Perturbative Quantum Field Theory

Jack Dinsmore

June 15, 2023

Contents

Abstract

Quantum Field Theory, or QFT, is our most fundamental description of the universe. It combines the lessons of quantum mechanics with the laws of special relativity to create a theory consistent with both. Except for gravity, quantum field theory predicts every phenomenon yet observed, with only twenty or so free parameters to be determined by experiment. Below, we provide a **JTD: number** page summary of QFT, with links to the chapters in which each topic is described.

The need for QFT arose from a variety of sources. Theorists were bothered by the fact that quantum mechanics is inconsistent with relativity (for example, the Schrödinger equation is not Lorentz invariant), while relativity is inconsistent with quantum mechanics. Experimentally, new particles such as the positron were starting to appear and a theory that could create, destroy, and propagate these particles was necessary. With the benefit of hindsight, we present in chapter 1 a principle of least action and argue that it is consistent with relativity and quantum mechanics.

The left hand side of this principle of least action is an n point correlation functions of particles (nPCFs), where n is a positive integer. The right hand side is an integral over every path the n particles could take of e^{iS}, where S is the action. An nPCF may seem like an esoteric object at first, but we show that it is directly connected to scattering (chapter 2). For example, the 4PCF encodes the probability for two particles to scatter off each other. The 2PCF is special; it encodes many of the characteristic properties of a particle, including its mass, its ability to scatter with and decay into other particles, and even the energies of its bound states (chapter 3).

Now that the nPCFs are understood, we tackle the task of solving for an nPCF using the QFT principle of least action. The result depends on the action S used, so we suggest a simple action for a particle with no spin, called a scalar particle (chapter 5). We first solve for the nPCFs in a theory without interactions. This can be done exactly, and using our understanding of how nPCFs are related to scattering, we confirm some of the properties we expect a relativistic theory such as QFT to have.

(For example, we confirm that the theory is causal.)

Next, we solve for the nPCFs when the scalar particles are allowed to interact with each other. This cannot be done exactly, so we use perturbation theory. The mathematics of this process are long, but are greatly simplified by the use of Feynman diagrams. After a few calculations, we are able to compute our first scattering amplitudes for an interacting theory. This process leads to a remarkable intuition which holds for any weakly interacting theory: all interactions come about by the creation of virtual particles, which are like a temporary excited quantum state used by the interacting particles to tunnel into another state. Just as an electron can tunnel through an energy barrier in non-relativistic quantum mechanics, now a relativistic particle uses the relativistic equation $E = mc^2$ to create virtual particles as a means of tunneling through an energy barrier.

We also show that a light scalar particle is able to carry information about the presence of other particles over long distances. To an observer who cannot see the light scalar particle, this may look like a long-range force. For a light enough scalar, we derive that this force is proportional to $1/r^2$, like Coulomb's force. However, we also suggest that it is unlikely that scalar particles can carry forces like this in Nature as it theories which have light enough scalars are rare.

Having treated the simplest case of a scalar particle, we consider particles with spin one-half (chapter 6), or fermion. Immediately, we notice that these fermions must behave differently from scalars because of the Pauli exclusion principle, which says that two fermions cannot exist in the same quantum state. Fortunately, the spin statistics theorem allows us to encode this Pauli exclusion principle in the action, so we can derive nPCFs for fermions using the same techniques as with scalars while remaining consistent with the Pauli exclusion principle.

With these nPCFs, we propose a new model that binds the neutron to the proton in a nucleus.The model treats protons and neutrons as fermions and proposes a light scalar known as a pion to transfer force between them. We compute several predictions of this model which are consistent with experiment to first order, though experiments also tell us that protons, neutrons, and pions have substructure which this model does not contain. Emboldened, we suggest a similar model for the Coulomb force between a positron and an electron. However, this model cannot explain magnetic fields, so we reject it. **JTD: Symmetry breaking**

In order to pursue magnetic fields, we further consider spin-one particles (chapter 7). Like scalars, these particles are bosons so they do not obey the Pauli exclusion

principle and can carry forces. However, a spin-one particle has too many degrees of freedom to be detectable in four dimensions. We propose the removal of one of these degrees of freedom by suggesting that the spin-one particle obeys some self-symmetry, in that a certain transformation of the spin-one field leaves the action invariant. This removes one physical degree of freedom from the particle, making it observable. It also limits the number of possible actions down to a very small number. We choose the simplest action, allow the spin-one particle to interact with the electron, and we see that the spin-one particle has all the properties of a photon. This spin one-half electron plus a spin-one photon theory is named Quantum Electrodynamics, or QED. We spend the rest of the chapter showing that it reproduces magnetism and predicts that an electron has a magnetic dipole moment of about two, consistent with experiment. We also show that this theory guarantees the photon to be massless.

We would like to compute the predictions of QED to see how well they hold. We start with electron to electron scattering, but **JTD: infrared** . We move on to compute the magnetic moment of the electron g, but again we find an infinite prediction. We interpret this infinity as an unphysical prediction coming from the fact that the virtual particles that the electron uses to tunnel between states can be very high energy. We don't necessarily expect QFT to be valid at very high energies where we have not tested the theory. We therefore invent a technique known as renormalization to subtract off these high-energy predictions, leaving us only with the predictions of lower-energy QFT. Using this technique, we show that QED agrees exactly with experiment in the magnetic moment of the electron, the spectrum of hydrogen, and the probability for electrons and positrons to annihilate. We also show a few additional calculations which reveal new properties of electrons and photons which non-relativistic quantum mechanics was ignorant of.

JTD: Non-abelian

Notation

Units

In quantum mechanics, observable quantities have units of mass, length, time, and charge. This book collapses these four units down to one by defining Planck's reduced constant, the speed of light, and permittivity of free space to be $\hbar = c = \epsilon_0 = 1$. It's customary to keep mass as the remaining unit and define the other fundamental units in terms of mass:

$$[\text{length}] = [\text{mass}]^{-1}, \qquad [\text{time}] = [\text{mass}]^{-1}, \qquad [\text{charge}] = [\text{mass}]^{0}. \quad (1)$$

in four spacetime dimensions. Such a unit system is called *natural units*. It's also common just to quote the exponent of [mass], saying for example that length has dimension -1.

To convert a number from natural units back to a more standard unit system, one need only multiply by the correct combination of $\hbar$, c, and ϵ_0. As an example, the fine-structure constant $\alpha = e^2/(4\pi)$ is $\alpha = e^2/(4\pi\epsilon_0\hbar c)$. Other important dimensions are

$$\begin{gathered}[\text{Action}] = 0, \qquad [\text{Lagrangian}] = d, \qquad [\partial_\mu] = 1, \\ [G_{\text{Newton}}] = -2, \qquad [G_{\text{Fermi}}] = -2\end{gathered} \quad (2)$$

where d is the dimension of spacetime.

Special Relativity

We will use the mostly-minus form of the metric tensor for special relativity:

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2 \qquad \text{or} \qquad \eta_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{3}$$

which is equivalent to $\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$. A result of this choice is that a timelike vector has positive norm:

$$\eta_{\mu\nu} p^\mu p^\nu = p \cdot p = p^\mu p_\mu = p^2 > 0. \tag{4}$$

If this momentum represents a particle, then the mass of the particle is defined by $p^2 = m^2$. This is known as the on-shell condition. When we refer to the spacial part of a four vector, we use boldface:

$$\boldsymbol{p} = (p_1, p_2, p_3) \tag{5}$$

so that we can also write a particle mass as

$$m^2 = p_0^2 - \boldsymbol{p}^2. \tag{6}$$

The energy-mass relation of special relativity, $E^2 = m^2 + p^2$, indicates that $|p_0|$ is the energy of the particle. We therefore often refer to p_0 as E_p and require that $p_0 = E_p > 0$.

Fourier Transforms

Our normalization for the Fourier transform and its inverse is as follows:

$$f(k) = \int d^d x \, e^{-ik\cdot x} f(x) \qquad f(x) = \int \frac{d^d k}{(2\pi)^d} \, e^{ik\cdot x} f(k) \tag{7}$$

where d is the number of space(time) dimensions used. To prove that the Fourier transform and inverse Fourier transform are indeed inverses, the following formula is useful:

$$(2\pi)^d \delta(k) = \int d^d x e^{ik\cdot x}. \tag{8}$$

Chapter 1

Principles of Quantum Field Theory

Historical Introduction

The first three decades of the 20th century marked an unparalleled advance in fundamental physics. Sparked by the observation black body radiation — specifically its sharp frequency cutoff which suppresses high-energy radiation — Max Planck postulated in 1900 that the energy of light had to come in *quanta*. Inspired, Albert Einstein used this idea of light quanta to explain the photoelectric effect, and Niels Bohr and Ernest Rutherford generalized the idea of quanta by suggesting in 1913 that angular momentum was also quantized in order to explain atomic emission lines. With Louis de Broglie's proposal in 1924 that matter was also quantized, it became rare to find properties of nature that were *not* quantized — a statement made rigorous by the discovery of the Schrödinger equation in 1926, and the foundation of non-relativistic quantum mechanics.

Independently, Maxwell's prediction and the following experimental confirmation that the speed of light be constant had initiated its own cascade of discoveries. In 1905, Einstein reshaped our definitions space and time by uniting them into one entity — spacetime — all in order to apply Maxwell's equations of electromagnetism on moving bodies. His new theory is now called special relativity. Ten years later, he unified special relativity with gravity to produce general relativity, which explained the slow precession of Mercury's orbit, and now explains much more exotic phenomena such as black holes and the history of the universe.

Though these two theories of relativity and quantum mechanics were each incredibly successful, they were completely in conflict. The famous Schrödinger equation, though in agreement with experiment, contained one time derivative and

two position derivatives:[1]

$$i\partial_t\psi = \left(-\frac{1}{2m}\partial_x^2\right)\psi. \tag{1.1}$$

We've set potential energy equal to zero assuming that the particle is in vacuum. According to special relativity, a Lorentz boost to a new reference frame with velocity v must introduce the substitutions $\partial_t \to \partial_t + v\partial_x$ and $\partial_x \to \partial_x + v\partial_t$. But that would introduce additional space and time derivatives which no longer fit the form of (1.1). Thus, the Schrödinger equation is inconsistent with relativity.

A physicist in the 20s could have guessed at the qualitative behavior of a fully complete theory of physics; it should both inherit the core principles of both relativity and quantum mechanics, including the possibility of quantum tunneling and the famous statement of mass-energy equivalence — $E = mc^2$. Together, these two imply that a particle could "tunnel" from energy E and mass m to a different energy E' and mass m', effectively switching to an entirely new particle type on its own. Thus, a relativistic quantum theory should predict particle creation, annihilation, and decay. It would take another forty years for the machinery to generate these predictions was worked it.

Aside from this theoretical problem with late 1920s-era physics, a few experimental anomalies had begun to appear. It was becoming increasingly apparent that electron spin, largely pioneered by Pauli, needed to be introduced to quantum mechanics. The spectrum of hydrogen, once the poster child of the quantization of angular momentum, was showing a small discrepancy with the theoretical prediction of s orbital energy. Separately, Compton scattering experiments where high energy photons were scattered off electrons showed that light had important quantum mechanical properties which could not be understood with the Schrödinger equation.

There were two examples of equations of motion already introduced in physics that were consistent with relativity: one was Maxwell's equations (in vacuum), written in terms of the vector potential A^μ

$$\partial_\mu\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) = 0 \tag{1.2}$$

and the other was the wave equation, also called the Klein-Gordon equation, writ-

[1] This is the first time we have used natural units, wherein $\hbar = c = 1$.

ten for a wavefunction ϕ:

$$\partial^\mu \partial_\mu \phi^2 - m\phi^2 = 0. \tag{1.3}$$

Both of these equations are Lorentz invariant, though they are not good replacements for the Schrödinger equation since they do not have the same non-relativistic limit. For the same reason, they also don't resolve the new experimental anomalies. Instead, Paul Dirac resolved them in 1928 with his famous Dirac Equation. Like the Schrödinger equation, the Dirac equation is a one-derivative equation and consequentially reproduces the Schrödinger equation in the non-relativisitic limit. In order to contract away the index of the single derivative ∂_μ, the Pauli matrices σ_1, σ_2, and σ_3 had to be combined with the identity into a four-vector $\sigma^\mu = (\mathbb{1}, \sigma_1, \sigma_2, \sigma_3)$ which represents the spin operator.[2]

$$i\left(\partial_\mu \sigma^\mu\right)\psi = 0. \tag{1.4}$$

This was the first fundamental theory to describe particle spin. It also allowed the anomalous energy of the s orbitals to be understood as a relativistic correction to the Schrödinger equation, due to the electron's high velocity. The same theoretical tools addressed the Compton effect, with its relativistic electrons.

However, the Dirac equation had one glaring flaw. It predicted the existence of a new particle of equal mass to the electron but with opposite charge — a particle which somehow had never been seen despite the recent surge of particle physics experiments. Ashamed of his prediction, Dirac postulated that the proton was this new electron-like particle, even though experiment had already shown that the proton was much more massive than the electron.

Four years later, it appeared Dirac should have stuck to his guns. The positron was discovered by Carl David Anderson and confirmed to have the same mass and opposite charge of the electron, though it decayed not long afterwards. However,

[2](1.4) only applies to massless particles, which is why m is not present in the equation. It can be generalized to apply to massive particles such as electrons with some extra mathematical machinery, which is done in the main text.

The fact that spin and Lorentz invariance were resolved with the same blow in (1.4) is no accident; spin, also known as intrinsic angular momentum, is only a valid particle property because it is conserved, and angular momentum is only conserved because physics is invariant under rotations. Lorentz invariance implies invariance under rotations, so any theory which is Lorentz invariance should also come with a complementary interpretation of angular momentum (or spin).

the new particles didn't stop there. The neutron was discovered in the same year, and muon not long after in 1936. Together with these new particles was the ever-growing theory of nuclear decay, describing how protons, neutrons, and electrons conspired to turn into each other with fixed rates. No theory yet written allowed particles to change identity or split in such a way.

Keeping pace with experiment, many physicists had embarked on several methods of promoting the Dirac equation and Maxwell's equations to a "Quantum Field Theory," or QFT, which could handle the wavefunctions of several particles at once, allow them to interact and, crucially, to decay. Progress was slowed by the temperamental behavior of energy; the vacuum energy predicted by the Schrödinger became infinite in a relativistic version, as did other quantities such as particle masses. Some physicists began to suggest abandoning the new QFT framework altogether. However, the occasional success kept many physicists on the QFT track. For example, a new anomaly of the Hydrogen s orbital known as the Lamb shift was discovered in 1947 and explained by Hans Bethe using QFT, and a correction to the magnetic moment of the electron was derived in 1948 by Julian Schwinger and experimentally confirmed.

The theoretical troubles of QFT were redressed over the next decade through a number of procedures largely developed by Schwinger, Richard Feynman, and Shinichiro Tomonaga, for which they one the 1965 Nobel Prize in Physics. Their theory, named Quantum Electrodynamics (QED), explains all the phenomena of classical mechanics, non-relativistic quantum mechanics, and special relativity, as well as confirming all the new discoveries mentioned above. As of the 2000s, its predictions have been confirmed to as many as 12 decimal places in multiple experiments. The theory also addressed several philosophical questions, such as the origin of particle spin and the definitions of fermions and bosons. This book is largely devoted to understanding this theory and its consequences.

In the years since, many new particles have been discovered and the laws of QED have been generalized to the Standard Model, which is the most successful theory ever created. It is consistent with all scattering experiments performed in any collider, confirms the mass of nucleons such as the proton and neutron, and agrees with all measured particle properties except the neutrino masses. In the rest of this chapter, we summarize the physical principles underlying the standard model before introducing them in mathematical detail in the preceding chapters.

The Principle of Least Action

Despite the careening history of physics development, with its many different and inconsistent theories, every fundamental theory before QFT (classical mechanics, relativity, and quantum mechanics) had a principle of least action. It seems reasonable to pursue such a principle for QFT.

In classical mechanics, action S is defined as the integral of some function L with dimensions of energy, known as the Lagrangian, with respect to time. The integral is computed over the path C taken by a particle:

$$S = \int_C dt\, L. \tag{1.5}$$

This definition is combined with the physical law that Nature choses the path C such that the action is minimized. Thus, any small change to C does not lower the action:

$$\delta S = 0. \tag{1.6}$$

(1.6) is known as the principle of least action.

The advantages of the action formulation are many. Firstly, it reduces all the properties of a physical theory to one variable: the Lagrangian. All a physicist need do to create a theory is write down a new form of L. Furthermore, a physicist is aided in the process of writing a new Lagrangian by the fact that all the symmetries of S are symmetries of the theory in general. For example, we have observed that physics is invariant under translation — that is, that experiments done in different locations yield the same results. Thus, S cannot depend explicitly on position. Other examples of symmetry are invariance under orientation, and time translation. Relativity introduces Lorentz boosting as a symmetry, while QFT will introduce even more. All of these symmetries cut down so much on the number of forms S can take that Lagrangians tend to be quite simple in practice. In classical mechanics, the Lagrangian is

$$L_{\text{Classical}} = K - V \tag{1.7}$$

where K and V are kinetic and potential energy. All other terms are forbidden because they do not have the right units or the right symmetries. For example, a term proportional to momentum is not invariant under rotation.

In general relativity, the purpose of the Lagrangian is to dictate how the metric of spacetime $g_{\mu\nu}$. This requires us to change our thinking slightly, since the classical Lagrangian (1.7) was for a particle and $g_{\mu\nu}$ is not a particle. The solution is to think in terms of degrees of freedom. In classical mechanics, we minimized the action with respect to the position $x(t)$ and momentum $p(t)$ of a particle at each time. Now we change our degrees of freedom to $g_{\mu\nu}(x)$ — the metric tensor at every point in space — and employ the same procedure.

The time integral in the definition of action is not Lorentz-invariant, but the full spacetime volume element $d^4x\sqrt{-g}$ is.[3] Thus we must define the Lagrangian as

$$L = \int d^3x\, \mathcal{L} \implies S = \int d^4x\, \mathcal{L}(x). \tag{1.8}$$

where $\mathcal{L}(x)$ is called the *Lagrangian density*. Here we introduce a further principle, namely that general relativity is a *a local theory*. That is, the physics of one point in spacetime cannot be affected by a that of a distant point. Since $\mathcal{L}(x)$ encapsulates the physics at a point, it must only depend on x and derivatives with respect to x. In practice, the way information is communicated across spacetime is by waves, much like the electromagnetic waves of Maxwell's equations, which is also a local theory.

Putting together the notion of a Lagrangian density with the requirement of Lorentz invariance and the principle that it must only depend on local variables, the general relativity Lagrangian density is

$$\mathcal{L}_{\text{GR}} = \sqrt{-g}\,(R - 2\Lambda) + \mathcal{L}_{\text{Matter}}. \tag{1.9}$$

Here, R is the Ricci scalar (the simplest non-constant Lorentz-invariant scalar), and $\mathcal{L}_{\text{Matter}}$ is the Lagrangian density of matter that happens to be present. Λ is known as the cosmological constant which roughly represents the energy of the vacuum.

We often call $g_{\mu\nu}$ a field because it takes on a value at every point x in spacetime. This then represents the first generalization of the principle of least action to a field — that is, our first *field theory*. We could also treat the Klein-Gordon equation (1.3) or the Dirac equation (1.4) as classical equations of motion for fields ϕ and ψ, and

[3] In the general relativistic Lagrangian density, g is the determinant of the metric tensor $g_{\mu\nu}$, but we won't discuss it further because this book deals with flat spacetime.

we would get Lagrangian densities

$$\mathcal{L}_{\text{KG}} = \phi(\partial_\mu \partial^\mu - m)\phi, \qquad \mathcal{L}_{\text{Dirac}} = i\psi^\dagger(\sigma^\mu \partial_\mu)\psi$$
$$\mathcal{L}_{\text{EM}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{1.10}$$

where $F^{\mu\nu}$ is the Faraday tensor $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$. However, these field theories are still classical because they do not produce quantized particles. Covering the principle of least action for a quantum theory is the next step.

Quantum Principle of Least Action

The success of quantum mechanics is largely due to the presence of the new constant $\hbar$, which has units of energy times time. $\hbar$ is necessary to quantize the energy of light, angular momentum, and to write the Schrödinger equation, but $\hbar$ poses a problem for the principle of least action. Action has the same units as $\hbar$, so varying the path of a quantum particle could in principle change the action by any amount $\delta S = x\hbar$, where x is a unitless number. Furthermore, a true quantum particle doesn't have a path. It has a wavefunction, and may exist anywhere within that wavefunction at any time.

Richard Feynman developed a simple solution to both these problems in the 40s. To compute the probability that a quantum particle could go from position $|x(0)\rangle$ at time $t = 0$ to position $|y(t)\rangle$ at time t, one need only calculate

$$\langle \boldsymbol{x}(t)\boldsymbol{y}(0)\rangle = \sum_{\text{paths } C} \exp\left[-\frac{i}{\hbar}\int_C dt\, L\right]. = \sum_{\text{paths } C} e^{-iS/\hbar}. \tag{1.11}$$

The sum over paths runs over all possible paths from $\boldsymbol{x}$ to $\boldsymbol{y}$, with each weighted by a phase of $e^{-iS/\hbar}$, where S is computed over the path in question. This new action formulation acknowledges that quantum particles don't follow paths by equivalently stating that they follow *every* path. It also resolves the question of units by dividing action by $\hbar$. We'll now go back to natural units where $\hbar = 1$.

