A Very Short Introduction to Quantum Field Theory

A. W. Stetz

November 21, 2007

Contents

1 Introduction **5**

2 Second Quantization and Relativistic Fields **7**
2.1 Introduction to Field Theory 7
2.2 A Quick Review of Continuum Mechanics 9
2.3 Field Operators 14
2.3.1 The Quantum Mechanics of Identical Particles 14
2.3.2 Boson States 15
2.3.3 Field Operators 18
2.3.4 Momentum Representation 19
2.3.5 Fermions 20
2.4 Introduction to Second Quantization 21
2.5 Field Theory and the Klein-Gordon Equation 29
2.6 The Propagator 31

3 The Interaction Picture and the S-Matrix **35**
3.1 The Interaction Picture 36
3.2 Interactions and the S Matrix 39
3.2.1 Two-Particle Scattering in the φ^4 Theory 40
3.3 The Wick Expansion 42
3.4 New Example – φ^3 Theory 45
3.5 Feynman Diagrams 46
3.6 The Problem of Self Interactions 49
3.7 The LSZ Reduction Scheme 52
3.8 Correlation Functions 56
3.9 Two Examples 60

4 The Trouble with Loops **63**
4.1 Doing the Integrals 65

4.2 Renormalization 71
4.3 Appendix 76

5 Cross Sections and Decay Rates **79**
5.1 Classical Scattering 79
5.2 Quantum Scattering 81
5.3 Phase Space 86
5.4 Two-Particle Scattering 89
5.5 The General Case 91

6 The Dirac Equation **93**
6.1 The Equation 93
6.2 Plane Wave Solutions 96
6.3 Charge Conjugation and Antiparticles 98
6.4 Quantizing the Field 103
6.5 The Lorentz Group 107
6.6 Spinor Representations 111
6.7 The Dirac Propagator 114

7 The Photon Field **117**
7.1 Maxwell's Equations 117
7.2 Quantization in the Coulomb Gauge 120

8 Quantum Electrodynamics **125**
8.1 Gauge Invariance 125
8.2 Noether's Theorem 127
8.3 Feynman's Rules for QED 128
8.4 The Reaction $e^- + e^+ \to \mu^- + \mu^+$ 135
8.4.1 Trace Technology 137
8.4.2 Kinematics 137
8.5 Introduction to Renormalization 138

Chapter 1

Introduction

Quantum electrodynamics, QED for short, is the theory that describes the interactions of photons with charged particles, particularly electrons. It is the most precise theory in all of science. By this I mean that it makes quantitative predictions that have been verified experimentally to remarkable accuracy. Its crowning achievement is the calculation of the corrections to the anomalous magnetic moment of the electron and muon, which agree with experiment to seven or eight significant figures! Its classical limit reduces to and indeed explains the origin of Maxwell's equations. Its non-relativistic limit reduces to and justifies the approximations inherent in the conventional quantum-mechanical treatment of electromagnetic interactions. It has also provided us with a way of thinking about the interactions of particles by representing them pictorially through Feynman diagrams. Finally, it provides the formalism necessary to treat low-energy, many-particle systems such as superfluids and superconductors.

QED is also the starting point for all theoretical treatment of elementary particles. The strong and weak interactions are modeled after the interactions of electrons and photons. This is not quite such a tidy body of theory. There are many parameters that must be taken from experiment without any understanding of their origin, and many things that simply can't be calculated because of the limitations of perturbation theory. Taking one consideration with another, however, it's still an impressive body of knowledge.

QED is so accurate and all-encompassing that it can't be all wrong, but it does leave us with a number of puzzles and paradoxes.

- Truly elementary particles have mass, spin, and other additive quantum numbers like charge, baryon number, lepton number, etc., but

they have no size; they are point-like objects. How, for example, can something with no size – spin?

- If we regard the mass of a particle as fixed, then its interactions violate the laws of conservation of momentum and energy. In order to make this consistent with relativity, we say that these things are conserved, but the particle acts as if it has some unphysical mass.

- Some interaction terms that appear inevitably when doing perturbation theory lead to divergent integrals. We have learned to "subtract infinity" to get finite answers that agree with experiment, so our understanding of these divergences can't be all wrong, but they still are an embarrassment. The finite energy portion of divergent electron-positron pair production diagrams, for example, should contribute to the mass-energy density of the universe. This effect is inconsistent with cosmological estimations by roughly one hundred orders of magnitude.

- Finally, there is no way to reconcile standard quantum field theory with general relativity. String theory, which treats particles as partially curled-up strings in a higher-dimension space, promises a solution to this problem. But string theory is like the intelligent design hypothesis in as much as it has been unable to make any prediction that can be falsified.

So on one hand, QED is a sturdy computational tool that should be part of the knowledge base of any real physicist. On the other, it is a doorway to many of the unsolved problems of modern physics.

The purpose of this course is to provide a good working knowledge of QED. Chapter 2 introduces the subject by first reviewing classical continuum mechanics. It then develops the massive Klein-Gordon field as a kind of "toy model" in which to study second quantization, perturbation theory, renormalization, and scattering theory without the complications of spin and gauge invariance. Chapter 3 develops perturbation theory from the Dyson expansion and the LSZ reduction scheme. From this we are able to derive Feynman's rules and practice the art of Feynman diagrams. Chapter 4 explores the issues of divergent integrals and renormalization. Chapter 5 shows how to calculate actual scattering cross sections and decay rates from the S-matrix elements. Chapter 6 introduces relativistic electron theory as originally proposed by Dirac. Chapter 7 deals with the electromagnetic field and the problems posed by gauge invariance. Finally, Chapter 8 does some complete calculations of electron and photon scattering cross sections.

Chapter 2

Second Quantization and Relativistic Fields

2.1 Introduction to Field Theory

Imagine that space is like a rubber sheet. If I put a bowling ball on the sheet, it will create a depression, and nearby objects will roll into it. This is an imperfect analogy for an attractive potential. We could describe the attraction in one of two ways: we could say that there is an attractive potential between any pair of point-like masses, or we could introduce a continuous variable, $\phi(x, y)$ which describes the displacement of the sheet as a function of position. Such a continuous displacement variable is a *field* in the strict mathematical sense: it assigns a numerical value (or set of values) to each point in space. The quantum mechanics of such fields is called quantum field theory. Now suppose that instead of using a bowling ball I jump up and down on the sheet. The sheet will oscillate in response. My activity becomes a *source* of energy, which propagates outward in the form of waves. This is the rubber-sheet analogy to the propagation of particles.

This analogy can easily be misleading. For one thing, I don't want you to think we are doing general relativity. The rubber sheet is not intended as an analogy for ordinary space-time as it is often used in explaining general relativity. The field $\phi(x, y)$ describes a displacement, and I know you want to ask, "Displacement of what?"

The same question comes up in classical electromagnetic theory. When an electromagnet wave is propagating through space, what is waving? Folks in the 19'th century thought it must be some sort of mechanical medium, which they called the ether. According to the textbooks, Michaelson and

Morley proved that wrong with their famous interferometer. But just saying that the ether does't exist doesn't answer the question, it just makes it impossible to answer! Let's bite the bullet and agree for the purposes of this course that space is pervaded by a medium, which for lack of a better name, we will call the ether. Well, actually the *ethers*. Each species of particle corresponds to a set of vibrations in it's own specific ether. Electrons are all vibrations in the electron ether, etc. Space-time points in the ether can be labeled with Lorentz four-vectors or $(\boldsymbol{x}, t)$ as usual, and these points obey the usual rules for Lorentz transformations. This much is required by the M-M experiment. Ordinary bulk media have elastic properties that are described by two parameters, the density and Young's modulus. These parameters are not themselves relevant to our formalism, but their ratio gives the velocity of propagation, which is what we really care about.

I am fond of saying, "When correctly viewed, everything is a harmonic oscillator." Now you see that this is profoundly true. Each point on the rubber sheet or ether acts like a harmonic oscillator! *Quantum field theory is a theory about harmonic oscillators.*

Well – I have to modify that slightly. If each point on the sheet behaved like a *simple* harmonic oscillator with a quadratic potential, the waves propagating on the sheet would never interact. The principle of linear superposition would hold everywhere. This is a theory of free particles. If our theory is to describe interactions, then we must modify the potential so that it becomes anharmonic. Unfortunately, the anharmonic oscillator cannot be solve exactly in quantum mechanics. (If you think of a way to do it, tell me and I'll be famous.) We have to resort to approximations. The generic name for these approximations is perturbation theory.

There are two quite different schemes for doing perturbation theory. One is the path integral formulation. We will not have time to cover this important and relatively new formalism, but I should tell you a little about it. Suppose a particle starts out at the spacetime point (x_0, t_0). The quantum-mechanical probability amplitude for it to cross the point (x_f, t_f) is called the propagator $K(x_f, t_f; x_0, t_0)$. According to the path integral hypothesis, this amplitude is found as follows.

1. Draw all causal path in the $x-t$ plane connecting (x_0, t_0) with (x_f, t_f). By "causal" I mean that the paths must not loop back in time. There are no other restrictions. The paths can be wildly unphysical.

2. Find the classical action $S[x(t)]$ for each path $x(t)$.[1]

[1]The notation $S[x(t)]$ indicates that S is a functional of $x(t)$. It returns a single number

3. Perform the following sum.

$$K(x_f, t_f; x_0, t_0) = C \sum_{\text{paths}} e^{iS[x(t)]/\hbar}$$

The constant C is a normalization factor. The real question is how to do the sum over paths, and *a fortiori*, what does this mean anyhow. I can't begin to explain this a paragraph except to say that it involves doing a literally infinite number of integrals! The point here is that this sum lends itself to making systematic approximations, which constitute a kind of perturbation theory. This scheme is physically attractive (though mathematically bizarre) because the action is a classical quantity without any quantum mechanical operators. It is also based on the Lagrangian (rather than the Hamiltonian), which makes it easy to discuss the invariance properties of the theory. It is paradoxically a way of doing quantum field theory *without any quantum mechanics!*

There is an alternative way of dealing with interaction involving the creation and annihilation of particles. It is the older way, sometimes called canonical quantization or second quantization. The path integral formalism, seeks to banish all operators from the theory. Second quantization goes in the other direction. It turns the wave functions themselves into operators by imbedding creation and annihilation operators into them; *but they are the raising and lowering operators of the harmonic oscillator!* The universe, according to second quantization, is an infinity of harmonic oscillators. This approach is complementary to path integrals in other ways as well. One needs to master both.

Continuum mechanics is not covered in most graduate mechanics classes. There is a good discussion in the last chapter of Goldstein, but we never make it that far. What follows is a brief introduction.

2.2 A Quick Review of Continuum Mechanics

The rules of continuum mechanics are derived by starting with a system with a finite number of degrees of freedom and then passing to the limit in which the number becomes infinite. Let's do this with the simplest possible system, a long chain of masses connected by springs. It's a one-dimensional problem. The masses can only oscillate along the chain. We will use φ_i,

for each distinct path.

the displacement of the i-th particle from its equilibrium position, as the generalized coordinate. The Lagrangian is constructed in the obvious way.

$$T = \frac{1}{2}\sum_i m\dot{\varphi}_i^2 \tag{2.1}$$

$$V = \frac{1}{2}\sum_i k(\varphi_{i+1} - \varphi_i)^2 \tag{2.2}$$

$$L = T - V = \frac{1}{2}\sum_i a\left[\frac{m}{a}\dot{\varphi}_i^2 - ka\left(\frac{\varphi_{i+1} - \varphi_i}{a}\right)^2\right] \tag{2.3}$$

The equilibrium separation between masses is a. The spring constant is k. The Euler-Lagrange equations of motion are obtained from

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\varphi}_i} - \frac{\partial L}{\partial \varphi_i} = 0 \tag{2.4}$$

If there are N masses, then there are N coupled equation of this sort. They look like

$$\frac{m}{a}\ddot{\varphi}_i - ka\left(\frac{\varphi_{i+1} - \varphi_i}{a^2}\right) + ka\left(\frac{\varphi_i - \varphi_{i-1}}{a^2}\right) = 0 \tag{2.5}$$

We need different parameters to describe the continuum limit:

$$m/a \to \mu \qquad \text{mass per unit length}$$

$$ka \to Y \qquad \text{Young's modulus}$$

The index i points to the i-th mass, and φ_i gives its displacement. In the continuum limit, the index is replaced by the coordinate x. In elementary mechanics, x would be the displacement of a particle. Here $\varphi(x)$ is the displacement of the string *at the point* x. In the continuum limit

$$\frac{\varphi_{i+1} - \varphi_i}{a} \to \frac{\varphi(x+a) - \varphi(x)}{a} \to \frac{d\varphi}{dx}$$

$$L \to \frac{1}{2}\int dx\left[\mu\dot{\varphi}^2 - Y\left(\frac{d\varphi}{dx}\right)^2\right] \equiv \int dx\mathcal{L}(\varphi, \dot{\varphi}) \tag{2.6}$$

The last integral implicitly defines the *Lagrangian density* . The continuum version of the Euler-Lagrange equation is[2]

$$\frac{d}{dt}\left[\frac{\partial \mathcal{L}}{\partial\left(\frac{d\varphi}{dt}\right)}\right] + \frac{d}{dx}\left[\frac{\partial \mathcal{L}}{\partial\left(\frac{d\varphi}{dx}\right)}\right] - \frac{\partial \mathcal{L}}{\partial \varphi} = 0 \tag{2.7}$$

[2]See Goldstein Chapter 13 for a derivation of this important equation.

Use the Lagrangian density from (2.6) in (2.7).

$$\frac{\partial^2\varphi}{\partial x^2} = \left(\frac{\mu}{Y}\right)\frac{d^2\varphi}{dt^2} \tag{2.8}$$

(2.4) and (2.5) represent a set of N coupled equations for N degrees of freedom. (2.7) is one equation for an infinite number of degrees of freedom. In this sense, continuum mechanics is much easier that discrete mechanics.

Equation (2.8) should remind you of the equation for the propagation of electromagnetic waves.

$$\left(\frac{\partial^2\varphi}{\partial x^2}\right) + \left(\frac{\partial^2\varphi}{\partial y^2}\right) + \left(\frac{\partial^2\varphi}{\partial z^2}\right) = \frac{1}{c^2}\left(\frac{\partial^2\varphi}{\partial t^2}\right) \tag{2.9}$$

As you know, photons are *massless* particles. Notice that a string of *massive* particles yields a wave equation that when quantized describes the propagation of *massless* particles. (With a different velocity, of course.) This is worth a brief digression.

What does it mean to say that a wave function describes the propagation of a particle of a particular mass? The wave function $\psi = e^{i(kx-\omega t)}$ might describe a wave in classical E&M, or a massive particle in non-relativistic or relativistic quantum mechanics. The question is, what is the relation between k and ω? The relationship between the two is called a *dispersion relation*. It contains a great deal of information. In the case of EM waves in vacuum, $k = \omega/c$. Frequency and wave number are simply proportional. This is the hallmark of a massless field. The velocity is the constant of proportionality, so there can only be one velocity. In Schrodinger theory

$$\frac{\hbar^2 k^2}{2m} = \hbar\omega \tag{2.10}$$

The relationship is quadratic. The relativistic wave equation for a spin-zero particle is called the Klein-Gordon equation.

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\varphi - \frac{m^2c^2}{\hbar^2}\varphi = 0 \tag{2.11}$$

The dispersion relation is

$$(c\hbar k)^2 + m^2c^4 = (\hbar\omega)^2, \tag{2.12}$$

or in other words, $p^2c^2 + m^2c^4 = E^2$. All these equations can be obtained from (2.7) with the appropriate Lagrangian density. They are all three-dimensional variations of our "waves on a rubber sheet" model. What does

this have to do with the particle's mass? It's useful to plot (2.10) and (2.12), i.e. plot ω versus k for small values of k. In both cases the curves are parabolas. This means that in the limit of small k, the group velocity,

$$v_{\text{group}} = \frac{d\omega}{dk} \approx \frac{\hbar k}{m} \tag{2.13}$$

In other words, the group velocity is equal to the classical velocity for a massive particle $v = p/m$. All the wave equations I know of fall in one of these two categories; either ω is proportional to k, in which case the particle is massless and its velocity $v = \omega/k$, or the relationship is quadratic, in which case

$$m = \lim_{k\to 0} \left(\hbar k \frac{dk}{d\omega} \right). \tag{2.14}$$

So long as we are talking about wave-particle duality, *this is what mass means.*

One of the advantages of using Lagrangians rather than Hamiltonians is that Lagrangians have simple transformation properties under Lorentz transformations. To see this, let's rewrite (2.7) in relativistic notation. Construct the contravariant and covariant four-vectors

$$x^\mu \equiv (x^0, x^1, x^2, x^3) = (ct, x, y, z) \tag{2.15}$$

$$x_\mu = (x_0, x_1, x_2, x_3) = (ct, -x, -y, -z) \tag{2.16}$$

and the corresponding contravariant and covariant derivatives

$$\partial^\mu \equiv \frac{\partial}{\partial x_\mu} \qquad \partial_\mu \equiv \frac{\partial}{\partial x^\mu}. \tag{2.17}$$

This puts the Euler-Lagrange equation in tidy form

$$\partial^\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial^\mu \varphi)} \right) - \frac{\partial \mathcal{L}}{\partial \varphi} = 0 \tag{2.18}$$

This is slightly amazing. Equation (2.7) was derived without reference to Lorentz transformations, and yet (2.18) has the correct form for a scalar wave equation. We get relativity for free! If we can manage to make $\mathcal{L}$ a Lorentz scalar, then (2.18)) will have the same form in all Lorentz frames. Better yet, the action

$$S = \int dt\, L = \int dt \int d^3x\, \mathcal{L} = \frac{1}{c} \int d^4x\, \mathcal{L} \tag{2.19}$$

is also a Lorentz scalar. We can do relativistic quantum mechanics using the canonical formalism of classical mechanics.

Here's an example. Rewrite (2.6) in 3-d

$$\mathcal{L} = \frac{1}{2}\left\{\mu\left(\frac{\partial\varphi}{\partial t}\right)^2 - Y\left[\left(\frac{\partial\varphi}{\partial x}\right)^2 + \left(\frac{\partial\varphi}{\partial y}\right)^2 + \left(\frac{\partial\varphi}{\partial z}\right)^2\right]\right\} \tag{2.20}$$

This would be the Lagrangian density for oscillations in a huge block of rubber. Take

$$\frac{\mu}{Y} = \frac{1}{c^2}. \tag{2.21}$$

Obviously $\mathcal{L}$ can be multiplied by any constant without changing the equations of motion. Rescale it so that it becomes

$$\mathcal{L} = \frac{1}{2}\left\{\frac{1}{c^2}\left(\frac{\partial\varphi}{\partial t}\right)^2 - \left[\left(\frac{\partial\varphi}{\partial x}\right)^2 + \left(\frac{\partial\varphi}{\partial y}\right)^2 + \left(\frac{\partial\varphi}{\partial z}\right)^2\right]\right\} \tag{2.22}$$

Substituting (2.22) into (2.18) yields the usual equation for EM waves, $\Box\varphi = 0$.

Notice how the Lagrangian for oscillations a block of rubber (2.20) turns into the Lagrangian for oscillations in the ether (2.22). We don't have to worry about the mechanical properties of the ether, because μ and Y are scaled away. Despite what you may have been told, the Michelson-Morley experiment proves the *existence* of the ether. When correctly viewed, everything is a bunch of harmonic oscillators, even the vacuum!

Using Einstein's neat notation, we can collapse (2.22) into one term

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\varphi)(\partial^\mu\varphi) \equiv \frac{1}{2}(\partial\varphi)^2 \tag{2.23}$$

The last piece of notation $(\partial\varphi)^2$, is used to save ink. The fact that we can write $\mathcal{L}$ like this is proof that it is a Lorentz scalar. This is an important point; we can deduce the symmetry properties of a theory by glancing at $\mathcal{L}$.

Now you can make up your own field theories. All you have to do is add scalar terms to (2.23). Try it. Name the theory after yourself. Here's a theory that already has a name. It's the Klein-Gordon theory.

$$\mathcal{L} = \frac{1}{2}\left[(\partial\varphi)^2 - m^2\varphi^2\right] \tag{2.24}$$

(I have set $c = 1$ and $\hbar = 1$.) Using our new notation, the equation of motion is

$$(\partial_\mu\partial^\mu + m^2)\varphi = 0 \tag{2.25}$$

If we assume that $\varphi(x)$ (x is a 4-vector in this notation.) is a one-component Lorentz scalar, then this describes a spinless particle with mass m propagating without interactions. Spin can be included by adding more components to φ. More about this later.

2.3 Field Operators

All interactions as seen by quantum field theory come about because particles are created and annihilated. Two charged particles attract or repel one another, for example, because photons are radiated by one particle and absorbed by the other. This happens on the appropriate quantum-mechanical time scale, but during this brief interval, the number of photons in the system changes. The formalism for describing this arises naturally from the quantum mechanics of nonrelativistic many-particle systems, so let's take a brief detour through this body of theory.

2.3.1 The Quantum Mechanics of Identical Particles

Let's write a many-particle wave function as follows:

$$\psi = \psi(1, 2, \ldots, N) \tag{2.26}$$

In this notation "1", is shorthand for $\mathbf{x_1}, \sigma_\mathbf{1}$, referring to the position and spin or particle number 1. Of course identical particles don't have numbers on them like billiard balls. That's the whole point, but suppose they did. Then $\psi(1, 2)$ means that the particle numbered 1 was at the point $\mathbf{x_1}$ and it's z-component of spin was σ_1. The wave function $\psi(2, 1)$ means that the number-one ball has components $\mathbf{x_2}$ and σ_2. Our job is to construct a theory in which the sequence of numbers in ψ has no observable consequences. That is what we mean by indistinguishable particles.

It is well known that wave functions describing identical bosons must be symmetric with respect to the exchange of any pair of particles. Functions describing identical fermions must be antisymmetric in the same sense. There is a vast amount of experimental evidence to corroborate this. There is also a deep result known as the *spin-statistics theorem*, which shows that it is virtually impossible to construct a covariant theory that does not have this property.

One way to make completely symmetric or antisymmetric states is simply to multiply single-particle states in all possible combinations. We'll call the basic ingredient $|i\rangle_\alpha$. By this I mean that the ball that wears the number

α is in the quantum state given by i. We assume these are orthonormal, $\langle i|j\rangle = \delta_{ij}$. We can write an N-particle state

$$|i_1, i_2, \ldots, i_N\rangle = |1\rangle_1 |2\rangle_2 \cdots |i_N\rangle_N \tag{2.27}$$

We construct totally symmetric or antisymmetric states with the help of a permutation operator. The following facts are relevant.

- Any permutation of N objects can be achieved by interchanging pairs of the objects. There are of course an infinite number of ways to reach any given permutation in this way, *but* permutations can be reached by an odd number of interchanges or an even number. There is no permutation that can be reached by both. Therefore we can speak unambiguously of a given permutation as being even or odd.
- There are $N!$ distinct permutations of N objects.

The symmetrized and antisymmetrized basis states are then written

$$S_\pm |i_1, i_2, \ldots, i_N\rangle \equiv \frac{1}{\sqrt{N!}} \sum_{j=1}^{N!} (\pm)^P P_j |i_1, i_2, \ldots, i_N\rangle \tag{2.28}$$

The sum goes over all of the $N!$ distinct permutations represented by P_j. Equation (2.27) defines the symmetric- and antisymmetric-making operators $S_\pm$. The symbol $(\pm)^P = 1$ for symmetric states and $(\pm)^P = -1$ for antisymmetric states. Of course, the upper sign refers to bosons and the lower, to fermions.

2.3.2 Boson States

We must allow for the possibility that there are several particles in one quantum state. If, for example, there are n_i particles in the i-th state, there will be $n_i!$ permutations that leave the N-particle state unchanged. In this case (2.28) will not be normalized to unity. A properly normalized state can be constructed as follows:

$$|n_1, n_2, \ldots\rangle = S_+ |i_1, i_2, \ldots, i_N\rangle \frac{1}{\sqrt{n_1! n_2! \cdots}} \tag{2.29}$$

The numbers n_1, n_2, etc. are called *occupation numbers*. The sum of all occupation numbers must equal to the total number of particles:

$$\sum_i n_i = N \tag{2.30}$$

All this assumes that there are exactly N particles. We are interested in precisely the situation in which the total number of particles is not fixed. We allow for this by taking the basis states to be the direct sum of the space with no particles, the space with one particle, the space with two particles, etc. A typical basis element is written $|n_1, n_2, \ldots\rangle$. There is no constraint on the sum of the n_i. The normalization is

$$\langle n_1, n_2, \ldots | n_1', n_2', \ldots\rangle = \delta_{n_1,n_1'}\delta_{n_2,n_2'}\cdots \tag{2.31}$$

and the completeness relation

$$\sum_{n_1,n_2,\ldots} |n_1, n_2, \ldots\rangle\langle n_1, n_2, \ldots| = \mathbf{1} \tag{2.32}$$

Since there are physical processes that change the number of particles in a system, it is necessary to define operators that have this action. The basic operators for so doing are the *creation and annihilation operators.* As you will see, they are formally equivalent to the raising and lowering operators associated with the harmonic oscillator. For example, suppose a state has n_i particles in the i'th eigenstate. We introduce the creation operator $a_i^\dagger$ by

$$a_i^\dagger|\ldots, n_i, \ldots\rangle = \sqrt{n_i+1}|\ldots, n_i+1, \ldots\rangle, \tag{2.33}$$

ie. $a_i^\dagger$ increases by one the number of particles in the i'th eigenstate. The adjoint of this operator reduces by one the number of particles. This can be seen as follows: Take the adjoint of (2.33) and multiply on the right by $|\ldots, n_i+1, \ldots\rangle$.

$$\begin{aligned}&\langle\ldots, n_i, \ldots|a_i|\ldots, n_i+1, \ldots\rangle \\ &\qquad = \sqrt{n_i+1}\langle\ldots, n_i+1, \ldots|\ldots, n_i+1, \ldots\rangle = \sqrt{n_i+1}\end{aligned}$$

Now replace n_i everywhere by $n_i - 1$.

$$\begin{aligned}&\langle\ldots, n_i-1, \ldots|a_i|\ldots, n_i, \ldots\rangle \\ &\qquad = \sqrt{n_i}\langle\ldots, n_i, \ldots|\ldots, n_i, \ldots\rangle = \sqrt{n_i}\end{aligned} \tag{2.34}$$

The effect of a_i on the state $|\ldots, n_i, \ldots\rangle$ has been to produce a state in which the number of particles in the i'th state has been reduced by one. Eqn. (2.34) also tells us what the normalization must be. In short

$$a_i|\ldots, n_i, \ldots\rangle = \sqrt{n_i}|\ldots, n_i-1, \ldots\rangle \text{ for } n_i \geq 1 \tag{2.35}$$

Of course if $n_i = 0$, the result is identically zero.

$$a_i|\ldots, n_i = 0, \ldots\rangle = 0$$

The commutation relations are important. Clearly all the a_i's commute among one another, since it makes no difference in what order the different states are depopulated, and by the same argument, the $a_i^\dagger$'s commute as well. a_i and $a_j^\dagger$ also commute if $i \neq j$, since

$$\begin{aligned} a_i a_j^\dagger|\ldots, n_i, \ldots, n_j, \ldots\rangle \\ &= \sqrt{n_1}\sqrt{n_j+1}|\ldots, n_i - 1, \ldots, n_j + 1, \ldots\rangle \\ &= a_j^\dagger a_i|\ldots, n_i, \ldots, n_j, \ldots\rangle \end{aligned}$$

Finally

$$\begin{aligned} \left(a_i a_i^\dagger - a_i\dagger a_i\right)|\ldots, n_i, \ldots, n_j, \ldots\rangle \\ &= \left(\sqrt{n_i+1}\sqrt{n_i+1} - \sqrt{n_i}\sqrt{n_i}\right)|\ldots, n_i, \ldots, n_j, \ldots\rangle \end{aligned}$$

In summary

$$[a_i, a_j] = [a_i^\dagger, a_j^\dagger] = 0, \qquad [a_i, a_j^\dagger] = \delta_{ij} \tag{2.36}$$

If it were not for the i and j subscripts, (2.36) would be the commutation relations for the harmonic oscillator, $[a, a] = [a^\dagger, a^\dagger] = 0$, $[a, a^\dagger] = 1$. (As usual I have set $\hbar = 1$.) In this context they are called *ladder operators* or *raising and lowering operators.* This is the essence of second quantization. Try to imagine a quantum system as an infinite forest of ladders, each one corresponding to one quantum state labelled by an index i. The rungs of the i'th ladder are labelled by the integer n_i. The entire state of the system is uniquely specified by these occupation numbers. The effect of a_i and $a_i^\dagger$ is to bump the system down or up one rung of the i'th ladder. There are several important results from harmonic oscillator theory that can be taken over to second quantization. One is that we can build up many-particle states using the $a_i^\dagger$'s. Starting with the *vacuum state* $|0\rangle$ with no particles, we can construct single-particle states, $a_i^\dagger|0\rangle$, two-particle states

$$\frac{1}{\sqrt{2!}}\left(a_i^\dagger\right)^2|0\rangle \qquad \text{or} \qquad a_i^\dagger a_j^\dagger|0\rangle,$$

or in general, many-particles states.

$$|n_1, n_2, \ldots\rangle = \frac{1}{\sqrt{n_1! n_2! \cdots}}\left(a_1^\dagger\right)^{n_1}\left(a_2^\dagger\right)^{n_2}\cdots|0\rangle \tag{2.37}$$

Another familiar result is that the number operator

$$N_i = a_i^\dagger a_i \tag{2.38}$$

is a Hermitian operator whose eigenvalue is the number of particles in the i'th quantum state.

$$N_i|\dots,n_i,\dots\rangle = n_i|\dots,n_i,\dots\rangle \tag{2.39}$$

Here is a useful result that you can prove by brute force or induction.

$$\left[a_i,\left(a_i^\dagger\right)^n\right] = n\left(a_i^\dagger\right)^{n-1}$$

Use this to do the following exercises.

- Show that (2.37) has the normalization required by (2.31).
- Prove (2.39).
- Show that the mysterious factor of $\sqrt{n_i+1}$ in (2.33) is in fact required by (2.37).

