambda|\Psi|^4$ interaction term. As we can see from the Hamiltonian, this term is repulsive. If we were to give a single boson $\hbar\mathbf{k}$ momentum,

it would collide with other bosons, affecting the entire system. We know that the relevant low energy field is the θ field, and so our boson will give rise to a θ wave. This is a a highly non-trivial effect, arising from the interactions of numerous bosons.

8.5 Superconductors

Superconductors are to superfluids what the Higgs mechanism is to the linear sigma model — it is what happens when a superfluid is coupled to a gauge field. In fact, the idea of superconductivity predates the use of spontaneous symmetry breaking in particle physics. Coupling the field in (8.52) to the magnetic field, our new Lagrangian becomes

$$\mathcal{L} = \Psi^* i \frac{\partial}{\partial t} \Psi - \frac{1}{2m}(\nabla + i2e\mathbf{A})\Psi^*(\nabla - i2e\mathbf{A})\Psi - \frac{\lambda}{4}(\Psi^*\Psi - \rho_0)^2 + 2e\phi\Psi^*\Psi + \frac{1}{2}\mathbf{E}^2 - \frac{1}{2}\mathbf{B}^2. \tag{8.83}$$

Since we are working with a non-relativistic system, we have used non-relativistic notations for electromagnetism, with $\phi = A^0$.

Before studying the implications of this Lagrangian, we should probably explain where it comes from. In a real metal, the relevant degrees of freedom at low energies are the free electrons, in a background of positive ions. Because electrons are fermions, we cannot describe them using a bosonic field Ψ.

Nevertheless, superconductivity, and its close friend superfluidity, can occur in both fermionic and bosonic systems. This is because two electrons can couple together to form a *Cooper pair*. The details of this mechanism are described by *BCS theory*, and belong in a condensed matter text. For our purposes, the important feature is that *Cooper pairs* are spinless and are therefore bosons, not fermions. For this reason, at sufficiently low energy cooper pairs can exhibit all the features of bosonic systems, such as superfluidity and Bose-Einstein condensation.

Since the Cooper pairs have charge $+2e$, the Lagrangian (8.83) describes their behaviour at low energies. It is known as Ginzburg–Landau theory, and was discovered prior to BCS theory using phenomenological considerations. Nowadays we can of course derive it from BSC theory, but again this is a task for a condensed matter physics course.

Relabelling our fields

$$\Psi = \sqrt{\rho}e^{i\theta}, \tag{8.84}$$

we get the Lagrangian

$$\mathcal{L} = \rho\left(2e\phi - \frac{\partial\theta}{\partial t}\right) - \frac{1}{2m}\left[\frac{1}{4\rho}(\nabla\rho)^2 + \rho(\nabla\theta - 2e\mathbf{A})^2\right] - \frac{\lambda}{4}(\rho - \rho_0)^2 + \frac{1}{2}\mathbf{E}^2 - \frac{1}{2}\mathbf{B}^2. \tag{8.85}$$

Just like for the abelian Higgs mechanism, we can now absorb the θ field into the gauge field

$$\phi \to \phi - \frac{1}{2e}\frac{\partial\theta}{\partial t}, \quad \mathbf{A} \to \mathbf{A} - \frac{1}{2e}\nabla\theta. \tag{8.86}$$

You should check that this leaves the electromagnetic Lagrangian $\mathbf{E}^2 - \mathbf{B}^2$ unchanged. This redefinition gives us the new Lagrangian

$$\mathcal{L} = 2e\rho\phi - \frac{1}{2m}\left[\frac{1}{4\rho}(\nabla\rho)^2 + 4e^2\rho\mathbf{A}^2\right] - \frac{\lambda}{4}(\rho - \rho_0)^2 + \frac{1}{2}\mathbf{E}^2 - \frac{1}{2}\mathbf{B}^2. \tag{8.87}$$
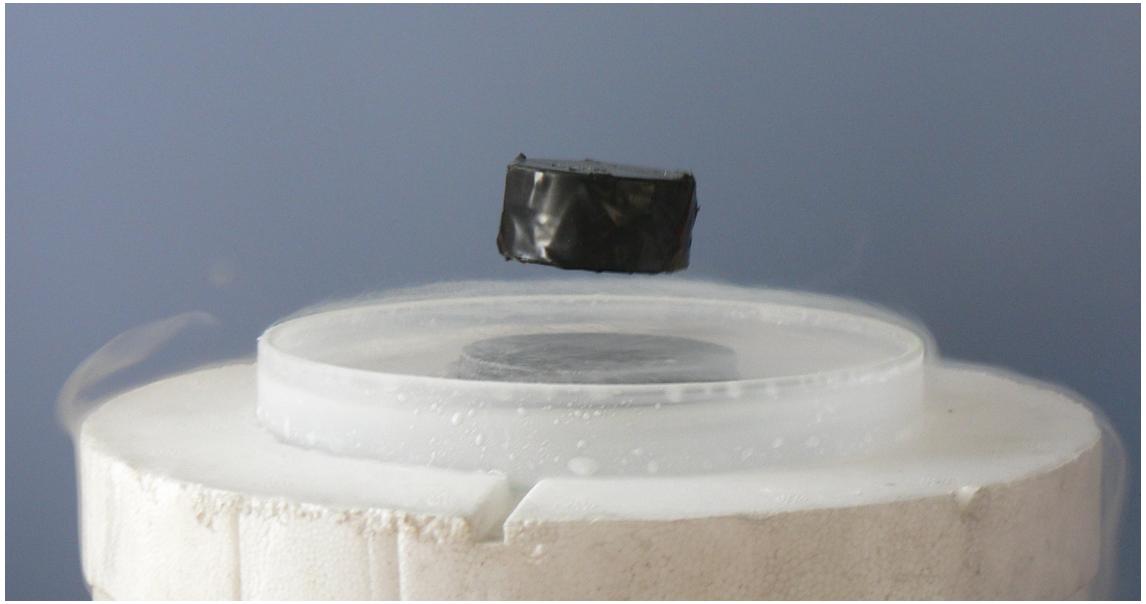
Figure 8.3: A magnet floating above a superconductor. Because magnetic fields cannot penetrate the superconductor, a force is generated which repels the magnet. Image from *https://commons.wikimedia.org/wiki/File:Meissner_effect_p1390048.jpg*.