Surprisingly, (1.11) is almost equivalent to the classical principle of least action $\delta S = 0$. Consider a specific C, which corresponds to action S_C. This path contributes a term of e^{iS_C} to $\langle x(0)|y(t)\rangle$, and all paths that differ only by a small action δS contribute contribute $e^{iS_C}e^{i\delta S}$. When added, the positive δS of some

terms destructively interfere with the negative δS of other terms, and the interference grows worse as larger and larger δS are considered. However, if C is the path that minimizes action, then there are no paths with negative δS, and there is no destructive interference. Thus, the paths that contribute the most to $\langle x(0)|y(t)\rangle$ are those that minimize or nearly minimize the action.[4]

Quantum Field Theory Lagrangian

Quantum field theory — the relativistic version of quantum mechanics — needs to share the field theory properties of relativity and the quantum properties of non-relativistic quantum mechanics. We'll start by borrowing the Lagrange density L of general relativity and the notion of minimizing action with respect to fields — except this time we replace the field $g_{\mu\nu}$ with a wavefunction $\psi(x)$. The we interpret the notion of summing over paths as integrating over all possible values of $\psi(x)$. This integration, called *path integration* deserves a shorthand:

$$\prod_{x \,\in\, \text{spacetime}} \int \psi(x) \equiv \int \mathcal{D}\psi. \tag{1.12}$$

The last step is to interpret the left hand side of (1.11) in a Lorentz invariant way. We start with the vacuum $|0\rangle$, which must be Lorentz invariant. If we define an operator $\hat{\psi}(x)$ which creates a particle at four-position x, then $\hat{\psi}(x)\,|0\rangle$ is a Lorentz-invariant, single-particle state. The non-relativistic quantum mechanics notation for such a state was $|\boldsymbol{x}(t)\rangle$x, where $x = (t, \boldsymbol{x})$. Thus, the natural relativistic analog of the non-relativistic $\langle \boldsymbol{x}(t)|\boldsymbol{y}(0)\rangle$ in (1.11) is $\langle 0|\hat{\psi}(\boldsymbol{x})\hat{\psi}(\boldsymbol{y})|0\rangle$.

Given these insights, we extend (1.11) to the relativistic case as follows

$$\langle 0|\hat{\psi}(x)\hat{\psi}(y)|0\rangle \propto \int \mathcal{D}\psi\, \psi(x)\psi(y) \exp\left[-i \int d^4x\, \mathcal{L}\right]. \tag{1.13}$$

The proportionality is present because we have not yet normalized the wavefunction, but we can easily do this by dividing by $\langle 0|0\rangle$:

$$\langle 0|\hat{\psi}(x)\hat{\psi}(y)|0\rangle = \frac{\int \mathcal{D}\psi\, \psi(x)\psi(y)e^{-iS}}{\int \mathcal{D}\psi\, e^{-iS}}, \tag{1.14}$$

where we've used the definition of action $S = \int d^4x\, \mathcal{L}$. Generalized to arbitrary expectation values,

[4] The rigorous version of this statement is known as the "stationary phase approximation."

$$\langle 0|\hat{\psi}(x_1)\cdots\hat{\psi}(x_n)|0\rangle = \frac{\int \mathcal{D}\psi\, \psi(x_1)\cdots\psi(x_n)e^{-iS}}{\int \mathcal{D}\psi\, e^{-iS}}. \tag{1.15}$$

This law, due to Feynman, is the heart of QFT. It can be used to compute the outcome of scattering experiments, particle properties such as magnetic moments, and bound states like nuclei. The first two in particular can be done with the perturbative framework known as Feynman diagrams, which this book focuses on. Along with the quantitative predictions, (1.15) gives deep insight as to the nature of particles, since it can be used to derive the spin-statistics theorem which explains why spin-$1/2$ particles obey the Pauli exclusion principle.

Only a Few Lagrangians are Possible

(1.15) also suppresses the influence of complex terms in the Lagrangian have on experiments, leaving only a few terms which we can observe. This can be seen as follows. Consider a scattering experiment, where two particles ψ were collided with energy E. If $\mathcal{L}$ had a term like $\lambda\psi^4$, the constant λ would need to have mass dimension -2 to satisfy unit analysis. Whenever this constant appeared in the scattering cross section, it would need to be accompanied with an energy to allow the units to work: $\lambda/E^{-2} = \lambda E^2$. Thus, it is measurable at high energies but not at low energies. "High" energies here should be compared to the mass energy of a particle, so that E is high only for highly relativistic scenarios. Thus, these complex terms do not influence our low-energy environment.

There are only three terms which are Lorentz-invariant and low-dimension enough not to be suppressed. They are precisely the Lagrangians for Maxwell's equations, the Dirac equation, and the Klein-Gordon equation (1.10). Thus, the theory of quantum electrodynamics which simultaneously explains Maxwell's equations and the Dirac equation in the language of QFT, falls out naturally from (1.15). The only free parameters in the entire theory are the particle masses and value of the fine structure constant, $\alpha \approx 1/137$, which encapsulates the strength of the electric charge.

The fact that QFT narrows down the possible Lagrangian terms to only the ones Nature uses represents an incredible achievement of modern physics. We have boiled down incredibly complex phenomena to a mere handful of parameters using only the assumption of translational and Lorentz symmetry, locality, and the simple insights taken from the least action principles of other theories. The fantastic

predictiveness of QFT is not lost when we extend beyond QED to the broader standard model; even taking all particles into account in the Lagrangian, nearly every term allowed by QFT exists in Nature's Lagrangian.

The next part of this book is dedicated to understanding the n-point correlation functions, which are the left hand side of (1.15). The second and third parts are dedicated to computing the right hand side depending on which Lagrangian is used. This book discusses perturbative methods, known as Feynman diagrams, for calculating this right hand side, though non-perturbative methods also exist under the name of Lattice QCD.

JTD: Antiparticles and CPT symmetry. Also put in the daggers correctly for the QFT PLA

Part I

Non-Perturbative Techniques

Chapter 2

Scattering

This part is devoted to understanding n-point correlation functions, which are the output of QFT predictions. These functions have two major uses: one is to compute scattering cross-sections and decay probabilities (this chapter) while another is to compute the spectrum of a quantum particle (chapter 3). In these two chapters we will introduce important tools, but save actual predictions for later parts, where we introduce new Lagrangians and use them to compute the n-point correlation functions.

2.1 S Matrix and Scattering Amplitudes

A scattering experiment consists of shining a beam of particles with fixed four-momentum p_1 on one or more other beams with their own four-momenta p_2, p_3, $\ldots$, p_n. We call this initial state $|\boldsymbol{p}\rangle$. The momenta of outgoing particles q_1, q_2, $\ldots$, q_m are measured as the final state $|\boldsymbol{q}\rangle$. We label these states only with three-vectors $\boldsymbol{p}$ rather than four-vectors p because of the constraint $p^2 = m^2$, which makes the fourth component redundant. These states will be defined precisely in terms of correlation functions later, but for now we assert that they are Lorentz invariant. Then the probability of scattering from $|\boldsymbol{p}\rangle$ to $|\boldsymbol{q}\rangle$ is $|\langle \boldsymbol{q}|\boldsymbol{p}\rangle|^2$ and is also Lorentz invariant. We therefore define

$$iS_{q\to o} = \langle \boldsymbol{q}|\boldsymbol{p}\rangle \tag{2.1}$$

and $|S|^2$ is the observable probability of the scattering. This quantity is called the S-matrix. The coefficient i is included due to tradition.

If no interactions occur, we expect $n = m$ and $p_j = q_j$ for all j. Then $iS \propto \prod_i \delta(p_j - q_j)$. Even when interactions are possible, many particles will still not

interact and will pass through the beam unchanged. We therefore define a new quantity M such that

$$iS_{p\to q} = \prod_i \delta(p_j - q_j) + iT_{p\to q} \tag{2.2}$$

where T is called the T-matrix, which quantifies the interactions that occur during scattering.

Due to conservation of momentum, $\sum_i p_i = \sum_j q_j$ even when interactions are present. To extract this fact we make one more definition:

$$iT_{p\to q} = (2\pi)^4 \delta\left(\sum_i p_i - \sum_j q_j\right) iM_{p\to q} \tag{2.3}$$

where M is called the *scattering amplitude* and is Lorentz invariant. Unlike S and T it has finite matrix elements. Whenever we compute a prediction for a scattering experiment in this book, calculating M will be our goal.

2.2 Initial and Final Scattering States

Until now, we have made no reference to QFT. The M amplitude can be defined in a non-relativistic theory or even classical mechanics. Rather, QFT comes in when we define the Lorentz-invariant $|\boldsymbol{p}\rangle$ and $|\boldsymbol{q}\rangle$ states in terms of the $\hat{\psi}(x)$ operator introduced in chapter 1.

The definition of $\hat{\psi}(x)$ was that it creates a particle at position x when acting on the vacuum $|0\rangle$. We need to now define a $\hat{\psi}(p)$ which creates a particle with momentum p but undefined position. Inspired by non-relativistic quantum mechanics, we'll define $\hat{\psi}(p)$ via the Fourier transform.

$$\hat{\psi}(p) = \int d^4x\, e^{-ik\cdot x}\hat{\psi}(x). \tag{2.4}$$

The dot product in the exponent is the dot product of four-vectors: $k_\mu x^\mu = Et - \boldsymbol{k}\cdot\boldsymbol{x}$ where $E = k^0$ is the particle kinetic energy and t is the time. Thus, (2.4) states that $\hat{\psi}(p)$ is just a plane wave of $\hat{\psi}(x)$ operators. This is consistent with our intuitive understanding of momentum.

The physics of n particles with momentum $p_1, \ldots p_n$ scattering to m particles

with momentum $q_1, \dots q_n$ is contained in the correlation function

$$\langle \hat{\psi}(x_1) \dots \hat{\psi}(x_n) \hat{\psi}^\dagger(y_1) \dots \hat{\psi}^\dagger(y_m) \rangle \,.$$

Writing the correlation function in terms of momenta using (2.4),

$$\langle \boldsymbol{p} | \boldsymbol{q} \rangle = \langle \hat{\psi}(p_1) \dots \hat{\psi}(p_n) \hat{\psi}^\dagger(q_1) \dots \hat{\psi}^\dagger(q_m) \rangle = \prod_i^n \int d^4x_i \, e^{ip_i \cdot x_i} \prod_j^m \int d^4y_j \, e^{-iq_j \cdot y_j} \, \langle \hat{\psi}(x_1) \dots \hat{\psi}(x_n) \hat{\psi}^\dagger(y_1) \dots \hat{\psi}^\dagger(y_n) \rangle \,. \tag{2.5}$$

A strange aspect of the above equation is that it does not enforce the so-called *on-shell* constraint, that a particle with mass m and momentum p satisfies $p^2 = m^2$ always. The name on-shell stems from the fact that $p^2 = m^2$ is the equation of the "shell" of a hyperbola in $\mathbb{R}^4$. In fact, the on-shell constraint is effectively enforced because the spacetime integral in the second line constructively interferes over different values of x_i and y_j as long as all the momenta are on shell. This causes (2.5) to rise to infinity for physical momenta. Otherwise, destructive interference will occur and (2.5) remains finite. If this assertion seems doubtful, the reader can check that (2.5) is infinite when we compute scattering amplitudes in later chapters.

This concludes our discussion of the connection between correlation functions and scattering amplitudes, because (**??**) is valid now even when interactions are present. The reasoning is simple: the observer should not be able to tell based on the final state $|q\rangle$ whether it resulted from a sufficiently ancient scattering event or whether the state was always $|q\rangle$. Therefore, if (**??**) was valid for no scattering, it's also valid for a scattering in the past. The "sufficiently ancient" caveat will never matter for an experiment, since it is not possible to measure a particle as it is undergoing a scattering event. This is for the same reason it is not possible to observe a particle in the act of tunneling, or to observe Schrödinger 's cat die; part of the nature of quantum mechanics is that events such as scattering or tunneling can occur, even if they are never observed directly.

In some cases, a particle contains more information that just its momentum. From non-relativistic quantum mechanics, we know that fermions also each have spin, and from Maxwell's equations we know light has polarization. In these cases, $|\boldsymbol{p}\rangle$ will contain all the initial state's properties — momentum, spin, polarization, etc — and ditto for $|\boldsymbol{q}\rangle$. We will address this in greater detail when we apply (**??**) to

the case of fermions.

2.3 Decay Widths and Cross Sections

It's worth making the connection between scattering amplitudes and experiment more explicit. The connection is quite simple except for the fact that scattering amplitudes are Lorentz invariant and experiments, which are done at fixed velocities and orientations, are not. Scattering amplitudes and the outcome of scattering experiments can therefore differ by a energy-dependent coefficient, which we will work out by computing the normalization of the wavefunction $\psi(\boldsymbol{x})$

Let's consider a single particle of mass m scattering into n final particles, with momentum q_j. A single-particle scattering event is also known as particle decay, and the interesting quantity is the decay width Γ, which is the probability of scattering per unit time in the rest frame of the particle. The lifetime τ of the particle is related to Γ by $\tau = 1/\Gamma$.

Assume we know $|\boldsymbol{p}\rangle$ in terms of correlation functions $\langle\psi(p)\ldots\rangle$ as outlined in the previous section. We'll convert these fixed-momentum fields back to position fields via the inverse Fourier transform:

$$\psi(x) = \int \frac{d^4k}{(2\pi)^4}\, e^{-ik\cdot x}\delta(k^2 - m^2)\psi(k). \tag{2.6}$$

The delta function can be inserted with no penalty because of the on-shell constraint.

Since experiments measure the spatial part of momentum, we explicitly compute the k_0 integral. Using the fact that

$$\int dx\, \delta(f(x) - f(z)) = \frac{2\pi}{|f'(z)|}\delta(x - z), \tag{2.7}$$

we have

$$\psi(x) = \int \frac{d^3k}{(2\pi)^3}\, e^{-ik\cdot x}\frac{1}{\sqrt{2E_{\boldsymbol{k}}}}\psi(k) \tag{2.8}$$

where $E_{\boldsymbol{k}} = k_0 = \sqrt{m^2 - \boldsymbol{k}^2}$ is the energy of the particle at momentum $\boldsymbol{k}$. Integrating over all space to find the normalization gives

$$1 = \int d^3x\, |\psi(x)|^2 = \int \frac{d^3k}{(2\pi)^3}\, \frac{1}{2E_{\boldsymbol{k}}}|\psi(\boldsymbol{k})|^2 \tag{2.9}$$

The lesson to take from this is that the Lorentz-invariant Fourier transform $|\psi(\boldsymbol{k})|^2$, differs from the physical observable by a factor of $1/2E_{\boldsymbol{k}}$. In (**??**), we wrote scattering amplitudes in terms of Fourier transformed wavefunctions, so we should maintain this factor of $1/2E_{\boldsymbol{k}}$. Thus, the probability to decay into the final state $|\boldsymbol{q}\rangle$ is

$$\begin{aligned} d\Gamma &= \left(\frac{1}{2E_p}\right)\left(\prod_j \frac{d^3q_j}{(2\pi)^3}\frac{1}{2E_{q_j}}\right)\int \frac{d^3p'}{(2\pi)^3}\,\langle \boldsymbol{q}|\boldsymbol{p}\rangle\,\langle \boldsymbol{p}'|\boldsymbol{q}\rangle \\ &= \left(\frac{1}{2E_p}\right)\left(\prod_j \frac{d^3q_j}{(2\pi)^3}\frac{1}{2E_{q_j}}\right)|M_{p\to q}|^2(2\pi)^4\delta\left(p-\sum_j q_j\right). \end{aligned} \tag{2.10}$$

We integrated over $\boldsymbol{p}'$ in the first line because we don't know the initial momentum of an arbitrary particle drawn from the particle beam; we only now the final states as detected.

Similar techniques can be used to find the probability of interaction for multiple to multiple particle scattering; in this case, the probability is called a cross section σ. The differential cross section $d\sigma$ is very similar to the equation for $d\Gamma$, except that there may be additional coefficients due to the fact that multiple momenta $\boldsymbol{p}_i$ are now unknown. For example, for two incoming particles with velocity v_1 and v_2, the differential cross section is

$$\begin{aligned} d\sigma =& \frac{1}{|\boldsymbol{v}_1-\boldsymbol{v}_2|}\left(\prod_i \frac{d^3p_i}{(2\pi)^3}\frac{1}{2E_{\boldsymbol{p}_i}}\right)\left(\prod_j \frac{d^3q_j}{(2\pi)^3}\frac{1}{2E_{q_j}}\right) \\ &\times |M_{p\to q}|^2(2\pi)^4\delta\left(\sum_u p_i-\sum_j q_j\right). \end{aligned} \tag{2.11}$$

These cases of 1 and 2 particles cover most scattering events. To evaluate the total probability of scattering, $d\Gamma$ or $d\sigma$ may be integrated over all momenta. If identical particles A and B are present in the initial and final states, the states where A and B are switched should count only once. Thus, for k types of m_j identical particles, we should divide the cross section by $\prod_{j=1}^k m_j!$.

JTD: Need to fix this up, including defining σ, doing more math, and making it less hand-wavy

Chapter 3

Two Point Correlation Function Spectrum

The nPCFs were connected to scattering in the previous chapter, but the 2PCF in particular has an intriguing property known as the Källén-Lehmann spectral representation which we will remark in this chapter. Namely, the 2PCF $\langle\hat{\psi}(p)\hat{\psi}^\dagger(q)\rangle$ summarizes all of the important properties of the particle ψ, including its mass, decay rate, bound states, and scattering.

Let's build intuition for $\langle\hat{\psi}(p)\hat{\psi}^\dagger(q)\rangle$ by treating it in a scattering scenario as an S-matrix element. It would represent the probability for a particle to scatter from momentum p to momentum q, meaning $q = p$ by conservation of momentum. Also, due to Lorentz invariance, $\langle\hat{\psi}(p)\hat{\psi}^\dagger(p)\rangle$ cannot depend on the four-vector p^μ. It can only depend on the Lorentz invariant quantity p^2 and constants — that is, $\langle\hat{\psi}(p)\hat{\psi}^\dagger(p)\rangle$ is a function of one variable p^2.

We expect the on-shell constraint to force $p^2 = m^2$, meaning that $\langle\hat{\psi}(p)\hat{\psi}^\dagger(p)\rangle$ should blow up at $p^2 = m^2$ in order to suppress $p^2 \neq m^2$ effects. To show this, we define the *spectral-density* $\rho(\mu^2)$ such that

$$\langle\hat{\psi}(p)\hat{\psi}^\dagger(p)\rangle = \int d(\mu^2)\rho(\mu^2)\frac{i}{p^2-\mu^2} \tag{3.1}$$

and the on-shell constraint requires $\rho(\mu^2) = \delta(\mu^2 - m^2)$. That is, the spectrum has a Dirac spike at m^2.

We have produced this equation only using our interpretation of $\langle\hat{\psi}(p)\hat{\psi}^\dagger(q)\rangle$ as an S matrix element and requiring the on-shell constraint. In the next chapter, we will prove that (3.1) is true via the QFT principle of least action in a non-interacting theory.

In the case of interactions, the story is complicated however. Interactions allow

bound states between particles to exist, so that the ψ and $\psi^\dagger$ particles might not need be interpreted as plane wave states but as constituent particles in the bound state. Then the corresponding pole in ρ will be shifted by the binding energy of the bound state. For example, an electron and a positron can exist together in a bound state known as positronium, which has a ground state energy of $E = 2m - 6.8\,\text{eV}$ where $m = 511\,\text{keV}$ is the mass of the electron.

Positronium would place an additional pole in $\rho(p^2)$ at $p^2 = E^2$, except for the fact that positronium decays usually into two gamma rays with a lifetime of $\tau = 0.12$ ns. This decay has the effect of blurring out the pole into a *Breit-Wigner* distribution

$$\rho(p^2) \propto \frac{1}{(p^2 - E^2)^2 + E^2\Gamma^2} \tag{3.2}$$

where $\Gamma = \frac{1}{\tau} = \frac{1}{0.12\,\text{ns}}$ is the decay width. When the ψ particle itself is unstable, the $p^2 = m^2$ pole is similarly broadened. When detecting particles in a collider, this Breit-Wigner distribution is observed and its peak used to define the mass of the particle.

JTD: scattering

Chapter 4

Lattice Quantum Field Theory

The past few chapters dealt with connecting observables such as scattering amplitudes and particle creation to the n-point correlation functions. The most common way to actually evaluate these functions is to use the perturbative techniques discussed in the rest of this book, but in this chapter we briefly mention the alternative methods in case the perturbative approach should fail.

Part II

Quantum Electrodynamics & Abelian Gauge Theories

Quantum Electrodynamics, or QED, is the theory of the electron and the photon, and can also be extended to approximately describe the proton. It is the first example of a successful quantum field theory and remains one the most accurate theory measured.

QED contains two types of particles: fermions and gauge bosons. This part will be dedicated to discussing both types of particle, as well as a simpler particle known as a scalar boson. The difference between these particles is codified in what *quantum numbers* they hold.

A quantum number is a quantity which is conserved as the particle propagates. From Newtonian mechanics, we are aware of momentum as a conserved property — and therefore a quantum number — that particles should have. However, quantum mechanics adds other possibilities such as spin for electrons, polarization for photons, helicity for massive particles, and more. Any particle theory must provide a complete description for the evolution of each of these quantum numbers.

Chapter 5

Spin Zero

Our particle will be the simplest kind of particle: a *scalar boson* or *spin-zero particle*, whose only quantum number is momentum. The standard model — the complete theory of particle physics — contains one of these scalars in the form of the Higgs particle, though there are additional particles such as mesons which can be approximated as scalars, as we will later discuss.