2.3.3 Field Operators

I have used the symbol $|i\rangle$ to indicate a one particle "quantum state." For example, if one were dealing with hydrogen, i would stand for the discrete eigenvalues of some complete set of commuting operators, in this case n, l, m, and m_s. The creation operators $a_i^\dagger$ create particles in such a state. Whatever these operators might be, however, **none** of them is the position operator. An important question is what the creation operator formalism looks like in coordinate space. First consider two basis systems based on two alternative sets of commuting observables. Use i to index one set and λ to index the other. Then

$$|\lambda\rangle = \sum_i |i\rangle\langle i|\lambda\rangle. \tag{2.40}$$

Since this is true for states, it must also be true of creation operators.

$$a_\lambda^\dagger = \sum_i \langle i|\lambda\rangle a_i^\dagger \tag{2.41}$$

So far we have tacitly assumed that indices like i and λ refer to discrete quantum numbers. Let's broaden our horizons and consider the possibility that $|\lambda\rangle$ might be an eigenstate of the position operator $|\lambda\rangle \to |\mathbf{x}\rangle$, where

$$\mathbf{X}|\mathbf{x}\rangle = \mathbf{x}|\mathbf{x}\rangle \tag{2.42}$$

Remember that what we call a wave function in elementary quantum mechanics is really a scalar product on Hilbert space of the corresponding state and eigenstates of the position operator, i.e.

$$\langle \mathbf{x}|i\rangle = \varphi_i(\mathbf{x}). \tag{2.43}$$

We assume that the φ_i are a complete set of orthonormal functions, so that

$$\int d^3x\ \varphi_i^*(\mathbf{x})\varphi_j(\mathbf{x}) = \delta_{ij} \tag{2.44}$$

and

$$\sum_i \varphi_i^*(\mathbf{x})\varphi_i(\mathbf{x}') = \delta^{(3)}(\mathbf{x}-\mathbf{x}') \tag{2.45}$$

So far so good, but what are we to make of $a_\lambda^\dagger$? This is the creation operator in coordinate space, which I will write $\psi(\mathbf{x})^\dagger$. Combining (2.41) and (2.43) gives

$$\psi^\dagger(\mathbf{x}) = \sum_i \varphi_i^*(\mathbf{x}) a_i^\dagger \tag{2.46}$$

and its adjoint

$$\psi(\mathbf{x}) = \sum_i \varphi_i(\mathbf{x}) a_i \tag{2.47}$$

$\psi^\dagger(\mathbf{x})$ creates a particle at $\mathbf{x}$, and $\psi(\mathbf{x})$ destroys it. $\psi^\dagger$ and ψ are called *field operators*. Their commutation relations are important.

$$[\psi(\mathbf{x}), \psi(\mathbf{x}')]_\pm = [\psi^\dagger(\mathbf{x}), \psi^\dagger(\mathbf{x}')]_\pm = 0 \tag{2.48}$$

$$[\psi(\mathbf{x}), \psi^\dagger(\mathbf{x}')]_\pm = \delta^{(3)}(\mathbf{x}-\mathbf{x}')$$

I have used the symbol $[\cdots,\cdots]_\pm$ to allow for both boson (commutation) and fermion (anticommutation) rules. The first line of (2.48) is more or less obvious. The second line follows from (2.45)

2.3.4 Momentum Representation

It's usually easier to formulate a theory in position space and easier to interpret it in momentum space. In this case we work exclusively in a finite volume with discrete momentum eigenvalues. The basic eigenfunctions are

$$\varphi_{\mathbf{k}}(\mathbf{x}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{x}} \tag{2.49}$$

We assume the usual periodic boundary conditions force the momentum eigenvalues to be

$$\mathbf{k} = 2\pi \left(\frac{n_x}{L_x}, \frac{n_y}{L_y}, \frac{n_z}{L_z} \right) \tag{2.50}$$

Where each of the n's can take the values $0, \pm 1, \pm 2, \cdots$ independently. With this normalization, the eigenfunctions are orthonormal,

$$\int_V d^3x \; \varphi^*_{\mathbf{k}}(\mathbf{x}) \varphi_{\mathbf{k}'}(\mathbf{x}) = \delta_{\mathbf{k},\mathbf{k}'} \tag{2.51}$$

Combining (2.47) and (2.49) we get the all-important result

$$\psi(\mathbf{x}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} a_{\mathbf{k}} \tag{2.52}$$

2.3.5 Fermions

We will no be dealing with electrons until much later in the quarter, but this is a good place to look at the difference between boson and fermion operators. The fact that fermion wave functions are antisymmetric introduces a few small complications. They are easy to explain looking at two-particle states. When I write $|i_1, i_2\rangle$, I mean that particle 1 is in state i_1, which is to say that the left-most entry in the ket refers to particle 1, the second entry on the left refers to particle 2, etc. Antisymmetry then decrees that $|i_1, i_2\rangle = -|i_2, i_1\rangle$. If both particles were in the same state $|i_1, i_1\rangle = -|i_1, i_1\rangle$, so double occupancy is impossible. If I describe this state in terms of occupation numbers $|n_1, n_2\rangle$, the left-most entry refers to the first quantum state (rather than the first particle), but which state is the first state? You have to decide on some convention for ordering states and then be consistent about it.

These states are constructed with creation and annihilation operators as in the case of bosons, but now we must be more careful about ordering. Consider the following two expressions.

$$a^\dagger_{i_1} a^\dagger_{i_2} |0\rangle = |i_1, i_2\rangle$$

$$a^\dagger_{i_2} a^\dagger_{i_1} |0\rangle = |i_2, i_1\rangle = -|i_1, i_2\rangle$$

I have decreed, and again this is a convention, that the first operator on the left creates particle 1, etc. Obviously

$$a^\dagger_{i_1} a^\dagger_{i_2} + a^\dagger_{i_2} a^\dagger_{i_1} = 0 \tag{2.53}$$

We say that fermion operators *anticommute*. The usual notation is

$$[A, B]_+ = \{A, B\} = AB + BA \tag{2.54}$$

Of course, fermions don't have numbers painted on them any more than bosons do, so we must use occupation numbers. Here the convention is

$$|n_1, n_2, \cdots\rangle = \left(a_i^\dagger\right)^{n_1} \left(a_2^\dagger\right)^{n_2} \cdots |0\rangle \qquad n_i = 0, 1 \tag{2.55}$$

The effect of the operator $a_i^\dagger$ must be

$$a_i^\dagger |\ldots, n_i, \ldots\rangle = \eta |\ldots, n_i + 1, \ldots\rangle, \tag{2.56}$$

where $\eta = 0$ of $n_i = 1$ and otherwise $\eta = +/-$, depending on the number of anticommutations necessary to bring the $a_i^\dagger$ to the position i. The commutation relations of the a_i's and $a_i^\dagger$'s are obtained by arguments similar to those leading to (2.36). The results are

$$\{a_i, a_j\} = \{a_i^\dagger, a_j^\dagger\} = 0 \qquad \{a_i, a_j^\dagger\} = \delta_{ij} \tag{2.57}$$

2.4 Introduction to Second Quantization

In Section 2.2 I derived the Klein-Gordon equation (2.25) by considering deformations of a continuous medium. This is classical field theory, ie. the fields are ordinary functions, c-numbers as they are called, and not operators. In the Section 2.3 I introduced the field operator

$$\hat{\psi}(\mathbf{x}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} \hat{a}_{\mathbf{k}} \tag{2.58}$$

I said that $\hat{a}_{\mathbf{k}}^\dagger$ and $\hat{a}_{\mathbf{k}}$ were creation and annihilation operators and that all this was necessary to treat systems in which the number of particles was not constant.[3] In this section I would like to examine the motivations behind (2.58) more carefully and also investigate the many subtleties that arise when we apply these ideas to relativistic wave equations. We will eventually derive a completely relativistic generalization of (2.58), which will be our starting point for doing relativistic field theory.

We have encountered so far three similar wave equations, the Schrodinger equation (2.59), the Klein-Gordon equation (2.60), and the equation for

[3] In this section I will be careful to use hats to identify operators since the distinction between "operator" and "non-operator" will be very important.

massless scalar particles, which is just the K-G equation with $m = 0$ (2.61). I will write the free-particle versions of them in a peculiar way to emphasize the complementary roles of time and energy.

$$i\frac{\partial}{\partial t}\psi(\mathbf{x},t) = \frac{\hat{\mathbf{k}}^2}{2m}\psi(\mathbf{x},t) \tag{2.59}$$

$$\left(i\frac{\partial}{\partial t}\right)^2\psi(\mathbf{x},t) = (\hat{\mathbf{k}}^2 + m^2)\psi(\mathbf{x},t) \tag{2.60}$$

$$\left(i\frac{\partial}{\partial t}\right)^2\psi(\mathbf{x},t) = \hat{\mathbf{k}}^2\psi(\mathbf{x},t) \tag{2.61}$$

I have used $\hat{\mathbf{k}} = -i\nabla_{\mathbf{x}}$ and $\hbar = c = 1$. The operator on the right side of (2.59) is the kinetic energy. Einstein's equation $E^2 = \mathbf{k}^2 + m^2$ suggests that the operators on the right side of (2.60) and (2.61) are the total energy squared. Suppose that the ψ's are eigenfunctions of these operators with eigenvalue $\omega(\mathbf{k})$. (Each of these three equations will define a different functional relation between $\mathbf{k}$ and ω, of course.) The equations become

$$i\frac{\partial}{\partial t}\psi_\omega(\mathbf{x},t) = \omega(\mathbf{k})\psi_\omega(\mathbf{x},t) \tag{2.62}$$

$$\left(i\frac{\partial}{\partial t}\right)^2\psi_\omega(\mathbf{x},t) = \omega^2(\mathbf{k})\psi_\omega(\mathbf{x},t) \tag{2.63}$$

$$\left(i\frac{\partial}{\partial t}\right)^2\psi_\omega(\mathbf{x},t) = \omega^2(\mathbf{k})\psi_\omega(\mathbf{x},t) \tag{2.64}$$

Although we don't usually use this language, we could think of $i\partial/\partial t$ as a kind of energy operator whose eigengvalues are the total energy of the particle. Suppose now that the ψ_ω's are also momentum eigenstates so that $\hat{\mathbf{k}}\psi_\omega = \mathbf{k}\psi_\omega$. The simplest solutions of (2.59) and (2.62) with $\omega = \mathbf{k}^2/2m$ are

$$\psi_{\mathbf{k}}(\mathbf{x},t) = \frac{1}{\sqrt{V}}e^{i(\pm\mathbf{k}\cdot\mathbf{x}-\omega t)} \tag{2.65}$$

whereas the simplest solutions of (2.60) and (2.63) with $\omega^2 = \mathbf{k}^2 + m^2$ or (2.61) and (2.64) with $\omega^2 = \mathbf{k}^2$ are

$$\psi_{\mathbf{k}}(\mathbf{x},t) = \frac{1}{\sqrt{V}}e^{i(\pm\mathbf{k}\cdot\mathbf{x}\mp\omega t)} \tag{2.66}$$

(The $1/\sqrt{V}$ is a normalization factor put in for later convenience.) Evidentally the solutions to (2.63) and (2.64) comprise a larger family than those of (2.62), and it is this larger family that I want to investigate.

To avoid ambiguity, I will assume that the *symbol* ω refers to a positive quantity. Then

$$i\frac{\partial}{\partial t}e^{\mp i\omega t} = \pm\omega e^{\mp i\omega t} \tag{2.67}$$

Since $\hbar = 1$, ω has the units of energy. Schrodinger's equation does not have negative energy solutions. This is the clue that the upper sign in (2.66) and (2.67) gives positive-energy solutions and the lower sign gives negative-energy solutions *whatever that may mean!* What about the other sign ambiguity? Think about the solution $e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}$. Pick out some point on the wave where the phase of the exponential is ϕ. As time goes by, this point will move so that $\mathbf{k}\cdot\mathbf{x} - \omega t = \phi$, or

$$\mathbf{k}\cdot\mathbf{x} = \omega t + \phi.$$

This point is moving in the general direction of $\mathbf{k}$. We call this a *positive-frequency* solution. If ω and $\mathbf{k}\cdot\mathbf{x}$ have opposite signs, the solution has positive frequency (in the sense above). If the signs are the same, one gets the negative-frequency solution.

Now take an arbitrary time-independent wave function and expand it in a Fourier series. Assume periodic boundary conditions so that $\mathbf{k}$ is discretized as in (2.50).

$$\psi(\mathbf{x}) = \frac{1}{\sqrt{V}}\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}a_{\mathbf{k}} \tag{2.68}$$

At this point $a_{\mathbf{k}}$ is a Fourier coefficient and nothing more. We can make ψ time dependent by building the time dependence into the $a_{\mathbf{k}}$'s, $a_{\mathbf{k}} \to a_{\mathbf{k}}(t)$. In order that (2.63) and (2.64) be satisfied, the $a_{\mathbf{k}}$'s should satisfy

$$\ddot{a}_{\mathbf{k}} + \omega_{\mathbf{k}}^2 a_{\mathbf{k}} = 0 \tag{2.69}$$

This is the differential equation for the harmonic oscillator, except for two peculiar features. First, the $a_{\mathbf{k}}$'s are complex functions, and second, the frequency (and hence the "spring constant") is a function of $\boldsymbol{k}$. In some sense, each term in (2.68) has a harmonic oscillator associated with it. We can tie into the usual harmonic oscillator formalism and avoid the complex coordinates at the same time by defining the real generalized coordinate,

$$q_{\mathbf{k}}(t) = \frac{1}{\sqrt{2\omega_{\mathbf{k}}}}\left[a_{\mathbf{k}}(t) + a_{\mathbf{k}}^*(t)\right]. \tag{2.70}$$

The conjugate momentum is given by $\mathbf{p}(t) = \dot{\mathbf{q}}(t)$, but before we take the derivative, we must decide on whether we are dealing with the positive- or

negative-energy solution. In order that each term in (2.68) has the form (2.66), the time derivitave of $a_{\mathbf{k}}(t)$ must be $\dot{a}_{\mathbf{k}}(t) = \mp i\omega a_{\mathbf{k}}$. For the time being take positive energy (upper sign)

$$p_{\mathbf{k}}(t) = -i\sqrt{\frac{\omega_{\mathbf{k}}}{2}}\left[a_{\mathbf{k}}(t) - a^*_{\mathbf{k}}(t)\right] \tag{2.71}$$

These are real variables oscillating with frequency ω. We know that the Hamiltonian for simple harmonic motion is

$$H_{\mathbf{k}} = \frac{1}{2}\left[p^2_{\mathbf{k}} + \omega^2_{\mathbf{k}} q^2_{\mathbf{k}}\right]. \tag{2.72}$$

You can verify with the definitions (2.70) and (2.71) that $H_{\mathbf{k}}$ is time-independent, that $p_{\mathbf{k}}$ is canonically conjugate to $q_{\mathbf{k}}$, and that Hamilton's equations of motion are satisfied. We can turn (2.72) into a quantum-mechanical Hamiltonian by simply making $a_{\mathbf{k}}$, $a^*_{\mathbf{k}}$, $q_{\mathbf{k}}$ and $p_{\mathbf{k}}$ into operators. We know that $\hat{a}_{\mathbf{k}}$ must be an annihilation operator with the commutation relations (2.36). The operators in (2.36), however, are time-independent, Schrodinger-picture operators as is the field operator (2.58). We will want to work in the Heisenberg representation, so we must be careful about the time dependence. The natural assumption is

$$a_{\mathbf{k}}(t) \to \hat{a}_{\mathbf{k}} e^{-i\omega t} \qquad a^*_{\mathbf{k}} \to \hat{a}^\dagger_{\mathbf{k}} e^{+i\omega t} \tag{2.73}$$

In (2.73) $\hat{a}_{\mathbf{k}}$ and $\hat{a}^\dagger_{\mathbf{k}}$ are time-independent, Schrodinger-picture operators. I'll argue presently that these are consistent and reasonable assumptions. The commutation relations are then,

$$[\hat{a}_{\mathbf{k}}, \hat{a}^\dagger_{\mathbf{k}'}] = \delta_{\mathbf{k},\mathbf{k}'} \qquad [\hat{a}^\dagger_{\mathbf{k}}, \hat{a}^\dagger_{\mathbf{k}'}] = [\hat{a}_{\mathbf{k}}, \hat{a}_{\mathbf{k}'}] = 0 \tag{2.74}$$

Since $\hat{p}_{\mathbf{k}}$ and $\hat{q}_{\mathbf{k}}$ don't have this simple time dependence, the commutation relations must be taken at equal times.

$$[\hat{q}_{\mathbf{k}}(t), \hat{p}_{\mathbf{k}'}(t)] = i\delta_{\mathbf{k},\mathbf{k}'} \qquad [\hat{q}_{\mathbf{k}}(t), \hat{q}_{\mathbf{k}'}(t)] = [\hat{p}_{\mathbf{k}}(t), \hat{p}_{\mathbf{k}'}(t)] = 0 \tag{2.75}$$

With this substitution (2.72) becomes

$$\hat{H}_{\mathbf{k}} = \frac{1}{2}\omega_{\mathbf{k}}\left[\hat{a}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}} + \hat{a}_{\mathbf{k}}\hat{a}^\dagger_{\mathbf{k}}\right] = \omega_{\mathbf{k}}\left[\hat{a}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}} + \frac{1}{2}\right] \tag{2.76}$$

The same replacement turns (2.68) into (2.58). Almost by definition, the Hamiltonian must have the same form in the Schrodinger and Heisenberg pictures. The Hamiltonian in (2.76) clearly has that property.

The last factor of 1/2 in (2.76) presents something of a dilemma. This $\hat{H}_{\mathbf{k}}$ is just the Hamiltonian for a single $\mathbf{k}$ value. The complete Hamiltonian is a sum over all values.

$$\hat{H} = \sum_{\mathbf{k}} \hat{H}_{\mathbf{k}} \tag{2.77}$$

An infinite number of 1/2's is still infinity. It is customary to discard the constant with some weak argument to the effect that in defining energy, additive constants are meaningless. Since this problem will appear again and again in different contexts, it is useful to have some formal procedure for sweeping it under the rug. To this end we introduce the concept of "normal ordering." We will say that an operator has been normal ordered if *all creation operators are placed to the left of all annihilation operators.* The usual symbol to indicate that an operator has been normal ordered is to place it between colons, so for example,

$$: \hat{H}_{\mathbf{k}} := \omega_{\mathbf{k}} \hat{a}^{\dagger}_{\mathbf{k}} \hat{a}_{\mathbf{k}} \tag{2.78}$$

To put it another way, (2.78) was obtained from (2.76) by commuting the $\hat{a}_{\mathbf{k}}$ past the $\hat{a}^{\dagger}_{\mathbf{k}}$ in the second term *and discarding the commutator.* Whenever we use a Hamiltonian in a practical calculation, we will assume that it has been normal ordered.

We can check that this Hamiltonian is consistent with the time dependence assumed in (2.73) First note that $[\hat{a}_{\mathbf{k}}, \hat{H}] = \omega_{\mathbf{k}} \hat{a}_{\mathbf{k}}$, so

$$\hat{H}\hat{a}_{\mathbf{k}} = \hat{a}_{\mathbf{k}}(\hat{H} - \omega_{\mathbf{k}}) \tag{2.79}$$

hence

$$\hat{H}^n \hat{a}_{\mathbf{k}} = \hat{a}_{\mathbf{k}}(\hat{H} - \omega_{\mathbf{k}})^n \tag{2.80}$$

as a consequence

$$\begin{aligned} \hat{a}_{\mathbf{k}}(t) &= e^{i\hat{H}t} \hat{a}_{\mathbf{k}} e^{-i\hat{H}t} = \hat{a}_{\mathbf{k}} e^{-i\omega_{\mathbf{k}} t} \\ \hat{a}^{\dagger}_{\mathbf{k}}(t) &= e^{i\hat{H}t} \hat{a}^{\dagger}_{\mathbf{k}} e^{-i\hat{H}t} = \hat{a}^{\dagger}_{\mathbf{k}} e^{i\omega_{\mathbf{k}} t} \end{aligned} \tag{2.81}$$

The deeper question is why this quantization procedure makes any sense at all. The justification lies in the canonical quantization procedure from elementary quantum mechanics. It uses the Poisson bracket formalism of classical mechanics and then replaces the Poisson brackets with commutator brackets to get the corresponding quantum expression. A Poisson bracket is defined as

$$\{F, G\} \equiv \sum_{k=1}^{N} \left(\frac{\partial F}{\partial q_k} \frac{\partial G}{\partial p_k} - \frac{\partial F}{\partial p_k} \frac{\partial G}{\partial q_k} \right) \tag{2.82}$$

where q_k and p_k are any pair of conjugate variables, and F and G are any two arbitrary functions of q_k and p_k. The sum runs over the complete set of generalized coordinates. Obviously

$$\begin{aligned} \{q_n, p_m\} &= \delta_{mn} \\ \{q_n, q_m\} = \{p_n, p_m\} &= 0 \end{aligned} \tag{2.83}$$

This looks like the uncertainty relation in Quantum Mechanics, $[x, p] = i\hbar$. We get the quantum expression from the classical expression by the replacement

$$\{F, G\} \to [\hat{F}, \hat{G}]/i\hbar, \tag{2.84}$$

where $\hat{F}$ and $\hat{G}$ are the quantum mechanical operators corresponding to the classical quantities F and G, and $[\hat{F}, \hat{G}] = \hat{F}\hat{G} - \hat{G}\hat{F}$. In the case where $\hat{F}$ and $\hat{G}$ depend on time, the commutator is taken at equal times. This seems like a leap of faith, but it is valid for all the familiar operators in quantum mechanics.[4] Now inverting (2.70) and (2.71) gives

$$a_{\mathbf{k}} = \frac{ip_{\mathbf{k}} + \omega_{\mathbf{k}} q_{\mathbf{k}}}{\sqrt{2\omega_{\mathbf{k}}}} \qquad a^*_{\mathbf{k}} = \frac{-ip_{\mathbf{k}} + \omega_{\mathbf{k}} q_{\mathbf{k}}}{\sqrt{2\omega_{\mathbf{k}}}} \tag{2.85}$$

Substituting (2.85) into (2.82) gives $\{a_{\mathbf{k}}, a^\dagger_{\mathbf{k}'}\} = -i\delta_{\mathbf{k},\mathbf{k}'}$ and $\{a_{\mathbf{k}}, a_{\mathbf{k}'}\} = \{a^\dagger_{\mathbf{k}}, a^\dagger_{\mathbf{k}'}\} = 0$, so that

$$\begin{aligned} [\hat{a}_{\mathbf{k}}, \hat{a}^\dagger_{\mathbf{k}}] &= \delta_{\mathbf{k},\mathbf{k}'} \\ [\hat{a}_{\mathbf{k}}, \hat{a}_{\mathbf{k}}] = [\hat{a}^\dagger_{\mathbf{k}}, \hat{a}^\dagger_{\mathbf{k}}] &= 0 \end{aligned} \tag{2.86}$$

(with $\hbar = 1$).

The value of the Poisson bracket $\{F, G\}$ is independent of the choice of canonical variables. That is a fundamental theorem. Since (2.70) and (2.71) are together a canonical transformation, (2.75) and (2.86) are identical. *Any* choice of variables will do so long as they are related to $q_{\mathbf{k}}$ and $p_{\mathbf{k}}$ by a canonical transformation. We simply chose $q_{\mathbf{k}}$ so that it was real and had convenient units. The rest followed automatically. The fact that the resultant Hamiltonian is that of harmonic oscillators is simply a consequence of the fact that we choose to expand $\psi(\mathbf{x}, t)$ in a Fourier series.

I can now write Equation (2.58) as

$$\hat{\psi}^{(+)}(\mathbf{x}, t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} e^{i(\mathbf{k}\cdot\mathbf{x} - \omega(\mathbf{k})t)} \hat{a}_{\mathbf{k}} \tag{2.87}$$

[4]Much of the formal structure of quantum mechanics appears as a close copy of the Poisson bracket formulation of classical mechanics. See Goldstein, Poole and Safko, *Classical Mechanics* Third Ed., Sec. 9.7

The superscript (+) means the positive energy solution. The functional form of $\omega(\mathbf{k})$ is determined by the wave equation it represents, but I want to concentrate on solutions of the Klein-Gordon equation. Suppose we had chosen the negative energy solution.

$$\hat{\psi}^{(-)}(\mathbf{x},t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} e^{i(\mathbf{k}\cdot\mathbf{x}+\omega(\mathbf{k})t)} \hat{c}_{\mathbf{k}} \tag{2.88}$$

(This automatically becomes a negative frequency solution as well.) The operators $\hat{c}_{\mathbf{k}}$ and $\hat{c}^{\dagger}_{\mathbf{k}}$ annihilate and create these new negative energy particles. Everything goes through as before except that $\hat{p}_{\mathbf{k}}(t) = \dot{\hat{q}}_{\mathbf{k}}(t) = +i\omega\hat{q}_{\mathbf{k}}(t)$ changes sign, so that (2.71) becomes

$$\hat{p}_{\mathbf{k}} = i\sqrt{\frac{\omega(\mathbf{k})}{2}} \left[\hat{c}_{\mathbf{k}} - \hat{c}^{\dagger}_{\mathbf{k}}\right]. \tag{2.89}$$

The counterpart of (2.74) is

$$[\hat{c}^{\dagger}_{\mathbf{k}}, \hat{c}_{\mathbf{k}'}] = \delta_{\mathbf{k},\mathbf{k}'} \tag{2.90}$$

It seems that the new creation operator $c^{\dagger}_{\mathbf{k}}$ stands in the place of the old annihilation operator $\hat{a}_{\mathbf{k}}$. This is not just a mathematical accident. It points to an important piece of physics. To see this we define another pair of creation and annihilation operators.

$$\hat{d}_{\mathbf{k}} = \hat{c}^{\dagger}_{-\mathbf{k}} \qquad \hat{d}^{\dagger}_{\mathbf{k}} = \hat{c}_{-\mathbf{k}} \tag{2.91}$$

Substituting this in (2.88) and changing the sign of the summation variable from $\mathbf{k}$ to $-\mathbf{k}$ gives

$$\hat{\psi}^{(-)}(\mathbf{x},t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega(\mathbf{k})t)} \hat{d}^{\dagger}_{\mathbf{k}} \tag{2.92}$$

What is it that the $\hat{d}$'s are creating and destroying, and what is the significance of the change in the sign of the momentum? If these were low energy electrons we could invoke the notion of the Fermi sea, a set of low-lying energy levels all occupied by fermions. Removing one particle from the sea leaves a "hole," which behaves in some ways like a real particle. If the hole moves in the positive $\mathbf{k}$ direction, a real particle must move in the $-\mathbf{k}$ direction to backfill the hole. Dirac proposed this mechanism to explain the negative energy states that appear in relativistic electron theory. The correspondence between holes moving forward and real particles moving

backward is a good way of visualizing the significance of (2.91). Unfortunately, the Klein-Gordon equation describes bosons, so there is no Fermi sea.[5] Nowadays, we regard these negative-energy solutions as representing real positive-energy antiparticles. There are two lines of evidence for this. For one thing, the states created by $\hat{d}_{\mathbf{k}}$ have momentum $\mathbf{k}$ (rather than $-\mathbf{k}$). This can be proved formally, but it is almost obvious from (2.92), which is a positive-frequency solution. Later on when we discuss the interaction of the electromagnetic field with bosons, we will show that $\hat{d}^{\dagger}_{\mathbf{k}}$ creates a particle of the *opposite* charge to that created by $\hat{a}^{\dagger}_{\mathbf{k}}$. The complete operator-valued wave function is the sum of (2.87) and (2.92).

$$\hat{\psi}(\mathbf{x},t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} \left[e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \hat{a}_{\mathbf{k}} + e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \hat{d}^{\dagger}_{\mathbf{k}} \right] \tag{2.93}$$

There are several profound reasons why the positive- and negative-energy solutions must be added in just this way. These will emerge as we go along.

Let's note in passing that there are several neutral spin-zero particles such as the π^0 that have no non-zero additive quantum numbers. Such particles are thereby identical to their antiparticles. If $\hat{a}^{\dagger}_{\mathbf{k}}$ creates a π^0, then $\hat{d}_{\mathbf{k}}$ destroys the *same* particle. In this case there is no point in distinguishing between $\hat{a}_{\mathbf{k}}$ and $\hat{d}_{\mathbf{k}}$. We write (2.93)

$$\hat{\psi}(\mathbf{x},t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} \left[e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \hat{a}_{\mathbf{k}} + e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \hat{a}^{\dagger}_{\mathbf{k}} \right] \tag{2.94}$$

Fields corresponding to neutral particles are Hermitian. Those corresponding to charged particles are not.

In some ways (2.93) and (2.94) are relics of our nonrelativistic fields from Chapter 5. Because they are based on discrete $\mathbf{k}$ values and periodic boundary conditions they behave under Lorentz transformations in a most awkward fashion. We are accustomed to passing to the continuum limit through the replacement

$$\frac{1}{V} \sum_{\mathbf{k}} \rightarrow \int \frac{d^3k}{(2\pi)^3},$$

but this may be a bad idea for relativistic fields. The trouble is that the integration measure d^3k does not transform like a scalar under Lorentz transformations. A better choice might be

$$\frac{1}{V} \sum_{\mathbf{k}} \rightarrow \int \frac{d^4k}{(2\pi)^3}\, \delta(k^2 - m^2), \tag{2.95}$$

[5]The idea is only relevant to low-temperature conductors anyway.

which is clearly invariant. (The symbol k here refers to the usual four-vector $k^\mu \to (k^0, \mathbf{k})$.) The dk^0 integration is done as follows.

$$\int dk^0\, \delta(k^2 - m^2) = \int d(k^0)^2 \left(\frac{dk^0}{d(k^0)^2}\right) \delta((k^0)^2 - \omega_k^2)$$

$$= \int \frac{d(k^0)^2}{2k^0} \delta((k^0)^2 - \omega_k^2) = \frac{1}{2\omega_k}$$

Equation (2.95) becomes

$$\frac{1}{V}\sum_{\mathbf{k}} \to \int \frac{d^3k}{(2\pi)^3 2\omega_k} \tag{2.96}$$

Although this convention is sometimes used, it is somewhat simpler to use the following

$$\boxed{\hat{\varphi}(\boldsymbol{x}, t) = \int \frac{d^3k}{\sqrt{(2\pi)^3 2\omega_k}} \left[\hat{a}(\boldsymbol{k}) e^{-ikx} + \hat{d}^\dagger(\boldsymbol{k}) e^{ikx}\right]} \tag{2.97}$$

where $kx = \omega t - \mathbf{k}\cdot\mathbf{x}$. The point is that we must also consider the transformation properties of the creation and annihilation operators. The natural generalization of (2.74) is

$$\boxed{[\hat{a}(\mathbf{k}), \hat{a}^\dagger(\mathbf{k}')] = \delta^{(3)}(\mathbf{k} - \mathbf{k}') \qquad [\hat{a}(\mathbf{k}), \hat{a}(\mathbf{k}')] = [\hat{a}^\dagger(\mathbf{k}), \hat{a}^\dagger(\mathbf{k}')] = 0} \tag{2.98}$$

and and similarly for the $\hat{d}$'s. Although $\delta^{(3)}(\mathbf{k} - \mathbf{k}')$ by itself is not a Lorentz scalar, the field definition (2.97) together with (2.98) does have the right transformation properties. This will be apparent once we have calculated the propagator. Notice finally that I have switched the notation for the field operator from ψ to φ. The reason is that at this point we are really committed to the second-quantized version of the Klein-Gordon equation, and φ (or sometimes ϕ) is the universally-used symbol for this field.