From this we calculate the Hamiltonian

$$H = \int d^3x \; 2e\rho\phi + \frac{1}{2m}\left[\frac{1}{4\rho}(\nabla\rho)^2 + 4e^2\rho\mathbf{A}^2\right] + \frac{\lambda}{4}(\rho - \rho_0)^2 + \frac{1}{2}\mathbf{E}^2 + \frac{1}{2}\mathbf{B}^2. \qquad (8.88)$$

8.5.1 The Meissner Effect

Now we will assume that there is no electric field, and that the magnetic field is weak. The energy will then be minimised when $\rho(\mathbf{x}) = \rho_0$. We hence expand $\rho = \chi + \rho_0$, and find that

$$H = \int d^3x \; \frac{1}{8m}\frac{(\nabla\chi)^2}{\rho_0 + \chi} + \frac{2e^2\rho_0}{m}\mathbf{A}^2 + \frac{2e^2}{m}\chi\mathbf{A}^2 + \frac{\lambda}{4}\chi^2 + \frac{1}{2}\mathbf{B}^2. \qquad (8.89)$$

Usually, the energy of a constant magnetic field is $V\mathbf{B}^2$, where V is the volume of the system. But in a superconductor, we have an additional term which is proportional to $\mathbf{A}^2$. Since

$$\nabla \times \mathbf{A} = \mathbf{B} \qquad (8.90)$$

the potential $\mathbf{A}$ grows linearly in a region where $\mathbf{B}$ is constant. In this situation, the total energy will grow much faster than V. For this reason the magnetic field is expelled from the superconductor. This is known as the *Meissner effect*, and as can be seen in *Figure 8.3*, it is not difficult to demonstrate experimentally.

London Penetration Depth and Coherence Length

Say we applied a magnetic field to our superconductor. The magnetic field will not discontinuous vanish within the superconductor; rather, we expect the field to decay to zero with some characteristic length scale. For weak fields the variation in field density χ will

be negligible, and we can simply minimize

$$H = \int_V \frac{2e^2\rho_0}{m}\mathbf{A}^2 + \frac{1}{2}\mathbf{B}^2 \tag{8.91}$$

over the volume of the superconductor V, subject to our boundary conditions on the magnetic field. This gives us a characteristic length scale

$$l_P = \sqrt{\frac{m}{4e^2\rho_0}}, \tag{8.92}$$

known as the *London penetration length*. With a bit more work varying (8.91), we can show that the magnetic field satisfies the equation

$$l_P^2 \nabla^2 \mathbf{B} = \mathbf{B}. \tag{8.93}$$

If we have a superconductor on the half-plane $x > 0$, and a constant magnetic field $\mathbf{B}_0$ in the region $x < 0$, then the magnetic field will exponentially decay in the superconductor

$$\mathbf{B}(x) = \mathbf{B}_0 e^{-x/l_P} \qquad \text{for } x > 0. \tag{8.94}$$

Likewise we can consider variations of the density χ. When there is no magnetic field, the Hamiltonian becomes

$$H = \int d^3x \, \frac{1}{8m}\frac{(\nabla\chi)^2}{\rho_0 + \chi} + \frac{\lambda}{4}\chi^2 \approx \int d^3x \, \frac{1}{8m\rho_0}(\nabla\chi)^2 + \frac{\lambda}{4}\chi^2. \tag{8.95}$$

Notice the similarity in form to (8.91). The characteristic length for density fluctuations is known as the *coherence length*:

$$l_C = \frac{1}{\sqrt{2\lambda m\rho_0}}. \tag{8.96}$$

Varying (8.95), we discover that χ satisfies the differential equation

$$l_C^2 \nabla^2 \chi = \chi. \tag{8.97}$$

Hence if we have a superconductor in the half-plane $x > 0$, and if at $x = 0$ the density was χ_0, then

$$\chi(x) = \chi_0 e^{-x/l_C}. \tag{8.98}$$

These two length scales, l_P and l_C, govern the behaviour of a superconductor. Their ratio has a particularly simple form:

$$\frac{l_P}{l_C} = \sqrt{\frac{m}{4e^2\rho_0}} \times \sqrt{2\lambda m\rho_0} = \frac{m}{e}\sqrt{\frac{\lambda}{2}}. \tag{8.99}$$