The first step in creating this scalar is to write a Lagrangian for a field ϕ defined at every spacetime point x. Let's assume for now that the field takes on real values. Symmetries place the following severe limit the terms this Lagrangian may contain:

1. Translations symmetry forbids terms that contain coordinates like x or t
2. Lorentz symmetry forces forbids uncontracted spacetime indices
3. The Hamiltonian (and Lagrangian) must have a minimum energy so the field does not fly off to infinity
4. If ϕ were complex, the fact that action is real would forbid any complex-valued term, which includes any odd power of ϕ

Furthermore, in the low energy regime where we might expect ϕ to be small (few particles present) and $\partial_\mu \phi$ to be small (low momentum), we should not include terms which involve too many powers of ϕ. The exact cutoff for how many powers is "too many" will be discussed later.

Given the above constraints on the Lagrangian, the only available terms to the Lagrange density $\mathcal{L}$ are

$$\phi \partial_\mu \partial^\mu \phi, \qquad (\partial_\mu \phi)(\partial^\mu \phi), \qquad \phi^2, \qquad \phi^3, \qquad \phi^4,$$

and higher order terms. However, the first two terms are identical. This is because the QFT principle of least action (1.15) only involves $\int d^4x\mathcal{L}$ rather than $\mathcal{L}$ directly. When under an integral, integration by parts guarantees that $\phi\partial_\mu\partial^\mu\phi = -(\partial_\mu\phi)(\partial^\mu\phi)$. For this reason, we drop the first term with no loss of generality. We also usually drop the last term, but we will add it back in for a later section.

This discussion leaves us with only one possible action at low order:

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu\phi)^2 - \frac{m^2}{2}\phi^2 + \frac{\lambda}{4!}\phi^4. \tag{5.1}$$

The constants m^2 and λ were inserted to preserve generality, but no constant was inserted before the first term because such a constant can always be set to $1/2$ by rescaling ϕ. Each term in (5.1) as an intuitive meaning which we will briefly describe now, and the later sections will elucidate them in much greater detail.

1. $-\frac{1}{2}(\partial_\mu\phi)^2$ is known as the *kinetic term*, defined to be negative because, as will be seen later, if it were positive the n-point correlation functions would all diverge. This term encapsulates the kinetic energy of the particle because derivatives of fields represent momentum in much the same way that the quantum mechanical definition of the momentum operator $\hat{\boldsymbol{p}}$ is $i\boldsymbol{\nabla}$. The presence of two derivatives is akin to the $\boldsymbol{p}^2$ of kinetic energy.

2. $\frac{m^2}{2}\phi^2$ is known as the *mass term*. Since ϕ^2 is positive for any nonzero phi and is larger than any other power of ϕ, this term controls the amount of energy that a particle contains merely by existing, and thereby increasing the local value of ϕ. This inherent energy is defined as its mass. These first two terms of the Lagrangian are very similar to the Hamiltonian quantum harmonic oscillator, which has energies of $E = m(n + \frac{1}{2})$, where n is an integer. Thus, energy can only increase in integer multiples of mass, which is an intuitive result. We will see that the same result occurs in quantum field theory.

3. $\frac{\lambda}{4!}\phi^4$ is an *interaction term*. Since it is proportional to the square of the mass term, it represents two-body interactions, whereas ϕ^5 would represent three-body interactions, etc. The presence of this term ruins the harmonic oscillator-like spectrum of $E = m(n + \frac{1}{2})$ because it generates potential energy when two particles are present in the same system. This is reminiscent one of the

$1/r$ Coulomb potential of a charged particle, and indeed we will see that $\lambda\phi^4$ produces a $1/r$ potential between two ϕ particles.

Encouragement to pursue (5.1) comes from the fact that, if we set $\lambda = 0$ and treat $\mathcal{L}$ as ϕ as a classical field theory, it produces the equation of motion

$$\partial^\mu \partial_\mu \phi^2 - m\phi^2 = 0 \tag{5.2}$$

which is the famous Klein-Gordon equation mentioned in the introduction (1.3). It was already known to satisfy the laws of special relativity, and if $m = 0$, it reduces to the familiar wave equation. The scalar Lagrangian (5.1) is therefore a largely familiar beast.

5.1 Solving the Scalar Path Integral Without Interactions

In this section, we will solve for n-point correlation functions $\langle\hat{\phi}(x_1)\cdots\hat{\phi}(x_n)\rangle$ using the QFT principle of least action (1.15). We neglect the ϕ^4 term by setting $\lambda = 0$ in this section, but we will restore λ in the following section. As such, the result of this section will be correlation functions for a non-interacting scalar theory.

Even with $\lambda = 0$, evaluating the path integral is an intricate process that requires several new mathematical techniques. We introduce these techniques in this section, which ends with a formula for the nPCFs of non-interacting scalar theory. The physical meaning of this formula is discussed in the following section.

Plugging the scalar Lagrangian into (1.15),

$$\langle 0|\hat{\phi}(x_1)\cdots\hat{\phi}(x_n)|0\rangle = \frac{\int \mathcal{D}\phi\, \phi(x_1)\cdots\phi(x_n) e^{-\frac{i}{2}\int d^4x\left[-(\partial_\mu\phi)^2 - m\phi^2\right]}}{\int \mathcal{D}\phi\, e^{-\frac{i}{2}\int d^4x\left[-(\partial_\mu\phi)^2 - m\phi^2\right]}}. \tag{5.3}$$

The ϕ fields in the exponent are functions of x, but the (x) has been suppressed for clarity.

It appears that if we can calculate the numerator for any $x_1 \ldots x_n$, then we can calculate the denominator too and therefore any nPCF. We'll start with the case where $n = 0$:

$$\int \mathcal{D}\phi\, e^{-\frac{i}{2}\int d^4x\left[-(\partial_\mu\phi)^2 - m\phi^2\right]} \tag{5.4}$$

Note that both terms in the exponent are quadratic in ϕ. We can therefore separate the ϕs as follows:

$$\int \mathcal{D}\phi \, e^{-\frac{i}{2}\int d^4x \, \phi^* \left[-\partial^2 - m\right]\phi}. \tag{5.5}$$

(5.5) is a Gaussian integral because the exponent is quadratic in ϕ. This makes it one of the only path integrals which can be solved analytically with perfect accuracy, and we will therefore spend considerable time on it. The next subsection is dedicated to evaluating it, while the following generalizes the result to $n \neq 0$. Following this, we will return to computing nPCFs.

Infinite Dimensional Gaussian Integrals

The Gaussianity of (5.5) can be seen more clearly if we consider a one-dimensional spacetime, broken up into a lattice with N sites ans spacing a. In this case, $\phi(x_i)$ becomes ϕ_i, $\int d^4x$ becomes $\sum_i a^4$ and $\partial^2\phi(x_i)$ becomes $(\phi_{i-1} + \phi_{i+1} - 2\phi_i)/a^2$. Furthermore, the path integral breaks down as an integral over each ϕ_i, so that the entire function becomes

$$\prod_i^N \int d\phi_i \, e^{-\frac{i}{2}\sum_{jk} \phi_j^* M_{jk} \phi_k} = \prod_i^N \int d\phi_i \, e^{-\frac{i}{2}\boldsymbol{\phi}^\dagger M \boldsymbol{\phi}} \tag{5.6}$$

where

$$M_{ij} = -ma^4\delta_{ij} + a^2(\delta_{i,j+1} + \delta_{i,j-1} - 2\delta_{ij}) \tag{5.7}$$

is a matrix. In the second equality (5.6), we have switched to matrix notation, where $\boldsymbol{\phi}$ is a vector of ϕ_i. Inspection reveals that M is Hermitian and therefore can be diagonalized as $M = U^\dagger \Lambda U$, where U is a unitary matrix and Λ is a diagonal matrix of eigenvalues λ_i. Thus, the path integral becomes

$$\int d^N\boldsymbol{\phi} \, e^{-\frac{i}{2}(U\boldsymbol{\phi})^\dagger \Lambda (U\boldsymbol{\phi})} = \int d^N\boldsymbol{\phi} \, e^{-\frac{i}{2}\boldsymbol{\phi}^\dagger \Lambda \boldsymbol{\phi}} = \prod_i^N \left[\int d\phi_i \, e^{-\frac{i}{2}\phi_i^* \lambda_i \phi_i} \right]. \tag{5.8}$$

In the first equality we changed variables $\boldsymbol{\phi} \to U\boldsymbol{\phi}$, but since $\det U = 1$, doing so does not change the value of the integral. In the second equality we took advantage of the fact that Λ is diagonal to integrate each component separately, leaving us

with N single-variable Gaussian integrals. It's a common fact that

$$\int dx\, e^{-\alpha x^2} = \sqrt{\pi \alpha^{-1}} \qquad \text{whenever} \qquad \operatorname{Re}\alpha > 0. \tag{5.9}$$

Here, $\alpha = i\lambda_i/2$, and since M is Hermitian, λ is real and α is purely imaginary. We are therefore stuck on the edge of the validity of (5.9). A common choice is to add a small imaginary number $i\epsilon$ to every eigenvalue so that Re α is always slightly positive. Then our integral becomes

$$\prod_i^N \int d\phi_i\, e^{-\frac{i}{2}\sum_{jk}\, \phi_j^* M_{jk} \phi_k} = \frac{\pi^{N/2}}{\prod_i^N \sqrt{\lambda_i}} = \frac{\pi^{N/2}}{\sqrt{\det M}}. \tag{5.10}$$

The next step is to remove the discretization we performed at the beginning, by dividing spacetime into a lattice. This involves sending $N \to \infty$ and also changing det M such that the value of the above integral may become infinite. This is worrying, but not catastrophic, since the nPCFs of (5.3) are all *ratios* of gaussian integrals. If the $\pi^{N/2}/\sqrt{\det M}$ is common to both the numerator and the denominator, it will cancel out regardless of whether it is infinite.

The next step is therefore to generalize our Gaussian integral result to $n \neq 0$ so that we can confirm this cancellation in the nPCFs.

Wick's Theorem

The last barrier preventing us from evaluating the properties of non-interacting scalar theory is the integral

$$\int \mathcal{D}\phi\, \phi(x_1) \ldots \phi(x_n) e^{-\frac{i}{2}\int d^4x\, \phi^* [-\partial^2 - m]\phi}. \tag{5.11}$$

We'll use a trick to solve this problem. First, we'll discretize spacetime as we did in the previous section:

$$\prod_i^N \int d\phi_i\, \phi_{j_1} \ldots \phi_{j_n} e^{-i\, \boldsymbol{\phi}^\dagger M \boldsymbol{\phi}}. \tag{5.12}$$

where ϕ_{j_1} is the discretized $\phi(x_1)$. This integral can be solved more easily via the *generating function*

$$Z(\boldsymbol{J}) = \prod_i^N \int d\phi_i \, e^{-i\,\boldsymbol{\phi}^\dagger M \boldsymbol{\phi} + \boldsymbol{J}\cdot\boldsymbol{\phi}} \tag{5.13}$$

where $\boldsymbol{J}$ is an arbitrary vector. This function exists only to be differentiated, since

$$\left.\frac{d^2 Z}{dJ_{j_1} dJ_{j_2}}\right|_{\boldsymbol{J}=0} = \prod_i^N \int d\phi_i \, \phi_{j_1}\phi_{j_2} e^{-i\,\boldsymbol{\phi}^\dagger M \boldsymbol{\phi}} \tag{5.14}$$

which can be seen by explicitly differentiating (5.13). This derivative is clearly useful, since it is the integral need for a 2PCF. We therefore go about finding $Z(\boldsymbol{J})$ via the same method as in the previous section:

$$\begin{aligned}
Z(\boldsymbol{J}) &= \int d\boldsymbol{\phi}\, e^{-i\left[(U\boldsymbol{\phi})^\dagger \Lambda (U\boldsymbol{\phi}) + (U\boldsymbol{J})^\dagger (U\boldsymbol{\phi})\right]} \\
&= \prod_i \left[\int d\phi_i \, e^{-i\left[\phi_i^* \lambda_i \phi_i + (U\boldsymbol{J})_i^* \phi_i\right]}\right] \\
&= \prod_i \left[\int d\phi_i \, e^{-i\left|\phi_i\sqrt{\lambda_i} + \frac{(U\boldsymbol{J})_i^*}{2\sqrt{\lambda_i}}\right|^2} e^{i\frac{|U\boldsymbol{J}|_i^2}{4\lambda_i}}\right] \\
&= \prod_i \left[\sqrt{\frac{\pi}{\lambda_i}} e^{i\frac{|U\boldsymbol{J}|_i^2}{4\lambda_i}}\right] \\
&= \frac{\pi^{N/2}}{\det M} e^{\frac{i}{4}\sum_i \Lambda_i^{-1} |U\boldsymbol{J}|_i^2} = \frac{\pi^{N/2}}{\det M} e^{\frac{i}{4}\boldsymbol{J}^\dagger U^\dagger \Lambda^{-1} U \boldsymbol{J}} \\
&= \frac{\pi^{N/2}}{\det M} e^{\frac{i}{4}\boldsymbol{J}^\dagger M^{-1} \boldsymbol{J}}
\end{aligned} \tag{5.15}$$

where the first equality is true by diagonalizing the Hermitian matrix M, the second by substitution of variables, the third by completing the square, the fourth by substitution of variables again and evaluating the Gaussian integral, and the rest by rearranging the equation. This calculation revealed two crucial properties: (1) that $Z(\boldsymbol{J})$ has the same infinite prefactor that the original Gaussian integral in the last section had, and (2) this prefactor is multiplied by a simple term containing M^{-1}.

This term is simple enough that its derivative can be easily computed:

$$\left.\frac{d^2 Z}{dJ_{j_1} dJ_{j_2}}\right|_{\boldsymbol{J}=0} = Z(0) M^{-1}_{j_1 j_2}. \tag{5.16}$$

Higher derivatives are also simple: two more derivatives multiplies the above by$M^{-1}_{j_3 j_4}$. However, since derivatives commute, $Z(0)M^{-1}_{j_1 j_2} M^{-1}_{j_3 j_4}$ cannot be the only term. The value of the integral must be symmetric under all permutations of the indices. Therefore,

$$\left.\frac{d^2 Z}{dJ_{j_1} dJ_{j_2} dJ_{j_3} dJ_{j_4}}\right|_{\boldsymbol{J}=0} = Z(0)\left(M^{-1}_{j_1 j_2} M^{-1}_{j_3 j_4} + M^{-1}_{j_1 j_3} M^{-1}_{j_2 j_4} + M^{-1}_{j_1 j_4} M^{-1}_{j_2 j_3}\right). \tag{5.17}$$

For n-many derivatives with n even, there are $(n)!$ terms where each term represents one possible way to pair all the indices up. This rule for evaluating Gaussian integrals is called *Wick's theorem*.[1] If n is odd, there are no ways to pair up the indices and the value of the Gaussian integral is zero. This can also be seen by noticing that the Gaussian is an even function, and an odd number of derivatives applied to an even function is an odd function which will take the value zero at $\boldsymbol{J} = 0$.

Remembering the connection between Z and the PCFs, as in (5.14), Wick's theorem states that

$$\langle 0|\hat{\phi}(x_{j_1}) \cdots \hat{\phi}(x_{j_{2n}})|0\rangle = \sum_{\text{pairs } (a,b)_i \text{ of indices}} M^{-1}_{a_1 b_1} M^{-1}_{a_2 b_2} \ldots M^{-1}_{a_n b_n}. \tag{5.18}$$

We have successfully computed our first nPCF! The next section is devoted to returning to a continuous, four-dimensional spacetime, after which we will discuss the consequences of this nPCF.

Continuous Spacetime Gaussian Integrals

When spacetime is continuous, the matrix M used profusely in the preceding sections does not make sense, meaning that the M^{-1} of Wick's theorem is not defined. For scalar the free theory, we had $\mathcal{L} = \phi^*(-\partial^2 - m^2)\phi/2$, so the analog of the matrix M will be the differential operator $O_x = (-\partial^2 - m^2)/2$, where the sub-

[1] Problem 1 is to prove this fact by induction.

script x indicates that O contains derivatives with respect to x. The analog of M^{-1} is the function $G(x, y)$ which satisfies

$$O_x G(x, y) = \delta(x - y). \tag{5.19}$$

This is analogous to the definition of a matrix inverse, which is that

$$M_{ij} M^{-1}_{jk} = \delta_{ik}. \tag{5.20}$$

The function $G(x, y)$ is called the Green's function for operator O and must be computed for each O that may appear. Computing a Green's function for an arbitrary operator is generally difficult, but our operator $O_x = (-\partial^2 - m^2)$ has translation symmetry, which makes its Green's function easier to compute. Translation symmetry is manifested in the fact that O_x not a function of x, but we could have also guessed it would appear from the fact that the Lagrangian density from which O_x is derived had to be translation invariant. This symmetry requires that G must also be translation invariant, so that $G(x, y) = G(x - y_0, y - y_0)$ for any y_0. Choosing $y_0 = y$ allows us to write $G(0, x - y)$, or $G(x - y)$ if we drop the first argument. Thus, the new definition of the Green's function can be simplified to

$$O_z G(z) = \delta(z) \tag{5.21}$$

for $z = x - y$.

Translational symmetry allows us to solve for G_x via the Fourier transform. Let's define the Fourier transform of $G(z)$ as

$$G(k) = \int d^4 z \, e^{-ik\cdot z} G(z), \qquad G(z) = \int \frac{d^4 k}{(2\pi)^4} e^{ik\cdot z} G(k), . \tag{5.22}$$

We must also transform O_x such that it can operate on $G(k)$. Constants such as m are unaffected by a Fourier transform, but derivatives such as ∂_μ do transform:

$$\partial_\mu f(x) = \int \frac{d^4 k}{(2\pi)^4} \partial_\mu e^{ik\cdot x} f(k) = \int \frac{d^4 k}{(2\pi)^4} i k_\mu e^{ik\cdot x} f(k). \tag{5.23}$$

These two substitutions for $G(x)$ and ∂_μ transform (5.21) to

$$\int \frac{d^4k}{(2\pi)^4} \left(k^2 - m^2\right) G(k) = \delta(z). \tag{5.24}$$

A property of the δ function is

$$\delta(z) = \int \frac{d^4k}{(2\pi)^4} e^{ik\cdot z}. \tag{5.25}$$

Substituting this property into (5.24) causes the k integral to appear on both sides of the equation. Removing it, we get

$$(k^2 - m^2)G(k) = 1 \implies G(k) = \frac{1}{k^2 - m^2}. \tag{5.26}$$

This concludes our calculation of the Green's function. The analog of Wick's theorem (5.18) for Green's functions is

$$\langle 0|\hat{\phi}(x_1)\cdots\hat{\phi}(x_{2n})|0\rangle = \sum_{\text{pairs } (a,b)_i \text{ of indices}} G(x_{a_1} - x_{b_1})\ldots G(x_{a_n} - x_{b_n}). \tag{5.27}$$

and nPCFs with n odd are zero.

Putting the Pieces Together

In order to use our momentum-space Green's function (5.26) we must Fourier transform the nPCF:

$$\begin{aligned}\langle 0|\hat{\phi}(k_1)\cdots\hat{\phi}(k_{2n})|0\rangle &= \int d^4x_1\ldots d^4x_{2n}\,\left(e^{ik_1\cdot x_1}\ldots e^{ik_{2n}\cdot x_{2n}}\right)\\ &\quad\times\langle 0|\hat{\phi}(x_1)\cdots\hat{\phi}(x_{2n})|0\rangle\\ &= \int d^4x_1\ldots d^4x_{2n}\,\left(e^{ik_1\cdot x_1}\ldots e^{ik_{2n}\cdot x_{2n}}\right)\\ &\quad\times\sum_{\text{pairs }(a,b)_i} G(x_{a_1}-x_{b_1})\ldots G(x_{a_n}-x_{b_n})\\ &= \sum_{\text{pairs }(a,b)_i}\int d^4x_1\ldots d^4x_{2n}\,\left(e^{ik_1\cdot x_1}\ldots e^{ik_{2n}\cdot x_{2n}}\right)\\ &\quad\times G(x_{a_1}-x_{b_1})\ldots G(x_{a_n}-x_{b_n})\end{aligned} \tag{5.28}$$

where the second equality came from applying Wick's theorem and the third commuted the sum over pairs to the front of the equation. We can now perform a unitary change of variables. For each Wick pairs of indices $(a,b)_i$, we define $y_i = (x_{a_i}+x_{b_i})/2$ and $z_i = x_{a_i}-x_{b_i}$, so that $d^4x_{a_i}d^4x_{b_i} = d^4y_id^4z_i$. The advantage of this transformation is that $G(x_{a_i}-x_{b_i})$ is now merely $G(z_i)$ and does not depend on y_i. The exponents also transform:

$$e^{ik_{a_i}\cdot x_{a_i}}e^{ik_{b_i}\cdot x_{b_i}} = e^{ip_i\cdot y_i}e^{i\ell_i\cdot z_i} \tag{5.29}$$

where $p_i = k_{a_i}+k_{b_i}$ and $\ell_i = (k_{a_i}-k_{b_i})/2$. The result is

$$\begin{aligned}\langle 0|\hat{\phi}(k_1)\cdots\hat{\phi}(k_{2n})|0\rangle &= \sum_{\text{pairs }(a,b)_i}\prod_{i=1}^{n}\int d^4y_id^4z_i\,e^{ip_i\cdot y_i}e^{i\ell_i\cdot z_i}G(z_i)\\ &= \sum_{\text{pairs }(a,b)_i}\prod_{i=1}^{n}(2\pi)^4\delta(p_i)G(\ell_i).\end{aligned} \tag{5.30}$$

The $\delta(p_i)$ term is nonzero only when $k_{a_i} = -k_{b_i}$, in which case $\ell_i = (k_{a_i} - k_{a_i})/2 = k_{a_i}$. Thus,

$$\langle 0|\hat{\phi}(k_1)\cdots\hat{\phi}(k_{2n})|0\rangle = \sum_{\text{pairs } (a,b)_i} \prod_{i=1}^{n} (2\pi)^4 \delta(k_{a_i} + k_{b_i}) G(k_{a_i}). \tag{5.31}$$

This result is actually valid for any quadratic Lagrangian, even ones which are not the free scalar lagrangian. Plugging in the Green's function for a scalar field,

$$\langle 0|\hat{\phi}(k_1)\cdots\hat{\phi}(k_{2n})|0\rangle = \sum_{\text{pairs } (a,b)_i} \prod_{i=1}^{n} (2\pi)^4 \frac{\delta(k_{a_i} + k_{b_i})}{k_{a_i}^2 - m^2}. \tag{5.32}$$

This is our first ever nPCF, computed exactly assuming the QFT principle of least action (1.15) and the scalar Lagrange density (5.1). It depends only on the mass of the ϕ particle m (included in the form of $G(k)$) and the k values chosen to evaluate the nPCF at. In the next section, we will interpret the physical implications of this nPCF formula and show that it has both quantum mechanical and relativistic properties, confirming our hope that the QFT principle of least action would lead to a relativistic quantum theory. We will also discuss the spectrum of the 2PCF using the framework discussed in chapter 3, showing that the 2PCF indeed contains arbitrarily many particles of mass m.