2.5 Field Theory and the Klein-Gordon Equation

The classical Klein-Gordon equation and its associated Lagrangian were discussed briefly in Section 2.2. The Lagrangian density (2.24) is

$$\mathcal{L} = \frac{1}{2}\left[(\partial\varphi)^2 - m^2\varphi^2\right] \tag{2.99}$$

Suppose we regard φ as a generalized "coordinate" perhaps referring to the displacement of some hypothetical continuous medium. The conjugate field is

$$\pi(x) = \frac{\delta \mathcal{L}}{\delta \dot{\varphi}} = \partial_0 \varphi(x) \tag{2.100}$$

We can use the Poisson bracket approach to quantize these fields just as we quantized the $a_{\mathbf{k}}$'s in the previous section. Classically,[6]

$$\{\varphi(\mathbf{x}, t), \pi(\mathbf{x}', t)\} = \delta^{(3)}(\mathbf{x} - \mathbf{x}') \tag{2.101}$$

Our quantized version of this is

$$\boxed{[\hat{\varphi}(\mathbf{x}, t), \hat{\pi}(\mathbf{x}', t)] = i\delta^{(3)}(\mathbf{x} - \mathbf{x}')} \tag{2.102}$$

It is easy to verify that (2.97) satisfies (2.102) so long as the creation operators are normalized according to (2.98). It's crucial that the two fields in (2.102) are evaluated at the same time. As such they are called *equal time commutation relations.* Many texts in fact consider (2.102) to be the fundamental postulate of field theory and use it to derive the properties of the creation and annihilation operators.

Before we leave the subject of commutation relations, there is an issue that has been hanging fire since (2.93). In that equation the positive- and negative-energy solutions appear with equal magnitude. Nothing I have said so far requires the presence of the negative-energy term, let alone that it have the same normalization as the positive-energy part. In fact, we have no choice in the matter. The argument goes like this. Consider two space-like separated points $(\mathbf{x}, t)$ and $(\mathbf{x}', t')$. There will always be a Lorentz frame such that $t = t'$. If $\mathbf{x}$ and $\mathbf{x}'$ are distinct points, then a signal can propagate from one to the other only by travelling at infinite velocity. We believe this to be impossible, so our theory must not allow it, even in principle. We call this the requirement of *causality.* It can be argued that a necessary requirement is that the fields $\hat{\varphi}(x)$ and $\hat{\varphi}^\dagger(x')$ also commute at equal times.[7] Let's rewrite (2.97) with an additional parameter α that we can tweak at our pleasure.

$$\hat{\varphi}(\boldsymbol{x}, t) = \int \frac{d^3k}{\sqrt{(2\pi)^3 2\omega_k}} \left[\hat{a}(\boldsymbol{k})e^{-ikx} + \alpha \hat{a}^\dagger(\boldsymbol{k})e^{ikx}\right] \tag{2.103}$$

[6] See J.V. Jose and E. J. Saletan, *Classical dynamics: a contemporary approach*, Sec 9.3.1 for a derivation of this rather non-trivial result.

[7] See Paul Teller, *An Interpretive Introduction to Quantum Field Theory*, Chapter 4, for a careful discussion of this point.

A simple calculation now gives

$$[\hat{\varphi}(\mathbf{x},t),\hat{\varphi}^{\dagger}(\mathbf{x}',t)]=\int\frac{d^3k}{2(2\pi)^3\omega_k}(1-|\alpha|^2)e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} \tag{2.104}$$

This is not zero (because of the ω in the denominator) unless $|\alpha|=1$. Relativity and causality together require an equal admixture of negative- and positive-energy states. This argument takes on additional significance when spin is taken into account. It can be shown that the requirement proved above only holds for *integer* spin. In the case of half-odd integer spin, the result only holds if the corresponding creation and annihilation operators *anticommute*.

2.6 The Propagator

At this point I must anticipate some developments from the next few chapters. It will turn out that one of the key ingredients of any perturbation calculation is the Feynman propagator defined by

$$G(x,y)=iD(x-y)=\langle 0|T[\hat{\varphi}(x),\hat{\varphi(y)}]|0\rangle \tag{2.105}$$

Where $T[\ ,\ ]$ is the "time-ordered product" defined by

$$T[\hat{\varphi}(x)\hat{\varphi}(y)]=\theta(x^0-y^0)\hat{\varphi}(x)\hat{\varphi}(y)+\theta(y^0-x^0)\hat{\varphi}(y)\hat{\varphi}(x) \tag{2.106}$$

In a time-ordered product the time-dependent operators are ordered so that later times stand to the left of earlier times. Time-ordered products read like Hebrew, right to left. There are several ways of interpreting $D(x-y)$. From a mathematical point of view, it is the Green's function of the Klein-Gordon equation, i.e.

$$(\partial_\mu\partial^\mu+m^2)D(x-y)=-\delta^{(4)}(x-y) \tag{2.107}$$

From a physical point of view, it is the probability amplitude for a particle to propagate from y to x. I need to prove the central result that

$$\boxed{iD(x-y)=i\int\frac{d^4k}{(2\pi)^4}\frac{e^{-ik(x-y)}}{k^2-m^2+i\epsilon}} \tag{2.108}$$

Each kind of particle has its own propagator, but the $+i\epsilon$ term in the denominator is ubiquitous. The ϵ stands for an infinitesimal *positive* quantity. It's job is to get the boundary conditions right as you will see in the derivation.

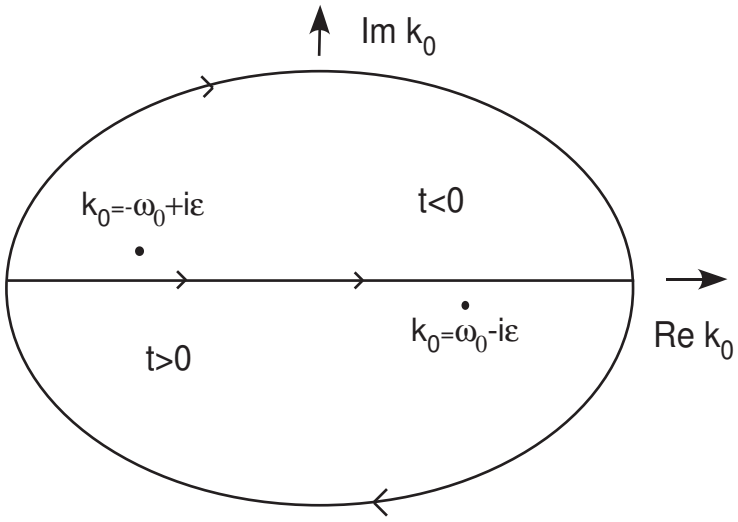


Figure 2.1: The complex k_0 plane

We know that this will be a function of $x-y$, so we can make the algebra a bit simpler by setting $y = 0$. Just remember to replace $x \to x - y$ at the end of the calculation. Substituting the fields from (2.97) into (2.105) and taking the vacuum expectation value gives

$$\begin{aligned} iD(x) &= \langle 0|T[\hat{\varphi}(\boldsymbol{x},t)\hat{\varphi}(0,0)]|0\rangle \\ &= \int \frac{d^3k}{(2\pi)^3 2\omega_k}\left[\theta(t)e^{-i(\omega_k t-\boldsymbol{k}\cdot\boldsymbol{x})}+\theta(-t)e^{i(\omega_k t-\boldsymbol{k}\cdot\boldsymbol{x})}\right] \end{aligned} \tag{2.109}$$

Equations (2.108) and (2.109) are really the same result, though this is far from obvious. In principle, we could derive either form from the other, but it's probably easier to start from (2.108).

$$iD(x) = i\int \frac{d^3k}{(2\pi)^4}\, e^{i\mathbf{k}\cdot\mathbf{x}} \int dk_0 \frac{e^{-ik_0 t}}{(k_0-\omega_k+i\epsilon)(k_0+\omega_k-i\epsilon)} \tag{2.110}$$

Notice how the denominator is factored. Multiplying the two factors and making the replacements, $2i\omega_k\epsilon \to i\epsilon$ and $\epsilon^2 \to 0$, gives the same denominator as (2.108). The dk_0 integration is now performed as a contour integration in the complex k_0 plane as shown in Figure 2.1. For $t < 0$ the contour is completed in the upper half-pane enclosing the point $k_0 = -\omega_k + i\epsilon$, and for $t > 0$ the contour is completed in the lower half-plane enclosing the

point $k_0 = \omega - i\epsilon$. The result is identical to (2.109). You see how the $i\epsilon$ in the denominator displaces the poles so as to pick up the right integrand depending on whether t is positive or negative. Notice finally that (2.106) is a Lorentz scalar since kx, k^2 and d^4k are all scalar quantities. You will see how $D(x-y)$ becomes a key player in perturbation theory via the interaction picture in the next chapter.

Chapter 3

The Interaction Picture and the S-Matrix

Most of what we know about subatomic physics comes from two kinds of experiments: decay experiments and scattering experiments. In a decay experiment, one starts with some system such as an atom or nucleus or "elementary" particle and observes the spontaneous transitions that it undergoes. One can determine the lifetime of the system, the identity of the decay products, the relative frequency of the various decay modes, and the distribution of momentum and energy among the resulting particles. In a scattering experiment, one starts with a stable system and bombards it with another particle. One measures the distribution of momenta among the various particles produced by the reaction and determines the probability that the scattering will lead to a particular final state. One common feature of both these experiments is that the particles are detected when they are not interacting with one another. They are detected on a scale of distance that is many orders of magnitude larger than the characteristic de Broglie wavelength and at times that are vastly longer than the time of interaction. In non-relativistic quantum mechanics these processes are calculated with first-order, time dependent perturbation theory using a bit of hocus pocus called "Fermi's golden rule." This is not altogether wrong, but it is inadequate for several reasons: it can't accommodate the creation of new particles, it's hopelessly non-relativistic, and it only works to first order.

Real scattering theory is difficult. There are many subtle issues involved. Much of the material in advanced quantum books relates to scattering in one way or another. I say this because it's easy to lose sight of the goal amidst all the technical difficulties. Roughly speaking, there are two basic issues: how

do the quantum fields $\varphi(x)$ evolve in time, and given this information, how can we calculate the results of these experiments in terms of the momenta of particles measured in the asymptotic regime as explained above? The first question is answered neatly by the interaction picture formalism first developed by Freeman Dyson. The second question is much more difficult and will require the remainder of this chapter.

3.1 The Interaction Picture

Path integrals use the Lagrangian; the interaction picture uses the Hamiltonian; so I should say a few words about Hamiltonians in general. In classical mechanics, the two are related by

$$H = p\dot{q} - L \tag{3.1}$$

In our variety of continuum mechanics the Hamiltonian density becomes

$$\mathcal{H} = \pi\dot{\varphi} - \mathcal{L} \tag{3.2}$$

For the free Klein-Gordon field

$$\mathcal{L} = \frac{1}{2}\left[(\partial_\mu\varphi)(\partial^\mu\varphi) - m^2\varphi^2\right] = \frac{1}{2}\left[\dot{\varphi}^2 - (\nabla\varphi)^2 - m^2\varphi^2\right] \tag{3.3}$$

$$\mathcal{H} = \frac{1}{2}\left[\dot{\varphi}^2 + (\nabla\varphi)^2 + m^2\varphi^2\right] \tag{3.4}$$

We get the Hamiltonian by integrating (3.4) over all space.

$$H = \int d^3x\mathcal{H} \tag{3.5}$$

We assume that the Hamiltonian can be split up into two pieces: $H = H_0 + H_{\text{int}}$. H_0 is the Hamiltonian of the free field, and H_{int} is everything left over. We assume the H_{int} is "small" in some sense; so that the perturbation series converges. We also have to assume that H_{int} is a polynomial of the fields and their derivatives. The reason for this technical assumption will appear later.

You will recall the relation between the Schrodinger and Heisenberg pictures is

$$\begin{aligned} |\psi(t)\rangle_S &= e^{-iHt}|\psi\rangle_H \\ Q_H(t) &= e^{iHt}Q_S e^{-iHt} \end{aligned} \tag{3.6}$$

This is the usual convention at any rate. $|\psi\rangle_H$, which does not depend on time, is a "snapshot" of $|\psi(t)\rangle_S$ taken at $t = 0$. This is not necessarily the most convenient time to take the picture as we will see.

We define the interaction state by

$$|\psi(t)\rangle_I \equiv e^{iH_0t}|\psi(t)\rangle_S \tag{3.7}$$
$$Q_I(t) = e^{iH_0t}Q_S\, e^{-iH_0t}$$

Some comments:

- It's clear from (3.6) the Hamiltonian H is the same in the Schrodinger and Heisenberg pictures, and equally clear from (3.7) that H_0 is the same in the interaction picture as well. This is not true of the interaction piece; since in general, $[H_{\rm int}, H_0] \neq 0$. I will write

  $$H^I_{\rm int}(t) \equiv e^{iH_0t}H^S_{\rm int}e^{-iH_0t} \tag{3.8}$$

  Naturally, $H^S_{\rm int}$ is the interaction Hamiltonian in the Schrodinger picture. Notice that $H^I_{\rm int}(t)$ *depends on time.*

- If there were no interactions, $|\psi(t)\rangle_I = |\psi\rangle_H$, so the interaction picture state would be stationary. Now think of a scattering experiment. When the particles are far apart, there is no interaction between them. In this regime, the interaction picture and Heisenberg pictures are identical. As the particles begin to interact, $H_{\rm int}$ "turns on" and $|\psi(t)\rangle_I$ begins to change. Eventually, after the particles are separated it settles down to a new Heisenberg state.

  Actually, this business of "turning on" and " turning off" skirts a profound and difficult problem. The trouble is that $H_{\rm int}$ never really turns off. Even when the particles are not interacting with one another, they are still interacting with themselves by emitting and reabsorbing virtual particles. Worse yet, even the vacuum interacts with itself by creating particle-antiparticle pairs. This is a difficult problem, but one that can be dealt with rigorously.

Define the time evolution operator in the interaction picture, $U(t, t_0)$, by

$$|\psi(t)\rangle_I = U(t, t_0)|\psi(t_0)\rangle_I. \tag{3.9}$$

Since H is Hermitian and the norm of $|\psi\rangle_I$ is conserved, we feel entitled by the rules of quantum mechanics to the following basic relations:

$$U^\dagger(t, t_0)U(t, t_0) = \mathbf{1} \tag{3.10}$$

$$U(t,t) = \mathbf{1} \tag{3.11}$$

$$U(t,t_0) = U^{-1}(t_0,t) \tag{3.12}$$

$$U(t_2,t_0) = U(t_2,t_1)U(t_1,t_0) \tag{3.13}$$

We can derive an equation of motion for U by starting with Schrodinger's equation

$$i\frac{\partial}{\partial t}|\psi(t)\rangle_S = H|\psi(t)\rangle_S. \tag{3.14}$$

A quick calculation with (3.7) and (3.9) yields

$$i\frac{\partial}{\partial t}U(t,t_0) = H^I_{\rm int}(t)U(t,t_0). \tag{3.15}$$

We know how to solve equations like this.

$$U(t,t_0) = 1 - i\int_{t_0}^{t} dt'\; H^I_{\rm int}(t')U(t',t_0) \tag{3.16}$$

Well – maybe solve is too strong a word, since U appears on both sides of (3.16). We would like to claim that

$$U(t,t_0) = (?)\exp\left\{-i\int_{t_0}^{t} dt'\; H^I_{\rm int}(t')\right\}$$

This would be the obvious solution if U and H were not operators. The flaw in the reasoning here is that factors of $H_{\rm int}$ don't commute at different times, $[H_{\rm int}(t), H_{\rm int}(t')] \neq 0$. We can come up with a valid solution by iterating (3.16) paying careful attention to the time dependence.

$$\begin{aligned} U(t,t_0) &= 1 - i\int_{t_0}^{t} dt_1 H^I_{\rm int}(t_1) \\ &+(-i)^2\int_{t_0}^{t} dt_1\int_{t_0}^{t_1} dt_2 H^I_{\rm int}(t_1)H^I_{\rm int}(t_2) + \cdots \end{aligned} \tag{3.17}$$

The entire series can be summarized by

$$U(t,t_0) = \sum_{n=0}^{\infty}(-i)^n\int_{t_0}^{t} dt_1\int_{t_0}^{t_1} dt_2\cdots\int_{t_0}^{t_{n-1}} dt_n H^I_{\rm int}(t_1)H^I_{\rm int}(t_2)\cdots H^I_{\rm int}(t_n) \tag{3.18}$$

The series (3.18) is more difficult than it looks. Since the Hamiltonians don't commute, we must be meticulously careful to keep **later times to**

the left of earlier times. This is called *time ordering.* We need some machinery to do this for us. Define the time ordering operator,

$$T[H(t_1)H(t_2)] \equiv H(t_1)H(t_2)\theta(t_1 - t_2) + H(t_2)H(t_1)\theta(t_2 - t_1) \tag{3.19}$$

The generalization to three or more Hamiltonians is obvious. You should convince yourself that the following things are true:

$$\begin{aligned}\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 H(t_1)H(t_2) &= \int_{t_0}^{t} dt_2 \int_{t_0}^{t_2} dt_1 H(t_2)H(t_1) \\ &= \frac{1}{2}\int_{t_0}^{t}\int_{t_0}^{t} dt_1 dt_2 T[H(t_1)H(t_2)],\end{aligned}$$

and in general

$$\int_{t_0}^{t} dt_1 \cdots \int_{t_0}^{t_{n-1}} dt_n H(t_1)\cdots H(t_n) = \frac{1}{n!}\int_{t_0}^{t} dt_1 \cdots \int_{t_0}^{t} dt_n T[H(t_1)\cdots H(t_n)]$$

So our final result for the U operator is

$$U(t,t_0) = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!}\int_{t_0}^{t}\cdots\int_{t_0}^{t} dt_1 dt_2 \cdots dt_n T[H^I_{\text{int}}(t_1)H^I_{\text{int}}(t_2)\cdots H^I_{\text{int}}(t_n)], \tag{3.20}$$

which can be written in shorthand notation as

$$\boxed{U(t,t_0) = T\exp\left\{-i\int_{t_0}^{t} dt' H^I_{\text{int}}(t')\right\}} \tag{3.21}$$

The reason that this is a workable approximation is that the fields that make up H^I_{int} are interaction-picture fields, which according to (3.8), transform as free fields. Equation (3.20) is a "recipe." You just insert the formulas for the free fields and do the algebra.

3.2 Interactions and the S Matrix

A theory with just one kind of particle doesn't present much opportunity for interactions. We consider these "toy theories." We learn to play with the scattering formalism before we face up to the much greater difficulties of spin. There are two such models worthy of our attention, the so-called "phi-cubed" and "phi-to-the-fourth" theories.

$$\mathcal{H}^I_{\text{int}} = \frac{\lambda}{3!} : \varphi^3(x) : \tag{3.22}$$

and

$$\mathcal{H}^I_{\rm int} = \frac{\lambda}{4!} : \varphi^4(x) : \tag{3.23}$$

These represent particles interacting with themselves. In (3.22) they interact in groups of three as you might expect. In (3.23) they interact in groups of four. As mentioned in Section 2.4, all Hamiltonians must be normal ordered. Let's use these to do some sample calculations of scattering precesses. From a quantum-mechanical perspective, one starts with an initial state consisting of some particles with well-defined momentum, say $|k_1, k_2, \cdots\rangle$ prepared in the lab at a time $t \to -\infty$ and observes the probability that this state evolves to a state $\langle p_1, p_2, \cdots|$ as $t \to +\infty$. A plausible, though not entirely correct, way to calculate this is as follows. We start by evaluating (3.21) at infinite times.

$$S = T \exp\left\{-i \int_{-\infty}^{\infty} dt' H^I_{\rm int}(t')\right\} \tag{3.24}$$

and simply take the expectation value of this operator between free-particle momentum eigenstates constructed with our creation and destruction operators. So for example, the two-particle scattering process, $k_1 + k_2 \to p_1 + p_2$, would be described by the matrix element

$$S_{\beta\alpha} = \langle\beta|S|\alpha\rangle \tag{3.25}$$

where $|\alpha\rangle = a^\dagger(\mathbf{k}_2)a^\dagger(\mathbf{k}_1)|0\rangle$ and $\langle\beta| = \langle 0|a(\mathbf{p}_1)a(\mathbf{p}_2)$. The reason this is not quite correct is that in addition to interacting with other particles, particles also interact with themselves, and these self-interactions are partly responsible for the mass and other physical properties we observe in the lab. The LSZ reduction scheme to be described later in this chapter deals with this problem in a systematic way (and leads to simpler calculations actually) at the expense of adding another conceptual layer of complexity to the theory. So long as one is working in the approximation that these self-interactions are ignored (the so-called tree level), the approach described above is quite satisfactory.

3.2.1 Two-Particle Scattering in the φ^4 Theory

The simplest possible scattering process to calculate is $k_1 + k_2 \to p_1 + p_2$ to lowest order in the φ^4 theory. In this case

$$S = -\frac{i\lambda}{4!} \int_{-\infty}^{\infty} d^4x \ : \varphi^4(x) : \tag{3.26}$$

$$S_{\beta\alpha} = \langle 0|a(\mathbf{p}_1)a(\mathbf{p}_2)Sa^\dagger(\mathbf{k}_1)a^\dagger(\mathbf{k}_2)|0\rangle \tag{3.27}$$

Since $:\varphi^4:$ is already normal-ordered, there is no time ordering to be done. If we write $\varphi = \varphi^+ + \varphi^-$ where φ^+ is the positive frequency part with annihilation operators and φ^- is the negative frequency part with creation operators, normal ordering implies that all φ^-'s stand to the left of all φ^+'s. Only those terms with two creation operators and two annihilation operators will give a non-zero result when sandwiched between the initial and final states. The expression $:(\varphi^+ + \varphi^-)^4:$ contains six such terms. A simple calculation (to be done as an exercise) gives

$$[\varphi^\dagger(x)]^2|k_1k_2\rangle = \frac{2e^{-i(k_1+k_2)x}}{\sqrt{2\omega_{k_1}2\omega_{k_2}(2\pi)^6}}|0\rangle \tag{3.28}$$

$$\langle p_1p_2|[\varphi^-(x)]^2 = \frac{2e^{+i(p_1+p_2)x}}{\sqrt{2\omega_{p_1}2\omega_{p_2}(2\pi)^6}}\langle 0|$$

where

$$kx = \omega t - \mathbf{k}\cdot\mathbf{x} \qquad \omega = \sqrt{m^2+\mathbf{k}^2} \tag{3.29}$$

$$\langle p_1p_2| :\varphi^4(x): |k_1k_2\rangle = 6\times 2\times 2\prod_i \frac{1}{\sqrt{2\omega_i(2\pi)^3}}e^{-i(k_1+k_2-p_1-p_2)x} \tag{3.30}$$

$$S_{\beta\alpha} = -i\lambda\prod_i \frac{1}{\sqrt{2\omega_i(2\pi)^3}}(2\pi)^4\delta^{(4)}(k_1+k_2-p_1-p_2) \tag{3.31}$$

The index i on the product sign refers to the momenta, k_1, k_2, p_1, and p_2. Notice that the 4! in the denominator of (3.23) cancels the other numerical factors that appeared in the course of the calculation. It was put there for that purpose.

The result is surprisingly simple and easy to understand. This is a simple instance of a general principle first discovered by Feynman; S-matrix elements can be written down without doing any calculations by just following a set of simple rules. We will derive the complete set of rules as we go along. Here I will just indicate the specific rules that (3.31) illustrates. First draw a picture in which the initial and final particles are represented by lines labeled with their momenta. The lines can intersect, but each intersection point must have exactly four lines. The relevant picture is shown in Figure 3.1. The S-matrix will consist of a product of the following terms obtained from the picture.

- A factor $(2\pi)^4\delta^{(4)}(\cdots)$. The argument of the delta function, here represented by a row of dots, insures that total four-momenta is conserved.
- A factor of $(-i\lambda)$ for each vertex.

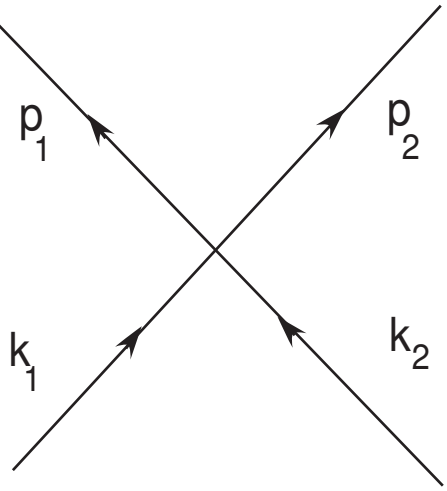


Figure 3.1: The reaction $k_1 + k_2 \to p_1 + p_2$

- A factor $1/\sqrt{2\omega_i(2\pi)^3}$ for each of the external momenta. By external momenta I mean the momenta of particles observed in the initial and final states. This particular diagram has no internal momenta. We'll get to that later.

This particular diagram has only one vertex and one factor of $(-i\lambda)$ because we are using the lowest-order term in the perturbation theory expansion. For the same reason we did not have to deal with the complications of time ordering. First-order perturbation theory doesn't cut much mustard, however. The φ^3 theory as well as QED require at least second order to calculate anything. We will have to evaluate expressions like $T[:\varphi^4(x_1)::\varphi^4(x_2):\cdots]$. This is done with the help of Wick's theorem.

3.3 The Wick Expansion

Suppose we were to calculate the scattering $k_1 + k_2 \to p_1 + p_2$ in φ^3 theory to lowest non-trivial order perturbation theory. The Feynman diagrams now have three particles meeting at each vertex, so it requires at least two vertices to scatter anything. We will have to evaluate the time-ordered product $T[:\varphi^3(x_1)::\varphi(x_2):]$. To do this by brute force we would split φ as before $\varphi = \varphi^- + \varphi^+$, and then rearrange the terms to that all φ^-'s stood to the left of all φ^+'s. There would be $2^6 = 64$ terms, so this is already

a formidable calculation. Fortunately there is a theorem from G. C. Wick that makes it possible to do this calculation without much work.

Take a look at the derivation of the propagator at the end of Chapter 2

$$\langle 0|T[\varphi_I(x_1)\varphi_I(x_2)]|0\rangle = iD(x_1 - x_2) \tag{3.32}$$

I am going to describe the steps that led from (2.105) to (2.108) in rather formal terms that make it easy to generalize to situations in which there are more than two fields. In order to keep the notation under control, I will write $\varphi(x) = \varphi^+(x) + \varphi^-(x)$, where φ^+ is the positive-energy piece containing an annihilation operator, and φ^- is the negative-energy term with the creation operator.

- Multiply the fields together with the time-ordering θ functions. This gives eight terms, each one containing a product of two of the creation and/or annihilation operators.

$$\begin{aligned} T[\varphi(x_1)\varphi(x_2)] = (\varphi^+(x_1) + \varphi^-(x_1))(\varphi^+(x_2) + \varphi^-(x_2))\theta(t_1 - t_2) \\ + (\varphi^+(x_2) + \varphi^-(x_2))(\varphi^+(x_1) + \varphi^-(x_1))\theta(t_2 - t_1) \end{aligned} \tag{3.33}$$

- Use the commutation relations (2.98) to rewrite each term (if necessary) in such a way that **all creation operators stand to the left of all annihilation operators.** This is called *normal ordering.* The usual notation for this is to enclose the expression between colons, i.e. $:\varphi(x_1)\varphi(x_2):$ means that the terms between the colons have already been placed in normal order. In the course of normal ordering there will appear terms without operators since, for example

$$\varphi^+(x_1)\varphi^-(x_2) =: \varphi^-(x_2)\varphi^+(x_1) : + \int \frac{d^3k}{(2\pi)^3 2\omega} e^{-ik(x_1 - x_2)} \tag{3.34}$$

  The first term on the right is normal ordered. The second term is a c-number. Combining all eight terms like this gives

$$T[\varphi(x_1)\varphi(x_2)] =: \varphi(x_1)\varphi(x_2) : + iD(x_1 - x_2) \tag{3.35}$$

- The vacuum expectation value of any normal ordered product is zero. As a consequence

$$\langle 0|T[\varphi(x_1)\varphi(x_2)]|0\rangle = \langle 0| : \varphi(x_1)\varphi(x_2) : |0\rangle + iD(x_1 - x_2) \tag{3.36}$$

Of course, the first term on the right is zero. We say that $iD(x_1-x_2)$ is the *contraction* of $\varphi(x_1)$ and $\varphi(x_2)$. I will use the overbrace to indicate contraction. In this notation (3.35) is

$$T[\varphi(x_1)\varphi(x_2)] =: \varphi(x_1)\varphi(x_2) : + \overbrace{\varphi(x_1)\varphi(x_2)} \tag{3.37}$$

I am now in a position to state Wick's theorem. The time ordered product of n fields $T[\varphi(x_1)\varphi(x_2)\cdots\varphi(x_n)]$ can be written as a sum of terms as follows:

- A single term in which all n fields are normal-ordered.
- A sum of terms, each one of which consists of one pair of contracted φ's. The $n-2$ remaining fields are normal-ordered. There is one term in the sum for each possible pairing of two of the fields.
- A sum of terms, each one of which consists of two pairs of contracted φ's. The remaining $n-4$ remaining fields are normal-ordered. There is one term in the sum for each way in which two distinct pairs can be made out of n fields.
- You see the pattern? One term with no contractions plus all possible terms with one contraction plus all possible terms with two contractions plus all possible terms with three contractions and so forth. If n is even, repeat this procedure $n/2$ times. The last group of terms will consist of $n/2$ contractions. There will be one term in this group for each way of choosing $n/2$ pairs out of n objects. If n is odd, repeat this procedure $(n-1)/2$ times. The last terms will have one normal-ordered field left over

This is all much easier to understand that to explain. The following example with $n=4$ should make this clear.

$$\begin{aligned}
&T[\varphi(x_1)\varphi(x_2)\varphi(x_3)\varphi(x_4)] =: \varphi(x_1)\varphi(x_2)\varphi(x_3)\varphi(x_4) : \\
&+ \overbrace{\varphi(x_1)\varphi(x_2)} : \varphi(x_3)\varphi(x_4) : + \overbrace{\varphi(x_1)\varphi(x_3)} : \varphi(x_2)\varphi(x_4) : \\
&+ \overbrace{\varphi(x_1)\varphi(x_4)} : \varphi(x_2)\varphi(x_3) : + \overbrace{\varphi(x_2)\varphi(x_3)} : \varphi(x_1)\varphi(x_4) : \\
&+ \overbrace{\varphi(x_2)\varphi(x_4)} : \varphi(x_1)\varphi(x_3) : + \overbrace{\varphi(x_3)\varphi(x_4)} : \varphi(x_1)\varphi(x_2) : \\
&+ \overbrace{\varphi(x_1)\varphi(x_2)}\overbrace{\varphi(x_3)\varphi(x_4)} + \overbrace{\varphi(x_1)\varphi(x_3)}\overbrace{\varphi(x_2)\varphi(x_4)} \\
&+ \overbrace{\varphi(x_1)\varphi(x_4)}\overbrace{\varphi(x_2)\varphi(x_3)}
\end{aligned} \tag{3.38}$$

You should try working out one of these time ordered products by brute force starting with (2.97) and (2.98). You will be convinced that the theorem is true. The proof is more work that it's worth.[1]

3.4 New Example – φ^3 Theory

Let's redo our elastic, two-particle scattering problem with the Hamiltonian (3.22). To lowest non-zero order of the perturbation expansion

$$S = \frac{(-i)^2}{2!}\left(\frac{\lambda}{3!}\right)^2 \int d^4x_1 \int d^4x_2 T[:\varphi^3(x_1)::\varphi^3(x_2):] \tag{3.39}$$

We need to calculate $S_{\beta\alpha}$ from (3.27), consequently the only non-vanishing terms in the Wick expansion are those with four normally-ordered fields and one pair of contracted fields. There are nine ways of contracting two groups of three fields. Consequently

$$T[:\varphi^3(x_1)::\varphi^3(x_2):] = 9\overbrace{\varphi(x_1)\varphi(x_2)}:\varphi^2(x_1)\varphi^2(x_2): + \cdots \tag{3.40}$$

$$\begin{aligned} S_{\beta\alpha} =& \langle \mathbf{p}_1\mathbf{p}_2|S|\mathbf{k}_1\mathbf{k}_2\rangle = \frac{(-i)^2}{2!}\left(\frac{\lambda}{3!}\right)^2 \int d^4x_1 \int d^4x_2 \\ & \times 9\,\overline{\varphi(x_1)\varphi(x_2)}\langle \mathbf{p}_1\mathbf{p}_2|:\varphi^2(x_1)\varphi^2(x_2):|\mathbf{k}_1\mathbf{k}_2\rangle \end{aligned} \tag{3.41}$$

A straightforward if time-consuming calculation yields

$$\langle \mathbf{p}_1\mathbf{p}_2|:\varphi^2(x_1)\varphi^2(x_2):|\mathbf{k}_1\mathbf{k}_2\rangle = 4\prod_i \frac{1}{\sqrt{2\omega_i(2\pi)^3}} \tag{3.42}$$

$$\begin{aligned} \times\Big\{ & \exp[-i(k_1+k_2)x_2 + i(p_1+p_2)x_1] + \exp[-i(k_1+k_2)x_1 + i(p_1+p_2)x_2] \\ & + \exp[-i(k_1-p_2)x_1 + i(p_1-k_2)x_2] + \exp[-i(k_1-p_1)x_1 + i(p_2-k_2)x_2] \\ & + \exp[-i(k_2-p_2)x_1 + i(p_1-k_1)x_2] + \exp[-i(k_2-p_1)x_1 + i(p_2-k_1)x_2]\Big\} \end{aligned}$$

Our strategy will be to combine (3.41) and (3.42) and do the integrals over d^4x_1 and d^4x_2. This will give six terms which collapse into three after a bit of algebra. The result is rather simple, so let's just look at the integrals

[1] If you want a proof see J. D. Bjorken and S. D. Drell, *Relativistic quantum Fields*, Section 17.4

over the first term in (3.42). First I need to remind you of some definitions. From (2.108),

$$\begin{aligned}\overbrace{\varphi(x_1)\varphi(x_2)} =& iD(x_1-x_2) = i\int \frac{d^4x}{(2\pi)^4}\frac{e^{-ik(x_1-x_2)}}{k^2-m^2+i\epsilon} \\ \equiv& i\int \frac{d^4x}{(2\pi)^4} e^{-ik(x_1-x_2)}\Delta_F(k)\end{aligned} \tag{3.43}$$

The last line implicitly defines the Feynman propagator,

$$\Delta_F(k) \equiv \frac{1}{k^2-m^2-i\epsilon} \tag{3.44}$$

Notice that $\Delta_F(k) = \Delta_F(-k)$. Now substitute (3.40), (3.42), and (3.43) into (3.41). Do the integrals, all of which are trivial, starting with the integrals over d^4x. The result is

$$\begin{aligned}S_{\beta\alpha} =& (-i\lambda)^2 \prod_i \frac{1}{\sqrt{2\omega_i(2\pi)^3}} (2\pi)^4\delta^{(4)}(k_1+k_2-p_1-p_2) \\ & \times [i\Delta_F(k_1-p_1) + i\Delta_F(k_1+k_2) + i\Delta_F(k_1-p_2)]\end{aligned} \tag{3.45}$$

Compare this with the example (3.31). The new feature is the Feynman propagators. Their role is easy to understand as I will explain in the next section.

3.5 Feynman Diagrams

Perturbation theory is hard because of all the algebra. The example I have just worked is the second-easiest scattering problem in QFT, but (3.42), even with the help of Wick's theorem, took me nine pages of scratch paper. And yet, the final answer (3.45) is very simple, and as we will see, can be represented by three simple diagrams. Feynman diagrams, whatever else they may be, are part of a "cookbook" recipe for calculating scattering amplitudes. You draw some diagrams according to a precise set of topological rules. Each diagram corresponds to one term in the perturbation-theory expansion of the scattering amplitude. Each feature of the diagram contributes a specific factor to the term. Once you are in possession of the rules, you can calculate things like (3.45) *without any thought whatsoever.* Here then are (most of) the rules.

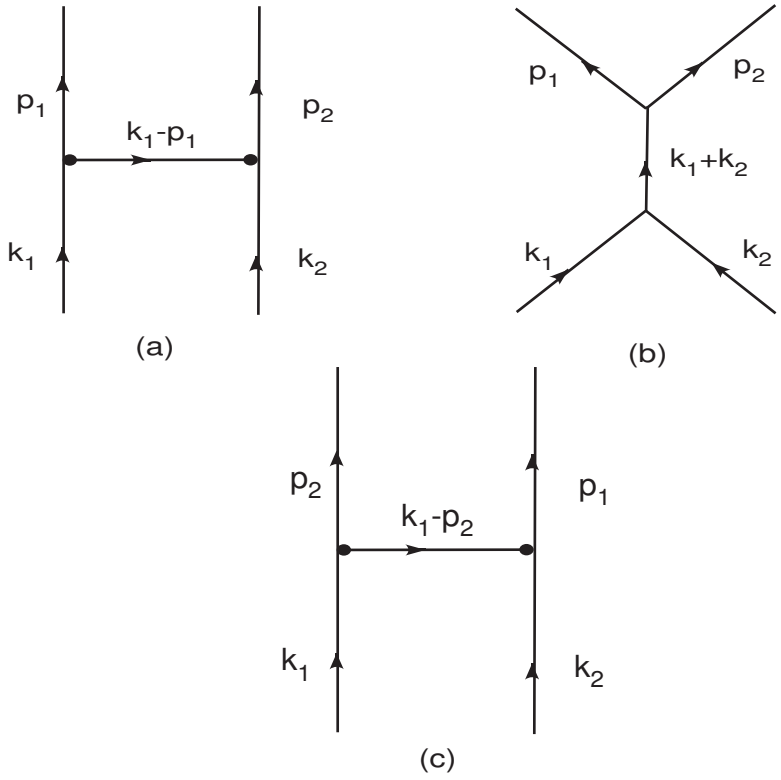


Figure 3.2: The reaction $k_1 + k_2 \to p_1 + p_2$

- The diagrams consist of lines that intersect at points called vertices. The number of lines (and the identity of the corresponding particles) depend on the structure of the interaction Hamiltonian. Four lines meet at each vertex in the φ^4 theory and three in the φ^3 theory. Diagrams arising from the n'th-order terms in the perturbation expansion have n vertices. For each vertex there is a factor of $(-i\lambda)$ in the corresponding term.

- The lines representing the initial- and final-state particles are called "external" lines. Each external line contributes a factor $1/\sqrt{2\omega_i(2\pi)^3}$ to the amplitude. I always draw the diagrams so that the lines representing initial-state particles enter the diagram from below and the final-state lines leave from above. The diagram represents a sequence of events in which "time runs up." Not all authors use the same convention, and it's important to be consistent. I also put arrows on the lines indicating the flow of momentum. Although this has not been relevant so far, lines representing antiparticles carry arrows that point in the *opposite* direction to the flow of momentum.

- Suppose you are doing the n-th order term in the perturbation expansion. Draw all topologically distinct diagrams with n vertices connected by the prerequisite number of lines and with the external lines entering and leaving the diagram appropriately. In addition to the external lines that begin or terminate at a vertex, there will be lines that connect two vertices. These are called internal lines.

- Label all lines with their momentum. External lines carry the momentum that is observed in the lab; in the previous example, k_1, k_2, p_1, and p_2. Label the internal lines as you wish, perhaps q_1, q_2, q_3 $\cdots$ would be a good choice. Put an arrow on each of the internal lines indicating the flow of momentum. The direction you choose will be arbitrary, but you must make a choice. Momentum is conserved at each vertex, i.e. the total momentum flowing in (according to the arrows you have just assigned) equals the total momentum flowing out. *If there are no loops* all the internal momenta are determined uniquely by the external momenta. Diagrams without loops are called tree diagrams. Life is simple in the trees. For every loop, however, there is an integral over an undetermined momentum. These integrals usually diverge and are generally deep trouble. We'll get to them in the next chapter.

- For each internal line with momentum q there is a corresponding factor of $i\Delta_F(q)$ in the scattering amplitude.
- There is an all-over factor of $(2\pi)^4\delta^{(4)}(\sum k_i - \sum p_j)$ enforcing conservation of total four-momentum.
- Not all diagrams are relevant. For example, if one or more of the external lines don't interact with other particles the diagram represents a process in which you are probably not interested. There are other issues that will be discussed later.
- There is one last point that cannot be decreed automatically with rules. In each calculation there will be a number of combinatorial factors such as the 9 in (3.40) and the 4 in (3.42). In this particular calculation they all canceled out, *but you can't count on it in general.* The LSZ reduction scheme in the next section is a big help in deciding what these factors are.

The rules are easier to understand than to explain. Just look at the diagrams in Figure 3.2. If I tell you that we are doing second-order two-particle scattering in φ^3 theory, then there are only three possible diagrams consistent with the rules above, and these are those! You are invited to review the rules and see how they apply to each of the diagrams and how they lead to the formula (3.45).

3.6 The Problem of Self Interactions

We can imagine quantum scattering as taking place in three phases. In the limit $t \to -\infty$ the particles are separated and not interacting *with one another*. As the particles approach one another, the interaction between them "turns on" as they exchange virtual particles. At some later time $t \to +\infty$, this interaction turns off and they again become free particles. This is difficult to treat theoretically, since the interaction Hamiltonian itself in no way turns on or off. So far as theoretical difficulties are concerned, this is only the tip of the iceberg. As soon as we allow particles to interact with other particles, they inevitably interact with themselves. When this happens, the notions of free particles, charge, and vacuum become profoundly subtle.

From the point of view of theory, a free particle is a boiling stream of self-interactions. From the point of view of experiment, every electron in the universe is an unremarkable little thing with a well-defined mass, spin, charge, and magnetic moment. These two views can exist simultaneously

because the self-interactions are virtual processes, and as such they are contained within a region roughly the size of the Compton wavelength. So long as one does not probe inside this region, the electron has it's usual physical properties, and probing inside the region means doing a high-energy scattering experiment.

The same sort of ambiguities also complicate our understanding of the vacuum state. According to perturbation theory, virtual particle-antiparticle pairs are spontaneously emitted from and reabsorbed into empty space. This fact is established experimentally (to some extent) by an experiment measuring the so-called Casimir effect,[2] but apart from this, these vacuum fluctuations have very little consequence, and at any rate, are not well understood.[3]

Finally consider the charge of an electron (or more generally, the coupling constant for any sort of interaction). The charge is the measure of the strength with which photons couple to electrons, but this can only be measured by doing a scattering experiment, and the experiment must be interpreted in light of the virtual processes involved.

We believe that all these problems can be addressed, at least for quantum electrodynamics, but this is a long story, which will occupy us for the next few chapters. For the time being, let us say that the theory recognizes two kinds of mass and two kinds of charge. There are the physical or "dressed" masses and coupling constants – those that are measured in the lab – and the "bare" masses and coupling constants, which are the parameters that go into the interaction Lagrangians and Hamiltonians. Somehow the effect of all virtual processes formulated in terms of the bare particles is to produce the physical particles. What then are the numerical values of these bare masses and charges? Alas, we do not know. There is no way to measure them directly. They could be calculated in principle, *but every such calculation yields infinity!* Despite this, we are able to calculate everything that can be measured (except mass and charge, of course) using only the physical mass and charge. The fact that we have learned to skirt infinities so violent as to give sci-fi fans nightmares, is one of the ironic triumphs of modern physics.

In view of this, the interaction picture formulation in the previous section is a bit too naive. What masses are we to put in H_{int}, the bare or the physical mass? It seems that either way we are stuck. We need to back up and look at the basic description of scattering processes.

Consider the state of a scattering experiment before the particles have

[2]Two charged plates suspended parallel to one another in vacuum experience a very small repulsive force because of virtual electron-positron pairs created between them.

[3]One can calculate the gravitational mass of all the virtual particle-antiparticle pairs in the universe. The result come out too large by many orders of magnitude.

interacted. For mathematical purposes we say this corresponds to a time "$t \to -\infty$" (The actual time for a relativistic particle could be as small as $t \to -10^{-23}$ seconds.) At this time the particles are free in the sense explained above. It is customary to call such states "in states," and write them $|\alpha, \text{in}\rangle$. The symbol α stands for all the quantum numbers required to completely specify the state. We will be particularly concerned with the momentum. If there are n particles in the initial state, we can write $|k_1, k_2, \ldots, k_n, \text{in}\rangle$. The field operators introduced in the previous chapter can be written

$$\varphi_{\text{in}}(x) = \int \frac{d^3x}{\sqrt{2E_k(2\pi)^3}} \left[a_{\text{in}}(k)e^{-ikx} + a^\dagger_{\text{in}}(k)e^{ikx}\right] \tag{3.46}$$

It is understood that in calculating with this expression, one uses the physical mass of the particle. In the same way, one defines "out states" $\langle\beta, \text{out}|$, "out fields" $\varphi_{\text{out}}(x)$, and "out" creation operators $a^\dagger_{\text{out}}(k)$ corresponding to the limit $t \to \infty$. A scattering experiment can be idealized as follows. The initial target and beam particles are prepared so that they are in a unique eigenstate of all the relevant operators. We call this state $|\alpha, \text{in}\rangle$. The detectors of the scattered particles are so refined that they can specify that the final state was in the exact eigenstate $\langle\beta, \text{out}|$. The object of the experiment is to find what fraction of particles in $|\alpha, \text{in}\rangle$ make the transition to $\langle\beta, \text{out}|$. In other words, we are measuring the probability

$$P_{\beta\alpha} = |S_{\beta\alpha}|^2 \tag{3.47}$$

where the so-called S matrix is defined by

$$S_{\beta\alpha} = \langle\beta, \text{out}|\alpha, \text{in}\rangle \tag{3.48}$$

If we assume that the in states and out states are both complete sets, there must be some operator S that transforms in states into out states.

$$\langle\beta, \text{out}| = \langle\beta, \text{in}|S \tag{3.49}$$

$$S_{\beta\alpha} = \langle\beta, \text{in}|S|\alpha, \text{in}\rangle \tag{3.50}$$

The goal of scattering theory is to calculate the S matrix and the corresponding S operator.

Here are some common-sense properties of S.

1. Nothing in – nothing out, i.e. $|S_{00}|^2 = 1$. Put it another way

$$\langle 0, \text{in}|S = \langle 0, \text{out}| = e^{i\theta}\langle 0, \text{in}| \tag{3.51}$$

The vacuum is stable, but we must allow the possibility that phase of the vacuum state evolves with time.

2. One-particle states are also stable.

$$\langle p, \text{out}|p, \text{in}\rangle = \langle p, \text{in}|S|p, \text{in}\rangle = \langle p, \text{in}|p, \text{in}\rangle = 1 \tag{3.52}$$

3. $$\varphi_{\text{in}} = S\varphi_{\text{out}}S^{-1} \tag{3.53}$$

Proof:

$$\langle\beta, \text{out}|\varphi_{\text{out}}|\alpha, \text{in}\rangle = \langle\beta, \text{in}|S\varphi_{\text{out}}|\alpha, \text{in}\rangle$$

But $\langle\beta, \text{in}|\varphi_{\text{out}}$ is itself an in state, so we can write

$$\langle\beta, \text{in}|\varphi_{\text{out}}|\alpha, \text{in}\rangle = \langle\beta, \text{out}|\varphi_{\text{in}}S|\alpha, \text{in}\rangle$$

Comparing these two equations gives

$$S\varphi_{\text{out}} = \varphi_{\text{in}}S$$

4. S is unitary.

Proof:

$$\langle\beta, \text{in}|S = \langle\beta, \text{out}| \qquad S^\dagger|\alpha, \text{in}\rangle = |\alpha, \text{out}\rangle$$
$$\langle\beta, \text{in}|SS^\dagger|\alpha, \text{in}\rangle = \langle\alpha, \text{out}|\beta, \text{out}\rangle = S_{\beta\alpha}$$

3.7 The LSZ Reduction Scheme

What is the relation between the fully-interacting field φ and φ_{in}? The natural guess is

$$\lim_{t\to-\infty}\varphi(x) = \varphi_{\text{in}}(x)$$
$$\lim_{t\to\infty}\varphi(x) = \varphi_{\text{out}}(x)$$

On second thought though, this can't be right. The complete field $\varphi(x)$ is capable of creating states out of the vacuum with two or more virtual particles. A more careful analysis suggests

$$\lim_{t\to-\infty}\varphi(x) = \sqrt{Z}\,\varphi_{\text{in}}$$

$$\lim_{t\to\infty}\varphi(x) = \sqrt{Z}\,\varphi_{\text{out}} \tag{3.54}$$

Z is understood as the probability for $\varphi(x)$ to produce a single-particle state out of the vacuum. Even this is not quite correct as an operator equation for some subtle mathematical reasons that are discussed in the original article.[4] It *is* correct in terms of matrix elements, for example

$$\lim_{t\to-\infty} \langle\alpha|\varphi(x)|\beta\rangle = \sqrt{Z}\langle\alpha|\varphi_{\rm in}(x)|\beta\rangle$$

where $\langle\alpha|$ and $|\beta\rangle$ are arbitrary states. Since we are only interested in matrix elements in the long run, we will assume (3.54) without any further comment. Furthermore, the processes in which $\varphi(x)$ produces multiparticle states out of the vacuum only occur in higher orders of perturbation theory. When we are working to lowest non-trivial order we can (and will) set $Z = 1$.

Now we have the apparatus at hand for defining and studying the S matrix defined in (3.48). We start from an initial state of a system with n noninteracting physical particles, denoted by

$$|k_1 \cdots k_n \text{ in}\rangle = |\alpha \text{ in}\rangle, \tag{3.55}$$

and a final state in which m particles emerge denoted by

$$\langle p_1 \cdots p_m \text{ out}| = \langle\beta \text{ out}| \tag{3.56}$$

Our agenda is to convert (3.48) into an expression involving vacuum expectation values of interacting fields. This is the LSZ reduction technique.

Let us write the in-field in a way analogous to (2.97)

$$\varphi_{\rm in}(x) = \int d^3k \left[f_k(x)a_{\rm in}(k) + f_k^*(x)a_{\rm in}^\dagger(k)\right] \tag{3.57}$$

where

$$f_k(x) = \frac{e^{-ikx}}{\sqrt{(2\pi)^3 2\omega_k}} \tag{3.58}$$

Equation (3.57) can be inverted to yield

$$a_{\rm in}(k) = i\int d^3x\ f_k^*(x) \overset{\leftrightarrow}{\partial_0}\ \varphi_{\rm in}(x)$$

$$a_{\rm in}^\dagger(k) = -i\int d^3x\ f_k(x) \overset{\leftrightarrow}{\partial_0}\ \varphi_{\rm in}(x) \tag{3.59}$$

[4] See also *Quantum Theory of Point Particles and Strings*, Brian Hatfield, Chapter 7.

The notation $a \overset{\leftrightarrow}{\partial_0} b$ means $a\partial_0 b - (\partial_0 a)b$. Since the in-state in (3.48) $|\alpha \text{ in}\rangle$ is made up with a product of creation operators as in (3.27), we can write

$$S_{\beta\alpha} = \langle \beta \text{ out}|a^\dagger_{\text{in}}(k)|\alpha - k \text{ in}\rangle \tag{3.60}$$

where $|\alpha - k \text{ in}\rangle$ represents the initial assemblage of particles represented by α with one particle of momentum k removed. The procedure is now to use (3.59) to replace $a^\dagger_{\text{in}}(k)$ with an integral over $\varphi_{\text{in}}(x)$. The procedure is then repeated until all the particles have been removed from the in- and out- states leaving the vacuum expectation value of the fields.

$$\begin{aligned} S_{\beta\alpha} &= \langle \beta \text{ out}|a^\dagger_{\text{out}}(k)|\alpha - k \text{ in}\rangle \\ &+\langle \beta \text{ out}|\left[a^\dagger_{\text{in}}(k) - a^\dagger_{\text{out}}(k)\right]|\alpha - k \text{ in}\rangle \\ &= \langle \beta - k \text{ out}|\alpha - k \text{ in}\rangle - i\langle \beta \text{ out}|\left[\int d^3x\, f_k(x) \overset{\leftrightarrow}{\partial_0} (\varphi_{\text{in}}(x) - \varphi_{\text{out}}(x))\right]|\alpha - k \text{ in}\rangle \end{aligned} \tag{3.61}$$

I have added and subtracted $a^\dagger_{\text{out}}(k)$ and used (3.59) to eliminate $a^\dagger_{\text{in}}$ and $a^\dagger_{\text{out}}$. The first term in the last line vanishes unless the initial and final states are identical. This is the unscattered wave, which we will neglect for the time being. Now use (3.54) to replace the in- and out- fields with the complete interacting field. The scattered part can be written as follows:

$$S_{\beta\alpha} \sim \frac{i}{\sqrt{Z}}\left(\lim_{x_0\to\infty} - \lim_{x_0\to-\infty}\right)\int d^3x\, f_k(x) \overset{\leftrightarrow}{\partial_0} \langle \beta \text{ out}|\varphi(x)|\alpha - k \text{ in}\rangle \tag{3.62}$$

(The symbol $\sim$ means that I have temporarily neglected the forward scattering term.) The limits have a nasty non-covariant look about them. I will clean up the expression with a series of elegant mathematical tricks. For any two functions such as $f_k(x)$ and $\varphi(x)$,

$$\begin{aligned} \left(\lim_{x_0\to\infty} - \lim_{x_0\to-\infty}\right)\int d^3x\, f_k \overset{\leftrightarrow}{\partial_0} \varphi &= \int_{-\infty}^{\infty} d^4x\, \partial_0\left[f_k \overset{\leftrightarrow}{\partial_0} \varphi\right] \\ &= \int_{-\infty}^{\infty} d^4x\left[f_k \frac{\partial^2}{\partial x_0^2}\varphi - \varphi\frac{\partial^2}{\partial x_0^2} f_k\right] \end{aligned} \tag{3.63}$$

The fact that f_k satisfies the Klein-Gordon equation allows us to write

$$\varphi\frac{\partial^2}{\partial x_0^2} f = \varphi(\nabla^2 - m^2)f \to f(\nabla^2 - m^2)\varphi \tag{3.64}$$

The sense of the right arrow is that the subsequent expression is obtained by integrating by parts twice and discarding the surface terms. Finally substituting (3.63) and (3.64) into (3.62) gives the final result

$$S_{\beta\alpha} = \langle \beta \text{ out}|\text{in } \alpha\rangle = \langle \beta - k \text{ out}|\alpha - k \text{ in}\rangle \tag{3.65}$$

$$+\frac{i}{\sqrt{Z}}\int d^4x\ f_k(x)(\Box + m^2)\langle \beta \text{ out}|\varphi(x)|\alpha - k \text{ in}\rangle$$

You see the pattern? We have "reduced" an asymptotic particle from the in-state and replaced it with an interacting field $\varphi(x)$. The price we pay for this is an integral, a wave function $f_k(x)$, and a differential operator $(\Box + m^2)$. We will eventually Fourier transform this expression, whereupon all these complications will disappear leaving only the inverse of the momentum space propagator $\Delta(k)$.

As an exercise (I am getting tired of typing) you should take the next step and reduce a particle of momentum p out of the out-state. Convince yourself that the result is

$$S_{\beta\alpha} = \langle \beta - p \text{ out}|\alpha - k \text{ in}\rangle \tag{3.66}$$

$$\sim \left(\frac{i}{\sqrt{z}}\right)^2 \int d^4x \int d^4y\ f_k(x) f_p^\dagger(y)(\Box_x + m^2)(\Box_y + m^2)$$
$$\times \langle \beta - p, \text{out}|T[\varphi(x)\varphi(y)]|\alpha - k, \text{in}\rangle$$

The $\sim$ sign again means that we have dropped the forward scattering terms. The new thing here is the time ordered product. You should do the calculation carefully to see how this comes about. Now suppose there are m particles in the initial state and n particles in the final state. Just repeat this procedure $n + m$ times to get the following result:

$$\boxed{\begin{aligned} S_{\beta\alpha} =& \langle p_1 \cdots p_n \text{ out}|k_1 \cdots k_m \text{ in}\rangle \\ \sim \left(\frac{i}{\sqrt{Z}}\right)^{m+n} & \prod_{i=1}^{m}\int d^4x_i f_{k_i}(x_i)(\Box_{x_i} + m^2) \prod_{j=1}^{n} d^4y_j f_{p_j}^*(y_j)(\Box_{y_j} + m^2) \\ & \times \langle 0|T[\varphi(y_1)\cdots\varphi(y_n)\varphi(x_1)\cdots\varphi(x_m)]|0\rangle \end{aligned}} \tag{3.67}$$

Equation (3.67), to quote Bjorken and Dell, "serves as a cornerstone for all calculation of scattering amplitudes in modern quantum field theory." (This was before the days of path integrals.)

Feynman rules are usually expressed in momentum space, so introduce the Fourier transform,

$$G(x_1 \cdots x_m y_1 \cdots y_n) = \prod_{i=1}^{m} \int d^4k_i \, \frac{e^{ik_i x_i}}{(2\pi)^4} \prod_{j=1}^{n} \int d^4p_j \, \frac{e^{-ip_j y_j}}{(2\pi)^4} \, \tilde{G}(k_1 \cdots k_m p_1 \cdots p_n) \tag{3.68}$$

There is a possible confusion about this equation. Since

$$G(x_1 \cdots x_m y_1 \cdots y_n) = \langle 0|T[\varphi(y_1) \cdots \varphi(y_n)\varphi(x_1) \cdots \varphi(x_m)]|0\rangle \tag{3.69}$$

There is no distinction between x's and y's, but in (3.68), x's are clearly associated with incoming particles, and y's are associated with outgoing particles. The point is that $G(x_1 \cdots x_m y_1 \cdots y_n)$ is a "general purpose" tool that can be used in many situations including those on a quantum-mechanical time scale in which there is no clear-cut distinction between past and future. In scattering experiments, on the other hand, past and future are unambiguous. We therefore associate some coordinates with in-states and some with out-states. We do this by assigning them to k's (for incoming particles) or with p's (for outgoing particles). In this simple theory with self-interacting scalar fields, the only difference is the $+$ or $-$ signs in the exponentials in (3.68), but this is necessary to get the conservation of momentum right. When there are several different kinds of particles involved, this choice becomes even more significant. For example, if we were looking at the process $e^+ + e^- \to \gamma + \gamma$, (positron annihilation) we would need to assign x's to the electron fields and y's to the photon fields. After making this assignment, substitute (3.68) into (3.67) and do the necessary integrals.

$$\boxed{\begin{aligned} S_{\beta\alpha} =& \langle p_1 \cdots p_n \text{ out}|k_1 \cdots k_m \text{ in}\rangle \\ &\sim \left(\frac{i}{\sqrt{Z}}\right)^{m+n} \prod_{i=1}^{m} \frac{1}{\sqrt{(2\pi)^3 2E_{k_i}}}(k_i^2 - m^2) \\ &\times \prod_{j=1}^{n} \frac{1}{\sqrt{(2\pi)^3 2E_{p_j}}}(p_j^2 - m^2)\tilde{G}(k_1 \cdots k_m p_1 \cdots p_n) \end{aligned}} \tag{3.70}$$

3.8 Correlation Functions

The heart of (3.67) is the vacuum correlation function

$$G(x_1 \cdots x_n) = \langle 0|T[\varphi(x_1)\varphi(x_2) \cdots \varphi(x_n)]|0\rangle \tag{3.71}$$

The fields $\varphi(x)$ are the fully interacting fields in the Heisenberg picture. If we are to calculate anything in this formalism, we must somehow relate them to the "in" fields defined in (3.54). Let's hope there exists an operator $U(t)$ such that

$$\varphi(x) = U^{-1}(t)\varphi_{\text{in}}(x)U(t) \qquad \text{with } t = x^0 \tag{3.72}$$

I happen to know that the necessary operator is given by

$$U(t) = \lim_{t_0 \to -\infty} U(t, t_0) \tag{3.73}$$

where $U(t, t_0)$ is time time-development operator given by (3.21). This is plausible, but it need to be proved.