5.2 The Spin Zero Free Theory Is Quantum Mechanical and Relativistic

The nPCFs of (5.32) have several crucial properties which are necessary for a relativistic quantum theory, which we enumerate below.

Mass Energy Equivalence (5.32) enforces the relativistic $k^2 = m^2$ constraint due to the $k^2 - m^2$ in the denominator of the Green's function. This causes the nPCFs to be much greatest near $k^2 = m^2$ — that is, the probability for a particle with momentum k to be produced is greatest for $k^2 = m^2$ and much smaller elsewhere.

Quantized Amplitude The nPCFs also show that a field $\hat{\phi}(k)$ at a given momentum k cannot be arbitrarily rescaled. Since m is constant and $k^2 = m^2$ as discussed above, there is no freedom to choose an arbitrary amplitude for ϕ and obey the (5.32). That is, the amplitude of ϕ is quantized. This is unlike a classical

field theory such as electromagnetism, where the $\boldsymbol{E}$ field may take on whatever amplitude it wishes. But in any quantum mechanical theory, we expect the properties of particles to be quantized, with values such as field amplitude taking on only discrete values.

No Scattering in Free Theory Chapter 2 outline a toolkit to use nPCFs to compute scattering cross-sections. It interpreted $\langle 0|\hat{\phi}(p_1)\hat{\phi}(p_n)\cdots\hat{\phi}(-q_1)\hat{\phi}(-q_m)|0\rangle$ as the amplitude of scattering n particles from momentum p_i to m particles of momentum q_i. (5.32) shows that this amplitude is zero except when the momentum can be paired up as $p_i = q_i$. Each pair adds the same factor to the correlation function due to the quantization of amplitude. This $p_i = q_i$ requirement shows that the free theory nPCFs do not allow particles to change momentum, even when other particles are present. This is equivalent to requiring no interactions, which is the definition of a free theory.

Mass Spectrum of a Particle In chapter 3, we discussed the 2PCF in particular and showed that it is expected to have a pole at $k^2 = m^2$ where m is the mass of any particle in the theory. We also showed that branch cuts where the 2PCF is infinite represent interaction cross-sections. Our 2PCF contains a pole at $k^2 = m^2$, where m is the mass of the ϕ particle, exactly as expected. This is the only pole and there are no branch cuts, indicating again that this theory contains no interactions.

Causality A founding principle of relativity is that information cannot be translated faster than the speed of light. If (5.32) is to be taken seriously as a theory of physics, we should expect $\hat{\phi}(x)$ to be uncorrelated with $\hat{\phi}(y)$ as long as x and y are spacelike-separated. Thus,

$$\langle 0|\hat{\phi}(x)\hat{\phi}(y)|0\rangle = 0 \qquad \text{whenever} \qquad |\boldsymbol{x} - \boldsymbol{y}| > |x_0 - y_0|. \tag{5.33}$$

To check this, we will Fourier transform the 2PCF

$$\langle 0|\hat{\phi}(k)\hat{\phi}(-k)|0\rangle = (2\pi)^4 \frac{\delta(0)}{k^2 - m^2}$$

back into position space and check this fact. If we translate to $y = 0$ and rotate until $x^\mu = (t, x, 0, 0)$, then

$$\begin{aligned}\langle 0|\hat{\phi}(x)\hat{\phi}(0)|0\rangle &= \int \frac{d^4k}{(2\pi)^4} e^{ik\cdot x} \frac{(2\pi)^4}{k^2 - m^2} \\ &= \int d^3\boldsymbol{k} \int dk^0 \, \frac{e^{-ik_0 t + ik_1 x}}{k_0^2 - \boldsymbol{k}^2 - m^2}.\end{aligned} \tag{5.34}$$

The integrand contains a pole at $k_0^2 = \boldsymbol{k}^2 + m^2$, making the Fourier transform undefined. However, if we add a small $i\epsilon$ term into the denominator where $\epsilon > 0$ is small, we can move the pole off the real axis and out of the way of the integration without changing the value of $G(k)$ very much. Defining $E^2 = \boldsymbol{k}^2 + m^2 > 0$, we must now solve[2]

$$I = \int dk^0 \, \frac{e^{-ik_0 t}}{k_0^2 - E^2 - i\epsilon} = \int dk^0 \, \frac{e^{-ik_0 t}}{2E} \left(\frac{1}{k_0 - E + i\epsilon} - \frac{1}{k_0 + E - i\epsilon} \right). \tag{5.35}$$

This integral is performed over $z \in (-\infty, \infty)$, but we can always add more to the contour of integration (even extending it into the complex plane) as long as the value of the integrand on that contour is zero. For $t > 0$, $e^{-ik_0 t}$ is small whenever $\operatorname{Im} k^0 \ll 0$ and for $t > 0$ the condition is $\operatorname{Im} z \gg 0$. We therefore choose to extend the contour in a large arc around the complex plane, below the real axis for $t > 0$ and above it for $t < 0$ (Figure 5.1). Since this new contour is closed, we may use the Residue Theorem (see chapter 17) to evaluate it, picking up the $-E$ pole when closing the curve above the real line, and the $+E$ pole when closing the curve below the real line. Thus,

$$I = \begin{cases} \pi i \frac{e^{-itE}}{E} & \text{for } t > 0 \\ -\pi i \frac{e^{itE}}{E} & \text{for } t < 0 \end{cases} \tag{5.36}$$

The only difference between the $t > 0$ and $t < 0$ cases is an overall sign, which we ignore.

[2] $E^2 = \boldsymbol{k}^2 + m^2$ is actually the energy of a particle with mass m and momentum $\boldsymbol{k}$, but we will not need this fact.

Furthermore, in this equation we redefine ϵ on the right hand side to absorb its coefficients, which is allowed since ϵ is assumed to be small enough that it remains small even when multiplied by a finite number.

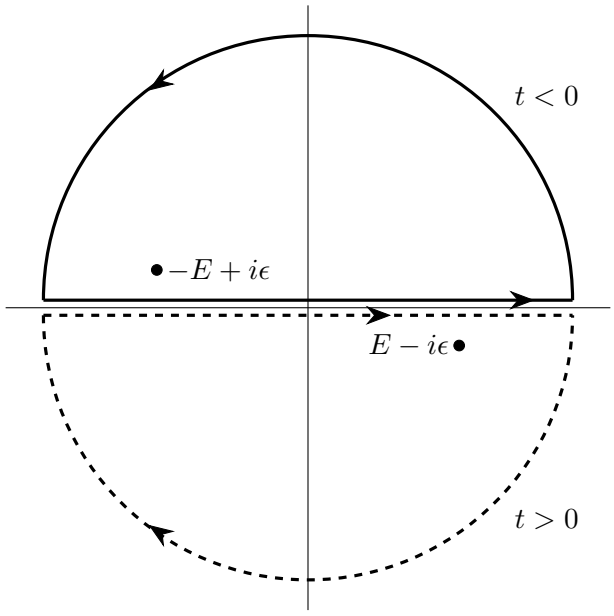


Figure 5.1: The contour to be used for $\int dk^0 \frac{e^{-ik_0 t}}{k_0^2 - A^2 - i\epsilon}$ depending on whether $t > 0$ or $t < 0$.

We can now use I to solve (5.34). Dropping proportionality constants which do not matter for this demonstration, and assuming $x > 0$ and $t > 0$,

$$\begin{aligned}\langle 0|\hat{\phi}(x)\hat{\phi}(0)|0\rangle &\propto \int d^3\boldsymbol{k}\, \frac{e^{i(xk_1 - t\sqrt{\boldsymbol{k}^2+m^2})}}{\sqrt{\boldsymbol{k}^2+m^2}} \\ &= \int dk_2 dk_3 \int dk_1\, \frac{e^{i(xk_1 - t\sqrt{k_1^2+F^2})}}{\sqrt{k_1^2+F^2}}\end{aligned} \tag{5.37}$$

where in the second line we defined $F^2 = k_2^2 + k_3^2 + m^2$. We now work on the k_1 integral in a similar fashion to the k_0 integral. Again, we have poles when the denominator is zero, at $k_1^2 = F^2$, or $k_1 = \pm iF$. But this time there is also a branch cut for all $k_1^2 + F^2 < 0$ due to the difficulty of defining the square root operator on negative real values. The complex plane of this function is depicted in Figure 5.2.

To work out whether to close the contour above or below the real line, we substitute $k_1 = Re^{i\theta}$ into the exponent, leaving a real part of $-R\sin\theta(x - t)$. We

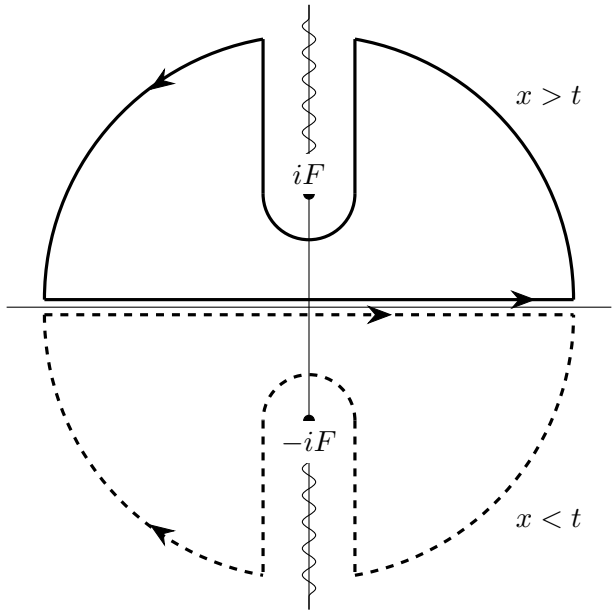


Figure 5.2: The contour to be used for $\int dk_1 \frac{e^{i(xk_1 - t\sqrt{k_1^2+F^2})}}{\sqrt{k_1^2+F^2}}$, with the branch cut.

close the contour in the region where this real part is small, which is above the real line for $x > t$ and below it for $x < t$. However, we must avoid the branch cut by looping the contour around it, which adds nonzero values where the contour nears the origin. The scale of this added amount is controlled by the value of the integrand near the end of the branch cut, where the integrand is largest. For $x > t$, this occurs for $k_1 = i(F - \delta)$ with δ small, and the integrand is

$$\sim \frac{e^{-xF}}{\sqrt{2F\delta}}. \tag{5.38}$$

Meanwhile, for $x < t$, the integrand near $k_1 = -i(F - \delta)$ is considerably larger:

$$\sim \frac{e^{xF}}{\sqrt{2F\delta}}. \tag{5.39}$$

It follows that the 2PCF is nonzero for $x < t$, which is inside the lightcone. But when the lightcone and correlations are forbidden by causality, the 2PCF decays exponentially. This is reminiscent of how a wavefunction decreases exponentially in a region where potential energy is greater than the energy of the particle in non-relativistic quantum mechanics, or of how a light wave exponentially decays in a medium which cannot support its frequency. Thus, the QFT principle of least action not only obeys causality, but does so in a manner similar to both quantum mechanical and classical field theories.

The exponential decrease outside the lightcone came from our choice to add $i\epsilon$ to the denominator of the Green's function. This shifted the $\pm E$ poles above and below the real line as shown in Figure 5.1. If we had shifted the poles in another way, we would not have gotten the cancellation we achieved, and the 2PCF would not have fallen outside the light cone. This choice of pole-shifting was made by Richard Feynman and is known as the *Feynman Propagator*, and we will nearly always use it in the future.

Due to the success of the QFT principle of least action in this free scalar field case, we are inspired to continue on to the $\lambda \neq 0$ case of an interacting theory, which contains a vast variety of novel phenomena.

5.3 Solving the Spin Zero Path Integral With Interactions

One of the key results of the previous section was that for a non-interacting theory where the Lagrangian density is Gaussian, particles cannot transfer momenta and cannot scatter. The proof of this is (5.30), where we separated an $2n$PCF attained by Wick's theorem into n multiplied factors, each of which enforced momentum conservation.

However, real particles scatter, so to make a true particle theory we will have to break the chain of logic that led us to the no-scattering conclusion. The solution is to add a small, non-quadratic term to the Lagrangian density, which adds small additional terms to the $2n$PCF which are not separable. These small terms allow a small amount of momentum transfer between the n particles, causing scattering. This is the motivation for us associating the $\frac{\lambda}{4!}\phi^4$ term in $\mathcal{L}$ with scattering. In this section, we'll compute the effect of this new term on the momentum-space nPCFs.

Our strategy will be to apply the momentum-transferring ϕ^4 term as a perturbation to the Gaussian free theory discussed above. In other words, we'll write

$$\begin{gathered} S = S_0 + \lambda S_1, \\ S_0 = \int d^4x \left(\frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2\right), \qquad S_1 = -\int d^4x \, \frac{1}{4!}\phi^4 \end{gathered} \tag{5.40}$$

Our goal will be to produce a power series in λ for nPCFs, of which the zeroth order term is (5.32) and the following terms are small momentum-mixing terms under the assumption that $\lambda \ll 1$. To do this, we start with the QFT principle of least action:

$$\langle 0|\hat{\phi}(x_1)\cdots\hat{\phi}(x_n)|0\rangle = \frac{\int \mathcal{D}\phi \, \phi(x_1)\cdots\phi(x_n)e^{iS_0+i\lambda S_1}}{\int \mathcal{D}\phi \, e^{iS_0+i\lambda S_1}}. \tag{5.41}$$

Obtaining a power series in λ is just as simple as expanding the exponent as a power series:

$$\langle 0|\hat{\phi}(x_1)\cdots\hat{\phi}(x_n)|0\rangle = \frac{\int \mathcal{D}\phi \, \phi(x_1)\cdots\phi(x_n)e^{iS_0}\left(1 + i\lambda S_1 - \frac{\lambda^2}{2}S_1^2 + \dots\right)}{\int \mathcal{D}\phi \, e^{iS_0}\left(1 + i\lambda S_1 - \frac{\lambda^2}{2}S_1^2 + \dots\right)}. \tag{5.42}$$

Now that only the quadratic term S_0 remains in the exponent of the integrand, (5.42) can be completely solved to arbitrary order with Wick's theorem! For example, when S_1 as defined in (5.40) is substituted into the above equation, the numerator expands to

$$\begin{gathered}\int \mathcal{D}\phi\, \phi(x_1)\cdots\phi(x_n)e^{iS_0} \\ -\frac{i\lambda}{4!}\int d^4x \int \mathcal{D}\phi\, \phi(x_1)\cdots\phi(x_n)e^{iS_0}\phi^*(x)\phi(x)\phi^*(x)\phi(x) + \mathcal{O}(\lambda^2).\end{gathered} \tag{5.43}$$

Both of these terms are Gaussian correlation functions. The first line is the same nPCF we found for the non-interacting theory, confirming our expectation that adding the small ϕ^4 perturbation did not change the first-order structure of the theory. The second is an $n+4$PCFs which will allow these n particles to interact and exchange momentum.

To see this momentum exchange clearly, we will switch to the momentum basis. Fortunately, the hard work of determining how nPCFs Fourier transform was done in (5.32), which reduced a momentum-space nPCF to products of Greens functions evaluated on Wick contractions of ϕs. The only trickiness introduced by interactions is the fact that the $\phi(x)$s that come from the perturbation S_1 are integrated over, as in the second line of (5.43). This position integral becomes

$$\begin{aligned}\int d^4x\, \langle\hat{\phi}(x)^4\rangle &= \int d^4x \int \frac{d^4k_1}{(2\pi)^4}\cdots\frac{d^4k_4}{(2\pi)^4}e^{ik_1\cdot x}\ldots e^{ik_4\cdot x}\,\langle\hat{\phi}(k_1)\ldots\phi(k_4)\rangle \\ &= \int d^4x \int \frac{d^4k_1}{(2\pi)^4}\cdots\frac{d^4k_4}{(2\pi)^4}e^{i(k_1+\cdots+k_4)\cdot x}\,\langle\hat{\phi}(k_1)\ldots\phi(k_4)\rangle \\ &= \int \frac{d^4k_1}{(2\pi)^4}\cdots\frac{d^4k_4}{(2\pi)^4}\delta(k_1+\cdots+k_4)\,\langle\hat{\phi}(k_1)\ldots\phi(k_4)\rangle\,.\end{aligned} \tag{5.44}$$

In other words, when some of the x values of an nPCF are the same in position space, then the corresponding k values in the momentum space have to sum to zero. This mathematical fact looks similar to momentum conservation, and in fact we'll see in the next section that (5.44) does enforce momentum conservation in scattering events.

In principle, we have now computed an interacting theory nPCF. (5.42) describes how to use perturbation theory to write an interacting nPCF in terms of

non-interacting mPCFs We computed these mPCFs in (5.32) using Wick's theorem, and (5.44) describes how to integrate over the ϕ values present in the perturbing action S_1. Actually using these formulas however is a complicated and error-prone task. Fortunately, there is a graphical technique to isolate the Wick pairs and reduce these three equations to one simple form. This technique is called a Feynman diagram, introduced by Richard Feynman in the 1960s, and is widely used across high energy physics.

Feynman Diagrams

To compute an nPCF in momentum space, (5.32) dictates that we must contract the $\hat{\phi}$ operators into pairs. The goal of a Feynman diagram is to do this visually. Let's represent every point in spacetime as a vertex in a graph, and Wick contractions are represented as edges (called propagators) between the vertices. In an interacting theory, some of the vertices come from the original nPCF to be computed. These vertices were denoted as k_i and are called *external vertices*. Others, called *internal vertices*, come from the perturbation S_1 and will be integrated over. Because each external vertex only has one ϕ operator attached to it, they will only connect to one edge. Each internal vertex will connect to four edges because every S_1 term contains four ϕs each, due to the ϕ^4 of the Lagrangian.

After the propagators are drawn in according to these rules, we can assign momentum values to the ϕ terms. Each Wick contraction consists of two operators $\phi(k_1)\phi(k_2)$ which may have different momenta. However, (5.32) showed that if $k_1 \neq k_2$, then the entire diagram will evaluate to zero. Thus, we can assume that $k_1 = k_2 = k$. As shorthand, we assign a momentum of k to the propagator that represents this Wick contraction.

For a 4PCF in the scalar interacting theory there are four external vertices. There is one internal vertex to first order in λ. Following the rules for drawing edges, there are ten ways of pairing the ϕs to evaluate the numerator of (5.42) which are shown in Figure 5.3. Diagrams A1-A6 are similar, as are B1-B3, so we will only calculate one diagram from each of these classes. For this reason, only the momenta of this diagram are labeled. The direction in which momentum is chosen to flow is a matter of preference.