Both φ and φ_{in} satisfy Heisenberg equations of motion.

$$\frac{\partial}{\partial t}\varphi(t) = i[H(\varphi), \varphi(t)] \tag{3.74}$$

$$\frac{\partial}{\partial t}\varphi_{\text{in}}(t) = i[H_0(\varphi_{\text{in}}), \varphi_{\text{in}}(t)] \tag{3.75}$$

I have written $H(\varphi)$ and $H_0(\varphi_{\text{in}})$ to emphasize that the Hamiltonians are functionals of their respective fields. I have also written $\varphi(t)$. I realize the φ is really a function of the four vector x^μ, but I am specifically concerned with the time development here, and $x_0 = t$. Note that

$$\frac{d}{dt}\left[UU^{-1}\right] = \frac{dU}{dt}U^{-1} + U\frac{dU^{-1}}{dt} = 0$$

In a condensed notation this is

$$U\dot{U}^{-1} = -\dot{U}U^{-1} \tag{3.76}$$

Now substitute (3.72) into (3.75) and use (3.74) to eliminate $\dot{\varphi}$.

$$\begin{aligned}
\frac{\partial}{\partial t}\varphi_{\text{in}} = \frac{\partial}{\partial t}\left[U\varphi U^{-1}\right] &= \dot{U}\varphi U^{-1} + U\dot{\varphi}U^{-1} + U\varphi\dot{U}^{-1} \\
= \dot{U}(U^{-1}\varphi_{\text{in}}U)U^{-1} + Ui\left[H(\varphi), \varphi\right]U^{-1} &+ U(U^{-1}\varphi_{\text{in}}U)\dot{U}^{-1} \\
= \dot{U}U^{-1}\varphi_{\text{in}} - \varphi_{\text{in}}\dot{U}U^{-1} &+ i\left[(UH(\varphi)U^{-1}), U\varphi U^{-1}\right] \\
= \left[\dot{U}U^{-1}, \varphi_{\text{in}}\right] + i\left[H(\varphi_{\text{in}}), \varphi_{\text{in}}\right] &= i\left[H_0(\varphi_{\text{in}}), \varphi_{\text{in}}\right]
\end{aligned} \tag{3.77}$$

I have used (3.76) to simplify the second line of (3.77). I have also used the fact that Hamiltonians are polynomials of φ (and possibly its derivatives), so for example, $UH(\varphi)U^{-1} = H(\varphi_{\rm in})$. Therefore

$$\left[\left\{\dot{U}U^{-1} + iH_{\rm int}(\varphi_{\rm in})\right\}, \varphi_{\rm in}\right] = 0$$

The operator in curly brackets commutes with all $\varphi_{\rm in}$. It must therefore be a c number. We could think of this constant as a number added to the Hamiltonian. We have had several occasions to discard such numbers before and will do so now. We conclude that (to within a meaningless additive constant)

$$i\frac{dU(t)}{dt} = H_{\rm int}(\varphi_{\rm in})U(t) \tag{3.78}$$

This is equation (3.15) with the important clarification that $H_{\rm int}$ is constructed with in fields, i.e. fields using the physical masses of the particles involved. We can take $U(t, t_0)$ to be given by (3.21) and $U(t)$ by (3.73). Using (3.12) and (3.13) we can argue that

$$\begin{aligned} U(t_2, t_1) &= \lim_{t\to-\infty} U(t_2, t_0)U(t_0, t_1) \\ &= \lim_{t\to-\infty} U(t_2, t_0)U^{-1}(t_1, t_0) = U(t_2)U^{-1}(t_1) \end{aligned} \tag{3.79}$$

Now consider a set of time-ordered points $x_1, \cdots, x_n$ satisfying $x_0^1 > x_2^0 > \cdots > x_n^0$. Using (3.79) we can write

$$\begin{aligned} G(x_1, \cdots, x_n) &= \langle 0|\varphi(x_1)\cdots\varphi(x_n)|0\rangle \\ &= \langle 0|U(t_0, t_1)\varphi_{\rm in}(x_1)U(t_1, t_2)\varphi_{\rm in}(x_2)\cdots U(t_{n-1}, t_n)\varphi_{\rm in}(x_n)U(t_n, t_0)|0\rangle \end{aligned} \tag{3.80}$$

The sequence of times reading from left to right runs from $t_0 \to -\infty$ to t_1 and back again to $-\infty$. We can deform the contour by writing $U(t_0, t_1) = U(t_0, -t_0)U(-t_0, t_1)$. An operator of the form $U(-t_0, t_0)$ runs from $-\infty$ to $+\infty$. It has the effect of transforming asymptotic initial states to asymptotic final states. This is just the role of the S operator defined in (3.74). We will simply call $U(t_0, -t_0) = S^{-1}$. With this change, our correlation function becomes

$$\begin{aligned} &G(x_1, \cdots, x_n) \\ &= \langle 0|S^{-1}U(-t_0, t_1)\varphi_{\rm in}(x_1)U(t_1, t_2)\cdots U(t_{n-1}, t_n)\varphi_{\rm in}(x_n)U(t_n, t_0)|0\rangle \end{aligned} \tag{3.81}$$

According to (3.51), the vacuum state is an eigenstate of S with a complex phase. We can write $\langle 0|S^{-1} = e^{-i\theta}\langle 0|$ and find the phase as follows.

$$\langle 0|S|0\rangle\langle 0|S^{-1} = \langle 0|$$

The point is that since the vacuum is stable, $|0\rangle\langle 0|$ constitutes a complete set of states. Consequently

$$\langle 0|S^{-1} = \frac{\langle 0|}{\langle 0|S|0\rangle} \tag{3.82}$$

Equation (3.81) becomes

$$G(x_1, \cdots, x_n) \tag{3.83}$$
$$= \frac{\langle 0|U(-t_0, t_1)\varphi_{\rm in}(x_1)U(t_1, t_2)\cdots U(t_{n-1}, t_n)\varphi_{\rm in}(x_n)U(t_n, t_0)|0\rangle}{\langle 0|S|0\rangle}$$

Consider an operator like $U(t_{i-1}, t_i)$. It contains an infinite sum of products of fields evaluated at various times, but every term in the sum contains fields that are time-ordered between t_{i-1} and t_i. The entire numerator of (3.83) consists of a vast array of fields, *but they are all time ordered.* Of course, we started out with the assumption that the sequence of coordinates $x_1, x_2, \cdots x_n$ was already time ordered, so this makes the "scaffolding" around which the U's are time ordered. The original definition of G, equation (3.71) already contains the time ordering operator, so we need not have made that assumption. We just need to write it

$$G(x_1, \cdots, x_n) \tag{3.84}$$
$$= \frac{\langle 0|T\left[U(-t_0, t_1)\varphi_{\rm in}(x_1)U(t_1, t_2)\cdots U(t_{n-1}, t_n)\varphi_{\rm in}(x_n)U(t_n, t_0)\right]|0\rangle}{\langle 0|S|0\rangle}$$

With the time-ordering operator in place, we are free to move the factors around in the numerator to suit our convenience. In particular, we can gather all the U's together.

$$U(-t_0, t_1)U(t_1, t_2)\cdots U(t_n, t_0) = U(-t_0, t_0) = S$$

Our final result for the correlation function function is

$$\boxed{G(x_1, \cdots, x_n) = \frac{\langle 0|T\left[\varphi_{\rm in}(x_1)\varphi_{\rm in}(x_2)\cdots\varphi_{\rm in}(x_n)S\right]|0\rangle}{\langle 0|S|0\rangle}} \tag{3.85}$$

Itzykson and Zuber[5] in a masterstroke of French understatement, call this derivation "heuristic." The fact is that we just don't know how to do rigorous mathematics with such badly-behaved operator expansions as (3.20), and there are many subtle physical questions as well. You might call this "experimental mathematics." We use this formula and it works. It is also true that equivalent formulas can be derived using the path integral formalism. This also goes way beyond the limits of conventional mathematics, but the same result is obtained in the end.

3.9 Two Examples

Let's take the interaction Hamiltonian $\mathcal{H}_{\rm int} = (\lambda/4!)\varphi^4(x)$ that we studied in Sections 4.3 and 4.4, and calculate the four-point Green's function $G(x_1, x_2, x_3, x_4)$ to first order in λ. The S-operator (3.24) to first order is

$$S = 1 - \frac{i\lambda}{4!}\int d^4x : \varphi^4(x) : + \cdots \tag{3.86}$$

Note that to first order $\langle 0|S|0\rangle = 1$, so (3.85) becomes

$$\begin{aligned} G(x_1, x_2, x_3, x_4) &= \langle 0|T[\varphi_{\rm in}(x_1)\varphi_{\rm in}(x_2)\varphi_{\rm in}(x_3)\varphi_{\rm in}(x_4)]|0\rangle \\ &- \frac{i\lambda}{4!}\langle 0|T\left[\varphi_{\rm in}(x_1)\varphi_{\rm in}(x_2)\varphi_{\rm in}(x_3)\varphi_{\rm in}(x_4)\int d^4x : \varphi_I^4(x) :\right]|0\rangle + \cdots \end{aligned} \tag{3.87}$$

The zeroth-order term can be read off of (3.37)

$$G(x_1, x_2)G(x_3, x_4) + G(x_1, x_3)G(x_2, x_4) + G(x_1, x_4)G(x_2, x_3) \tag{3.88}$$

This is the sum of all possible ways two particles can propagate between two pairs of points. As you would expect in this zeroth-order term, the particles do not interact .

The first-order term has eight fields each with two terms. If you were to simply multiply them out there would be $2^8 = 256$ terms, so here's where Mr. G. C. Wick earns his keep. First notice that the vacuum expectation value of any normal-ordered product of fields vanishes. We need only be concerned with terms in which all eight fields are contracted. Because of the normal-ordered Hamiltonian, we don't contract the four $\varphi(x)$'s with one another. The only surviving terms are those in which each of $\varphi(x_1)$, $\varphi(x_2)$, $\varphi(x_3)$, and $\varphi(x_4)$ is contracted with one of the $\varphi(x)$'s. There are, of course,

[5] *Quantum Field Theory*, C. Itzykson and J.-B. Zuber, McGraw-Hill 1980

4! ways of doing this all leading to the same expression for the first-order contribution to $G(x_1, x_2, x_3, x_4)$.

$$G(x_1, x_2, x_3, x_4) = -i\lambda \int d^4x G(x_1, x)G(x_2, x)G(x_3, x)G(x_4, x) \tag{3.89}$$

Now suppose you want to calculate the S matrix element for the scattering of two particles $k_1 + k_2 \to p_1 + p_2$ again to first order. We will use (3.67) with $m = n = 2$. To first order there is no renormalization so $Z = 1$.(More about this in the next chapter.) We'll use (3.89) with x_3 set equal to y_1 and $x_4 = y_2$. Take advantage of the fact that G is a Green's function after all, so

$$(\Box_x + m^2)G(x, x') = i(\Box_x + m^2)D(x - x') = -i\delta^{(4)}(x - x') \tag{3.90}$$

You see, all the $(\Box_{x_i} + m^2)$ operators in (3.67) produce factors of $-i\delta^{(4)}(x_i - x)$ when they operate on $G(x_i, x)$. Now all the integrals are trivial.

$$\begin{aligned} S_{\beta\alpha} =& \langle p_1 p_2 \text{ out}|k_1 k_2 \text{ in}\rangle \\ =& (-i\lambda)(2\pi)^4 \delta^{(4)}(p_1 + p_2 - k_1 - k_2) \\ & \times \prod_{i=1}^{2} \frac{1}{\sqrt{(2\pi)^3 2E_{k_i}}} \prod_{j=1}^{2} \frac{1}{\sqrt{(2\pi)^3 2E_{p_j}}}, \end{aligned} \tag{3.91}$$

which is the same result as (3.31) obtained with our more naive approach. It's worth noting that the arguments leading up to this result are very subtle, but the actual calculation is very little effort.

Before leaving this example I should mention the factor $\langle 0|S|0\rangle$ in the denominator of (3.84). It is an infinite series of terms representing particles that spring spontaneously out of the vacuum and then disappear back into the vacuum, vacuum-to-vacuum graphs the're called. The same graphs appear in the numerator, but in our simple example, they are all multiplied by the S-matrix for $k_1 + k_2 \to p_1 + p_2$. If we carried out the calculation to infinite order, all the vacuum-to-vacuum graphs in the denominator would cancel those in the numerator leaving an infinite series for $k_1 + k_2 \to p_1 + p_2$ without any "disconnected" graphs. We are thus entitled to ignore the denominator so long as we remember to discard all the multiplicative vacuum-to-vacuum graphs in the numerator.

Finally let's work out the S-matrix for $k_1 + k_2 \to p_1 + p_2$ in the φ^3 theory. The correlation function (3.85) is

$$\begin{aligned} G(x_1, x_2, y_1, y_2) =& \frac{1}{2!}\left(\frac{-i\lambda}{3!}\right)^2 \int d^4x \int d^4y \\ \times \langle 0|T[\varphi_{\text{in}}(x_1)\varphi_{\text{in}}(x_2)\varphi_{\text{in}}(y_1)\varphi_{\text{in}}(y_2) &: \varphi_I^3(x) :: \varphi_I^3(y) :]|0\rangle \end{aligned} \tag{3.92}$$

Only those terms in the Wick expansion with fully contracted fields will survive, and only those in which all of the $\varphi_{\rm in}$'s are contracted with the φ_I's will be relevant. The surviving terms then are these

$$\begin{aligned}
&\langle 0|T[\varphi_{\rm in}(x_1)\varphi_{\rm in}(x_2)\varphi_{\rm in}(y_1)\varphi_{\rm in}(y_2) : \varphi_I^3(x) :: \varphi_I^3(y) :]|0\rangle = \cdots \\
&+ \overbrace{\varphi_{\rm in}(x_1)\varphi_I(x)}\overbrace{\varphi_{\rm in}(x_2)\varphi_I(y)}\overbrace{\varphi_{\rm in}(y_1)\varphi_I(x)}\overbrace{\varphi_{\rm in}(y_2)\varphi_I(y)}\overbrace{\varphi_I(x)\varphi_I(y)} \\
&+ \overbrace{\varphi_{\rm in}(x_1)\varphi_I(x)}\overbrace{\varphi_{\rm in}(x_2)\varphi_I(x)}\overbrace{\varphi_{\rm in}(y_1)\varphi_I(y)}\overbrace{\varphi_{\rm in}(y_2)\varphi_I(y)}\overbrace{\varphi_I(x)\varphi_I(y)} \\
&+ \overbrace{\varphi_{\rm in}(x_1)\varphi_I(x)}\overbrace{\varphi_{\rm in}(x_2)\varphi_I(y)}\overbrace{\varphi_{\rm in}(y_1)\varphi_I(y)}\overbrace{\varphi_{\rm in}(y_2)\varphi_I(x)}\overbrace{\varphi_I(x)\varphi_I(y)} \\
&+ \cdots
\end{aligned} \tag{3.93}$$

and another three terms that are identical except that the x's and y's (without subscripts) are interchanged. There are $(3!)^2$ ways of making each of these six terms, so all the combinatorial factors cancel. When G from (3.92) is substituted back into (3.67) all the $\varphi_{\rm in}$'s disappear since for example (see (3.90))

$$\int d^4x_1(\Box_{x_1} + m^2)\overbrace{\varphi_{\rm in}(x_1)\varphi_I(x)} = -i. \tag{3.94}$$

When the smoke has cleared the result is the same as (3.45).

Chapter 4

The Trouble with Loops

Suppose we were doing our standard two-body scattering problem to fourth order in φ^3 theory. The correlation function is

$$G(x_1,x_2,y_1,y_2) = \frac{1}{4!}\left(\frac{-i\lambda}{3!}\right)^4 \int d^4w \int d^4x \int d^4y \int d^4z$$
$$\times \langle 0|\varphi_{\rm in}(x_1)\varphi_{\rm in}(x_2)\varphi_{\rm in}(y_1)\varphi_{\rm in}(y_2) : \varphi^3(w) :: \varphi^3(x) :: \varphi^3(y) :: \varphi^3(z) : |0\rangle \tag{4.1}$$

The Wick expansion will consist of a vast number of terms, but there is one class of terms that concerns me now.

$$: \varphi_{\rm in}(x_1)\varphi_{\rm in}(x_2)\varphi_{\rm in}(y_1)\varphi_{\rm in}(y_2) : \overbrace{\varphi(w)\varphi(x)}\,\overbrace{\varphi(x)\varphi(y)}\,\overbrace{\varphi(x)\varphi(y)}\,\overbrace{\varphi(y)\varphi(z)} \tag{4.2}$$

The associated Feynman diagram is shown in Figure 4.1. Because of conservation of momentum at each vertex, $p = k_1 - p_1 = p_2 - k_2$. I have also labeled the loop momentum in such as way as to acknowledge this rule. If we were to finish the calculation with this diagram and obtain the corresponding term in the S-matrix expansion, it would be identical to the first term of (3.45) except for the replacement

$$i\Delta_F(p) \to i\Delta_F(p)\left[\int \frac{d^4q}{(2\pi)^4}\,(-i\lambda)i\Delta_F(p-q)(-i\lambda)i\Delta_F(q)i\Delta_F(p)\right] \tag{4.3}$$

There are three points to be made about this. The first is that the integral over q is ubiquitous whenever there are loops. This is in fact the last of the Feynman rules for this simple scalar field theory.

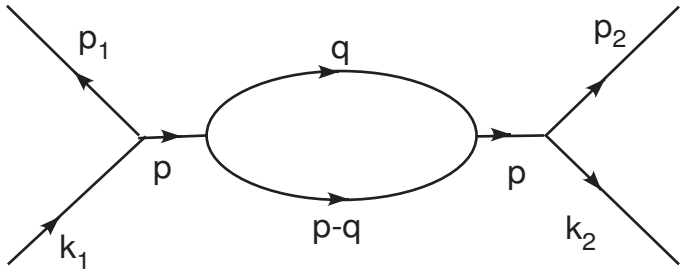


Figure 4.1: The one-loop correction to the propagator.

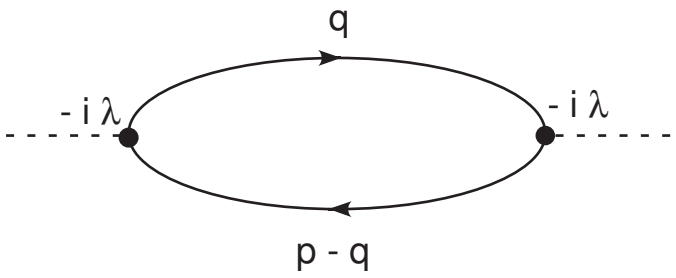


Figure 4.2: The truncated one-loop diagram.

- For every continuous loop there will be one undetermined momentum and an integral over it of the form

$$\int \frac{d^4q}{(2\pi)^4}. \tag{4.4}$$

The second point is that the effect of the loop is to make a multiplicative correction to the $\Delta_F(p)$ that appeared in second order. The correction factor is the quantity in the square brackets in (4.3). The correction is proportional to λ^2, so if λ is small, whatever that means, the correction itself might be small. As we explore higher orders in perturbation theory, there will be two-loop corrections proportional to λ^4, three-loop corrections proportional to λ^6 and so forth. If we are lucky the series will converge.

The third point is that we are not lucky. The integrals are all infinite. To see why this is so look at the "truncated" diagram Figure 4.2

$$\begin{aligned} &\int \frac{d^4q}{(2\pi)^4}(-i\lambda)i\Delta_F(q)(-i\lambda)i\Delta_F(p-q) \\ &= \lambda^2 \int \frac{d^4q}{(2\pi)^4}\frac{1}{q^2-m^2+i\epsilon}\ \frac{1}{(p-q)^2-m^2+i\epsilon} \end{aligned} \tag{4.5}$$

There are four powers of q in the numerator and four in the denominator, so the integral diverges like $\int dx/x$, i.e. logarithmically. Nonetheless, integrals of this sort can be unambiguously divided into a finite and an infinite part. We can show that the infinite part affects only the coupling constant λ and the mass m, whereas the finite part provides a real and measurable correction to the theory. The purpose of this chapter is to show how this comes about.

4.1 Doing the Integrals

The integral in (4.5) is typical of loop integrals in general, and over the years a number of tricks and identities have evolved for doing them. What follows is a brief primer on these techniques. The first trick is to make the integral five-dimensional by introducing the Feynman parameter z. You can readily verify that

$$\frac{1}{AB} = \int_0^1 \frac{dz}{[Az+B(1-z)]^2} \equiv \int_0^1 \frac{dz}{D^2} \tag{4.6}$$

or more generally

$$\frac{1}{A_1A_2\cdots A_n} = \int_0^1 dz_1\cdots dz_n\delta\left(\sum z_i - 1\right)\frac{(n-1)!}{[z_1A_1+z_2A_2+\cdots+z_nA_n]^n} \tag{4.7}$$

In this case set

$$\begin{aligned} A =&(p-q)^2-m^2+i\epsilon \qquad B = q^2-m^2+i\epsilon \\ &D = q^2-m^2+z(p^2-2pq)+i\epsilon \end{aligned} \tag{4.8}$$

Now complete the square by setting

$$\ell = q - pz \tag{4.9}$$

$$D = \ell^2 - m^2 + p^2z(1-z) + i\epsilon = \ell^2 - \Delta + i\epsilon \tag{4.10}$$

where

$$\Delta \equiv m^2 - p^2 z(1-z) \tag{4.11}$$

Equation (4.5) is now

$$\lambda^2 \int_0^1 dz \int \frac{d^4\ell}{(2\pi)^4} \frac{1}{(\ell^2 - \Delta + i\epsilon)^2} \tag{4.12}$$

Most loop integrals are done by combining denominators using (4.6) or (4.7) and then completing the squares as in (4.9) and (4.10). The result is an integral of the form

$$\int \frac{d^4 l}{(2\pi)^4} \frac{1}{(l^2 - \Delta + i\epsilon)^n} \tag{4.13}$$

There is a general procedure for evaluating such integrals which I will now describe. As it stands this is an integral in Minkowski space. The first step toward putting it in a manageable form is to convert it to Euclidean space. We therefore define a Euclidean momentum

$$\ell^0 \equiv i\ell_E^0 \qquad \boldsymbol{\ell}_E = \boldsymbol{\ell} \tag{4.14}$$

Recall from Section 2.6 and particularly equation (2.110) that the dk_0 integral in the propagator is to be understood as a contour integral in the complex k_0 plane. The transformation (4.14) amounts to rotating the contour 90^0 counterclockwise. You can see from Figure 2.1 that this is possible without crossing any poles. This maneuver is called a Wick rotation. The integral is now

$$\begin{aligned} & i(-1)^n \int \frac{d^4\ell_E}{(2\pi)^4} \frac{1}{(\ell_E^2 + \Delta)^n} \\ & = \frac{i(-1)^n}{(2\pi)^4} \int d\Omega_4 \int_0^\infty d\ell_E \, \frac{\ell_E^3}{(\ell_E^2 + \Delta)^n}. \end{aligned} \tag{4.15}$$

The $i\epsilon$ has no further role to play, so I have deleted it. The integral over $d\Omega_4$ gives the total solid angle in four-dimensional space. There is a simple formula for the solid angle in any dimension d.

$$\int d\Omega_d = \frac{2\pi^{d/2}}{\Gamma(d/2)} \tag{4.16}$$

The Γ function is defined as usual.

$$\Gamma(\alpha) = \int_0^\infty dt \; t^{\alpha-1} e^{-t} \tag{4.17}$$

with $\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$ and $\Gamma(n)=(n-1)!$ for n integer. The derivation of (4.16) is quite ingenious. I have included it as an appendix to this chapter so as not to distract from our topic. It's easy to verify that this gives the expected results. $\int d\Omega_2 = 2\pi$ and $\int d\Omega_3 = 4\pi$. As it turns out $\int d\Omega_4 = 2\pi^2$, a result that could have been obtained in other ways as well. The interesting point, however, is that the right side of (4.16) *is perfectly well defined for non-integer d.* Thus we can imagine doing the integral

$$\int \frac{d^d\ell_E}{(2\pi)^d}\frac{1}{(\ell_E^2+\Delta)^2} = \int \frac{d\Omega_d}{(2\pi)^d}\cdot\int_0^\infty d\ell_E \frac{\ell_E^{d-1}}{(\ell_E^2+\Delta)^2} \tag{4.18}$$

using (4.16) for $\int d\Omega_d$. As it turns out, the second integral is also doable for non-integer d. Again the derivation is quite remarkable. I have included a brief outline of it in the appendix. The final result is

$$\boxed{\int \frac{d^d\ell_E}{(2\pi)^d}\frac{1}{(\ell_E^2+\Delta)^n} = \frac{1}{(4\pi)^{d/2}}\frac{\Gamma(n-d/2)}{\Gamma(n)}\left(\frac{1}{\Delta}\right)^{n-d/2}} \tag{4.19}$$

Now if $n=2$ and $d=4$, $\Gamma(0)=\infty$, which tells us what we already knew, the integral in (4.5) diverges. Of course $\Gamma(x)$ has poles at zero and negative integers, but aside from that, it is smooth and well defined. We can get a finite answer by taking $d=4-\epsilon$, where ϵ is some small positive number. Figuratively speaking, we are integrating over slightly less than four dimensions! You will see the point in this in a few lines.

Before finishing the integral, we need one more theorem (see Appendix)

$$X^\epsilon = 1+\epsilon\ln X + \mathcal{O}(\epsilon^2) \tag{4.20}$$

and a more or less well known result

$$\Gamma(\epsilon) = \frac{1}{\epsilon} - \gamma + \mathcal{O}(\epsilon) \tag{4.21}$$

where γ is the Euler-Mascheroni constant ≈ 0.5772.

Now using (4.14), (4.19), (4.20) and (4.21), (4.12) becomes

$$\begin{aligned}
&\lambda^2\int_0^1 dz\int\frac{d^dl_E}{(2\pi)^d}\frac{i}{(l_E^2+\Delta)^2}\\
&= i\lambda^2\int_0^1 dz\frac{1}{(4\pi)^{d/2}}\frac{\Gamma(2-d/2)}{\Gamma(2)}\left(\frac{1}{\Delta}\right)^{2-d/2}\\
&= i\lambda^2\int_0^1\frac{dz}{(4\pi)^2}\left[\frac{2}{\epsilon}-\gamma+\ln 4\pi-\ln\Delta+\mathcal{O}(\epsilon)\right]
\end{aligned} \tag{4.22}$$

Now here's the point; it's still divergent as $\epsilon \to 0$, but so long as we keep ϵ finite, we can still do rigorous mathematics. Our hope is that when we put this diagram in context with the rest of the theory, we will arrive at a point where all the ϵ-dependent terms cancel out. We can then set $\epsilon = 0$, *because nothing depends on* ϵ.

I should mention that there are at least two other ways to separate out the divergence. One can simply put a cutoff Λ at some large value of ℓ_E^2 in (4.15) and then investigate the result in the limit $\Lambda \to \infty$. One can also modify the propagator so that it goes to zero more rapidly and helps the convergence of the loop integrals. For instance we could replace the propagator

$$\frac{1}{p^2 - m^2 + i\epsilon} \to \frac{1}{p^2 - m^2 + i\epsilon} - \frac{1}{p^2 - M^2 + i\epsilon} \tag{4.23}$$

where $M \gg m$ is the mass of some fictitious particle. For $p^2 \ll M^2$ the second propagator is essentially zero, but when $p^2 > M^2$ the two propagators partially cancel and the combination falls off like $1/p^4$. We would then take the limit $M \to \infty$ to separate out the divergence. This is called Pauli-Villars regularization. All three techniques yield consistent results for our current problem, but the cutoff method gets into trouble with gauge invariance where that is an issue, and the Pauli-Villars scheme is "messy" in that it involves introducing non-existent particles.

So what are we to make of these infinities? Perhaps the theory is somehow wrong at large momentum transfers. We could decide to live with our ignorance and regard it as a low-energy approximation to some as yet undiscovered formalism. There is a more ambitious hope that the shortcomings of the theory only affect our ability to calculate the mass and coupling constant. The idea here is that the m and λ that appear in the Lagrangian are not the real physical quantities but rather some "bare" mass and coupling constant. If all our integrals were finite, then the loop diagrams would modify the bare quantities. If we knew what to put in the Lagrangian in the first place and then carried out the perturbation expansion to arbitrarily high order, we would end up with the laboratory values of m and λ. A theory with these properties is said to be renormalizable. Our scalar field theories in fact are. We will explore this in the next section.

Before we do so, however, there are two other diagrams that we must evaluate. The first, Figure 4.3, is affectionately known as a tadpole diagram.

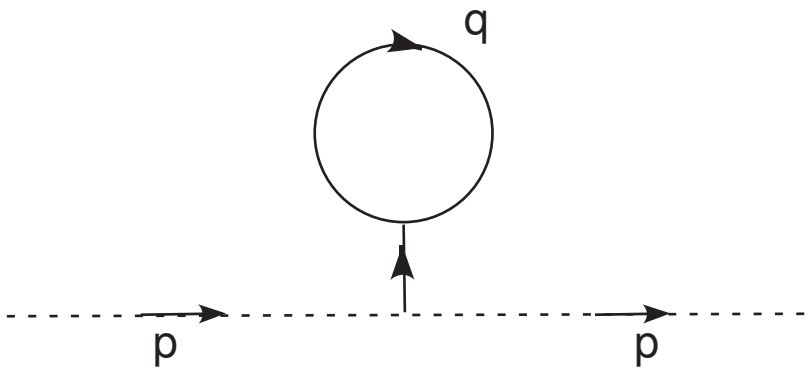


Figure 4.3: The tadpole correction to the propagator.

Truncating the diagram as in (4.5) leaves the terms

$$\begin{aligned}(i\lambda)i\Delta_F(0)(i\lambda)\int\frac{d^4q}{(2\pi)^4}i\Delta_F(q^2)\\ =\frac{\lambda^2}{m^2}\int\frac{d^4q}{(2\pi)^4}\frac{1}{p^2-m^2+i\epsilon}\end{aligned}\tag{4.24}$$

We can use (4.19) with $\Delta = m^2$ and $n = 1$.

$$=\frac{-i}{(4\pi)^{d/2}}\frac{\lambda^2}{m^2}\Gamma(1-d/2)(m^2)^{d/2-1}\tag{4.25}$$

Because the integral diverges quadratically, we need to generalize (4.21) a bit. For x near $-n$

$$\Gamma(x)=\frac{(-1)^n}{n!}\left(\frac{1}{x+n}-\gamma+1+\cdots+\frac{1}{n}+\mathcal{O}(x+n)\right)\tag{4.26}$$

In this case $x+1=2-d/2\equiv\epsilon$.

$$\Gamma(1-d/2)=-\left(\frac{1}{\epsilon}-\gamma+1+\mathcal{O}(\epsilon)\right)\tag{4.27}$$

Our final result for Figure 4.3 is[1]

$$\frac{i\lambda^2}{(4\pi)^2}\left(\frac{1}{\epsilon}-\gamma-\ln(m^2)+\ln(4\pi)\right)\tag{4.28}$$

[1]Both (4.28) and (4.22) have the odd feature of a logarithm of a dimensionfull quantity. The reasons for this are both subtle and important. By insisting that spatial integrals

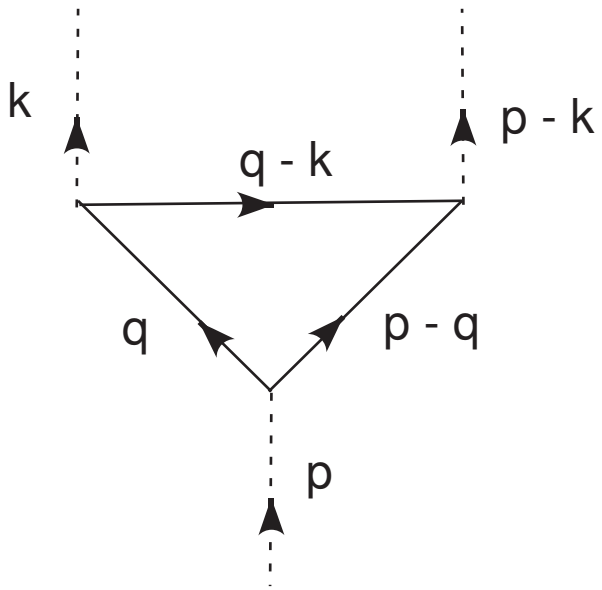


Figure 4.4: The third-order vertex correction.

The diagram in Figure 4.4 modifies the vertex where three lines meet. Neglecting the external lines, the triangle part of the diagram is

$$(i\lambda)\int \frac{d^4q}{(2\pi)^4}\frac{i}{q^2-m^2+i\epsilon}(i\lambda)\frac{i}{(q-k)^2-m^2+i\epsilon}(i\lambda)\frac{i}{(p-q)^2-m^2+i\epsilon} \tag{4.29}$$

Combine the denominators using (4.7)

$$= i\lambda^3\int_0^1 dz_2\int_0^1 dz_3\int \frac{d^4\ell_E}{(2\pi)^4}\frac{1}{(\ell_E^2+\Delta)^3} \tag{4.30}$$

Where

$$\Delta = (z_2k+z_3p)^2 - z_2k^2 - z^3p^2 + m^2 \tag{4.31}$$

have dimension $d \neq 4$ we have changed the dimensionality of φ and $\mathcal{L}$. In order to keep the action dimensionless, λ is no longer dimensionless. In order to keep it so it must be multiplied by μ^{3-d} where μ is an arbitrary mass. Of course, the result can't depend on the value of an arbitrary mass, which is my excuse for ignoring this point in the text. The fact that the theory *is* invariant under a rescaling of this mass is a key to an important development called the renormalization group. See Lewis Ryder, *Quantum Field Theory*, Section 9.1, for a good introduction to this subject.