To calculate the 4PCF $\langle\hat{\phi}(p_1)\hat{\phi}(p_2)\hat{\phi}(q_1)\hat{\phi}(q_2)\rangle$ from a diagram, we follow (5.32) by writing $G(k)$ for every propagator with momentum k. For internal vertices, (5.44) tells us to enforce momentum conservation by integrating over all mo-

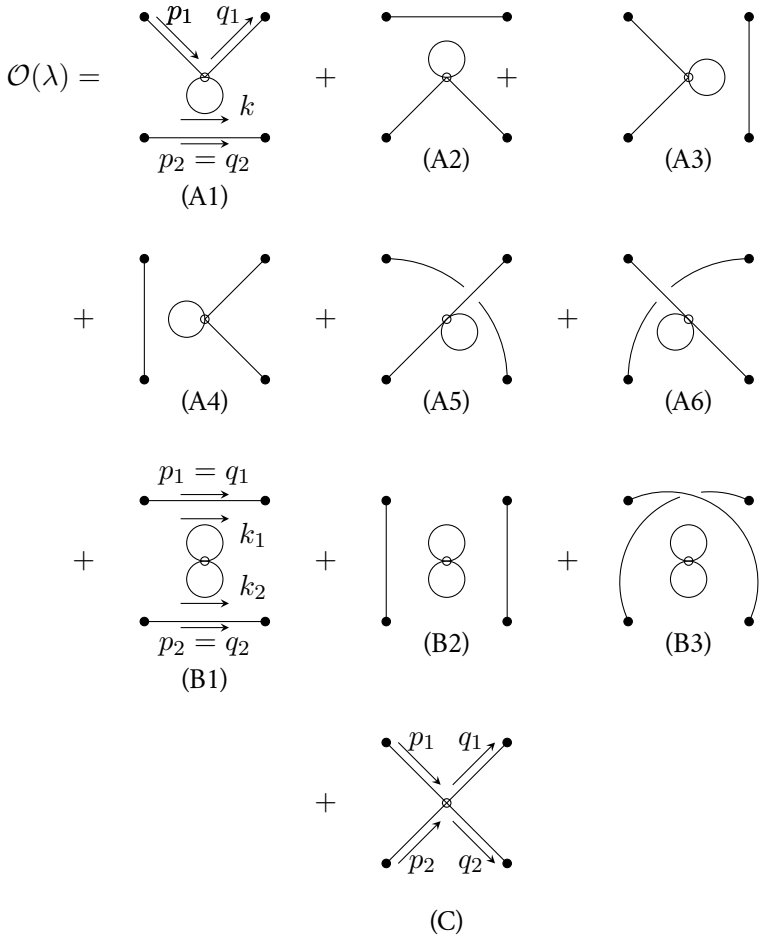


Figure 5.3: Feynman diagrams for a 4PCF $\langle\hat{\phi}(p_1)\hat{\phi}(p_2)\hat{\phi}(q_1)\hat{\phi}(q_2)\rangle$ in the interacting scalar theory. Each line is a Wick pair and each vertex is a point in spacetime.

menta subject to the constraint that the total momentum of all the propagators connected to an internal vertex is zero. Thus diagram A1 diagrams evaluates to

$$\begin{aligned}(\text{A1}) &\propto G(p_1)G(q_1)G(p_2)\delta(p_2-q_2)\int\frac{d^4k}{(2\pi)^4}\delta(p_1-q_1+k-k)G(k)\\ &\propto G(p_1)G(q_1)\delta(p_1-q_1)\delta(p_2-q_2)\int\frac{d^4k}{(2\pi)^4}G(k).\end{aligned} \tag{5.45}$$

The proportionality statements acknowledge that we are focusing on the structure of the Green's functions and have not yet put in the correct powers of λ or other coefficients that might appear.

The first delta function in line 1 is the delta function of (5.32), enforcing that momentum does not change along a propagator. The second delta function is the delta function of (5.44), enforcing that momentum cannot change at a vertex. For this diagram, these constraints together enforce that momentum of each must be conserved, and the double δ-function in the above equation falls out to prevent the momenta from being mixed. That is, the free theory did not lead to momentum mixing in this case. It only changed the value of the 4PCF by some value $\int\frac{d^4k}{(2\pi)^4}G(k)$ which we have not computed yet.

We can compute the diagrams (B1) and (C) in a similar way:

$$\begin{aligned}(\text{B1}) &\propto G(p_1)\delta(p_1-q_1)G(p_2)\delta(p_2-q_2)\int\frac{d^4k_1}{(2\pi)^4}\frac{d^4k_2}{(2\pi)^4}G(k_1)G(k_2)\\ (\text{C}) &\propto G(p_1)G(p_2)G(q_1)G(q_2)\delta(p_1+p_2+q_1+q_2).\end{aligned} \tag{5.46}$$

Diagram (B1) is like A1 in that it conserves momentum for each pair of particles. However, Diagram (C) allows the much-heralded momentum mixing. Even if p_1, p_2, q_1, and q_2 are all different, (C) will be nonzero as long as the momenta sum to zero. This diagram (and only this diagram) is the leading order contribution to $\phi\phi\rightarrow\phi\phi$ scattering, and we will use it a great deal throughout the rest of this chapter.

We have now verified that this ϕ^4 theory allows particles to interact. Before moving on to quantify this interaction by computing scattering cross sections, it's worth covering two more details. The first is to replace the above two proportion-

ality statements with actual equations. The second is to work out the denominator of (5.42) and divide the sum of the (A), (B), and (C) diagrams by that to have a full equation for the interacting theory nPCF.

Symmetry Factor and Feynman Rules

All of the diagrams drawn in figure 5.3 were first order, since they contained only one power of S_1 and therefore only one power of λ. If we had written the diagrams of order λ^2, they would have had two internal vertices (one for each S_1). Thus, every diagram should be multiplied by the constant $i\lambda$ for every internal vertex.

Another potential worry is the possibility of under-counting. S_1 contains four ϕs, which can each be contracted with another ϕ in the nPCF. Contracting with the first ϕ and contracting with the second represent different contractions and should be added individually. But this is not reflected in our Feynman diagram, where we draw a contraction with any one of the ϕs as a propagator running into the internal vertex and make no note of which ϕ it was contracted with. Another source of error enters when we compute second order diagrams: these involve two identical S_1 values connected in different ways to the exterior vertices. We should double each diagram to reflect the fact that the two S_1 values could be swapped without changing the diagram structure. There are more numerical factors we haven't taken care of as well, which are all represented in table 5.3.2.

To get the total correction to each diagram, we multiply every entry in the left hand column of table 5.3.2. Most of the entries cancel, but to keep track of those that don't, we define the *symmetry factor*

$$S_f = (1 + \text{identical vertices})!(1 + \text{reversible propagators})! \tag{5.47}$$

where identical vertices are internal vertices which are connected to the same vertices as each other, and reversible propagators are those that lead to and out of the same internal vertex. in which case we should multiply each diagram by the factor $(-i\lambda)^n/S_f$.

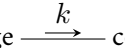
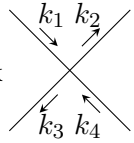
As an example, in figure 5.3, all the A diagrams have symmetry factor 2 due to the reversible propagator with momentum k. The B diagrams have symmetry factor 4 because they have two reversible propagators, and the C diagram has symmetry factor 1 because there are no reversible propagators or identical vertices.

The definition of the symmetry factor allows us to write down the value of any

Contribution	Source
$\frac{(-i\lambda)^n}{n!}$	Taylor series expansion of e^{-iS_1}
$\frac{1}{4!}$	Definition of the interacting term in the Lagrangian
$n!$	The number of ways to swap internal vertices
$\frac{1}{(\text{identical vertices})!}$	In the above row, we should not have swapped vertices which were already identical in the diagram
$4!$	The number of ways to connect 4 ϕs to 4 incoming propagators for each internal vertex
$\frac{1}{2^{\text{reversible propagators}}}$	In the above row, we should not have swapped propagators that both came from and led to the same vertex

Table 5.1: Factors we should multiply each Feynman diagram by due to expansion coefficients, over- or under-counting, and other sources. Here, n represents the number of internal vertices.

diagram exactly by following the following rules, called the *Feynman rules*:

1. The edge  contributes $G(k)$
2. A vertex  contributes $-i\lambda\delta(k_1 - k_2 - k_3 + k_4)$
3. Every internal propagator k is integrated over $\int \frac{d^4k}{(2\pi)^4}$
4. Diagrams with exchangeable vertices or reversible edges contribute a symmetry factor S_f.

Once the diagrams are computed, they may be added and the sum divided by the denominator of (5.42), discussed in the next section, to get the nPCF

The Denominator: Vacuum Diagram Cancellation

This section deals with the denominator of (5.42):

$$\int \mathcal{D}\phi e^{-iS_0} e^{-iS_1}$$

which we must divide the sum of all Feynman diagrams by to get an nPCF. This factor came from requiring the nPCF to be normalized. Fortunately, this integral will cancel perfectly with terms in the numerator.

To see how, consider the diagrammatic expansion of this denominator. It is essentially a 0PCF, if such a thing exists. For scalar field theory, the zeroth order term is 1 and the the first order diagram is

where we have left out the circle to indicate the internal vertex (this is often done since any intersection of propagators is clearly a vertex).

This kind of diagram which is unconnected to external vertices called a *vacuum diagram*. You may recognize it diagram as a part of all the (B) diagrams in figure 5.3. In fact, the vacuum diagrams in the denominator will cancel with all of the diagrams in the numerator like (B) that have vacuum sub-diagrams. The mathematical explanation for this fact is that for every diagram like (A) or (C) that has

no vacuum sub-diagrams, you can just multiply by a vacuum sub-diagram to get a diagram like (B), where both are combined. The reverse of this statement is that to divide by the vacuum diagrams, we need to remove all the diagrams like (B) that have vacuum sub-diagrams.

There is also physical intuition for removing by diagrams like (B) that have vacuum sub-diagrams. The vacuum diagrams have no external vertices, and external vertices generally represent particles. Thus, vacuum diagrams represent some change in energy of the vacuum, hence the name. Changes to the vacuum energy are generally undetectable, since we can only measure energy differences between systems and the vacuum is always present. For this reason, we should not count diagrams like (B) that measure the energy of the vacuum.[3]

Now that we know not to draw diagrams with vacuum sub-diagrams, we can fully compute any nPCF in the interacting theory with the Feynman rules. Our first and primary application for this newfound power will be to compute scattering amplitudes, which the next section is devoted to.

5.4 Spin Zero Scattering

In this section, we'll compute scattering amplitudes for $\phi\phi \to \phi\phi$ and $\phi \to \phi\phi\phi$ to get used to ϕ^4 theory and to practice with the Feynman rules. We'll follow up with a theory of two scalar particles and compute the probability of decay between the two.

Chapter 2 informed us that the scattering probability is the square of the S matrix, where the S matrix entry for m particles with momenta p_i scattering to n particles with momenta $-q_i$is given by the $(m+n)$PCF for momenta p_i and q_i. We noted however that the S matrix is mostly zero; only when the incoming total momentum equals the outgoing momentum can scattering occur. This appeared in our calculation in the previous section of the 4PCF as well: every diagram in (5.45) and (5.46) was multiplied by $\delta(p_1 + p_2 + q_1 + q_2)$.

We stripped off this delta function part by defining the scattering amplitude $-iM$ as the part of the S matrix with $p_1 + p_2 + q_1 + q_2 = 0$, but excluding the no-scattering, diagonal entries where $p_1 = p_2$ and $q_1 = q_2$. The (A) diagrams for the 4PCF shown in figure 5.3 had delta functions $\delta(p_1 - p_2)$ and $\delta(q_1 - q_2)$,

[3]This is fortunate for us, since the vacuum diagrams can be infinite. The internal vertices come with integrals over momenta due to the Feynman rules, but the lack of external vertices gives fewer delta functions. For many internal vertices, the diagrams can blow up.

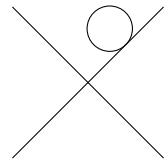
revealing these to be no-scattering diagrams. The only diagram that represents actual $\phi\phi \to \phi\phi$ scattering to first order is diagram C.

To extract a scattering amplitude from C, we must divide the corresponding nPCF by the propagator for each incoming and outgoing field, as **JTD: Equation number** dictates. The motivation for this was that we are not interested in the way the particles move before or after they interact, which is represented by the propagators connected to the external legs. We are interested in the actual interaction, which occurs in the middle of diagram (C). Once this division is complete, we get our first ever scattering amplitude:

$$-iM_{\phi\phi\to\phi\phi} = -i\lambda + \mathcal{O}(\lambda^2). \tag{5.48}$$

This equation is very simple, perhaps disappointingly so. Let's calculate the second-order contributions to the scattering amplitude to see if they are more interesting.

The first step is to draw new second-order diagrams. Many of these diagrams will look like this:

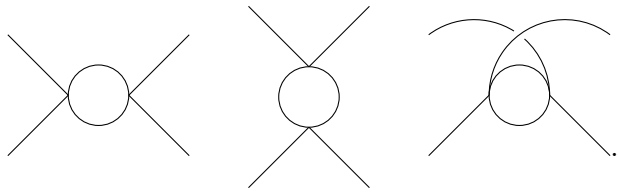


but these diagrams actually do not contribute to the scattering amplitude. Recall that the point of dividing by $G(p)$ for every external propagator p was to account for the motion of a particle before it interacts. The loop in the above diagram represents particle activity after it has interacted with the rest of the diagram. This loop represents self-interaction and is interesting in its own right. In fact, we'll study it in section 5.8. However, it still represents a particle propagating on its own and we must divide it out. Therefore, this diagram is just the same $-i\lambda$ first-order diagram we computed above.

This is an example of a generalized rule for calculating scattering amplitudes: the only diagrams that contribute to amplitudes are *truncated*. The definition of truncated is as follows:

In a truncated diagram, no internal propagators can be cut that sever exactly one external vertex from the diagram.

An internal propagator is one that does not connect directly to an external vertex. The above diagram was not truncated because the propagator between the central vertex and the loop could have been cut, and this would have severed the upper right external vertex from the diagram. However, the following diagrams *are* truncated, and do contribute to the matrix amplitude:



(5.49)

JTD: Put in momentum arrows

These represent the first non-trivial effect of QFT we have seen. Everything we have done so far has confirmed classical expectations, from $M = \lambda$ above to momentum conservation. These diagrams are fundamentally quantum and relativistic phenomena and represent a a true prediction of QFT. Let's compute them using the Feynman rules. Two propagators in each diagram are identical and can be exchanged, so they have symmetry factors of 2. Thus,

$$\begin{aligned}
\text{Left} &= -\frac{\lambda^2}{2}\int \frac{d^4k}{(2\pi)^4}\frac{1}{k^2-m^2}\frac{1}{(k-p_1-p_2)^2-m^2} \\
\text{Center} &= -\frac{\lambda^2}{2}\int \frac{d^4k}{(2\pi)^4}\frac{1}{k^2-m^2}\frac{1}{(k-p_1-q_1)^2-m^2} \\
\text{Right} &= -\frac{\lambda^2}{2}\int \frac{d^4k}{(2\pi)^4}\frac{1}{k^2-m^2}\frac{1}{(k-p_1-q_2)^2-m^2}.
\end{aligned} \tag{5.50}$$

where we have removed the external propagators in the spirit of truncation. It's common to deal with these diagrams by defining *Mandelstam variables*, which are a reparametrization for any two-to-two-particle interaction:

$$s = (p_1+p_2)^2 \qquad t = (p_1-q_1)^2 \qquad u = (p_1-q_2)^2. \tag{5.51}$$

There are three variables because momentum conservation makes a fourth one obsolete. They have the interpretation that s is the incoming momentum, equivalently the energy in the center of mass frame, and t and u represent the amount of momentum transfer. Calculation will reveal that each diagram is sensitive only to s, t,

or u respectively, so all three diagrams can be added up as follows:

$$-iM = -i\lambda - \lambda^2 \left(V(s) + V(t) + V(u)\right) + \mathcal{O}(\lambda^3) \tag{5.52}$$

where

$$V(s) = \frac{1}{2}\int \frac{d^4k}{(2\pi)^4} \frac{1}{k^2 - m^2} \frac{1}{(k - p_1 - p_2)^2 - m^2}. \tag{5.53}$$

This integral may strike you as a bit peculiar. Let's pretend for a moment that k^2 is the Euclidean norm $k_0^2 + \boldsymbol{k}^2$ rather than the Minkowski norm $k_0^2 - \boldsymbol{k}^2$. Then we could switch to spherical coordinates and d^4k would become $k^3 dk$ times a radial volume element $d\Omega$. For large k, this k^3 cancels with the integrand, which goes like k^{-4}, leading to an overall trend of $\int^\infty dk/k$. This integral is divergent.

The techniques for handling this divergence took a great deal of time to be developed historically. They necessitate a change in thinking which we will discuss in chapter **??**, so we'll postpone this particular integral until then. In three spacetime dimensions (two spatial dimensions), the integral trends to $\int^\infty dk/k^2$ which is a perfectly reasonable number. Figure 5.4 shows the plot of $V(s)$ in three dimensions.[4]

Something immediately apparent about Figure 5.4 is that the values are small. Multiplied by a small number λ^2 as 5.52 requires, and this $V(s)$ correction will be tiny except near the maximum value of $s = 4m^2$, where the function blows up. This is the hyper-relativistic limit, where the particles are colliding at enormous speeds. In this limit, the quantum corrections introduced by the second-order diagrams can become arbitrarily large. (5.53) is actually analytical computable:

$$V(s) = \frac{i}{16\pi\sqrt{s}} \ln\left(\frac{m + \sqrt{s/4}}{m - \sqrt{s/4}}\right). \tag{5.54}$$

We will discuss how to arrive at this value for the integral in chapter **??**.

Taking the limit as $s \to 4$ reveals that $V(s)$ blows up and scattering becomes extremely intense. This is a QFT correction to the classical expectation that $M = \lambda$; i.e., scattering amplitude is independent of energy.

[4]You might reasonably object to us computing predictions for an unphysical universe, like one with only three spacetime dimensions, in a book about the phenomena of this universe. However, many of the phenomena we'll see are qualitatively similar between dimensions, and the characteristics of dimension four specifically will be covered later.

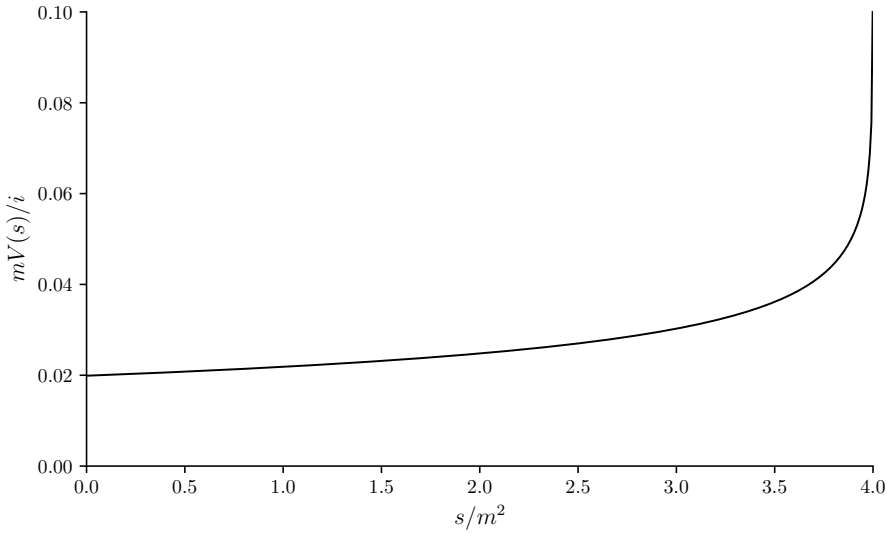


Figure 5.4: $mV(s)/i$, the unitless second order component of the scattering amplitude. This is computed for two scalar particles in 3 spacetime dimensions of center-of-mass energy s.

Intuition for Feynman Diagrams

In the derivation of this chapter, each propagator in a Feynman diagram represents a Wick contraction, but there is an intuitive picture for Feynman diagrams beyond the mathematical structure of Wick's theorem. We already know an external propagator — a particle that links to an external vertex — represents a particle entering the scattering experiment. Why not view *every* propagator as a particle? Then the ϕ^4 term of the Lagrangian, which corresponds to four propagators meeting at a vertex x in a diagram, really does represent an interaction at x. A diagram with a loop, such as those in (5.49), represent two particles colliding, only for the outgoing particles to collide with each other again. If one collision is rare, then this secondary collision should be rarer than one collision only, which explains why these diagrams are small. On the other hand, if a collision is common, we might expect every particle pair to collide not just once or twice but many times, in which case we would have to draw infinitely many diagrams. It is at this point that QFT is no longer perturbative, and alternative methods to Feynman diagrams must be taken to compute scattering amplitudes.

There is a small problem with this picture, however. The two internal propagators in (5.49) are not real particles in a key sense: they are not on-shell. This can be seen in (5.50): we integrate k over all four-dimensional space without requiring $k^2 = m^2$. Instead, we call these internal propagators *virtual particles*, which interact like normal particles but need not have all the physical characteristics of a true particle.

These virtual particles could have been predicted even before we saw the QFT principle of least action. One of the major results of special relativity was that mass is a form of energy, and a result of quantum mechanics is that a quantum particle may tunnel through a high-energy barrier in order to reach a low energy final state. Any theory of quantum mechanics and relativity, such as QFT, should combine these two well-understood phenomena. What better way to combine them than to allow particles to "tunnel" into an unphysical, perhaps higher-energy state before reaching a final state? The loop propagators represent this higher-energy state, and are unphysical in the same way that the electron should not classically exist in a high-energy barrier in quantum mechanics. Nevertheless, a particle can pass through the high-energy state in order to attain a different final state.

Since we needed to invoke quantum mechanics to interpret the loop diagrams of (5.49), we'll often treat any diagram with a loop as a "quantum correction," meaning a small correction quantum mechanics imposes on a classical theory. Figure 5.4 for the second-order $\phi\phi \to \phi\phi$ amplitude was our first example of a quantum corrections, but more will appear such as mass, energy, magnetic and electric dipole moment corrections.

The diagrams without loops, often called *tree-level diagrams*, generally represent the properties of a classical theory that could be attained by merely minimizing the action via the Euler-Lagrange equations. For example, a tree-level diagram delivered us $M_{\phi\phi\to\phi\phi} = \lambda$ as classical thinking would suggest.

Thinking of internal propagators as particles will be useful in the next section, where we discuss the ability of light scalars to carry long-range forces between other scalars.

5.5 Coulomb Potential for Scalars

JTD: Some Born approx shit

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m\phi^2 - \frac{1}{2}(\partial_\mu\chi)^2 - g\phi^2\chi \tag{5.55}$$

5.6 Complex Scalar Theory

So far, we've considered only real scalar fields where $\phi^* = \phi$, but one may well ask if complex scalar fields exist. After all, the fermions in Nature can all take on complex values, and the only fundamental scalar particle in Nature (the Higgs) is complex. Fortunately for us, generalizing the above results to a complex field will be mathematically easy. However, it will bring up a critical new physics concept which we will need to understand: the antiparticle.

Antiparticles

The Lagrangian of a complex scalar field is nearly the same as that of a real field. The only differences are coefficients to the term and the presence of complex conjugates:

$$\mathcal{L} = -|\partial_\mu \phi|^2 - m^2|\phi|^2 + \lambda|\phi|^4. \tag{5.56}$$

The reason why the coefficients have changed is so that the counting arguments of table 5.3.2 which allow computation of the symmetry factor still apply even though now $\phi^* \neq \phi$.

Immediately, we suspect that ϕ^* may be the antiparticle of ϕ. This is because, as mentioned in chapter 1, an antiparticle is the mirror image of a particle under CPT (charge-parity-time) reversal symmetry. Parity and time reversal do nothing on this ϕ scalar particle, but charge reversal is manifested in complex conjugation **JTD: why?** , so that CPT symmetry takes ϕ to ϕ^*. Thus, ϕ^* is the antiparticle of ϕ.

The antiparticle explanation may not be fully satisfying, however, since (5.56) still looks like the Lagrangian of only one particle. Let's therefore break up the ϕ field into components. Just as a complex number $z = (a + ib)\sqrt{2}$ can be broken into two real components (the $\sqrt{2}$ is for normalization), the complex field $\phi = (\phi_a + i\phi_b)/\sqrt{2}$ can be broken into real fields ϕ_a and ϕ_b. Plugging these into the Lagrangian, we get

$$\begin{aligned}\mathcal{L} = &-\frac{1}{2}(\partial_\mu \phi_a)^2 - \frac{m^2}{2}\phi_a^2 + \frac{\lambda}{4}\phi_a^4 \\ &-\frac{1}{2}(\partial_\mu \phi_b)^2 - \frac{m^2}{2}\phi_b^2 + \frac{\lambda}{4}\phi_b^4 \\ &+\frac{\lambda}{2}\phi_a^2\phi_b^2.\end{aligned} \tag{5.57}$$

We can now see that the complex scalar Lagrangian is just two real fields with the same mass m, same self-interaction λ, and an $a - b$ interaction term which is the last line. These properties were required for (5.56) to be written so simply.

The complex scalar Lagrangian (5.56) looks elegant, but when broken into its components it seems oddly contrived. Why should the real and imaginary parts have the same mass and interaction strengths? Why should they interact with each other in the exact way shown by (5.57)? These constraints are imposed by CPT symmetry. A particle's antiparticle must have the same mass and interactions. Mathematically speaking, they obey a *symmetry*, in that swapping ϕ_a and ϕ_b does not change the Lagrangian.

In fact, the symmetry is even deeper than this; any (normalized) linear combination of the two fields is also indistinguishable from ϕ_a and ϕ_b. For example, if we define $\chi_a = (3\phi_a + 4i\phi_b)/5$and $\chi_b = (4\phi_a - 3i\phi_b)/5$, we could rewrite (5.57) in terms of χ_a and χ_b and we would get the exact same form, just with ϕ swapped out for χ. If it weren't for the odd $\frac{\lambda}{2}\phi_a^2\phi_b^2$ interaction term, this would not be true.

JTD: Unitary group too

This particular symmetry is called SO(2), or *the Special Orthogonal Group of Order 2*. The name is motivated by the kinds of linear combinations which do not change the Lagrangian; "special" indicates that the new linear combinations χ must remain normalized, "orthogonal" indicates that the two fields χ_a and χ_b must have the same product as the original $\phi_a\phi_b$, "group" indicates that any linear combination that satisfies these two properties is allowed as a symmetry of the Lagrangian, and "order 2" means that the linear combination acts on two fields ϕ_a and ϕ_b. This is our first example of a *continuous symmetry* because the set of available linear combinations is continuous. Nature adores continuous symmetries. The Standard Model which is the best known description of the fundamental particles has three of these symmetries, as will be discussed in the next part of this book.

You might wonder where ϕ's antiparticle was for the real scalar field case. For real scalars, the particle is its own antiparticle. This also occurs for some Standard Model particles, such as the photon.

Feynman Rules for Complex Scalars

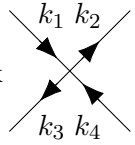
Now that we have shown that a complex scalar contains an antiparticle, it is time to discuss the Feynman rules of the new theory. Fortunately, we do not have to repeat the many pages of derivations necessary for the real field theory. The Lagrangian

is essentially the same, so the Green's function of a complex scalar field is identical to that of a real field. Wick's theorem and its application to the free theory and the interacting theory did not rely on anything other than that the perturbation λ is small. Thus, we can skip straight to the rules for how to draw Feynman diagrams.

The new Lagrangian restricts diagrams slightly. The ϕ^4 term became $\phi^*\phi\phi^*\phi$, meaning that only two antiparticles and two particles can meet at a vertex, not three and one, for instance. We'll have to enforce this in Feynman diagrams with a graphical rule. Suppose we draw an arrow on a propagator, where the arrow points one way for a particle and another way for an antiparticle. Then the $\phi^*\phi\phi^*\phi$ term means every vertex must have two arrows leading in and two leading out. (Since ϕ and $\overline{\phi}$ have opposite charge, you can think of these arrows as charge arrows.) Thus, the Feynman rules are

1. The edge 

contributes $G(k)$

2. A vertex 

contributes $-i\lambda\delta(k_1 - k_2 - k_3 + k_4)$

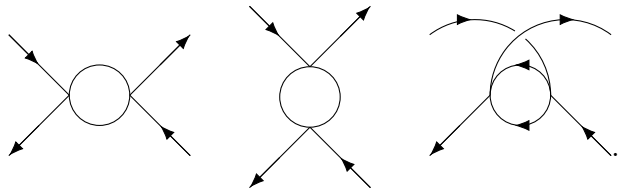
3. Every internal propagator k is integrated over $\int \frac{d^4k}{(2\pi)^4}$

4. Diagrams with exchangeable vertices or reversible edges contribute a symmetry factor S_f.[5]

We stopped drawing the momentum arrows in the diagrams to reverse clutter. There's no harm in just defining momentum to point in the same direction as the charge arrows.

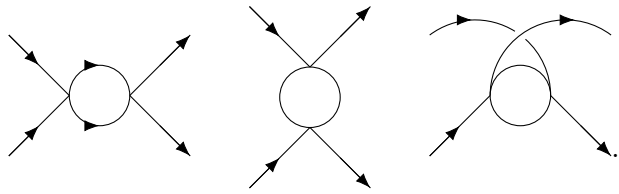
As an example, here are the diagrams of (5.49) for a complex scalar field, representing $\phi\overline{\phi} \to \phi\overline{\phi}$ scattering:

[5] You might recall that the definition of S_f derived from a counting problem for how much each diagram should be weighted by, and that problem depended on ϕ being real. The coefficients of the new Lagrangian (5.56) have been carefully chosen so that the definition of the symmetry factor is the same, so we don't have to worry about redoing that argument.



(5.58)

There are similar diagrams for $\phi\phi \to \phi\phi$ scattering:



(5.59)

When a loop is drawn without charge arrows, the charge arrows can go either way so long as they are antiparallel. This does *not* count as a reversible edge in the symmetry factor sense; a symmetry factor represents that we have overcounted a diagram, whereas if a loop can run either of two directions we are undercounting the diagrams. Thus, we should double each diagram where the loop can run either day, not half it.

Thus, a difference between complex scalar scattering and real scalar scattering is that the quantum corrections to complex scattering is more intense. Contributions from both particles and antiparticles can occur, increasing the total interaction strength. The exception is in some channels which forbid particle-antiparticle interactions. For example, for $\phi\phi \to \phi\phi$, the s channel is not amplified because only ϕ particles may mediate that interaction. In $\phi\overline{\phi} \to \phi\overline{\phi}$, the u channel is not amplified.

One of the most famous properties of matter and antimatter is that they explode on contact with each other. Electrons and positrons annihilate to two or three photons, as do neutral pions. Charged pions decay to muons and neutrinos, et cetera. In the next section, we'll introduce a real scalar field as a photon-like particle and compute amplitudes for a complex scalar field to decay in this theory.

5.7 Scalar Decay

5.8 Massless Scalars are Unprotected

Problems

5.1 Wick's Theorem. In section 5.1.2, we explained Wick's theorem, which is that $\frac{d^n Z}{dJ_{j_1}...dJ_{j_n}}$ is equal the sum of M^{-1}_{ab}, where a, b are pairs of $j_1, \ldots j_n$. This was shown for $n = 2$ in (5.16). Prove via induction that this fact holds for $n > 2$.

5.2 Momentum Conservation. In this question, we'll determine why momentum conservation came so naturally out of the QFT principle of least action.

1. Show that all Feynman diagrams, of any order and with any number of external vertices, conserve momentum in that if the external momenta do not sum to zero, then the nPCF is zero.
2. **JTD: Connect this to the real-ness of the Lagrangian via Fourier transform logic**
3. Show that the operator e^{iH} is unitary if and only if H is Hermitian. Since the integrand of the QFT principle of least action is $e^{i\mathcal{L}}$, where $\mathcal{L}$ is real, the conservation of momentum structure we got is just the requirement that $e^{i\mathcal{L}}$ be unitary. Unitarity is a property of non-relativistic quantum mechanics that was necessary for the probability to detect a particle not to change over time. The QFT principle of least action is written to preserve unitarity in the sense of momentum conservation, even though this theory can create and destroy particles.

5.3 Symmetry Factors. **JTD: Evaluate some diagrams and find some symmetry factors**

5.4 New Lagrangians. Suppose we designed a new scalar theory in which ϕ were real. This allows us to write a cubic term in the Lagrangian instead of a quartic term:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - m^2\phi^2 + \frac{g}{3!}\phi^3.$$

Write out the Feynman rules for this theory. What would be the Feynman rules of the Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - m^2\phi^2 + \frac{g'}{4}\phi^2(\partial_\mu \phi)^2?$$

Hint: this question requires very little math.

Chapter 6

Spin One Half

This chapter deals with extending QFT to cover particles with spin, also known as *fermions*. Spin is covered in non-relativistic quantum mechanics, though its properties seem sometimes mysterious and arbitrary. In particular, the fact that fermions obey the Pauli exclusion principle and bosons do not cannot be explained in non-relativistic quantum mechanics. This explanation is provided by relativity in QFT in the form of the spin-statistics theorem, which we explain in this chapter. Before arriving at this key result, however, we discuss spin in a generalized sense (section 6.1) and show that when relativity is taken into account, spin implies the existence of antiparticles (section 6.2).

Let's start with a simple definition: spin is the intrinsic angular momentum of a particle. In order to store a particle's spin state in QFT we'll need to graduate our scalar field particle operator $\hat{\phi}$ to a vector of operators $\hat{\psi} = (\hat{\psi}_0, \hat{\psi}_1, \dots)^T$. This vector is called a *spinor*.

The next section uses the definition of spin as intrinsic angular momentum to dictate how many entries the spinor must have and what operators can act on it. Our discussion will start with the non-relativistic case, where we do not consider boosts, and then we will add boosts to see how they change the discussion. Following this, we'll describe how to similarly graduate the ψ in the path integral of the QFT principle of least action to a spinor. With this complete, we will repeat the analysis of the previous chapter and derive Feynman rules for fermions. Finally, we will use these new fermions to understand electrons, the Coulomb potential, and atomic nuclei.

6.1 Spinors and Spin Operators

Spin is a quantum state, and in any quantum theory, a quantum state is acted on by operators. It will require the description of new mathematical tools which come from the field of group theory, but by the end we will have a complete description of any particle with spin.

To measure the spin of a spinor ψ, we act on ψ with a vector of operators $\boldsymbol{S}$. In order to be interpreted as an angular momentum vector, $\boldsymbol{S}$ must satisfy the defining feature of angular momentum:

Angular momentum generates rotations.

The notion of a *generator* is our first group theoretic term and we will define it in the following paragraph. But first, this definition of angular momentum is different from the classical definition of $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$. It is nevertheless a correct definition due to Noether's theorem, which states that an operator's generator is conserved if that operator is a symmetry of the Hamiltonian (appendix H). Rotations are a symmetry of spacetime, meaning that if angular momentum generates rotations then angular momentum is conserved. The cross product definition of the angular momentum will return later in this section.

Generators

Now to define what a generator is. Consider a rotation operation R which acts on a three-component vector. We know that R can be written as products of the three-dimensional rotation matrices:

$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \qquad R_y(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}$$
$$R_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{6.1}$$

This fact shows that there are many possible rotation matrices, but that they can each be traced down to the three matrices above. In fact, the above rotation matri-

ces can be simplified even more by writing $R_j(\theta) = e^{i\theta S_j}$, where

$$S_x = \frac{i}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \qquad S_y = \frac{i}{2}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$
$$S_z = \frac{i}{2}\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{6.2}$$

You can check for yourself that exponentiating these matrices gives back the rotation matrices. An arbitrary rotation can therefore be written as

$$R = e^{i\boldsymbol{\theta}\cdot\boldsymbol{S}} \tag{6.3}$$

where we have arranged S_x, S_y, and S_z into a vector called $\boldsymbol{S}$. The three components of the vector $\boldsymbol{\theta}$ are called Euler angles and they parameterize the rotation matrix R.

In mathematical language, we say that $\boldsymbol{S}$ *generates* the rotations because all rotations can be achieved by plugging different values of $\boldsymbol{\theta}$ into (6.3). The fact that angular momentum is generated by rotation could mean that we should just take $\boldsymbol{S}$ as the spin operator $\boldsymbol{S}$, and write the spinor as a three dimensional vector which $\boldsymbol{S}$ acts on.

Angular Momentum Algebra

This procedure does not work in practice because there are other ways to generate rotations. One can rotate a complex vector for instance, in which case we need a new set of $\boldsymbol{S}$ matrices which are themselves complex. There is an even finer point which must be made: it is necessary that $\boldsymbol{S}$ generate a three-dimensional rotation since spacetime is invariant under three-dimensional rotations. However, the matrix R need not act on a three-dimensional vector. R could act on a five-dimensional vector, for example, and rotate only in a three-dimensional subspace. In this case $\boldsymbol{S}$ would still contain three generators, but each generator would be a 5×5 matrix rather than 3×3.

We therefore need to write a more general definition of rotations than just the rotation matrices (6.1). One property rotations must have is that they must preserve

the length of vectors. That is, R must be unitary. Problem **JTD: cite** asks the reader to show that for $\boldsymbol{e}^{iM}$ to be unitary, M must be hermitian and determinant 1. Thus, the entries of $\boldsymbol{S}$ are hermitian.

The other required property of rotations comes from the way in which they anticommute. Consider rotating a vector aligned with the z-axis. If we rotate in the x direction by a small angle δ first and then y by the same $delta$, we get a different vector than rotating in the y direction first then y. This is because the rotation matrices anticommute. Using (6.1) or drawing pictures that the two rotations differ by a rotation about the z axis by angle δ^2. We can turn this fact into a constraint on $\boldsymbol{S}$ by using (6.3) to define the x, y, and z rotations. They show that $S_x S_y - S_y S_x = iS_z$.

In principle, we could have aligned our initial vector along any axis z, so that the more general version of this constraint is

$$[S_i, S_j] = i\epsilon_{ijk} S_k \tag{6.4}$$

where ϵ_{ijk} is the Levi-Civita symbol. Any set of hermitian, determinant one matrices $\boldsymbol{S} = (S_x, S_y, S_z)$ which satisfies this equation is a valid angular momentum operator. We usually refer to (6.4) as an *angular momentum algebra*, and matrices that satisfy it are *representations* of the angular momentum algebra

Using Lie groups to find representations of the angular momentum algebras

The task of finding matrices $\boldsymbol{S}$ which satisfy (6.4) (or in mathematical language, finding representations of the angular momentum algebra) has fortunately been solved using group theory. A Lie group is a closed, continuous set of operators, where closed signifies that two operators in the group U and V multiply to produce another operator W which is also in the group. We have unknowingly been working with the group of rotations in three dimensions, which is called the "special orthogonal group of order three," or SO(3). "Special" means that the group operators do not change the length of vectors, "orthogonal" means that the group operators are real and unitary, and the three indicates that the rotation occurs in three dimensions.

A Lie group satisfies the general result that any group element G can be written

as

$$G = e^{i\boldsymbol{\theta}\cdot\boldsymbol{S}} \tag{6.5}$$

where $\boldsymbol{S}$ are a special set of matrices known as the generators of the Lie group. We used this property in (6.3), but we did not know that it was valid for any Lie groups. The matrices $\boldsymbol{S}$ must satisfy special relations to each other which are dependent on the Lie group the generate; these relations are called a Lie algebra. An example is the angular momentum algebra we discussed above, which we now know as the Lie algebra of SO(3). Mathematicians sometimes label this algebra as $\mathfrak{so}(3)$ but physicists usually use SO(3) to refer to the algebra in addition to the group. We will therefore use the $\mathfrak{so}(3)$ notation in this section only.

Once we specify the dimensions of the spinor which the group elements act on — say the spinor has dimension n, then we can write group elements as $n \times n$ matrices. This is called a representation. Our three-dimensional rotation matrices (6.1) were a representation of SO(3), and our three-dimensional angular momentum operators (6.2) were representations of its Lie algebra. Matrix exponentiating the generators as in (6.5) can be used to produce the representation of the group if a representation of the Lie Algebra is known.

Physical consequences of the group theory explanation for spin

This new formalism connects directly to the definition used at the beginning of this section — "Angular momentum generates rotations." The connection is that "rotations" refers to the Lie group SO(3), which has the Lie algebra given in (6.4). The angular momentum operators we're looking for are the rotation generators, which are representations of this Lie algebra.

One reason why this group theory framework is necessary is because (6.4) is not just the Lie algebra of SO(3). It is also the Lie algebra of SU(2), which is the set of unitary matrices which rotate a two-dimensional complex vector, rather than a three-dimensional real vector. Valid angular momentum operators could therefore be generators of SO(3) or SU(2).

Another use of group theory is the mathematical result that all the faithful[1] representations of SO(3) have odd dimensions, and all the representations of SU(2)

[1] A faithful representation is one where every group element corresponds to exactly one matrix. For odd dimensions, SU(2) is not faithful because every group element corresponds to two matrices which can be generated using $e^{i\boldsymbol{\theta}\cdot\boldsymbol{S}}$.

have odd[2] dimensions. Therefore, the spinor which the spin operators act on can have any number of dimensions n. If n is even, then one uses representations of $\mathfrak{su}(2)$ as spin operators. If n is odd, one uses representations of $\mathfrak{so}(3)$.

This fact that the parity of n changes what group the spin operators come from has several fascinating properties. Firstly, it means that the SU(2) spins may have different properties than the SO(3) spins. This is manifested in the spin-statistics theorem, which is described in the next section. It also means that since SU(2) acts on complex vectors and SO(3) acts on real vectors. This is crucial because, as noted in the previous chapter, complex particles have an electric charge. Thus, all SU(2) particles are automatically charged. SO(3) particles can be neutral or charged, because one can multiply them by a complex phase.

Another result of group theory is that $\boldsymbol{S}^2 = \frac{n^2-1}{4}\mathbb{1}$, where $\boldsymbol{S}^2 = S_x^2+S_y^2+S_z^2$. This is interesting because we define the spin of a particle s such that the eigenvalue of the angular momentum operator $\boldsymbol{S}^2$ is $s(s+1)$ so that $s = \frac{n-1}{2}$. Thus, particles with half-integer spin s have spinors of even dimension and are therefore represented by SU(2) and we refer to them as fermions. Integer spin particles, called bosons, have odd dimension spinors and are represented by SO(3).

One final result of group theory is a list of what the spin operators actually are. For zero spin ($s = 0$), the spinor has dimension 1, and the operators are all the 1×1 matrices (0). This is an incredibly boring case from the perspective of spin, and it is the reason why we were able to complete all of the previous chapter on spin-zero particles.

Spin one-half particles have $n = 2$, so that the spin operators are 2×2 representations of $\mathfrak{su}(2)$ which are the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6.6}$$

Non-relativistic quantum mechanics courses usually skip the derivation of spin operators and merely state that they are equal to these Pauli matrices, partially because most of the particles dealt with in non-relativistic quantum mechanics are spin one-half so that the operators for other particles are unnecessary.

Spin one particles have $n = 3$, so that the spin operators are 3×3 represen-

[2] Except, that is, for the trivial representation, where all the generators are the 1×1 matrix with zero as the only element.

tations of $\mathfrak{so}(3)$. We have already written these matrices in (6.2) because $n = 3$ is the representation used to rotate normal three-dimensional geometric vectors. This means that for spin-one particles (and only spin-one particles), spin is a true geometric vector which rotates normally and need not be understood in a quantum mechanical way. Photons are spin-one particles whose spin corresponds to their polarization, and classical electromagnetism only works to describe photons because of this fact — their spin behaves like a classical vector and can be treated as such. For this reason, we sometimes call spin-one particles "vector bosons." Vector bosons are able to play a special role in Lagrangians because their spin can be used as a Lorentz index to turn vector-valued quantities into scalar-valued quantities that can go into the Lagrangian. This allows them to act as "gauge bosons," as described in the next chapter.

Throughout the rest of this chapter, we will explore only the spin one-half case, saving spin-one for the next chapter.

6.2 The γ Matrices and Relativistic Spin

In the last section, we used the invariance of spacetime under rotations to justify that spin is conserved. However, the full symmetry of spacetime is Lorentz symmetry, of which rotations is just one part. A general Lorentz transformation is

$$\Lambda^{\mu}{}_{\nu} = \left(\begin{array}{c|c} \text{Time dilation} & \text{Boosts} \\ \hline \text{Boosts} & \text{Rotations} \end{array}\right) \tag{6.7}$$

In this section, we will explore the influence of boosts on spin one-half particles. Even before this discussion, we can partially work out the answer due to the CPT theorem, which is also due to relativity.