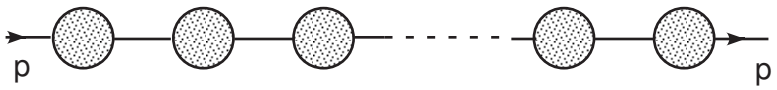


Figure 4.5: Iterated self-energy corrections to the propagator

The integral converges nicely, (This is a peculiarity of scalar field theories. Don't expect it in general.) so we don't have to worry about integrating over fractional dimensions. A straightforward application of (4.19) gives

$$= \frac{i\lambda^3}{2(4\pi)^2}\int_0^1 dz_2 \int_0^1 dz_3 \, \frac{1}{\Delta} \equiv \frac{i\lambda^3}{2(4\pi)^2} F(p^2, k^2, k\cdot p) \tag{4.32}$$

The z integrals are difficult;[2] the point is that the triangle diagram contributes a finite function of three scalar quantities, p^2, q^2, and $p \cdot k$.

4.2 Renormalization

Over the course of the last two chapters we have obtained the following.

$$\begin{aligned} G(x_1, x_2) &= \langle 0|T[\varphi(x_1)\varphi(x_2)]|0\rangle = iD(x_1 - x_2) \\ &= i\int \frac{d^4k}{(2\pi)^4} e^{-ik(x_1-x_2)} \Delta_F(k) \end{aligned} \tag{4.33}$$

$$\Delta_F(k) = \frac{1}{k^2 - m^2 + i\epsilon} \tag{4.34}$$

(See equations (3.71), (3.32), (3.44), and (2.108).) It is clear from (3.85) however, that the last line is only valid for $S = 1$, i.e. free, non-interacting fields. Since our purpose is to understand the effects of interactions, we must be more careful about notation. Let's use $G(x_1, x_2)$ and its Fourier transform $G(p)$ to represent the exact propagator for a particle to go from x_1 to x_2, and $G_0(x_1, x_2)$ and $G_0(p)$ to represent the zeroth-order propagators. Then $G_0(p) = i\Delta_F(p)$. The exact propagator calculated to all orders of perturbation theory would have an infinite number of diagrams, or course, but they could all be parceled up into an infinite sum of diagrams like Figure 4.5. The lines represent zeroth-order propagators, and each grey circle represents the sum of all diagrams with one particle in and one particle

[2] Actually, Maple and I find them impossible.

out, *which cannot be cut into two pieces by severing a single line.*[3] Such diagrams are said to be one-particle irreducible or 1PI for short. They are called proper self-energy parts for reasons that will become apparent shortly. In what is fairly standard notation, we denote them $-i\Sigma(p)$. The complete propagator is the sum of diagrams like Fig 4.5 with no self-energy part, one self-energy part, two parts, etc.

$$\begin{aligned} G(p) =& G_0(p) + G_0(p)\frac{\Sigma(p)}{i}G_0(p) + G_0(p)\frac{\Sigma(p)}{i}G_0(p)\frac{\Sigma(p)}{i}G_0(p) + \cdots \\ =& G_0\left(1 + \frac{\Sigma}{i}G_0 + \frac{\Sigma}{i}G_0\frac{\Sigma}{i}G_0 + \cdots\right) \\ =& G_0\left(1 - \frac{\Sigma}{i}G_0\right)^{-1} = \left[G_0^{-1}(p) - \frac{1}{i}\Sigma(p)\right]^{-1} \\ =& \frac{i}{p^2 - m^2 - \Sigma(p)} \end{aligned} \tag{4.35}$$

There is a deep result independent of perturbation theory that the exact propagator must have the form[4]

$$G(p) = \frac{i}{p^2 - m_R^2 + i\epsilon} + \mathcal{O}(p^2) \tag{4.36}$$

where m_R is the actual mass of the particle and $\mathcal{O}(p^2)$ is a function of p^2 with some branch points *but no poles*. Unless it should turn out that $\Sigma(p^2 = m^2) = 0$, which I assure you is not the case, the m we have been carrying around since Chapter 1, *is not the actual mass of the particle.* What is worse, both the single loop in Figure 4.2 and the tadpole in Fig 4.3 are 1PI and so are contained in Σ, and they are infinite or at least dependent on the dimensional parameter ϵ.

We interpret this as follows. The m in (4.35) originally appeared in the free-particle Lagrangian. As such it represents the mass of a particle without any interactions. There is no such thing of course, particles are always interacting, so it's a kind of meaningless theoretical abstraction. It is often called the "bare" mass in contrast to m_R, the renormalized mass. Therefore

$$0 = \left[p^2 - m^2 - \Sigma(p^2, \epsilon)\right]_{p^2 = m_R^2} \tag{4.37}$$

[3]Of course you can always make two pieces by cutting the head off a tadpole. That is not allowed in this definition.

[4]This is called the Lehman spectral representation. See M. E. Peskin and D. V. Schroeder, *An introduction to Quantum Field Theory*, Section 7.1

I have written $\Sigma(p^2, \epsilon)$ to emphasize the fact that so long as we keep ϵ finite, this is a reasonable mathematical expression. Now expanding Σ in the vicinity of $p^2 = m_R^2$,

$$\Sigma(p^2) \approx \Sigma(m_R^2) + Z^{-1}(p^2 - m_R^2) \tag{4.38}$$

where

$$Z^{-1} = -\left.\frac{d}{dp^2}\Sigma(p^2)\right|_{p^2=m_R^2} \tag{4.39}$$

Substituting (4.37), (4.38), and (4.39) into (4.35) gives

$$G(p^2) = \frac{iZ}{p^2 - m_R^2} + \cdots \tag{4.40}$$

This is not quite (4.36) because of the Z factor and for good reason. The field $\varphi(x)$ that appeared in our Lagrangian is also a free or bare field. It should also be renormalized by

$$\varphi_R(x) = Z^{-1/2}\varphi(x) \tag{4.41}$$

The factor $Z^{1/2}$ is called the wave function renormalization or field renormalization. We see from (4.41) that in terms of φ_R the propagator is the same as that of a free field with mass m_R, and therefore Z is the same factor that appeared in the LSZ reduction formula (3.67). In other words, Z disappears from the LSZ formula if, instead of using the bare field φ as we did in (3.67), we use the "physical" renormalized field. Thus after mass and wave function renormalization, there are no more bare quantities left in the propagator.

This is not the whole story however, because we are not given Σ ahead of time. We can only calculate it in perturbation theory. In order to do perturbation theory right we will need the renormalized mass, the field strength, and as it turns out, the renormalized coupling constant. It seems we are caught in another kind of loop. We can't find Σ until we know m_R, Z, and λ_R, and we can't find these things until we know Σ. It turns out that we can make sense of this order by order in perturbation theory. We proceed as follows. The lowest-order loop contributions come from Figures 4.2 and 4.3. Adding (4.22) and (4.28) we find

$$-i\Sigma = \frac{i\lambda^2}{(4\pi)^2}\frac{3}{\epsilon} + \text{ finite} \tag{4.42}$$

By "finite" I mean those terms that remain finite when $\epsilon = 0$. Let's say that in the one-loop approximation

$$\begin{aligned} G(p) &= \frac{i}{p^2 - m_R^2 + i\epsilon} + \text{ finite} \\ &= \frac{i}{p^2 - m^2 - \Sigma(p^2) + i\epsilon} + \text{ finite} \\ &\approx \frac{i}{p^2 - m_1^2 + i\epsilon} + \text{ finite} \end{aligned} \tag{4.43}$$

By m_1 I mean the one-loop approximation to the renormalized mass.

$$m^2 = m_1^2 + \frac{\lambda^2}{(4\pi)^2}\frac{3}{\epsilon} \tag{4.44}$$

We can also calculate Z from (4.39)

$$Z_1^{-1} = -\left.\frac{d}{dp^2}\Sigma(p^2)\right|_{p^2=m_1^2} \tag{4.45}$$

Notice that the constant is evaluated at $p^2 = m_1^2$ rather than m_R^2 as in (4.39). The point is that our approximation is correct to order λ^2. Using m_R rather than m_1 would introduce higher-order corrections.

The triangle diagram, Figure 4.4, modifies the three-particle vertex, but unlike the mass renormalization, it does so in a way that depends on momentum. We need to choose, somewhat arbitrarily, the kinematic point at which to make the correction. Simplest is the (unphysical) point $p^2 = k^2 = k{\cdot}p = 0$, in which case $\Delta = m^2$, $F(0,0,0) = 1/m^2$. Consequently the effect of the one-loop correction is to replace

$$-i\lambda \to -i\lambda_1 + \frac{i\lambda^3}{2(4\pi)^2 m_1^2} \tag{4.46}$$

or

$$\lambda = \lambda_1\left(1 + \frac{\lambda_1^2}{2(4\pi)^2 m_1^2}\right) \tag{4.47}$$

Again note where λ has been replaced by λ_1 and m by m_1 as appropriate. The implication is that if the one-loop approximation was adequate, λ_1 and m_1 would be the physical mass and coupling constant, and (4.44) and (4.47) could be used to remove all reference to m and λ in our calculations. If not, we need to go to two-loop corrections. Let us briefly consider what this would entail. Figure 4.6 shows a representative sample

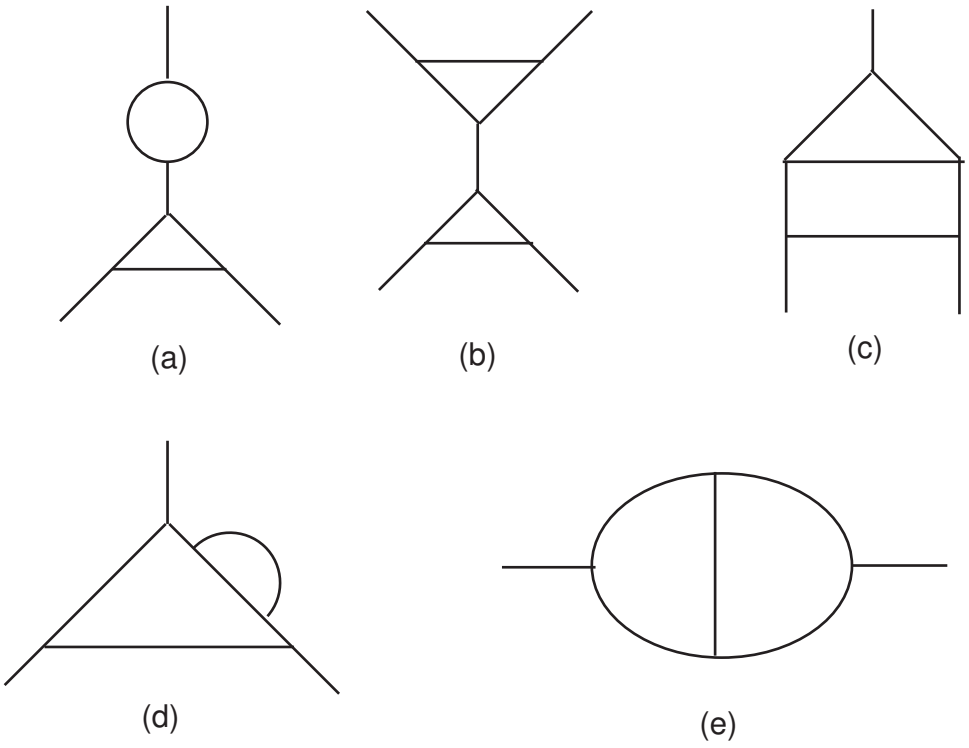


Figure 4.6: Some two-loop diagrams

of two-loop diagrams. Diagrams (a) and (b) are essentially trivial in that by cutting a single line they can be separated into two diagrams that we have already calculated. They are one-particle reducible. The other diagrams are genuinely new calculations. The λ^3 theory has the peculiarity that most higher-order diagrams are convergent. Consequently (c) and (e) simply give finite corrections to everything. Diagram (d) however, illustrates a notorious problem. It contains six propagators and four loop integrals, so the degree of convergence should be 8-12=-4, i.e. safely convergent. On the other hand, it contains within it the one-loop correction to the propagator, which we know diverges. It appears that one gets a finite or infinite answer depending on the order the two integrals are evaluated. This is an example of an "overlapping divergence." These are not so much a problem in QED because gauge invariance give a tool for dealing with them, but in λ^3 and especially in the λ^4 theory, they are a severe technical difficulty.

This concludes our study of scalar field theories. Except for the problem of overlapping divergences, they are child's play compared with the complexities of QED. They are a good introduction, however, for QED has all the same features with the additional complication of spin. The λ^4 theory, moreover, is of more that pedagogical interest. It forms the backbone of the electroweak theory, the closest thing we have to a theory of everything. Unfortunately, in this severely truncated course, we will have no time for that.

4.3 Appendix

Derivation of (4.16)

$$\begin{aligned}(\sqrt{\pi})^d &= \left(\int dx\ e^{-x^2}\right)^d = \int d^dx\ \exp\left(-\sum_{i=1}^{d} x_i^2\right) \\ &= \int d\Omega_d \int_0^\infty dx\ x^{d-1}e^{-x^2} = \left(\int d\Omega_d\right)\cdot\tfrac{1}{2}\int_0^\infty d(x^2)(x^2)^{\frac{d}{2}-1}e - x^2 \\ &= \left(\int d\Omega_d\right)\cdot\tfrac{1}{2}\Gamma(d/2)\end{aligned}$$

Second derivation

$$\begin{aligned}\int_0^\infty d\ell\ \frac{\ell^{d-1}}{(\ell^2+\Delta)^2} =& \frac{1}{2}\int_0^\infty d(\ell^2)\ \frac{(\ell^2)^{\frac{d}{2}-1}}{(\ell^2+\Delta)^2} \\ =& \frac{1}{2}\left(\frac{1}{\Delta}\right)^{2-\frac{d}{2}}\int_0^1 dx\ x^{1-\frac{d}{2}}(1-x)^{\frac{d}{2}-1}\end{aligned}$$

where I have substituted $x = \Delta/(\ell^2+\Delta)$ in the second line. Using the definition of the beta function

$$\int_0^1 dx\ x^{\alpha-1}(1-x)^{\beta-1} = B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

we can easily evaluate the integral over x.[5]

Proof of (4.20)

$$X^\epsilon = \exp\left\{\ln\left(X^\epsilon\right)\right\} = \exp\left\{\epsilon \ln X\right\} = 1 + \epsilon \ln X + \mathcal{O}(\epsilon^2)$$

[5]These derivations are taken from *An Introduction to Quantum Field Theory* by M. E. Peskin and D. V. Schroeder.

Chapter 5

Cross Sections and Decay Rates

The last step in the road from Lagrangians to cross sections is this: take the S-matrix computed in Chapter 7 and convert it to a scattering cross section or decay rate. There are two easy ways of doing this and one right way. One easy way is to assume that the initial and final state particles can be described by plane waves. Plane waves occupy infinite volume, unfortunately, so several key integrals diverge. This problem can be ducked with some judicious hand waving. The other easy way is to quantize the wave functions in a box of finite size. This avoids the infinite integrals, but of course, real experiments aren't done in a box. It is best to acknowledge that the beam and target particles should be represented by wave packets, but this makes for a long tedious derivation. We'll do it the ssecond easy way (at least in this edition).

5.1 Classical Scattering

From the point of view of classical mechanics, the cross section of a particle is just its projected area. If the particle is a sphere of radius R, the cross section $\sigma_T = \pi R^2$. This is usually called the "total" cross section. If the particle were very small, we might contrive to measure its cross section by means of a scattering experiment. First we prepare a target with n_t of these particles per unit volume, and bombard this target with N_i incident point-like particles. We assume that the target is so thin and the target particles so small that no one of them gets in front of another. We then count the number N of incident particles that are deflected as a result of a

collision. The fraction of particles that are scattered, N/N_i must be equal to the fraction of the target area covered by the projected area of the target particles, $n_t L\sigma_T$, where L is the target thickness. The cross section is then given by

$$\sigma_T = \frac{N}{N_i n_t L}$$

The idea works even if σ_T is arbitrarily small. We just have to make N_i and n_t large enough.

We can also determine the shape with a scattering experiment. Imagine that the target particles have concentric circles painted on them like an archery target. If one of the incident particles hits a particular area $d\sigma$ it will scatter into a solid angle $d\Omega$. OK – so your're not as good with a bow and arrow as Legolas. Just fire a gazillion arrows at random an observe the distribution of scattering angles. Your favorite theorist should then be able to infer the shape of the target particle from the distribution. The relevant equation is

$$\frac{d\sigma}{d\Omega} = \frac{1}{N_i n_t L}\frac{dN}{d\Omega} \tag{5.1}$$

If you are an experimentalist, you measure $dN/d\Omega$ in your lab, compute $d\sigma/d\Omega$, and pass it along to the theorist. This computed quantity is called the "differential cross section."

The quantities N_i, n_t, and L in (5.1) are directly measurable, but for purposes of comparing with theory, it is more convenient to express them in terms of *flux.* If the beam is moving at velocity v toward a stationary target, then the number of particles in the beam N_i is equal to the density of the beam n_i times the volume. If the beam is a pulse that is turned on for T seconds, then the volume of the beam is vTA, where A is the cross-sectional area of the beam (assumed smaller than the target.) Therefore $N_i = n_i vTA$. The cross section can be written as:

$$\begin{aligned}\frac{d\sigma}{d\Omega} =& \frac{1}{(n_i vTA)n_t L}\frac{dN}{d\Omega} \\ =& \frac{1}{I}\frac{dR}{d\Omega},\end{aligned} \tag{5.2}$$

where the transition rate R is the number of scattering events per unit time per unit target volume, $V = AL$.

$$R = N/VT, \tag{5.3}$$

and the flux I is

$$I = n_i n_t v \tag{5.4}$$

5.2 Quantum Scattering

In the quantum mechanics regime, quantities like size and shape don't have any direct meaning, since any putative size is bound to be smaller that the corresponding deBroglie wavelength; but the experiment can still be done exactly as I have described it, because all the measurements are macroscopic and asymptotic. The differential cross section (DCS for short) can't be interpreted directly in terms of shape, but it does contain most of the information we are entitled to have about particles and other subatomic systems.

The description of scattering in terms of archery targets is appropriate to low-energy potential scattering, but at higher energies new particles can be created and the total momentum shared among them. It requires more than a scattering angle to describe the final state. Rather we have to specify the momenta of each of the particles: $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n$. Not all values of the momenta are allowed of course, because of the conservation of momentum and energy. We can imagine the momenta as occupying some complicated-shaped region in a $3n$-dimensional space whose coordinates are the components of the individual momenta. This construct is called "phase space."[1] A typical multiparticle scattering experiment measures the transition probability into some region of phase space. We still use the differential cross section notation $d\sigma$, but now $d\sigma$ might be an $3n$-fold differential, depending on the details of the experiment. I will also use the notation $\Delta\Omega$ to indicate some small region of phase space (not just a region of solid angle).

Now let's put on our theorist's hats. We know how to calculate the S-matrix elements. How do we now find the cross sections? You will recall from the end of Chapter 3 that the first term of the perturbation theory expansion of the S-matrix describes particles that don't scatter. Feynman's rules decree that all other terms contain the factor $(2\pi)^4\delta^{(4)}(P_f - P_i)$, where P_i and P_f are the total momentum four-vectors of the initial and final state particles. It is customary to define a transition matrix $\mathcal{T}_{fi}$ as follows:

$$S_{fi} = \delta_{fi} - i(2\pi)^4\delta^{(4)}(P_f - P_i)\mathcal{T}_{fi} \tag{5.5}$$

The δ_{fi} just represents the particles that don't scatter. $\mathcal{T}$ is the non-trivial part of the S-matrix with the delta function "peeled off."

[1] Note that phase space in thermodynamics is $6n$-dimensional, since it includes space as well as momentum coordinates.

There is an alternative notation that is frequently used.

$$S_{fi} = \delta_{fi} + i(2\pi)^4\delta^{(4)}(P_f - P_i)\prod_i \frac{1}{\sqrt{2E_i(2\pi)^3}}\prod_f \frac{1}{\sqrt{2E_f(2\pi)^3}}\,\mathcal{M}_{fi} \quad (5.6)$$

In the case of the decay of a single particle, the index $i = 1$, and in the usual case of two-particle scattering $i = 1, 2$. The index f runs over all the final-state particles. Obviously, $\mathcal{M}_{fi}$ is just $\mathcal{T}_{fi}$ with the $1/\sqrt{2E(2\pi)^3}$ factors peeled off.

If should be clear from (3.24) that the S *operator* is unitary. The S-matrix in the current notation is $S_{fi} = \langle f|S|i\rangle$, where $|i\rangle$ and $\langle f|$ represent complete sets of states. As such we expect the S-matrix to be unitary as well. Very schematically, this means that

$$\sum_f S^*_{fi}S_{fk} = \delta_{ik} \quad (5.7)$$

I say "very schematically" because it's not at all clear what it means to sum over one of these indices. I'll get back to that.

Since S_{fi} is a probability amplitude, we should get the corresponding probability by squaring it. So for example,

$$P_{fi} = |S_{fi}|^2 \quad (5.8)$$

is the probability that the initial state $|i\rangle$ will scatter into the final state $|f\rangle$. Unfortunately this means that we have to square the delta function in (5.5), and this is bad news, since $\delta(P_f - P_i)\delta(P_f - P_i) = \delta(P_f - P_i)\delta(0)$ and $\delta(0)$ is infinite.

I'll deal with this last problem first. Squaring (5.5) gives

$$P_{fi} = (2\pi)^4\delta^{(4)}(0)(2\pi)^4\delta^{(4)}(P_f - P_i)|\mathcal{T}_{fi}|^2 \quad (5.9)$$

(I have ignored the non-scattering term.[2]) Here's the kludge. The factor $(2\pi)^4\delta^{(4)}(0)$ is interpreted as a volume of space-time, since in the limit $P \to 0$

$$(2\pi)^4\delta^{(4)}(P) = \int d^4x\; e^{iPx} \to \int d^4x = VT. \quad (5.10)$$

[2]All textbooks drop the non-scattering term. The reasons are obvious enough. No text I know of gives any explanation for dropping the *interference term* between the scattering and non-scattering terms. This can be partly justified by using wave packets to derive the scattering formalism, but as a matter of fact, interference effects should be observable at least in principle in small-angle scattering.

But the time and volume of what? We can use the derivation of (5.2) as a model. Let V be a small arbitrary volume inside the interaction region. Assume that it is small enough to fit inside the beam and target but very large compared the typical quantum-mechanical dimensions. T is the time required for a beam particle to pass through V. The analogy to (5.3) is

$$\begin{aligned} R_{fi} = \text{transition rate per unit volume } &= \frac{P_{fi}}{VT} \\ &= \frac{(2\pi)^4 VT\delta^{(4)}(P_f - P_i)|\mathcal{T}_{fi}|^2}{VT} \\ &= (2\pi)^4\delta^{(4)}(P_f - P_i)|\mathcal{T}_{fi}|^2 \end{aligned} \tag{5.11}$$

This is the first of several instances in which the arbitrary V's and T's cancel.

Now to the question of counting and normalizing states. This is not explained clearly in any book I know of, so I hope you will indulge me while I explain this boring technical detail in excruciating detail. Someone has to do it.

First of all, Klein-Gordon wave functions have a conserved norm, but it isn't $\int dV|\varphi|^2$. In order to have the right behavior under Lorentz transformations, the norm must transform like the time-like component of a conserved current. That is, there must be a probability density ρ and a probability current $\boldsymbol{j}$ such that

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla\cdot j} = 0 \tag{5.12}$$

The reason is that the rate of decrease of particles in a given volume is equal to the total flux of particles out of that volume, that is

$$-\frac{\partial}{\partial t}\int_V \rho\, dV = \int_S \mathbf{j}\cdot\hat{\mathbf{n}}\, dS = \int_V \nabla\cdot\mathbf{j}\, dV$$

The procedure for finding this current should be familiar from nonrelativistic quantum mechanics. Write the Klein-Gordon equation as

$$-\frac{\partial^2\varphi}{\partial t^2} + \nabla^2\varphi = m^2\varphi$$

Multiply this equation by $-i\varphi^*$ and the complex conjugate equation by $-i\varphi$ and subtract.

$$\frac{\partial}{\partial t}\underbrace{\left[i\left(\varphi^*\frac{\partial\varphi}{\partial t} - \varphi\frac{\partial\varphi^*}{\partial t}\right)\right]}_{\rho} + \nabla\cdot\underbrace{[-i(\varphi^*\nabla\varphi - \varphi\nabla\varphi^*)]}_{\mathbf{j}} = 0 \tag{5.13}$$

As a consequence the norm is

$$\rho = i \int d^3x \; \varphi^*(x) \overleftrightarrow{\partial_0} \varphi(x) \tag{5.14}$$

By $\varphi(x)$ I mean the wave function, not the quantized field, and

$$\varphi^* \overleftrightarrow{\partial_0} \varphi \equiv \varphi^* \partial_0 \varphi - (\partial_0 \varphi^*) \varphi.$$

We have been using the quantized field operator

$$\hat{\varphi} = \int \frac{d^3k}{\sqrt{(2\pi)^3 2E_k}} \left[\hat{a}_k e^{-ikx} + \hat{d}_k^\dagger e^{ikx} \right]$$

from (2.97). We obtain the classical wave function φ from the quantized field operator $\hat{\varphi}$ as follows

$$\begin{aligned} \varphi(x) = \langle 0|\hat{\varphi}(x)|\mathbf{p}\rangle &= \int \frac{d^3k}{\sqrt{(2\pi)^3 2E_k}} e^{-ikx} \langle 0|\hat{a}_k \hat{a}_p^\dagger|0\rangle \\ &= \frac{e^{-ipx}}{\sqrt{(2\pi)^3 2E_p}} \end{aligned} \tag{5.15}$$

This result depends on three normalization conventions that are to some extent arbitrary. First is the factor $1/\sqrt{(2\pi)^3 2E}$ in the field operator (2.97). Second is the commutation relations for the creation and annihilation operators (2.98). Finally, we have agreed to construct single-particle states as $|\mathbf{p}\rangle = \hat{a}_\mathbf{p}^\dagger |0\rangle$. Only the equal-time commutation relations (2.102) are nonnegotiable. We will finally arrive at formulas for the cross sections and decay rates that must be independent of all conventions. This will happen, but no two books do it the same way!

Our choice of normalization for the field operator has the advantage of making (2.98) as simple as possible, but it results in a rather odd-looking norm.

$$\begin{gathered} \rho = i(\varphi^* \overleftrightarrow{\partial_0} \varphi) = 2E\varphi^*\varphi = (2\pi)^{-3} \\ \int_V \rho \; dV = (2\pi)^{-3} V \end{gathered} \tag{5.16}$$

Yes that's right – there are $(2\pi)^{-3}$ particles per unit volume!

Equation (5.11) refers to scattering into a single final state $|f\rangle$. What does this mean? Since we already are committed to the imaginary box, we can impose periodic boundary condition to discretize the momenta. For

a single particle, quantum theory restricts the number of final states in a volume V with momenta in element d^3p to be $Vd^3p/(2\pi)^3$. Our V has $(2\pi)^{-3}V$ particles, however, so

$$N_f = \frac{Vd^3p}{(2\pi)^3(2\pi)^{-3}V} = d^3p \tag{5.17}$$

In (5.17) N_f counts the number of discrete quantum states available to a particle with momentum in the range between p and $p + d^3p$, where p in this case refers to the three-momentum of the particle. In the general case in which two particles go into N particles with initial and final states given by

$$\begin{aligned} |i\rangle &= |\boldsymbol{k}_2, \boldsymbol{k}_2\rangle \\ |f\rangle &= |\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_N\rangle \end{aligned} \tag{5.18}$$

the corresponding number is

$$N_f = \prod_{j=1}^{N} d^3p_j \tag{5.19}$$

Now turn to the incident flux (5.4). With our norm $n_i = n_t = (2\pi)^{-3}$, and v is the relative velocity of the colliding particles.

$$v = \left| \frac{\boldsymbol{k}_1}{E_1} - \frac{\boldsymbol{k}_2}{E_2} \right| = \frac{1}{E_1E_2}\sqrt{(k_1 \cdot k_2)^2 - (m_1m_2)^2} \tag{5.20}$$

The first expression for v is obviously correct for non-relativistic particles. The second expression is correct relativistically so long as $\mathbf{k}_1$ and $\mathbf{k}_2$ are collinear.

The generalization of (5.2) is

$$d\sigma = \frac{R_{fi}}{I} N_f$$

Using (5.11), (5.19) and (5.8) we get

$$\boxed{d\sigma = \frac{1}{v}(2\pi)^{10} \int \delta^{(4)}(P_f - P_i)|\mathcal{T}_{fi}|^2 \prod_{j=1}^{N} d^3p_j.} \tag{5.21}$$

Although $d\sigma$ is written as if it were a first-order differential, is (5.21) it is really a $3N$-fold differential. For reference we give the same formula using

the $\mathcal{M}$ matrix notation.

$$\boxed{d\sigma = \frac{1}{v}\frac{1}{4E_1E_2}\int \prod_f \frac{d^3p_f}{(2\pi)^3 2E_f}(2\pi)^4\delta^{(4)}(P_f - P_i)|\mathcal{M}_{fi}|^2} \tag{5.22}$$

Finally, we use this formalism to compute the probability of the decay of a single particle. The differential decay rate $d\Gamma$ of an unstable particle is defined by

$$\boxed{\begin{aligned} d\Gamma =& \frac{R_{fi}}{n_t}N_f \\ =& (2\pi)^7 \int \delta^{(4)}(P_f - P_i)|\mathcal{T}_{fi}|^2 \prod_{j=1}^{N} d^3p_j \\ =& \frac{1}{2E_i}\int \prod_f \frac{d^3p_f}{(2\pi)^3 2E_f}(2\pi)^4\delta^{(4)}(P_i - P_f)|\mathcal{M}|^2 \end{aligned}} \tag{5.23}$$

In this case n_t is the number of decaying particles per unit volume, which as usual is equal to $(2\pi)^{-3}$.

5.3 Phase Space

The outcome of a scattering experiment as described by (5.21) depends, generally speaking, on three kinds of things. First there are the fundamental dynamics, which are encoded in the Hamiltonian of the system. Second, there are the details of the experiment itself, e.g. the number of particles in the beam and target, the construction of the detectors, etc. There is a third category of influences that arise because momentum and energy are conserved and because the experiment must be describable in a relativistic invariant way. These last considerations are incorporated in the notion of phase space. Phase space in this context is the set of momentum states available to the particles after the interaction. Imagine leaving the parking lot in front of a movie theater after the show is over. If there is only one narrow road leading out of the lot, then it will take you a long time to escape. If there are freeways leading out in all directions, the lot will empty quickly. This is an analogy based on coordinate space. Try to imagine something like this in momentum space and you will have a feeling for what phase space means. If there are many momentum states available to the final-state particles, the scattering cross section will be large or the decay

will happen quickly. I will explain this again in the next paragraph using a semi-quantitative argument based on the uncertainty principle and finally present a version of the argument that is correct relativistically.

In classical statistic mechanics the state of motion of a particle is represented as a point in a six-dimensional manifold called phase space consisting of three position and three momentum coordinates, i.e. (x, y, z, p_x, p_y, p_z). Classical mechanics places no restrictions on the density of these points, since $\mathbf{x}$ and $\mathbf{p}$ are supposedly measurable with complete precision. Quantum mechanics on the other hand limits the precision by the uncertainty principle,

$$\Delta x_j \Delta p_j \geq 2\pi\hbar.$$

Phase space can be thought of as being subdivided into elementary cells of size $(2\pi\hbar)^3$. The particle's coordinates cannot be localized to any smaller region. The number of states available to one particle will therefore be equal to the total volume of phase space divided by the size of the elementary cell. Assuming that the particle is contained in a volume V, the number of available states is

$$N = \frac{1}{(2\pi\hbar)^3} \int d^3x \; d^3p = \frac{V}{(2\pi\hbar)^3} \int d^3p.$$

The limits of integration come about from the conservation of mass and energy. Since $p_\mu p^\mu = m^2$, the momentum integration is bounded by a sphere of radius $|\mathbf{p}|^2 = E^2 - m^2$, so that N depends on m and E. The factor $V/(2\pi\hbar)^2$ is part of the normalization discussed in the previous section. I am concerned here with the integral over d^3p.

Now suppose there are N such particles in the final state of a scattering experiment.

$$N_N = \int \prod_{j=1}^{N-1} d^3p_j \tag{5.24}$$

There are only $N-1$ integrals in (5.24), because the N $\mathbf{p}_j$'s are not all independent,

$$\sum_{j=1}^{N} \mathbf{p}_j \equiv \mathbf{P}_f = \mathbf{P}_i.$$

We can incorporate this constraint in a more flexible way by introducing a δ-function.

$$\int d^3p_N \; \delta^{(3)} \left(\mathbf{p}_N + \sum_{k=1}^{N-1} \mathbf{p}_k - \mathbf{P}_i \right) = 1$$

$$N_N = \int \prod_{j=1}^{N-1} d^3p_j = \int \prod_{j=1}^{N} d^3p_j \; \delta^{(3)}\left(\sum_{k=1}^{N} \mathbf{p}_k - \mathbf{P}_i\right)$$

Energy is also conserved, and this constraint can be included with a similar trick.

$$\int dE_i \; \delta(E_f - E_i) = 1$$

$$N_N = \int dE_i \prod_{j=1}^{N} d^3p_j \; \delta^{(3)}\left(\sum_{k=1}^{N} \mathbf{p}_k - \mathbf{P}_i\right) \delta\left(\sum_{k=1}^{N} E_k - E_i\right)$$

$$\frac{dN_N}{dE_i} = \int \prod_{j=1}^{N} d^3p_j \; \delta^{(4)}(P_f - P_i) \tag{5.25}$$

These are just the kinematic factors appearing in (5.21). It appears from this heuristic argument, that the reaction rate is proportional to the density of states dN_N/dE_i.[3] This is not the whole story, because (5.25) as it stands is not Lorentz invariant, but according to the Feynman rules from Section 3.5, the square of the scattering amplitude, $|\langle f|\mathcal{T}|i\rangle|^2$ in (5.21), contains a factor $1/2E_i(2\pi)^3$ for each particle in the final state. This is good, since

$$\int d^4p \; \delta(p^2 - m^2) = \int \frac{d^3p}{2E}$$

is clearly a Lorentz scalar. The point is that the factors

$$\left(\prod_{j=1}^{N} \int d^3p_j\right) (2\pi)^4 \delta^{(4)}(P_i - P_f)$$

from (5.21) together with the factors $1/2E_i$ from $|\langle f|\mathcal{T}|i\rangle|^2$ can be combined into a Lorentz-invariant quantity

$$\begin{aligned} \mathcal{P} &= \int \prod_{j=1}^{N} \frac{d^3p_j}{2E_j} \; \delta^{(4)}(P_f - P_i) \\ &= \int \prod_{j=1}^{N} d^4p_j \; \delta(p_j^2 - m^2) \; \delta^{(4)}(P_f - P_i) \end{aligned} \tag{5.26}$$

[3] You will recall from elementary quantum mechanics that the transition probability rate for a system to make a transition from an energy E_l to E_k is given by "Fermi's golden rule,"

$$w_{lk} = \frac{2\pi}{\hbar} g(E_k) |H_{kl}|^2$$

where $g(E_k)$ is the density of states whose energy is close to E_k. See R. L. Liboff, *Introductory Quantum Mechanics*.

The quantity $\mathcal{P}$ is called the Lorentz-invariant phase space. In general we are more interested in integrating some of the momenta in (5.26) over some portion of their allowed kinematic range, in which case

$$d\mathcal{P} = \int_{p\in\Delta} \prod_{j=1}^{N} d^4p_j \delta(p_j^2 - m^2)\delta^{(4)}(P_f - P_i). \tag{5.27}$$

You might say that (5.26) and (5.27) lie on the interface between theory and experiment. For example, if you were using (5.21) to predict the results of an experiment, the limits on the integrals would be set to represent the details of the experimental apparatus. If the apparatus did not detect particle j (it might be a neutral particle), then the corresponding momentum p_j would be integrated out. If the momentum of the k'th particle was measured, then the corresponding integral would not be done, and $\mathcal{P}$ would be left as a function of p_k.

For many-particle reactions, $\mathcal{P}$ will have a complicated dependence on the momentum variables.[4] It is often the case that the data from an experiment are determined mostly by $\mathcal{P}$ regardless of the dynamics. Such an experiment is a failure almost by definition. The general purpose of an experiment is always to find some deviation from phase space.

5.4 Two-Particle Scattering

To make this more concrete, let's calculate the cross section for the elastic scattering of two particles $k_1 + k_2 \to p_1 + p_2$ from the example in the previous chapter. Comparing (3.91) (with $Z = 1$) with (5.5) gives

$$\mathcal{T}_{fi} = \frac{\lambda}{(2\pi)^6\sqrt{2E_{k_1}2E_{k_2}2E_{p_1}2E_{p_2}}} \tag{5.28}$$

Then (5.21) gives

$$\begin{aligned} d\sigma =& \frac{\lambda^2}{4(2\pi)^2 v E_{k_1}E_{k_2}} \int \frac{d^3p_1}{2E_{p_1}} \frac{d^3p_2}{2E_{p_2}} \delta^{(4)}(P_f - P_i) \\ =& \frac{\lambda^2}{4(2\pi)^2 v E_{k_1}E_{k_2}} \int d\mathcal{P} \end{aligned} \tag{5.29}$$

The transition matrix $\mathcal{T}_{fi}$ is especially simple in this example. This is why there is nothing inside the integral except $d\mathcal{P}$. In general there will be some complicated function that we may or may not want to integrate.

[4] There are entire books written on the subject, e.g. *Kinematics and Multiparticle Systems*, M. Nikolic, Gordon and Breach (1968)

I will complete the calculation in the center of mass system assuming equal-mass particles. In this case $|\mathbf{k}_1| = |\mathbf{k_2}| = |\mathbf{p_1}| = |\mathbf{p_2}|$, $\mathbf{k_1} = -\mathbf{k_2}$, $\mathbf{p}_1 = -\mathbf{p}_2$, $E_{p_2} = E_{p_1} \equiv E_1$, and $E_{k_1} + E_{k_2} \equiv E_i$. For the remainder of this section only I will use the notation $|\mathbf{p}_1| = p_1$. Integrating out d^3p_2 gives

$$\mathcal{P} = \int \frac{1}{4E_1^2}\, p_1^2 dp_1 \delta(2E_{p_1} - E_i) d\Omega_1$$

It's convenient to change the variable of integration from p_1 to E_1.

$$p_1 dp_1 = E_1 dE_1 \qquad \delta(2E_1 - E_i) = \frac{1}{2}\delta(E_1 - E_i/2)$$

$$\mathcal{P} = \int \frac{p_1}{8E_1} dE_1 \delta(E_1 - E_i/2) d\Omega_1$$

There is still one δ-function, which will enforce energy conservation and hence fix the value of p_1.

$$\mathcal{P} = \int \frac{p_1}{8E_1}\, d\Omega_1 = \frac{\pi p_1}{4E_1}. \tag{5.30}$$

Notice that symmetry together with the conservation of momentum and energy largely predetermine the outcome of such a scattering experiment. If the particles are spinless there can be no dependence on the azimuthal angle. Only the scattering angle θ_1 is unconstrained, and the only way the underlying dynamics could manifest itself is by affecting the differential cross section as a function of scattering angle. Our example is trivial in the sense that there is no θ dependence, but let's go ahead and calculate the differential cross section anyhow. For this purpose it is best to leave (5.30) in the form of a differential

$$d\mathcal{P} = \frac{p_1}{8E_1}\, d\Omega_1 \tag{5.31}$$

We also need v from (5.20), which in this simple case gives

$$v = \frac{2p_1}{E_1} \tag{5.32}$$

Equation (5.29) becomes

$$\begin{aligned} d\sigma =& \frac{\lambda^2}{4(2\pi)^2 E_1^2}\, \frac{E_1}{2p_1}\, \frac{p_1}{8E_1} d\Omega_1 \\ \frac{d\sigma}{d\Omega_1} =& \frac{\lambda^2}{(16\pi E_1)^2} \end{aligned} \tag{5.33}$$

As anticipated, the scattering is isotropic. There is no θ_1 dependence.

5.5 The General Case

Here is a cookbook procedure for calculating scattering cross sections and decay rates for the general case in which there are N particles in the final state.

- Use Feynman's rules from Section 3.5 to calculate the S matrix.
- Find the transition matrix $\mathcal{T}_{fi}$ from (5.5).
- Find the relative velocity v from (5.20).
- The differential cross sections and decay rates are calculated from (5.21) and (5.23) integrated over the appropriate regions of phase space.

$$d\sigma = \frac{(2\pi)^{10}}{v} \int_{p\in\Delta\Omega} \prod_{j=1}^{N} d^3p_j \delta^{(4)}(P_f - P_i)|\mathcal{T}_{fi}|^2 \tag{5.34}$$

$$d\Gamma = (2\pi)^7 \int_{p\in\Delta\Omega} \prod_{j=1}^{N} d^3p_j \delta^{(4)}(P_f - P_i)|\mathcal{T}_{fi}|^2 \tag{5.35}$$

  When there are only two particles in the final state, one integrates over all the momentum components that are constrained by the delta function and leaves the remaining two angles in the form of a differential. When there are more than two particles in the final state, the decision regarding which variables to integrate and which to leave in the differential depends on the experiment that the theory is describing. That is the meaning of the notation $p \in \Delta\Omega$ that specifies the limits of the integrals

Because of our normalization conventions, $|\mathcal{T}_{fi}|^2$ will contain one factor of $1/(2\pi)^3 2E$ for each external particle in the reaction. The factors of 2π as well as the energies of the *initial-state* particles are constants that are gathered together outside the integrals. The energies of the *final-state* particles are functions of the variables of integration. It is best to think of them as part of the phase space defined by (5.26). Needless to say, these rules are tailored to the Feynman's rules from Section 3.5. If you combine formulas from different books, you are likely to get the wrong answer.

The rules so far only hold for spinless particles. The Dirac wave functions used for spin-$\frac{1}{2}$ scattering are normalized differently. I'll return to that point at the end of Chapter 6.

Chapter 6

The Dirac Equation

There is a story to the effect that Neils Bohr once asked Dirac what he was working on. Dirac, who was famous for not saying much, replied, "I'm trying to take the square root of the Klein-Gordon equation." The background to the story is this: the Klein-Gordon equation is invariant under Lorentz transformations, but it was known to have a strange pathology, it has negative energy solutions. This is inevitable with a second-order equation. The Schrodinger equation, on the other hand, is hopeless so far as relativity is concerned, but because it has only a first derivative with respect to time, it has only positive-energy solutions. (I'm talking about free-particle equations here. Bound state problems usually have negative total energy.) Dirac thought that he could combine the two equations in a way that would preserve the best features of each. The result is called the Dirac equation. Ironically, it has negative-energy solutions anyway, but it does provide a formalism for treating spin 1/2 particles in a way that is consistent with relativity.

6.1 The Equation

Let's follow in the footsteps of our taciturn master and see what we can come up with. Start with the Schrodinger equation.

$$i\frac{\partial\psi}{\partial t} = H\psi$$

Can we come up with a Hamiltonian that (1) is first order in space derivatives, (2) is Hermitian, and (3) leads to a covariant equation of motion? The

answer is *yes if ψ is a matrix.* Try

$$i\frac{\partial\psi}{\partial t} = H\psi = (\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta m)\psi \tag{6.1}$$

here $\boldsymbol{\alpha}$ and β are Hermitian matrices and $\boldsymbol{p} = -i\boldsymbol{\nabla}$. We will also use the replacement $E = i\partial/\partial t$. If E is the total energy of special relativity, it must satisfy $E^2 = p^2 + m^2$. Combine this with equation (6.1)

$$E^2\psi = \left(i\frac{\partial}{\partial t}\right)^2\psi = (\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta m)\left(i\frac{\partial\psi}{\partial t}\right) = (\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta m)^2\psi$$

We would like this to equal $(\boldsymbol{p}^2 + m^2)\psi$. This will be possible if the matrices have some special properties.

$$(\alpha_i p^i + \beta m)^2 = \beta^2 m^2 + (\alpha_i)^2(p^i)^2 + \{\alpha_i, \beta\}\, mp^i + \frac{1}{2}\{\alpha_i, \alpha_j\}_{i\neq j}\, p^i p^j \tag{6.2}$$

I have used the convention that repeated indices are summed. The curly brackets are endemic to Diracology. They represent the anticommutator.

$$\{A, B\} \equiv AB + BA$$

We would like (6.2) to boil down to $p^i p_i + m^2$. This will be possible if these matrices obey the following constraints:

$$\beta^2 = (\alpha_i)^2 = 1 \text{ for each } i \tag{6.3}$$

$$\{\alpha_i, \beta\} = \{\alpha_i, \alpha_j\}_{i\neq j} = 0 \tag{6.4}$$

We are trying to construct a theory of spin 1/2 particles, and we know that in non-relativistic QM Pauli spinors have two components. We expect therefore that α and β will be 2×2 matrices. Alas, it is not to be. They must be at least 4×4. Here is the argument:

1. $\beta\alpha_i + \alpha_i\beta = 0$ so $\beta\alpha_i\beta = -\alpha_i\beta^2 = -\alpha_i$. Take the trace of this equation. $Tr[\beta\alpha_i\beta] = Tr[\alpha_i\beta^2] = Tr[\alpha_i] = -Tr[\alpha_i]$. So α_i is traceless. In the same way we could prove that $Tr[\beta] = 0$.

2. We want α and β to be Hermitian. This together with (6.3) is sufficient to prove that their eigenvalues are ± 1

3. Arguments 1. and 2. lead us to conclude that the matrices must have even dimension.

4. There are only three independent traceless Hermitian 2×2 matrices, and you know what they are: the Pauli spin matrices.

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{6.5}$$

Unfortunately, we need four such matrices, so 2×2 won't do. We must have 4×4. The traditional choice is

$$\beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \qquad \alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \tag{6.6}$$

All this has a nasty, non-relativistic look about it. We therefore abandon Dirac's original notation for the following:[1]

$$\gamma^0 = \beta \qquad \gamma^i = \beta\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \tag{6.7}$$

Now (6.4) becomes

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \tag{6.8}$$

and (6.1) is

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \tag{6.9}$$

This is Dirac's equation. Of course ψ is a four-component column matrix, the so-called Dirac spinor. $g^{\mu\nu}$ is the usual metric tensor. There is one other convenient bit of notation.

$$\gamma_\mu B^\mu = \gamma^\mu B_\mu \equiv \not{B} \tag{6.10}$$

Dirac's equation finally is $(i\ \not{\partial} - m)\psi = 0$. It looks simple. It's not. It is however, covariant. This is far from obvious, but I will not prove it at this time.

Wave functions are supposed to have something to do with probability. This is true of ψ as well, but to see the connection, we need to derive the formula for a conserved current. Starting with (6.1) and taking the Hermitian conjugate, we have

$$-i\frac{\partial \psi^\dagger}{\partial t} = i\frac{\partial \psi^\dagger}{\partial x_i}\alpha_i + \psi^\dagger \beta m \tag{6.11}$$

[1] Some authors use different conventions. This is consistent with Zee, Gross, and many other texts.

Multiply on the right by $\beta^2 = 1$ and define

$$\overline{\psi} \equiv \psi^\dagger \beta = \psi^\dagger \gamma^0 \tag{6.12}$$

This object $\overline{\psi}$ is called the Dirac conjugate. It often appears in bilinear expressions in place of the usual Hermitian conjugate. Anyhow, the resulting equation is

$$-i\frac{\partial \overline{\psi}}{\partial t}\gamma^0 = -i\frac{\partial \overline{\psi}}{\partial x_i}\gamma_i + \overline{\psi} m$$

or

$$\overline{\psi}(i \overleftarrow{\not\partial} + m) = 0 \tag{6.13}$$

The arrow indicates that the derivative acts to the left. To find the current, multiply (6.9) on the left by $\overline{\psi}$ and (6.13) on the right by ψ and add the two. The result is

$$\partial_\mu \overline{\psi}\gamma^\mu \psi + \overline{\psi}\gamma^\mu \partial_\mu \psi = 0 = \partial_\mu(\overline{\psi}\gamma^\mu \psi) \tag{6.14}$$

So the conserved current is

$$j^\mu = \overline{\psi}\gamma^\mu \psi \tag{6.15}$$

The current density $j^0 = \psi^\dagger \psi$ is positive and has the same form $|\psi|^2$ as the probability density in non-relativistic QM.

6.2 Plane Wave Solutions

We know that electrons have two spin components. Why then are there four components to the Dirac spinor? The answer is that despite Dirac's best intentions, there are still negative energy solutions. This in turn is related to the existence of antiparticles. We will get to the particle interpretation later. For the time being we will just reconcile ourselves to negative energies. We start with

$$\psi^\pm = e^{\mp ipx} \begin{pmatrix} \chi \\ \eta \end{pmatrix} \tag{6.16}$$

As a convention, the symbol $E = p^0$ will always refer to a positive number. χ and η are two-component spinors. They are not independent; the Dirac equation imposes a relationship between the two. Our job is to find it. Start with the upper sign (positive energy) solution. Substitute (6.16) into (6.1).

$$\begin{pmatrix} E - m & -\boldsymbol{\sigma}\cdot\boldsymbol{p} \\ -\boldsymbol{\sigma}\cdot\boldsymbol{p} & E + m \end{pmatrix} \begin{pmatrix} \chi \\ \eta \end{pmatrix} = 0$$

This will only have a solution if the determinant is equal to zero, *i.e.* $E^2 - m^2 - (\boldsymbol{\sigma}\cdot\boldsymbol{p})^2 = 0$. Remember that

$$\sigma_i^2 = I \qquad \{\sigma_i, \sigma_j\}_{i\neq j} = 0 \tag{6.17}$$

so that

$$(\boldsymbol{\sigma}\cdot\boldsymbol{p})^2 = \boldsymbol{p}^2, \tag{6.18}$$

which is a useful identity to have at your disposal. We conclude that $E^2 = \boldsymbol{p}^2 + m^2$, which we knew anyhow. Unfortunately, this means that $E = \pm\sqrt{\boldsymbol{p}^2 + m^2}$, so we are stuck with negative energy solutions. At any rate

$$\eta = \left(\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\right)\chi$$

We can make χ into a Pauli spinor that satisfies $\chi^\dagger\chi = 1$.

$$\chi^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \chi^{(2)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{6.19}$$

The complete positive energy solution to the Dirac equation is written

$$\psi^+(x,s) = u(\boldsymbol{p},s)e^{-ipx} \tag{6.20}$$

where

$$u(\boldsymbol{p},s) = C\begin{bmatrix} \chi^{(s)} \\ \frac{\cdot\boldsymbol{p}}{E+m}\chi^{(s)} \end{bmatrix} \qquad s = 1,2 \tag{6.21}$$

Notice that in the non-relativistic limit, the third and fourth components become negligible.

The normalization constant C has to be chosen according to some convention. We will use $\overline{u}(\boldsymbol{p},s)u(\boldsymbol{p},s) = 1$. In order to appreciate this, it is necessary to use one of these group theory results that I have not yet proved. It is a fact that combinations like $\overline{\psi}\psi$ and $\overline{u}u$ are Lorentz scalars, so that if we choose the normalization in the rest frame of the particle, it will have the same value in all other frames. A quick calculation shows that

$$\overline{u}(\boldsymbol{p},s)u(\boldsymbol{p},s) = C^2\frac{2m}{E+m},$$

so our normalized spinor is

$$u(\boldsymbol{p},s) = \sqrt{\frac{E+m}{2m}}\begin{bmatrix} \chi^{(s)} \\ \frac{\cdot\boldsymbol{p}}{E+m}\chi^{(s)} \end{bmatrix} \qquad s = 1,2 \tag{6.22}$$

so that in the rest frame

$$u(0,1) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad u(0,2) = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{6.23}$$

Now back to the negative energy solutions, which I will provisionally call $u(\boldsymbol{p}, 3)$ and $u(\boldsymbol{p}, 4)$. Repeating the above arguments with the lower sign in (6.16) brings us to

$$\chi = -\left(\frac{\boldsymbol{\sigma} \cdot \boldsymbol{p}}{E+m}\right)\eta$$

and

$$u(\boldsymbol{p}, s) = \sqrt{\frac{E+m}{2m}} \begin{bmatrix} -\frac{\boldsymbol{\cdot p}}{E+m}\chi^{(s-2)} \\ \\ \chi^{(s-2)} \end{bmatrix} \qquad s = 3, 4 \tag{6.24}$$

It is easy to see that in the rest frame (and so everywhere else) (6.24) is normalized so that $\overline{u}(\boldsymbol{p}, s)u(\boldsymbol{p}, s) = -1$. The minus sign is forced on us. The complete negative-energy solution is

$$\psi^-(x, s) = u(\boldsymbol{p}, s)e^{ipx} \qquad s = 3, 4 \tag{6.25}$$

6.3 Charge Conjugation and Antiparticles

The appearance of negative-energy states was both an embarrassment and a disappointment; a disappointment because this wasn't supposed to happen with a linear wave equation and an embarrassment because of its physical implications. Let us consider an atomic electron. An excited electron will normally lose energy by emitting photons until it drops down to the lowest-energy or ground state. In the original version of Dirac theory there is no lowest-energy state since there exists an continuum of negative-energy states from $-mc^2$ to $-\infty$. The electron would fall into a bottomless pit emitting an infinite amount of energy in the process! To sidestep this difficulty, Dirac proposed that under normal circumstances, *all negative-energy states are completely filled.* This agglomeration of filled states is called the Dirac sea. Like all of our theories I suppose, it's a half truth. Unfortunately, the other half is very confusing. For one thing, it depends on the Pauli exclusion principle. The states are full because no two Fermions can occupy a single state. This is not true for integer spin particles, and the Klein-Gordon

equation also has negative energy states. Obviously the Dirac sea can't be the whole answer, but it contains a surprising amount of physical insight.[2]

Suppose one of the negative-energy electrons were to absorb a photon of energy $> 2mc^2$ and become an $E > 0$ state. As a result a "hole" is created in the Dirac sea. The *observable* energy of the Dirac sea is now the energy of the vacuum *minus* the *negative* energy of the vacated state, hence a *positive* quantity. In this way we expect that the absence of a negative-energy electron appears as the presence of a positive-energy particle. The same argument works with charge. Let Q_0 represent the charge of the vacuum and e the negative charge of an electron in the sea, then the total charge of the Dirac sea is

$$Q = Q_0 - e = Q_0 - (-|e|) = Q_0 + |e| \tag{6.26}$$

The observable charge is then $Q_{obs} = Q - Q_0 = |e|$. This is interpreted as a particle with positive energy and charge. This process actually happens. In the laboratory it looks like $\gamma \to e^- + e^+$. (It can't happen in empty space because of energy and momentum conservation. It does happen whenever high energy photons pass through matter.) When Dirac first proposed this in 1930, however, no such particle was known. For this reason it was assumed that the theory couldn't be right. The positron was discovered soon thereafter.

According to this theory, the total energy of the vacuum as well as the total charge is negative infinity. We have to take it on faith that these infinities are not observable. What is observable are the deviations from infinity corresponding to the energy and charge of a single electron! This is an unsatisfactory aspect of the theory of course, but it will continue to devil us long after we have abandoned the Dirac sea.

The argument made above regarding charge and energy also works for momentum and spin. The absence of momentum $\boldsymbol{p}$ in the Dirac sea appears as the presence of momentum $-\boldsymbol{p}$. Similarly the absence of a spin-up $E < 0$ electron is to be interpreted as the presence of a spin-down $E > 0$ positron.

This is all well and good, but we must face the fact that the Dirac sea is really just a metaphor. We need to describe antiparticles in a way that does not invoke negative energies. I once heard Edward Teller explain antiparticles like this: "Take a particle in one hand and an antiparticle in the other. Put them together, and what you have is – nothing"! Put less cryptically, all the additive quantum numbers sum to zero. What are the

[2]The best discussion of this entire subject is still the classic, *Advanced Quantum Mechanics*, by J. J. Sakurai, Addison-Wesley 1967, Sections 3.9-3.10

additive quantum numbers? Well, charge, baryon number, three kinds of lepton number, and strangeness for starters. Electrons have zero baryon number and strangeness anyhow. I will come to the issue of lepton number presently. Let's concentrate on the matter of charge. We can define a transformation called "charge conjugation." It has two effects. First, it replaces all particles with their antiparticles, and second, it changes the wave function in such a way that it correctly describes the new situation.

It's easy to modify the Dirac equation to include the electromagnetic potential. The procedure is called "minimal substitution." The rule is everywhere you see the derivative ∂_μ simply replace it with $\partial_\mu \to D_\mu = \partial_\mu + iqA_\mu$ where q is the charge of the particle. This rule is based on the idea of gauge invariance, which we will discuss in the next chapter. For the time being – just believe it. With this replacement the Dirac equation for an electron with charge $-e$ ($e > 0$) becomes

$$(i\ \not{\partial} + e\ \not{A} - m)\psi = 0 \tag{6.27}$$

Charge conjugation must change the charge of the electron so that Teller's dictum comes true. The transformed equation is

$$(i\ \not{\partial} - e\ \not{A} - m)\psi_c = 0. \tag{6.28}$$

Here ψ_c is the new "charge conjugated" wave function. Take the complex conjugate of (6.28) and multiply by C_0, a 4×4 matrix that operates on the space of the γ's.

$$\begin{aligned} C_0(i\gamma^{\mu *}\partial_\mu + e\gamma^{\mu *}A_\mu + m)\psi_c^* \\ = \left[(C_0\gamma^{\mu *}C_0^{-1})(i\partial_\mu + eA_\mu) + m\right] C_0\psi_c^* = 0 \end{aligned} \tag{6.29}$$

At this point we don't know what C_0 is. The challenge is to get (6.28) looking like (6.27). For this we will need $C_0\gamma^{\mu *}C_0^{-1} = -\gamma^\mu$ and $C_0\psi_c^* = \psi$. It doesn't take much trial and error to realize that all the γ's are pure real except γ^2. Try $C_0 = i\gamma^2$. (The i is necessary to make C_0 Hermitian.)[3] You can check that $C_0^2 = I$ and $C_0\gamma^{\mu *}C_0^{-1} = -\gamma^\mu$ as promised. The bottom line is

$$\psi_c = C_0\psi^* = i\gamma^2\psi^* \tag{6.30}$$

Let us see what this operation does to our prefab spinors. The following identities are useful.

$$\boldsymbol{\sigma}\cdot\boldsymbol{p} = \begin{pmatrix} p_z & p_- \\ p_+ & -p_z \end{pmatrix} \tag{6.31}$$

[3]There are several different representations of the gamma matrices that are useful in other contexts. Equation (6.30) is valid only in the the standard or Dirac-Pauli representation we are using.

$$C_0 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \tag{6.32}$$

$$\psi^+(x,1) = u(\boldsymbol{p},1)e^{-ipx} = \sqrt{\frac{E+m}{2m}} \begin{bmatrix} 1 \\ 0 \\ p_z/(E+m) \\ p_+/(E+m) \end{bmatrix} e^{-ipx} \tag{6.33}$$

$$\psi^+(x,2) = u(\boldsymbol{p},2)e^{-ipx} = \sqrt{\frac{E+m}{2m}} \begin{bmatrix} 0 \\ 1 \\ p_-/(E+m) \\ -p_z/(E+m) \end{bmatrix} e^{-ipx} \tag{6.34}$$

$$\psi^-(x,3) = u(\boldsymbol{p},3)e^{ipx} = \sqrt{\frac{E+m}{2m}} \begin{bmatrix} -p_z/(+m) \\ -p_+/(E+m) \\ 1 \\ 0 \end{bmatrix} e^{+ipx} \tag{6.35}$$

$$\psi^-(x,4) = u(\boldsymbol{p},4)e^{ipx} = \sqrt{\frac{E+m}{2m}} \begin{bmatrix} -p_-/(E+m) \\ p_z/(E+m) \\ 0 \\ 1 \end{bmatrix} e^{+ipx} \tag{6.36}$$

Let's find the charge-conjugated version of $\psi^+(x,1)$. Using (6.30), (6.32), and (6.33) gives

$$\begin{aligned} \psi_c^+(x,1) &= \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \sqrt{\frac{E+m}{2m}} \begin{bmatrix} 1 \\ 0 \\ p_z/(E+m) \\ p_+/(E+m) \end{bmatrix}^* e^{+ipx} \\ &= \sqrt{\frac{E+m}{2m}} \begin{bmatrix} -p_-/(E+m) \\ p_z/(E+m) \\ 0 \\ -1 \end{bmatrix} e^{+ipx} = -u(-\boldsymbol{p},4)e^{ipx} \end{aligned} \tag{6.37}$$

In the same way one can prove that

$$\psi_c^+(x,2) = u(-\boldsymbol{p},3)e^{ipx} \tag{6.38}$$

Thus the charge-conjugate wave function obtained from the positive-energy plane-wave solution is the wave function for a negative-energy plane wave

with the same $|E|$ and opposite momentum. The spin direction is also reversed since the spin-up spinor with $s = 1$ is changed into the spin-down spinor with $s = 4$. According to Dirac's hypothesis, the *observable* behavior of the charge-conjugate wave function should be a positron with $E > 0$ and the *same* momentum and spin. This is consequence of the *absence* of an electron with negative charge and energy and reversed momentum and spin.