The CPT theorem states that quantum mechanics is invariant under CPT symmetry, but a spin one-half spinor $\hat{\psi}_L$ is not CPT invariant because it is charged and because its spin operators σ^{μ} have parity[3]. In the previous chapter, we saw that the antiparticle of a scalar particle ϕ was its complex conjugate $\phi^{\dagger}$, and we should therefore expect $\hat{\psi}_L^{\dagger}$ to be the antiparticle of $\hat{\psi}_L$. However, if we keep $\hat{\psi}_L$ as a two-component spinor, this does not solve our parity problem.

[3]The angular momentum operators have parity due to the ϵ_{ijk} in the angular momentum algebra (6.4).

The solution is to create another particle $\hat{\psi}_R$, which is the parity reverse of $\hat{\psi}_L$. Then we can create a four-component spinor out of both $\hat{\psi}_R$ and $\hat{\psi}_L$, known as a Dirac spinor,

$$\hat{\psi} = \begin{pmatrix} \hat{\psi}_L \\ \hat{\psi}_R^\dagger \end{pmatrix}. \tag{6.8}$$

The spin operators on the larger spinor are

$$\gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{6.9}$$

which is a compact notation for a four-dimensional matrix. We also define the timelike component of γ as

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}. \tag{6.10}$$

By defining the four vectors $\sigma^\mu = (\mathbb{1}, \sigma^i)$ and $\overline{\sigma}^\mu = (\mathbb{1}, -\sigma^i)$, we can write the γ^μ matrices in a more compact form:

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \overline{\sigma}^\mu & 0 \end{pmatrix} \tag{6.11}$$

which are called the Dirac gamma matrices.

To tell apart the two spinors, we also define a fifth matrix

$$\gamma^5 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}. \tag{6.12}$$

because the eigenvectors of this matrix are $\hat{\psi}_L$ and $\hat{\psi}_R$, with eigenvalues $+1$ and -1 respectively. So this matrix separates the left and right components. In particular, $\frac{1+\gamma_5}{2}$ selects out the left-handed component and removes the right-handed component, which will be crucial when we discuss the weak force.

We define the antiparticle of $\hat{\psi}$ as $\overline{\psi} = \gamma^0 \hat{\psi}^\dagger$. Explicitly, this antiparticle is

$$\overline{\psi} = \begin{pmatrix} \hat{\psi}_R \\ \hat{\psi}_L^\dagger \end{pmatrix} \tag{6.13}$$

which is truly the CPT reverse of $\hat{\psi}$. Thus, we needed to create a second particle $\hat{\psi}_R$ in order to make $\hat{\psi}$ consistent with relativity. For convenience, we put $\hat{\psi}_L$ and $\hat{\psi}_R$ into the same spinor, similarly to how we put ϕ_a and ϕ_b into one particle ϕ in the case of a complex scalar field in the previous chapter.

Though this justification of $\hat{\psi}_L$ and $\hat{\psi}_R$ via CPT symmetry is easy to explain, it leaves several questions which are summarized below. Their answers are provided by the more lengthy derivation of the γ^μ matrices as generators of the Lorentz group, just as we derived the σ matrices as generators of the rotation group. We spend the rest of the section performing this derivation to satisfy curiosity, though the rest of this book can be understood even if the rest of this section is skipped.

The questions which are answered by investigating the generators of the Lorentz group are

1. In what sense are $\hat{\psi}_L$ and $\hat{\psi}_R$ left- and right-handed? (They have different helicities)

2. For a free particle, how does the spinor behave? (ψ_L and ψ_R encode the spin of the particle times factors that are determined by its velocity.)

3. Does γ^μ transform as a four-vector under Lorentz boosts as the μ index suggests? (Yes.)

4. Are there other 4×4 representations for γ^μ? (Yes; this one is called the Weyl representation.)

5. What are the properties of the γ^μ matrices? (These are summarized in appendix E.)

The separation of four-dimensional spinors into two parts was historically crucial for two reasons. Firstly, it allowed us to use the Pauli matrices instead of the γ^μ matrices for decades, because the difference between ψ_L and ψ_R is only apparent when one performs a relativistic boost, and measurements of spin are usually made in one reference frame, where one only measures one of the two spinors. Secondly, Paul Dirac's discovery of γ^μ matrices was the first theoretical prediction of antimatter, which was first observed several years after. It was this discovery that drew attention to his Dirac equation, and later, QFT.

Generators of the Lorentz Group

A full four-dimensional sketch of boosts is given in (6.7), and an example of a boost in two dimensions is

$$\Lambda = \begin{pmatrix} \cosh\gamma & \sinh\gamma \\ \sinh\gamma & \cosh\gamma \end{pmatrix} \tag{6.14}$$

Our first task is to add generators for boosts to our rotation generators in order to create the full set of generators of the Lorentz boosts Λ. A glance at the above example shows that the boosts are almost rotations, but they are symmetric and the trigonometric functions are replaced with hyperbolic trigonometric functions. A little more thought reveals that the additional generators for the boosts are

$$\begin{gathered} B_x = i\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad B_y = i\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ B_z = i\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \end{gathered} \tag{6.15}$$

which can be used to verify the example. To write an algebra that contains both the boost generators $\boldsymbol{B}$ and the rotation generators $\boldsymbol{S}$ we define an antisymmetric tensor of matrices Σ where the entries are generators. The timelike entries are boosts $\Sigma^{0i} = B^i$ and the spacelike components are rotations $\Sigma^{12} = S_z$, $\Sigma_{13} = -S_y$, and $\Sigma_{23} = S_x$[4]. The Σ tensor then satisfies the algebra

$$[\Sigma^{\mu\nu}, \Sigma^{\rho\sigma}] = i\left(\eta^{\nu\sigma}\Sigma^{\mu\rho} + \eta^{\mu\rho}\Sigma^{\nu\sigma} - \eta^{\mu\sigma}\Sigma^{\nu\rho} - \eta^{\nu\rho}\Sigma^{\mu\sigma}\right) \tag{6.16}$$

where $\eta^{\mu\nu}$ is the spacetime metric. This equation is called the *Lorentz algebra*.

We have done nothing except put the spin operators in relativistic language by adding in boost generators so far. Because there are now six generators, the Σ tensor can no longer be considered angular momentum. However, we can use

[4]This structure of the Σ tensor may seem odd, but it is very similar to the layout of the electromagnetic tensor, also known as the Faraday tensor, where the electric field is laid along the timelike components and the magnetic field along the spacelike components.

a trick due to Dirac to reduce the tensor to four components by defining a four-vector of matrices γ^μ such that

$$\Sigma^{\mu\nu} = -\frac{i}{4}\left[\gamma^\mu, \gamma^\nu\right]. \tag{6.17}$$

Then the Lie algebra of $\Sigma^{\mu\nu}$ implies the following Lie algebra for γ^μ:

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \tag{6.18}$$

where $\{\cdot, \cdot\}$ is the anticommutator. (6.18) is called the *Clifford algebra* or the Dirac algebra, and its representation γ^μ are called the *Dirac gamma matrices*.

We have finally arrived at a four-vector γ^μ which satisfies the properties of a relativistic spin operator. Its spacelike components should be related to the spin operators derived in the previous section because those operators were the spacelike components of $\Sigma^{\mu\nu}$, and the commutator (6.17) merely transforms two Pauli matrices γ^μ and γ^ν into the third Pauli matrix due to the angular momentum algebra (6.4). The only challenge is the that the boosts prevent the γ matrices from being odd-dimensional. In order to create them, we must double the dimensions of the old spin operators.

One can manually check that (6.11) is a valid representation of the γ^μ matrices — i.e., it satisfies th Clifford algebra. Any other representation that satisfies the Clifford algebra is also valid but we will not need them for this book. Likewise, there is no point in giving the additional 6×6 and higher-dimensional representations of the Clifford algebra because we will not use them, but they do exist.

Lorentz boosting the γ^μ matrices

We have defined the γ^μ matrices with a spacetime index, so far just because there were four γ^μ matrices and spacetime is four-dimensional. However, this index is not just for show. It is a true spacetime index in that it indicates that the γ^μ matrices transform as vectors under Lorentz transformations.

To see this, we will first act on the γ^μ matrices with the spin representation of a boost Λ. This is sometimes called an "internal boost" because it affects only the spin. We can write Λ as an exponential using the $\Sigma^{\mu\nu}$ generators of the Lorentz

algebra which we just worked out:

$$\Lambda = e^{-i\theta_{\mu\nu}\Sigma^{\mu\nu}}. \tag{6.19}$$

In principle, $\theta_{\mu\nu}$ can be any tensor of constants. However, because $\Sigma^{\mu\nu}$ is antisymmetric, only the antisymmetric part of $\theta_{\mu\nu}$ will matter. Acting with Λ on the γ^μ matrices gives the transformation

$$\gamma^\mu \to \Lambda^{-1}\gamma^\mu\Lambda. \tag{6.20}$$

Remember that the $\Sigma^{\mu\nu}$ matrices do not commute with γ^μ.

If we were to transform the γ^μ as a spacetime vector rather than as a set of matrices, then we would get

$$\gamma^\mu \to \Lambda^\mu{}_\nu\gamma^\nu. \tag{6.21}$$

where this time, $\Lambda^\mu{}_\nu$ is a boost of spacetime, meaning that its entries are numbers and not matrices. This is called an "external boost" because it explicitly does not affect spin.

It helps to write the $\Lambda^\mu{}_\nu$ matrix as an infinitesimal boost, in which case we can use the generators of the Lorentz algebra that we derived above in (6.15), which we now call $J^{\mu\nu}$ to distinguish them from the generalized spin operators $\Sigma^{\mu\nu}$. Then $\Lambda^\mu{}_\nu = \delta^\mu_\nu + \left(\frac{1}{2}\omega_{\alpha\beta}J^{\alpha\beta}\right)^\mu{}_\nu$. **JTD: Some factors of 1/2 are wrong** . Using this fact, a few calculations using the Lorentz algebra (6.16) and Clifford algebra (6.18) give that the internal and external boosts are identical: that is,

$$\Lambda^{-1}\gamma^\mu\Lambda = \Lambda^\mu{}_\nu\gamma^\nu. \tag{6.22}$$

This confirms that γ^μ transforms like a four-vector.

Helicity and the Free Particle Spinor

A last question we must answer is what the ψ_L and ψ_R components of a Dirac spinor are. The answer is that they both express the alignment of a particle's spin with its velocity — called helicity — which we will show by boosting a particle with spin from its rest frame into the lab frame and showing how its spinor behaves.

This boost is a good trick to learn since it connects a particle's Dirac spinor to the spin which is actually measured. To predict a spin measurement, we generate

a prediction for the Dirac spinor and then boost to the lab frame to predict what spin it truly measures. However, most particle detection methods are not sensitive to spin, in which case the spinor can be ignored and this boost technology is not necessary.

To boost a spinor, we act on it with $\Lambda = e^{i\omega_{\mu\nu}\Sigma^{\mu\nu}}$. Suppose for now that $\omega_{\mu\nu}$ is all zero except $\omega_{30} = -\omega_{03} = \epsilon$ which is small. This is a boost in the $+z$ direction, and has the effect of

$$\Lambda\psi = (\mathbb{1} + \epsilon B_z)\,\psi. \tag{6.23}$$

Suppose that the initial state ψ is just a ψ_R particle in the down state. $\psi = (0,0,0,1)^\dagger$. Then the boosted spinor is

$$\Lambda\psi = (\epsilon, 0, 0, 1). \tag{6.24}$$

It appears that the left-handed spinor ψ_L has gained a small component in the up state due to this boost. A stronger boost will increase the ψ_L component and decrease the ψ_R component. **JTD: Math is wrong, and I might be using inconsistent reps.**

This is an indication that ψ_R and ψ_L reflect how much a particle's spin is aligned with its velocity. As we boost and reduce the velocity of the particle, the alignment equalizes. If we boosted too much, the alignment would reverse. This alignment, more technically defined as the projection of a particle's spin onto its momentum, is called *helicity* and is conserved. Thus, a ψ_L with left-handed helicity will remain left-handed until interacted with. This is why we may insert it into the spinor.

JTD: Rewrite this section to actually perform this boost and show the spinor components as a function of velocity .

6.3 Spin Statistics Theorem: Fermions Anticommute

So far, we have discussed the structure of a spinor $\hat{\psi}$, which is used to define nPCFs. However, the other side of the QFT principle of least action requires us to integrate over numbers ψ, not operators. We will end this chapter by discussing how to define the number ψ, but first we'll start with a crucial result of QFT: the spin-statistics theorem.

Suppose

JTD: Spin statistics

JTD: Grassman variables.

6.4 Feynman Rules for Fermions

6.5 Yukawa Theory Potential

6.6 Pion Model of the Weak Force & Fermionic Spontaneous Symmetry Breaking

6.7 Yukawa Theory Electric Dipole Moment

Chapter 7

Spin One

7.1 Gauge Invariance

7.2 The Spin One Lagrangian

7.3 The Ward Identity

7.4 Coulomb Potential

7.5 Magnetic Dipole Moment

7.6 Massless Photons are Protected & Bosonic Spontaneous Symmetry Breaking

Chapter 8

Infrared divergences

8.1 Bremsstrahlung Diagrams

8.2 Massive photon renormalization

Chapter 9

Ultraviolet divergences

9.1 Wilsonian Renormalization

9.2 On-Shell Renormalization

9.3 Dimensional Regularization

Chapter 10

Effects of Ultraviolet Divergences

10.1 Electron Mass and Charge renormalization

10.2 Photon Self Energy

10.3 Electron Anomalous Magnetic Moment

10.4 Compton Scattering

10.5 Positronium Decay

10.6 Scattering to Hydrogen

10.7 Lamb Shift

10.8 (Other Hyperfine examples)

Part III

Non-Abelian Theories

Chapter 11

Yang-Mills Theory

11.1 Non-Abelian Symmetry Breaking

Chapter 12

Electroweak theory

12.1 Non-Abelian Symmetry Breaking

12.2 Feynman Rules for the Electroweak theory

12.3 Muon lifetime

12.4 Higgs Decay

Chapter 13

Standard Model

Part IV

Review

Chapter 14

Review of Special Relativity

14.1 Tenets of Special Relativity

14.2 Time Dilation, Length Contraction, and the Lorentz Group

14.3 Particle Momenta in Special Relativity

Chapter 15

Review of Electromagnetism

15.1 Maxwell's Equations

15.2 Covariant Electromagnetism: the Faraday Tensor

15.3 Gauge invariance

Chapter 16

Review of Non-relativistic Quantum Mechanics

16.1 Tenets of Quantum Mechanics

16.2 Schrödinger and Heisenberg Picture

16.3 Non-Relativistic Scattering

Since the predominant use of relativistic QFT is to compute scattering cross sections, it is worth reviewing non-relativistic scattering in quantum mechanics. Consider a particle propelled toward some potential $V(\boldsymbol{x})$ which tails off at large $\boldsymbol{x}$. For example, an electron incident on a proton with Coulomb attraction. When the particle is far from the potential, its wavefunction is a superposition of plane waves. For simplicity, suppose the initial state is just one plane wave of momentum $\boldsymbol{p}$:

$$|\psi(t \to -\infty)\rangle = |\boldsymbol{p}\rangle = \int dx\, e^{i\boldsymbol{p}\cdot x}\, |x\rangle\,. \tag{16.1}$$

The final state is also far from the potential, so it is a superposition of many plane waves

$$|\psi(t \to +\infty)\rangle = A\,|\boldsymbol{p}\rangle - i \int d^3\boldsymbol{k}\ |\boldsymbol{k}\rangle\, M(\boldsymbol{k}). \tag{16.2}$$

To understand what $M(\boldsymbol{p})$ is, use the fact that $\langle \boldsymbol{k}|\boldsymbol{p}\rangle = \delta(\boldsymbol{k}-\boldsymbol{p})$ so that

$$\langle \boldsymbol{k}|t \to +\infty\rangle = A\delta(\boldsymbol{k}-\boldsymbol{p}) - iM(\boldsymbol{k}). \tag{16.3}$$

Thus, $|A|^2$ is the probability for the particle not to be scattered at all, and $|M(\boldsymbol{k})|^2$ is the probability of scattering into momentum $\boldsymbol{k}$.

Our goal is to compute $M(\boldsymbol{k})$ using the Schrödinger equation, which states

that (for a static potential)

$$|\psi(t)\rangle = e^{-2iHt}\,|\psi(-t)\rangle\,. \tag{16.4}$$

Since our initial and final states are simple in the momentum basis, we'll write H in the momentum basis too. The kinetic energy part of H is simple:

$$K = \int d^3\boldsymbol{k}\;|\boldsymbol{k}\rangle\,\frac{k^2}{2m}\,\langle\boldsymbol{k}| \tag{16.5}$$

and the potential energy can be written in the momentum basis via its Fourier transform $V(\boldsymbol{k})$ and the fact that $\langle p|x\rangle = e^{i\boldsymbol{k}\cdot x}$.

$$\begin{aligned} V &= \int d^3\boldsymbol{x}\;|\boldsymbol{x}\rangle\,V(\boldsymbol{x})\,\langle\boldsymbol{x}| \\ &= \int d^3\boldsymbol{x}\,d^3\boldsymbol{k}\,d^3\boldsymbol{k}'\;|\boldsymbol{k}\rangle\,\langle\boldsymbol{k}|\boldsymbol{x}\rangle\,V(\boldsymbol{x})\,\langle\boldsymbol{x}|\boldsymbol{k}'\rangle\,\langle\boldsymbol{k}'| \\ &= \int d^3\boldsymbol{x}\,d^3\boldsymbol{k}\,d^3\boldsymbol{k}'\;|\boldsymbol{k}\rangle\,e^{-i(\boldsymbol{k}'-\boldsymbol{k})\cdot\boldsymbol{x}}V(\boldsymbol{x})\,\langle\boldsymbol{k}'| \\ &= \int d^3\boldsymbol{k}\,d^3\boldsymbol{k}'\;|\boldsymbol{k}\rangle\,V(\boldsymbol{k}'-\boldsymbol{k})\,\langle\boldsymbol{k}'|\,. \end{aligned} \tag{16.6}$$

We'll consider the case of weak scattering, where V is usually small[1] In this case, the particle will spend only a short time in the potential well, so we can strategically pick the time t such that $|\psi(t)\rangle \approx |\psi(\infty)\rangle$ (and $|\psi(-t)\rangle \approx |\psi(-\infty)\rangle$), but t is still small. We will therefore drop all terms of order $\mathcal{O}(t^3)$, $\mathcal{O}(t^2V)$, and $\mathcal{O}(tV^2)$ in our calculation for M:

$$\begin{aligned} \langle\boldsymbol{k}|\psi(t)\rangle &= \langle\boldsymbol{k}|\left[1 - 2iKt - 2K^2t^2 - 2iVt\right]|\psi(-t)\rangle \\ &= \left(1 - i\frac{k^2}{m} - \frac{k^4}{2m^2}\right)\langle\boldsymbol{k}|\boldsymbol{p}\rangle - 2it\int d^3k'\,V(\boldsymbol{k}'-\boldsymbol{k})\,\langle\boldsymbol{k}'|\boldsymbol{p}\rangle\,. \end{aligned} \tag{16.7}$$

Equation (16.3) connects the left hand side to the scattering amplitude M. Since $\langle\boldsymbol{k}|\boldsymbol{p}\rangle$ is the same $\delta(\boldsymbol{k}-\boldsymbol{p})$ as the coefficient of A, we identify A with the first term

[1] For example, the expectation value of the scattered particle's kinetic energy is much greater than its potential energy.

of the previous equation and iM with the second:

$$\begin{aligned} A &= 1 - it\frac{p^2}{m} - t^2\frac{p^4}{2m^2} \\ M(\boldsymbol{k}) &= 2tV(\boldsymbol{k}-\boldsymbol{p}). \end{aligned} \tag{16.8}$$

This is an interesting result: $|A|^2 = 1$ to order t^2, so that an unscattered particle is offset in phase by $\Delta\phi = -tp^2/m$. **JTD: Somehow this phase is -1** . It follows that our scattering amplitude is

$$M(\boldsymbol{k}) = 2\frac{m}{p^2}V(\boldsymbol{k}-\boldsymbol{p}). \tag{16.9}$$

For an inelastic position (i.e., when the potential contains no internal degrees of freedom which could steal energy form the particle), we expect the initial and final energies to be the same. This implies that $M(\boldsymbol{k}) \propto \delta(k^2 - p^2)$ so that the scattering amplitude lies on a sphere of radius p^2.

JTD: Alternate method

This time, instead of framing the problem as a single particle moving from late times to early times, we'll imagine a stream of particles continuously scattering from a distant source. At far distances, the wavefunction is a superpostion of the incoming plane wave and the scattered wavefunction

$$\psi(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{x}} + \phi(\boldsymbol{r}), \tag{16.10}$$

and this wavefunction satisfies the Schrödinger equation

$$H\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}). \tag{16.11}$$

Splitting the Hamiltonian into kinetic and potential energy $H = K + V$,

$$(E-K)\psi(\boldsymbol{r}) = V\psi(\boldsymbol{r}). \tag{16.12}$$

Suppose we have some function $G(\boldsymbol{x})$, which satisfies

$$(E-K)G(\boldsymbol{x}) = \delta(\boldsymbol{x}). \tag{16.13}$$

This equation is useful because it suggests that the following fact is true of $\psi(\boldsymbol{x})$

$$\psi(\boldsymbol{r}) - \psi_0(\boldsymbol{r}) = \int d^3\boldsymbol{x}'\, G(\boldsymbol{x} - \boldsymbol{x}')V(\boldsymbol{x}')\psi(\boldsymbol{r}) \tag{16.14}$$

where $\psi_0(\boldsymbol{r})$ is the wavefunction of free space, where $V = 0$. **JTD: Explain the fact that if we use an advanced Green's function, we can use $\psi_0 =$ plane wave.**

The left hand side of (16.14) is therefore the scattered wavefunction $\phi(\boldsymbol{r})$. A simple yet powerful approximation is to set $\phi(\boldsymbol{r})$ in the integrand equal to $\psi_0(\boldsymbol{r})$. This is known as the *Born approximation*, and it readily gives the scattered wavefunction.

Chapter 17

Complex Analysis

Complex Analysis is the branch of math that deals with calculus for complex functions. It is a very intuitive field, sharing many similarities with real analysis (calculus for real functions) and with vector calculus. It's techniques, particularly the Residue theorem, are very important in QFT, so we provide a summary of them.

We restrict ourselves to a subclass of complex functions called *analytical* or sometimes *holomorphic* functions. These functions are continuous and infinitely differentiable functions of a single complex variable z. Examples of *non*-analytical functions are

$$z^* z, \qquad |z|, \qquad \text{piecewise functions.}$$

We can express this requirement of analyticity in another way: if we write a complex number $z = a + bi$ as a vector $\boldsymbol{z} = (a, -b)^T$, then a function $f(z)$ being holomorphic is equivalent to the function $\boldsymbol{f}(\boldsymbol{z})$ being curl-free and divergence-free in the vector calculus sense:

$$\nabla \cdot \boldsymbol{f} = 0 \qquad \nabla \times \boldsymbol{f} = 0. \tag{17.1}$$

Despite this restriction, many functions are analytical and obey several very useful theorems concerning power series (section 17.1) and integration (section 17.2).

17.1 Series, Poles, and Branch Cuts

It is a fact that all analytical functions can be written as a Taylor series which converges in some disk of radius R centered on z_0:

$$f(z) = \sum_{n=0}^{\infty} \frac{(z-z_0)^n}{n!} \left.\frac{d^n f}{dz^n}\right|_{z=z_0} \text{ converges for all } |z-z_0| < R. \tag{17.2}$$

In some functions (called *entire* functions), that disk is the entire complex plane, but for most functions the disk can only grow so large before the series diverges. The maximum radius of the disk, called the *radius of convergence* R, is always the distance to the nearest singularity of the function.

A *singularity* is a location where the function is undefined. The most common singularity is a *pole*, where f blows up due to dividing by a small number. For example, the function $f(z) = 1/z$ has a pole at $z = 0$. $1/z^2$ also has a pole at zero, but we would call this a pole of multiplicity two due to the square in the denominator.

The relationship between series and poles is very useful, and sheds light on why some Taylor series of real functions diverge. For example, the function

$$f(x) = \frac{1}{x^2+1} \tag{17.3}$$

is defined for all real x, but its Taylor series centered on $x_0 = 0$ diverges at $x = \pm 1$. That is, it has a radius of convergence 1. We can see why by promoting x to a complex variable z. Then $f(z)$ has poles at $z = \pm i$, since $1/(i^2+1)$ diverges. The distance between the center of the Taylor series (0) and the pole (i) is one, so $R = 1$.

A series diverging is not always disastrous. Consider $f(z) = \ln z$, defined as the function such that $e^f(z) = z$. This function has a pole at $z = 0$, since $e^{-\infty} = 0$. For this reason, the Taylor series around $z_0 = 1$ has a radius of convergence of $R = 1$

$$\ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \cdots - (-1)^n \frac{z^n}{n}. \tag{17.4}$$

However, we can extend the domain of this function by evaluating the series at a new point, such as $z_1 = \ln((1+i)/\sqrt{2})$. Since z_1 is well within the radius of convergence of 17.4, we can also find all the derivatives of $\ln(z_1)$ and write out the

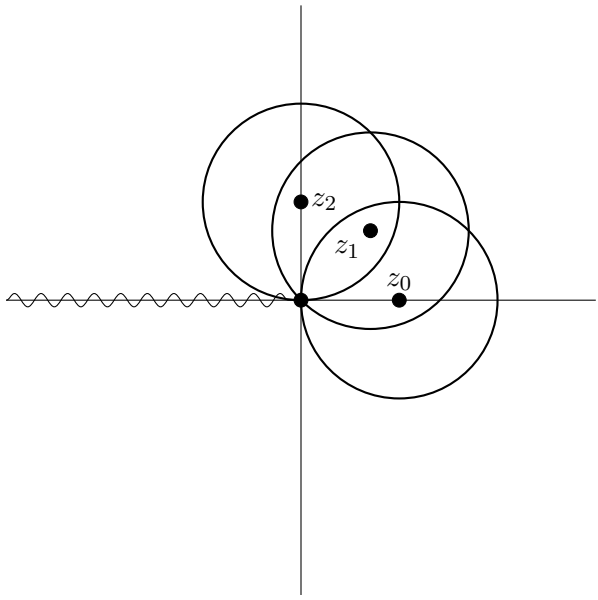


Figure 17.1: Analytic continuation of $\ln z$ via Taylor series centered on z_0, z_1, and z_2. The dot in the center is the pole and the wavy line to the left is the branch cut.

Taylor series:

$$\ln(z) = \sum_{n=0}^{\infty} a_n (z - z_1)^n \tag{17.5}$$

where a_n are new complex constants. Since the only pole of the logarithm is at $z = 0$, this function has a radius of convergence of 1, and we can repeat the same trick for a new point, such as $z_2 = i$. With this technique, $\ln z$ can be defined with a Taylor series at every point in the complex plane (except $z = 0$), despite the pole![1] This technique is called *analytic continuation*, and can be performed on any analytical function. A diagram is given in Figure 17.1.

Qualitatively speaking, we've shown that the divergence of (17.4) was a red herring except at $z = 0$. The function was perfectly well defined.

Or was it? There is still one property left to check; if we analytically continue in some path that starts at $z = 1$, encircles the origin, and comes back to $z = 1$, do we reproduce the same value of $\ln 1$? Perhaps surprisingly, the answer to this question

[1] Note that we do not have a single Taylor series which works for every point; you must chose the right Taylor series which is centered on a nearby point.

is no; originally $\ln 1 = 0$ but the analytically continued version states $\ln 1 = 2\pi$.[2] It cannot be that $\ln 1 = 0$ and 2π at the same time.

This problem illustrates the weakness of analytical continuation. It allows a smooth function to be defined anywhere except at a pole, but it may induce inconsistencies like constant offsets in the value of the function. We fix these by defining a "branch cut," which is an arbitrary path that connects to the simple pole and extends to infinity at which we say $\ln z$ is undefined. That way, the circular path used above to create an inconsistency is interrupted by the branch cut and the argument can no longer be made. This branch cut is represented with a wavy line in Figure 17.1 and chosen to lie along the negative real axis by convention. In general, any pole in a function generated by analytical continuation should have a branch cut attached to it.For example, $\sqrt{z}$ also has a branch cut, conventionally chosen to lie along the negative real axis as well.

Awareness of the poles and branch cuts in a complex function is a crucial trick when it comes to Taylor expanding or integrating a function. The connection between integration and poles is worked out in the next chapter.

17.2 Residue Theorem

This section is on the integrals of complex functions such as $f(z)$. Like in vector calculus, when we integrate a function $\int_\gamma dz\, f(z)$ we must specify a curve γ and we define the integral in the usual way using Riemann sums. The parametrization of the curve is unimportant, but the location of the curve in complex space is. In vector calculus a two-dimensional vector field $\boldsymbol{f}(\boldsymbol{z})$ obeys Green's theorem for integration around a closed contour γ:

$$\oint_\gamma \boldsymbol{d\ell} \cdot \boldsymbol{f}(\boldsymbol{z}) = \int_D d^2\boldsymbol{z}\, \nabla \times \boldsymbol{f}(\boldsymbol{z}) \tag{17.6}$$

where D is the interior of γ. If $\nabla \times \boldsymbol{f}(\boldsymbol{z}) = 0$ everywhere, then the line integral is zero. We would say that $\boldsymbol{f}(\boldsymbol{z})$ is *conservative*.

[2] In the case of the logarithm, it is simple to find the 2π offset which analytical continuation produces. It happens that $\ln(re^{i\theta})$, where r and θ are real, is $\ln r + i\theta$. Check this by exponentiating $\ln r + i\theta$ and confirming it is equal to $re^{i\theta}$. Then defining z_i along a circle of radius $r = 1$ keeps increasing the imaginary component of z_i as θ is increased until the circle is complete and the analytical continuation has accumulated a $2\pi i$ offset.

Analytical functions are also conservative in that

$$\oint_\gamma dz\, f(z) = 0 \tag{17.7}$$

when the complex function is defined everywhere inside γ. An equivalent statement is that one can move γ in the complex plane and not change the value of the integral, as long as f is defined everywhere that γ was moved through. Both these statements are due to the connection between analytical functions and curl-free vector fields mentioned at the beginning of this section. However, the function might not be defined at a few poles z_j within γ, and at these points we cannot use Green's theorem. Let us focus on these poles specifically now.

Suppose we have a function $f(z) = z^{-n}$, with a pole of multiplicity n at the origin. We can find the value of a closed integral around the origin by integrating in a nearly-closed contour in a circle around the pole and taking a limit as the ends approach each other. Note that the radius of this circle doesn't matter because if we shrink or expand the circle to another radius, we do not pass the curve through any poles, so that the value of the integral doesn't change.

Complex functions obey the fundamental theorem of calculus:

$$\int_a^b dz\, f(z) = F(b) - F(a) \qquad \text{where} \qquad \frac{dF}{dz} = f(z). \tag{17.8}$$

A function $F(z)$ is generally called an *antiderivative* of f. We changed notation here by not specifying γ and instead only specifying its endpoints. This is allowed because, as we just learned, γ may be moved through space where f is defined so its exact path doesn't matter.

An antiderivative of z^{-n} is $F(z) = z^{1-n}/(1-n)$, just like for real z, except when $n = 1$ in which case the antiderivative is $F(z) = \ln z$. For $n > 1$, $F(z)$ is continuous so that $F(b) - F(a) \to 0$ as $b \to a$. It follows that the closed integral around z^{-n} is zero for $n > 1$. However, for $n = 1$ we have to deal with the branch cut in the logarithm, over which $F(z)$ is *dis*continuous. We've already mentioned that $F(z)$ increases by $2\pi i$ when moving counterclockwise over the branch cut, so

that the value of $\int_a^b dz\, z^{-1}$ must likewise be $2\pi i$. To summarize,

$$\oint_a^b dz\, z^{-n} = \begin{cases} 0 & n \neq 1 \\ 2\pi i & n = 1 \end{cases}. \tag{17.9}$$

Returning to integrating an arbitrary function $f(z)$ with poles at z_j, by definition $f(z)$ can be expanded around a pole as follows:

$$f(z) = \frac{a_{-m}}{(z-z_j)^m} + \frac{a_{1-m}}{(z-z_j)^{m-1}} + \cdots = \sum_{n=-m}^{\infty} a_n (z-z_j)^n. \tag{17.10}$$

The integral of this function is then zero for each term except $n = -1$, where it is $2\pi i a_{-1}$. We generally write the coefficient a_{-1} of this Taylor series centered at z_j as $a_{-1} = \text{Res}\, f(z_j)$. Thus

$$\oint_\gamma dz f(z) = 2\pi i \sum_{z_j} \text{Res}\, f(z_j) \tag{17.11}$$

where z_j are the poles inside γ.

(17.11) is an incredibly powerful theorem. It can be used to integrate even very complicated functions over closed contours. It can even be used to perform integrals over non-closed contours, by extending the contour through some region of space where we know $f(z)$ is small until the contour becomes closed. For example, we can easily find that

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = \pi \tag{17.12}$$

via the Residue theorem. First we split the integrand into partial fractions to find the residue:

$$\frac{1}{2} \int_{-\infty}^{\infty} dx \left(\frac{i}{x+i} - \frac{i}{x-i} \right). \tag{17.13}$$

The function has poles at $z = \pm i$, with Res $f(i) = -i$ and Res $f(-i) = i$. To make a closed curve, we can close the contour in an enormous loop with very large $|z|$ from the $+\infty$ of the real axis to the $-\infty$ side. This can be done over the real line, in which case the curve will run counterclockwise and enclose the $+i$ pole, or it can be done under the real axis in which case it will run clockwise and enclose the $-i$ pole. Both these answers should be consistent, and they are. The two poles

have opposite residues, but integrating over a clockwise curve flips the sign of the integral so that the result for both is

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = 2\pi i \left(\frac{1}{2}(-i)\right) = \pi. \tag{17.14}$$

17.3 The Gamma Function

We conclude our function on special analysis with the Gamma function, $\Gamma(z)$. $\Gamma(z)$ is one of many special functions, much like the Riemann zeta function $\zeta(z)$, which are discussed in conjunction with complex analysis because their properties are so much easier to understand in the context of complex variables. Nevertheless, when $\Gamma(z)$ appears in physics, a real argument is usually used.

The Gamma function is a crucial function that appears in QFT due to its relation to the surface area of a sphere, discussed in the next chapter. It is also relevant to statistics and can be helpful in computing miscellaneous integrals as well. It is a generalization of the factorial:

$$\Gamma(n+1) = n! \tag{17.15}$$

for integer n. A consistent definition that works for all Re $z > 0$ is

$$\Gamma(z) = \int_0^{\infty} dt\, t^{1-z} e^{-t}. \tag{17.16}$$

One can check that this definition is consistent with the factorial by showing that

$$\Gamma(x+1) = x\Gamma(x) \tag{17.17}$$

and $\Gamma(0)$, which are also properties of factorials. The proof of this fact and other properties of the Gamma function are left to other books.

(17.16) defines a Gamma function is analytic and smoothly connects all the values of the factorial on the positive real axis. Using analytical continuation, we can extend the function to the negative real axis, where it has poles at every negative integer and at zero (Figure 17.2). At the zero pole, $\Gamma(z)$ has a convenient expansion:

$$\Gamma(\epsilon) = \frac{1}{\epsilon} - \gamma_E + \mathcal{O}(\epsilon) \tag{17.18}$$

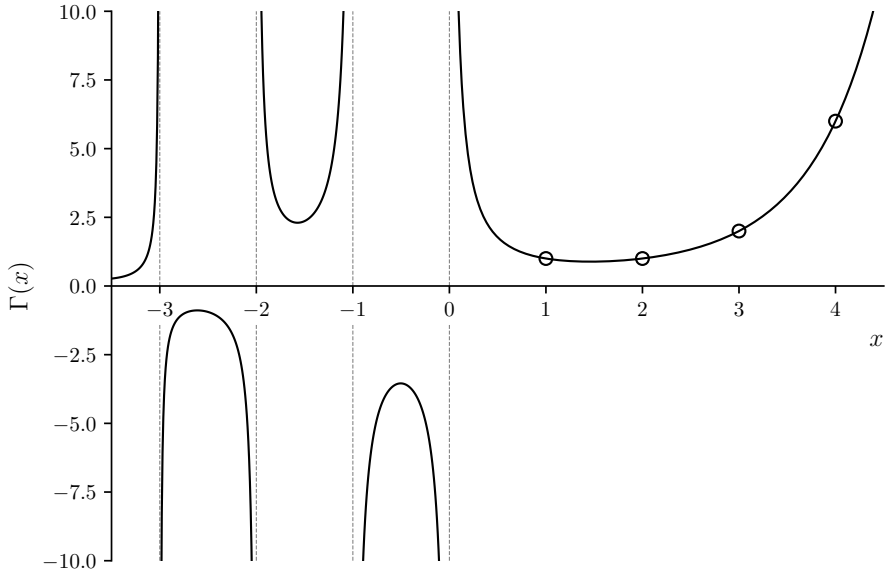


Figure 17.2: The Gamma function, $\Gamma(x)$ on the real line. Open circles represent the value of $(x+1)!$, and dotted vertical lines are the poles at every non-positive integer.

which generalizes to

$$\Gamma(-n+\epsilon) = \frac{(-1)^n}{n!}\left(\frac{1}{\epsilon} - \gamma_E + \sum_{j=1}^{n}\frac{1}{j}\right) + \mathcal{O}(\epsilon) \tag{17.19}$$

where $\gamma_E \approx 0.57721$ is the Euler-Mascheroni constant.

Another useful formula is the values of the Gamma function at half-integers:

$$\Gamma\left(z+\frac{1}{2}\right) = \frac{\sqrt{\pi}}{4^{z-\frac{1}{2}}}\frac{\Gamma(2z)}{\Gamma(z)} \tag{17.20}$$

which reduces to the special case of $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. The Gamma function can also be used to aid in evaluating the following integral

$$B(a,b) = \int_0^1 dt\, t^{a-1}(1-t)^{b-1} \tag{17.21}$$

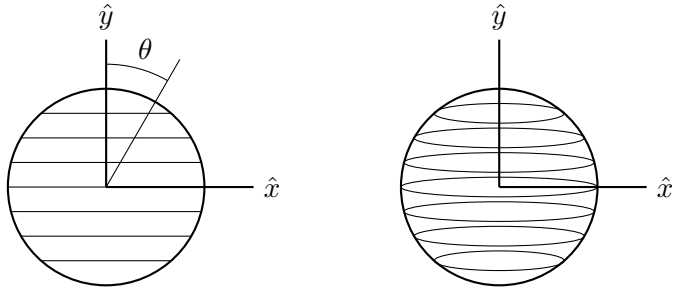


Figure 17.3: Stacking $d-1$-dimensional spheres to form a d-dimensional sphere for $d=2$ (*left*) and $d=3$ (*right*).

which is called the *Beta function*. It appears in QFT as well as in probability and various trigonometric functions. It is related to the Gamma function via

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}. \tag{17.22}$$

Below, we use the Beta and Gamma functions to compute the surface area of a d-dimensional sphere.

Surface Area of Spheres

In physics, integrals over quantities which are rotationally symmetrical are very common. Such integrals can be separated into a radial part and an angular part, even in an arbitrary d-dimensional space, by integrating over d-dimensional spheres like so

$$\int d^d x\, f(|x|) = \left(\int dr\, r^{d-1} f(r)\right) A_d \tag{17.23}$$

where A_d is the surface area of the d-dimensional unit sphere. For example, for $d=3$, we have $A_d = 4\pi$ and in $d=2$ we have $A_d = 2\pi$.

The task of this section is to find the surface area of this d-dimensional sphere as a function of d, which we will do via recursion. Suppose we know the area A_{d-1}. This knowledge allows us to ignore all the $d-1$ axes of this sphere and project them down into a single axis $\hat{x}$, with the dth axis $\hat{y}$ being perpendicular to $\hat{x}$. Now we can think two-dimensionally.

The d-dimensional unit sphere corresponds to a unit circle on the $\hat{x}$ and $\hat{y}$ axes,

which we parameterize by the angle θ between some point on the circle and the positive $\hat{y}$ axis. The d-dimensional unit sphere is created by stacking many $d-1$-dimensional spheres horizontally on top of each other like vertical lines inside the circle, meaning that the radius of one of these $d-1$-spheres is $r = \sin\theta$. It contributes an area of $dA = A_{d-1}\sin^{d-2}\theta d\theta$ to the total area. The $\sin^{d-2}\theta$ is present because surface area of the $d-1$ dimensional non-unit sphere is A_{d-1} (the unit sphere area) times the radius to the $d-2$th power. This integration scheme is shown in Figure 17.3. Thus, the d-sphere surface area is

$$\begin{aligned} A_d &= A_{d-1}\int_0^\pi d\theta\ \sin^{d-2}\theta \\ &= -A_{d-1}\int_0^\pi d(\cos\theta)\ \sin^{d-3}\theta \\ &= A_{d-1}\int_{-1}^1 dx\ \left(1-x^2\right)^{\frac{d-3}{2}} \\ &= A_{d-1}\int_0^1 dt\, t^{-\frac{1}{2}}\,(1-t)^{\frac{d-3}{2}} \end{aligned} \tag{17.24}$$

where in the second line we substituted variables from $\theta \to \cos\theta$, in the third line we expressed $\sin\theta = \sqrt{1-\cos^2\theta}$, and in the last line we substituted variables $x^2 \to t$. The reader may recognize the last integral as a Beta function $B(\frac{1}{2}, \frac{d-1}{2})$, so we may immediately write it in terms of Γ functions:

$$A_d = A_{d-1}\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{d-1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} = A_{d-1}\sqrt{\pi}\frac{\Gamma\left(\frac{d-1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}. \tag{17.25}$$

Now we must turn the above recursive equation into a formula for A_d, but fortunately this is not difficult due to the Γ function's own recursion relation (17.17). The following formula checks out:

$$A_d = \frac{2\pi^{d/2}}{\Gamma\left(\frac{d}{2}\right)}. \tag{17.26}$$

(17.26) draws a fascinating and unexpected connection between the Gamma function, which descended from factorials, and spheres. Without extending the factorials to the half-integers via the Gamma function, we could not have written (17.26) down.

Chapter 18

Review of Group Theory and Lie-Groups

Chapter A

Particles of the Standard Model

Chapter B

Encyclopedia of Quantum Field Theories

Chapter C

Important Formulas

Chapter D

Second Quantization

Chapter E

Properties of the γ^μ matrices

Chapter F

Lattice QCD

Chapter G

Spin 2 Particles and Gravitons

Chapter H

Noether's Theorem

Index