If you buy into this bizarre notion of the Dirac sea, the rest is logically consistent, but expectation values of the charge-conjugated wave functions do not give the results you would like in terms of antiparticles. Try for example, to calculate the expectation value of the charge of a positron. The electromagnetic current should be proportional to the zeroth component of the probability current (6.15).

$$Q = e \int j^0 d^3x = e \int \psi^\dagger \psi d^3x \tag{6.39}$$

The charge-conjugated wave functions give the same value for Q, since

$$Q_c = e \int \psi_c^\dagger \psi_c d^3x = e \int (C_0 \psi^*)^\dagger (C_0 \psi^*) d^3x = e \int (\psi^\dagger \psi)^* d^3x = Q \tag{6.40}$$

where I have used the fact that C_0 is Hermitian, $C_0^2 = I$, and currents are real functions. It's trivial to show that the expectation values of the momentum and spin do change sign. The reason is that ψ_c is not the positron wave function, it's the wave function that is *interpreted* as the positron wave function. Under the skin, it's still the wave function of a fictitious negative energy particle.[4] In order to get sensible expectation values we need to introduce second quantization with creation operators that create real positrons. In order to do that, it is customary to define spinors for real positrons as follows.

$$\begin{aligned} v(\mathbf{p}, 1) &\equiv -u(-\mathbf{p}, 4) \\ v(\mathbf{p}, 2) &\equiv +u(-\mathbf{p}, 3) \end{aligned} \tag{6.41}$$

With this definition

$$\begin{aligned} u_c(\mathbf{p}, s) &= C_0 u(\mathbf{p}, s)^* = v(\mathbf{p}, s) \qquad s = 1, 2 \\ v_c(\mathbf{p}, s) &= C_0 v(\mathbf{p}, s)^* = u(\mathbf{p}, s) \qquad s = 1, 2 \end{aligned} \tag{6.42}$$

$u(\mathbf{p}, 3)$ and $u(\mathbf{p}, 4)$ are properly called negative-energy spinors, while $v(\mathbf{p}, 1)$ and $v(\mathbf{p}, 2)$ are positron spinors.

[4] Again – thanks to Sakurai for finally making me understand this point.

Here is a summary of the important equations involving u and v:

$$\psi^+(x) = u(\boldsymbol{p}, s)e^{-ipx} \tag{6.43}$$

$$\psi^-(x) = v(\boldsymbol{p}, s)e^{+ipx} \tag{6.44}$$

$$\overline{u}(\boldsymbol{p}, s)u(\boldsymbol{p}, s') = -\overline{v}(\boldsymbol{p}, s)v(\boldsymbol{p}, s') = \delta_{ss'} \tag{6.45}$$

$$u^\dagger(\boldsymbol{p}, s)u(\boldsymbol{p}, s') = v^\dagger(\boldsymbol{p}, s)v(\boldsymbol{p}, s') = \frac{E}{m}\delta_{ss'} \tag{6.46}$$

$$\overline{u}(\boldsymbol{p}, s)v(\boldsymbol{p}, s') = \overline{v}(\boldsymbol{p}, s)u(\boldsymbol{p}, s') = 0 \tag{6.47}$$

$$u^\dagger(-\boldsymbol{p}, s)v(\boldsymbol{p}, s') = v^\dagger(\boldsymbol{p}, s)u(-\boldsymbol{p}, s') = 0 \tag{6.48}$$

$$\sum_s u(\boldsymbol{p}, s)\overline{u}(\boldsymbol{p}, s) = \left(\frac{\not{p} + m}{2m}\right) \tag{6.49}$$

$$\sum_s v(\boldsymbol{p}, s)\overline{v}(\boldsymbol{p}, s) = \left(\frac{\not{p} - m}{2m}\right) \tag{6.50}$$

These equations will all be useful later on. You can verify them by substituting the representations (6.22), (6.24), and (6.7).

6.4 Quantizing the Field

We repeat the procedure from Chapter 5 used to quantize the Kline-Gordon field. First we'll need a Lagrangian, which can be taken as

$$\mathcal{L} = \overline{\psi}(x)(i\ \not{\partial} - m)\psi(x) \tag{6.51}$$

The conjugate momentum is

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\psi^\dagger \tag{6.52}$$

and the Hamiltonian is

$$H = \int d^3x\ (\pi\dot{\psi} - \mathcal{L}) = \int d^3x\ \overline{\psi} i\gamma^0 \partial_0 \psi \tag{6.53}$$

Next we will need the plane wave solutions of the Dirac equation. These are $u(\mathbf{p}, s)e^{-ipx}$ and $v(\mathbf{p}, s)e^{ipx}$. We then expand $\psi(x)$ and $\pi(x)$ in a Fourier series in which creation and annihilation operators appear as Fourier coefficients. This should look like (2.93) with the following differences:

1. ψ is a four-component spinor rather than a scalar like φ.
2. $u(\mathbf{p}, x)$ and $v(\mathbf{p}, s)$ depend on spin, so in addition to integrating over all momentum components, we must also sum over the spin components.
3. We need two kinds of creation and annihilation operators, one set for electrons and one for positrons. In the usual notation

$$\begin{array}{ll} b(\mathbf{p}, s) & \text{Annihilates a positive energy electron} \\ b^\dagger(\mathbf{p}, s) & \text{Creates a positive energy electron} \\ d(\mathbf{p}, s) & \text{Annihilates a positive energy positron} \\ d^\dagger_(\mathbf{p}, s) & \text{Creates a positive energy positron} \end{array} \tag{6.54}$$

The analogs to (2.93) are

$$\begin{aligned} \psi(x) &= \int \sqrt{\frac{m}{E}} \frac{d^3p}{\sqrt{(2\pi)^3}} \sum_s [b(\mathbf{p}, s) u(\mathbf{p}, s) e^{-ipx} + d^\dagger(\mathbf{p}, s) v(\mathbf{p}, s) e^{ipx}] \\ \overline{\psi}(x) &= \int \sqrt{\frac{m}{E}} \frac{d^3p}{\sqrt{(2\pi)^3}} \sum_s [b^\dagger(\mathbf{p}, s) \overline{u}(\mathbf{p}, s) e^{ipx} + d(\mathbf{p}, s) \overline{v}(\mathbf{p}, s) e^{-ipx}] \end{aligned} \tag{6.55}$$

In analogy with (2.102) we would expect the equal-time commutation relations to be

$$[\psi_\alpha(\boldsymbol{x}, t), \psi^\dagger_\beta(\boldsymbol{y}, t)] = \delta_{\alpha\beta} \delta^3(\boldsymbol{x} - \boldsymbol{y})$$

$$[\psi_\alpha(\boldsymbol{x}, t), \psi_\beta(\boldsymbol{y}, t)] = [\psi^\dagger_\alpha(\boldsymbol{x}, t), \psi^\dagger_\beta(\boldsymbol{y}, t)] = 0.$$

Unfortunately, this leads to some wildly unphysical consequences. The origin of the problem is that electrons are fermions, *i.e.* particles with half-odd integer spin. Such particles obey the Fermi exclusion principle, no two particles can occupy the same quantum state. Mathematically, this means that creation and annihilation operators anticommute, and so as a consequence, must the fields themselves. We must have

$$\begin{aligned} \left\{\psi_\alpha(\boldsymbol{x}, t), \psi^\dagger_\beta(\boldsymbol{y}, t)\right\} &= \delta_{\alpha\beta} \delta^3(\boldsymbol{x} - \boldsymbol{y}) \\ \{\psi_\alpha(\boldsymbol{x}, t), \psi_\beta(\boldsymbol{y}, t)\} &= \left\{\psi^\dagger_\alpha(\boldsymbol{x}, t), \psi^\dagger_\beta(\boldsymbol{y}, t)\right\} = 0. \end{aligned} \tag{6.56}$$

Pay careful attention to the notation. $\psi\psi^\dagger$ is a 4×4 matrix, whereas $\psi^\dagger\psi$ is a 1×1 matrix, so $[\psi, \psi^\dagger]$ doesn't make any sense. The commutators in (6.56) refer to the α-th *component* of one ψ matrix with the β-th *component* of another. Put it another way, each commutator is a 4×4 matrix with rows and columns indexed by α and β. Be alert to this ambiguity in other contexts.

You should verify that (6.55) is consistent with (6.56) assuming the following anticommutation relations

$$\left\{b(\mathbf{p},s), b^\dagger(\mathbf{p}',s')\right\} = \delta_{ss'}\delta^3(\boldsymbol{p}-\boldsymbol{p}')$$

$$\left\{d(\mathbf{p},s), d^\dagger(\mathbf{p}'s')\right\} = \delta_{ss'}\delta^3(\boldsymbol{p}-\boldsymbol{p}') \tag{6.57}$$

All other combinations of b and d anticommute. The factor of $\sqrt{m/E}$ in (6.55) was chosen so that (6.56) would come out right.

> There are at least three places where we have to make a somewhat arbitrary choice of normalization. One is the normalization of u and v, equations (6.22) and (6.24). The second is the above-mentioned factor in the Fourier expansion of ψ. The third is the normalization of the creation and annihilation operators, (6.57). Only (6.56) is non-negotiable. It is only a slight exaggeration to say that no two books do it the same way. At least Zee, Michio Kaku (*Quantum Field Theory, A Modern Introduction*), and I are consistent.

It's easy to generalize charge conjugation (6.30) for the quantized fields. Define the charge conjugation *operator* $\hat{C}$ by

$$\hat{\psi}_c \equiv \hat{C}\hat{\psi}\hat{C}^{-1} = C_0\hat{\psi}^{\dagger T} = i\gamma^2\hat{\psi}^{\dagger T} \tag{6.58}$$

The point is that complex conjugation is not defined for the creation and annihilation operators. The combination $^{\dagger T}$ has the same effect as * on matrices and the same effect as † on $\hat{b}$ and $\hat{d}$. (I have put the hats back on the operators for emphasis. They are implied in the other operator equations in this section.) Now using (6.29), (6.55) becomes

$$\begin{aligned}
\psi(x) &= \int \sqrt{\frac{m}{E}}\frac{d^3p}{\sqrt{(2\pi)^3}}\sum_s [d(\mathbf{p},s)u(\mathbf{p},s)e^{-ipx} + b^\dagger(\mathbf{p},s)v(\mathbf{p},s)e^{ipx}] \\
\overline{\psi}(x) &= \int \sqrt{\frac{m}{E}}\frac{d^3p}{\sqrt{(2\pi)^3}}\sum_s [d^\dagger(\mathbf{p},s)\overline{u}(\mathbf{p},s)e^{ipx} + b(\mathbf{p},s)\overline{v}(\mathbf{p},s)e^{-ipx}]
\end{aligned} \tag{6.59}$$

The only change is that particle operators have been replaced by antiparticles operators and vice versa, $b(\mathbf{p},s) \leftrightarrow d(\mathbf{p},s)$ and $b^\dagger(\mathbf{p},s) \leftrightarrow d^\dagger(\mathbf{p},s)$ with *no change in spin*. This point will be central importance in the case of the weak interactions.

Now insert the Fourier decomposition back into the expression for the Hamiltonian, equation (6.53). Again – as an exercise in Diracology, you should verify that the following is obtained.

$$H = \int d^3pE \sum_s [b^\dagger(\mathbf{p},s)b(\mathbf{p},s) - d(\mathbf{p},s)d^\dagger(\mathbf{p},s)] \tag{6.60}$$

Here is the first of many places where commutation relations lead to unphysical results. If d and $d^\dagger$ commuted, the last term could be written $-d^\dagger(p,s)d(p,s)$, so by creating many particles with $d^\dagger$ we could create states of arbitrarily large negative energy. With the anticommutation rules we have

$$:H: = \int d^3pE \sum_s [b^\dagger(\mathbf{p},s)b(\mathbf{p},s) + d^\dagger(\mathbf{p},s)d(\mathbf{p},s)] \tag{6.61}$$

As in (2.78), the Hamiltonian must be normal ordered to avoid the infinite constant. When dealing with fermions, it is necessary to include the minus sign arising from the anticommutators. For example, if a, b, c, and d are fermion annihilation operators,

$$:(a+b^\dagger)(c+d^\dagger): = ac - d^\dagger a + b^\dagger c + b^\dagger d^\dagger \tag{6.62}$$

Take another look at the total charge calculated in (6.39)

$$Q = e\int d^3x\ \psi^\dagger(x)\psi(x)$$

where e is the electron charge, a negative number. The usual Dirac business gives us

$$Q = e\int d^3p \sum_s [b^\dagger(\mathbf{p},s)b(\mathbf{p},s) + d(\mathbf{p},s)d^\dagger(\mathbf{p},s)] \tag{6.63}$$

This operator also has to be normal ordered.

$$:Q: = e\int d^3p \sum_s [N^-(\mathbf{p},s) - N^+(\mathbf{p},s)] \tag{6.64}$$

Where N^+ and N^- are the number operators for positive and negative energy electrons respectively.

6.5 The Lorentz Group

We can't get much further with the Dirac equation without using some group theory. Here is a brief review of the Lorentz group.

The basic equation of Lorentz transformation is

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu} \tag{6.65}$$

$\Lambda^{\mu}{}_{\nu}$ is any 4×4 real matrix that satisfies

$$g_{\mu\nu} = \Lambda^{\sigma}{}_{\mu} g_{\sigma\gamma} \Lambda^{\gamma}{}_{\nu}, \tag{6.66}$$

or in matrix notation

$$g = \Lambda^{T} g \Lambda. \tag{6.67}$$

It is easy to show that this implies $\det \Lambda = \pm 1$. The identity transformation is given by $\Lambda = I$. Any transformation that can be obtained from this by a continuous variation of the transformation parameters (β for example) is called a *proper* transformation. There are three ways that a transformation might be not proper.

1. Try $\Lambda^{0}{}_{0} = -1$, $\Lambda^{i}{}_{i} = 1$, and all off-diagonal terms equal to zero. This simply makes time run backwards, and so it is called the *time reversal transformation* or T for short. It is easy to show from (6.66) that this is improper.

2. $\Lambda^{0}{}_{0} = 1$, $\Lambda^{i}{}_{i} = -1$, and all off-diagonal terms equal to zero. This reverses the spatial coordinates. We call it a *parity transformation* and refer to it as P.

3. We could have a product of the two, TP=PT.

These are examples of discrete transformations, and $T^2 = P^2 = I$. The most general Lorentz transformation consists of a product of a proper transformation and P, T, or PT.[5]

All this is preliminary to analyzing the proper Lorentz group and its associated Lie algebra.[6] Consider the basic Lorentz transformation

$$x' = \gamma(x - vt) \qquad y' = y \qquad t' = \gamma(t - vx/c^2) \tag{6.68}$$

[5]Later on we will use P and T to represent more general operators that have the effect of reversing $\boldsymbol{x}$ or t.

[6]I am following a particularly fine exposition in *Quantum Field Theory* by Michio Kaku.

If we make the standard replacement

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} = \cosh\zeta_1 \qquad \beta\gamma = \sinh\zeta_1 \qquad \beta = v/c \tag{6.69}$$

then this transformation can be written as:

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \cosh\zeta_1 & -\sinh\zeta_1 & 0 & 0 \\ -\sinh\zeta_1 & \cosh\zeta_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{6.70}$$

Transformations of this sort are called *boosts* and the angle ζ_1 that parameterizes the boost in (6.69) and (6.70) is called the *rapidity*. The 4×4 matrix above represents the tensor $\Lambda^\mu{}_\nu$ in (6.65). In matrix notation $x' = \Lambda x$.

According to the doctrine of Lie groups, we can calculate the group generators by differentiating Λ with respect to its parameters and then setting the parameters equal to zero.

$$K^1 \equiv \frac{\partial}{\partial\zeta_1}\Lambda(\zeta_1)\bigg|_{\zeta_1=0} \qquad \text{(provisional definition)}$$

The K's obtained in this way are constant matrices called the "generators" of the transformation. The group can then be "reconstituted" by exponentiating the generators.

$$\Lambda(\zeta_1) = e^{K^1\zeta_1}$$

At this point physicists often muddy the water by inserting i's in ways that appal the mathematicians. We usually write the above as

$$K^1 \equiv -i\,\frac{\partial}{\partial\zeta_1}\Lambda(\zeta_1)\bigg|_{\zeta_1=0} \tag{6.71}$$

and

$$\Lambda(\zeta_1) = e^{iK^1\zeta_1} \tag{6.72}$$

The reason for putting in the i and then taking it out will appear presently. With this convention

$$K^1 = -i\begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{6.73}$$

The generators of boosts in the y and z directions are

$$K^2 = i\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad K^3 = i\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{6.74}$$

Rotations are also Lorentz transformations, or to put it another way, the rotation group is a subgroup of the Lorentz group. The usual parameters are a set of three angles. For example, a rotation through an angle θ_1 around the 1 or x axis would yield

$$\Lambda(\theta_1) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta_1 & \sin\theta_1 \\ 0 & 0 & -\sin\theta_1 & \cos\theta_1 \end{pmatrix} \tag{6.75}$$

The corresponding generators are

$$J^1 = i\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \qquad J^2 = i\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \tag{6.76}$$

$$J^3 = i\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{6.77}$$

Notice that the J^i's are antisymmetric and because of the i's, Hermitian. The corresponding rotation matrices

$$\Gamma(\boldsymbol{\theta}) = e^{i\boldsymbol{J}\cdot} \tag{6.78}$$

are unitary. This is why physicists (who are fond of unitarity) put in the i's and mathematicians (who would like to keep the structure constants real) do not. The boost generators by contrast, are symmetric and boost matrices are not unitary. The generators have the following commutation relations:[7]

$$\begin{aligned} [K^i, K^j] &= -i\epsilon^{ijk}J^k \\ [J^i, J^j] &= i\epsilon^{ijk}J^k \\ [J^i, K^j] &= i\epsilon^{ijk}K^k \end{aligned} \tag{6.79}$$

[7]This is also worked out in Jackson, Sec. 11.7

The matrices above constitute the defining representations of the group and the algebra. That is to say, they operate on four-vectors. We are looking for a representation that transforms Dirac spinors, however, and for this purpose the commutation relations (6.79) must be put in a more convenient form. To this end we start by proving a simple theorem about the symmetry properties of the generators.

Consider infinitesimal Lorentz transformations of the form

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}_{\nu} + \epsilon M^{\mu}{}_{\nu} \tag{6.80}$$

with $M^{\mu}{}_{\nu}$ a constant matrix, *i.e.* a member of the algebra, and ϵ infinitesimal. Substitute (6.80) into (6.66) and discard the second-order terms.

$$M_{\mu\nu} + M_{\nu\mu} = 0 \tag{6.81}$$

This is equivalent in matrix notation to the statement that M (with both indices down) is an antisymmetric matrix. It is easy to check that the maximum number of independent 4×4 antisymmetric matrices is six and that the commutator of two antisymmetric matrices is itself antisymmetric. Thus this set of matrices constitutes a closed Lie algebra, the algebra of the Lorentz group. A particularly useful way of choosing such matrices is

$$(M^{\mu\nu})_{\alpha\beta} = i(\delta^{\mu}_{\alpha}\delta^{\nu}_{\beta} - \delta^{\mu}_{\beta}\delta^{\nu}_{\alpha}) \tag{6.82}$$

Just so there's no confusion, $\mu\nu$ is the *name* of the matrix and α and β index it's elements. It's easy to remember. The matrix named $M^{\mu\nu}$ contains all zeros except for two elements. The element that sits on the μ'th row and ν'th column is i and the element at the ν'th row and μ'th column is $-i$. One can derive the following commutation relations:

$$[M^{\mu\nu}, M^{\rho\sigma}] = i(g^{\nu\rho}M^{\mu\sigma} - g^{\mu\rho}M^{\nu\sigma} - g^{\nu\sigma}M^{\mu\rho} + g^{\mu\sigma}M^{\nu\rho}) \tag{6.83}$$

The M's defined by (6.82) constitute a matrix representation, but any set of objects that satisfy the commutation relations (6.83) also constitutes a valid representation of the Lie algebra. There are two other representations that are important. One is the differential operator

$$M^{\mu\nu} = i(x^{\mu}\partial^{\nu} - x^{\nu}\partial^{\mu}). \tag{6.84}$$

The other is the spinor representation, which I will come to shortly.

Referring again to (6.82), there are twelve non-zero 4×4 antisymmetric matrices related in a simple way to the K's and J's defined previously. It's

easy to verify that $gM^{01} = K^1$, $gM^{12} = J^3$, etc. Infinitesimal transformations are written

$$\Lambda^{\alpha}{}_{\beta}(\omega) = \delta^{\alpha}_{\beta} - \frac{i}{2}\omega_{\mu\nu}(M^{\mu\nu})^{\alpha}{}_{\beta}, \tag{6.85}$$

where $\omega_{\mu\nu}$ is an infinitesimal antisymmetric tensor. We can iterate (6.85) to build up finite transformations as usual.

$$\Lambda(\omega) = \exp\left\{-\frac{i}{2}\omega_{\mu\nu}M^{\mu\nu}\right\} \tag{6.86}$$

6.6 Spinor Representations

The anticommutation relations (6.8) define what is called a Clifford algebra. (This works in any even dimension.) These algebras can always be used to define a new sort of representation. Define

$$\sigma^{\mu\nu} \equiv \frac{i}{2}[\gamma^{\mu}, \gamma^{\nu}] \tag{6.87}$$

It is straightforward to show using (6.8) that the object $\frac{1}{2}\sigma^{\mu\nu}$ satisfies the commutation relations (6.83) and hence constitute a representation of the Lorentz group. The operator

$$S(\omega) \equiv \exp\left\{-\frac{i}{4}\omega_{\mu\nu}\sigma^{\mu\nu}\right\} \tag{6.88}$$

is exactly analogous to (6.86). This can also be written $S(\Lambda)$ to remind ourselves that given the parameters $\omega_{\mu\nu}$, we can calculate the defining representation $\Lambda^{\mu}{}_{\nu}$ as well as the spinor representation S. It is also straightforward to show that

$$[\gamma^{\mu}, \frac{1}{2}\sigma^{\rho\sigma}] = (M^{\rho\sigma})^{\mu}{}_{\nu}\gamma^{\nu} \tag{6.89}$$

where $M^{\rho\sigma}$ is defined by (6.82).

Before we apply all this to the Dirac equation, we should think about what it means to transform a vector field. The Dirac spinor $\psi(x)$, like the EM vector potential $A^{\mu}(x)$, describes some multi-component property (of the ether, if you like) at the space-time point x. I take the passive point of view, that a Lorentz transformed field describes the same phenomena at the same space-time point, but both the point and the components of the field are referred to a different set of axis, so

$$S(\Lambda)\psi(x) = \psi'(x'). \tag{6.90}$$

$x' = \Lambda x$ as usual, and the new field components ψ' will be determined by $S(\Lambda)$. If the Dirac equation is to make any sense, it must remain form invariant under this transformation, so that we must have both

$$(\gamma^\mu p_\mu - m)\psi(x) = 0 \tag{6.91}$$

and

$$(\gamma^\mu p'_\mu - m)\psi'(x') = 0 \tag{6.92}$$

(Here $p_\mu = -i\partial_\mu$.) This requirement leads to an important equation for $S(\Lambda)$. Substitute (6.90) and $p'_\mu = \Lambda_\mu{}^\nu p_\nu$ into (6.92).

$$S^{-1}(\Lambda)(\gamma^\mu \Lambda_\mu{}^\nu p_\nu - m)S(\Lambda)\psi(x) = 0$$

This will be identical to (6.91) if

$$S^{-1}(\Lambda)\gamma^\mu S(\Lambda) = \Lambda^\mu{}_\nu \gamma^\nu \tag{6.93}$$

This is interesting. The left side is a similarity transformation. This is how one transforms a matrix like γ^μ. The right side treats γ^μ like a four-vector. So is γ^μ a matrix like $S(\Lambda)$ or a four-vector like p^μ? The answer of course, is yes!

Or at least we hope so. The covariance of the Dirac equation stands or falls depending on (6.93). Here its proof: The first step is to prove that (6.93) is true for infinitesimal transformations

$$\begin{aligned}(1 + \frac{i}{4}\omega_{\rho\sigma}\sigma^{\rho\sigma})\gamma^\mu(1 - \frac{i}{4}\omega_{\rho\sigma}\sigma^{\rho\sigma}) &= (\gamma^\mu - \frac{i}{2}\omega_{\rho\sigma}[\gamma^\mu, \frac{1}{2}\sigma^{\rho\sigma}]) \\ &= (1 - \frac{i}{2}\omega_{\rho\sigma}M^{\rho\sigma})^\mu{}_\nu \gamma^\nu\end{aligned} \tag{6.94}$$

The last line makes use of (6.89). The left side of (6.94) is the infinitesimal form of (6.93) by construction. A glance at (6.86) shows that the last line of (6.94) is also the infinitesimal form of the right side of (6.93). Thus (6.93) is valid for infinitesimal transformations. It can be proved for finite transformations by iterating infinitesimal transformations. It's not a new argument, but it's very tedious to write out. I'll leave it as an exercise.

We are now in a position to prove the claim made earlier that $\overline{\psi}\psi$ is a Lorentz scalar. First take the Hermitian conjugate of (6.90) and right multiply by γ^0.

$$\psi'^\dagger(x')\gamma^0 = \psi^\dagger(x)\gamma^0\gamma^0 S^\dagger(\Lambda)\gamma^0$$

Remember that $(\gamma^0)^2 = 1$ and $(\gamma^0)^\dagger = \gamma^0$.

$$\overline{\psi}'(x') = \overline{\psi}(x)\gamma^0 S(\Lambda)^\dagger \gamma^0$$

$$= \overline{\psi}(x)S^{-1}(\Lambda) \tag{6.95}$$

The last line follows from (6.87) and (6.88). Evidentally, $\overline{\psi}'(x')\psi'(x') = \overline{\psi}(x)\psi(x)$. By the same token, $\overline{\psi}\gamma^\mu\psi$ is a genuine vector.

$$\overline{\psi}'(x')\gamma^\mu\psi'(x') = \overline{\psi}(x)S^{-1}\gamma^\mu S\psi = \overline{\psi}(x)\Lambda^\mu{}_\nu\gamma^\nu\psi(x) \tag{6.96}$$

The last step used (6.93).

Expressions like $\overline{\psi}\psi$ and $\overline{\psi}\gamma^\mu\psi$ are called Dirac bilinear forms. Notice that unlike ψ and γ^μ, they are not matrices but ordinary tensors. We can also show that $\overline{\psi}\sigma^{\mu\nu}\psi$ is a second-rank antisymmetric tensor. Since these forms are essential in describing interactions, we should enumerate the various possibilities. We need to define a new gamma matrix:

$$\gamma^5 = \gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{6.97}$$

The following properties can be verified with the help of the anticommutation relations.

$$(\gamma^5)^\dagger = \gamma^5 \tag{6.98}$$

$$(\gamma^5)^2 = 1 \tag{6.99}$$

$$\left\{\gamma^5, \gamma^\mu\right\} = 0 \tag{6.100}$$

With our conventions

$$\gamma^5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{6.101}$$

There are sixteen numbers in a 4×4 matrix, so there should be sixteen linearly independent 4×4 matrices with well-defined transformation properties. Here they are,

1	scalar	1
γ^μ	vector	4
$\sigma^{\mu\nu}$	tensor	6
$\gamma^\mu\gamma^5$	pseudo-vector	4
γ^5	pseudo-scalar	1

for a total of sixteen. The terms *pseudo-tensor* and *pseudo-scalar* are "pseudo" in the sense that they change sign under parity conjugation. I will return to this important subject later. One can show that these sixteen matrices are indeed linearly independent. Any product of gamma matrices can be reduced to a linear combination of these terms.

6.7 The Dirac Propagator

Consider a physical process in which an electron is created out of the vacuum at the point x and is subsequently reabsorbed at x'. This emission and reabsorption might take place because of sources, which we discussed in connection with path integrals, or it might come about because the electron interacted with other particles at x and x'. In any case, the amplitude for this process is called the propagator. It is the basic ingredient in all perturbation theory calculations. I have argued before that the field $\overline{\psi}(x)$ creates an electron at x and ψ destroys it, so the product $\langle 0|\psi(x')\overline{\psi}(x)|0\rangle$ must have something to do with the propagator. This can't be the whole story, however, because we cannot annihilate the electron before it has been created. True, simultaneity is relative, but it is logically impossible in any reference frame to make something disappear that doesn't exist.[8] Therefore, we should write out propagator as $\langle 0|\psi(x')\overline{\psi}(x)|0\rangle\theta(t'-t)$. This looks non-covariant, because or the $\theta(t'-t)$. The argument can be made in a completely covariant way, but this is just pedantry. We get a covariant result in the end anyway.

This is still not the complete propagator, because there is a distinct process that is physically equivalent. $\overline{\psi}$ doesn't just create electrons, *it reduces the charge*. So $\langle 0|\psi(x')\overline{\psi}(x)|0\rangle$ also include the process in which negative energy electrons travel backwards in time from x to x', or equivalently, positrons travel from x' to x. We must still insist that the positron be created before it is destroyed, so finally

$$iS_F(x',x) = \langle 0|\psi(x')\overline{\psi}(x)|0\rangle\theta(t'-t) - \langle 0|\overline{\psi}(x)\psi(x')|0\rangle\theta(t-t') \qquad (6.102)$$

So why subtract the two terms rather than adding them? The glib answer is that the spinor fields anticommute. We will get more insight into this presently. Just as we did in the case of the Kline-Gordon equation, eq. (5.33), we introduce the time ordering operator

$$T[\psi(x')\overline{\psi}(x)] = \begin{cases} \psi(x')\overline{\psi}(x) & t' > t \\ -\overline{\psi}(x)\psi(x') & t > t' \end{cases} \qquad (6.103)$$

Finally

$$i(S_F)_{\alpha\beta}(x',x) = \langle 0|T[\psi_\alpha(x')\overline{\psi}_\beta(x)]|0\rangle \qquad (6.104)$$

S_F is a 4×4 matrix indexed with α and β. It's easy to calculate. Substitute (6.55) into (6.102) and use (6.49) and (6.50) to do the spin sums. Zee

[8] "Yesterday upon a stair, / I met a man who wasn't there. / He wasn't there again today. / I wish that man would go away."

goes through this calculation in detail on pages 108-109. He arrives at the following:

$$iS(x) = i \int \frac{d^4p}{(2\pi)^4}\, e^{-ipx} \frac{\not{p} + m}{p^2 - m^2 + i\epsilon} \tag{6.105}$$

It turns out that $(\not{p} - m)(\not{p} + m) = p^2 - m^2$. We could say that the factors of $(\not{p} + m)$ in the numerator and denominator "cancel" leaving the rather whimsical equation,[9]

$$iS(x) = \int \frac{d^4p}{(2\pi)^4}\, e^{-ipx} \frac{i}{\not{p} - m + i\epsilon}. \tag{6.106}$$

In momentum space this is

$$iS(p) = \frac{i}{\not{p} - m + i\epsilon} \tag{6.107}$$

Causality requires that the propagator vanishes outside the light cone. You can show that (6.105) does have that property. It also reproduces the minus sign in (6.102). It can also be shown[10] that without that minus sign, the propagator *violates* causality. This is an illustration of a deep result, called the spin-statistics theorem. It says that in order to preserve causality, Fermion fields must anticommute, and boson fields must commute.

At this point in our treatment of the scalar field, we added some interactions and did perturbation theory, but electrons don't interact with themselves like our hypothetical scalar particle. They do interact with the electromagnetic vector potential and with the potential that carries the weak interactions. (The're the same field actually, but that's a story for another time.) Our next order of business is the electromagnetic field, to which we turn in the next chapter.

[9] Your instructor will become homicidal if you divide by a matrix under any other circumstances.

[10] Peskin and Schroeder, *An Introduction to Quantum Field Theory* page 56

Chapter 7

The Photon Field

7.1 Maxwell's Equations

In some sense Maxwell discovered both quantum electrodynamics and relativity. I mean by this odd remark that Maxwell's equations are consistent with both. Nothing about them needs to be changed. Here they are:

$$\begin{aligned} \boldsymbol{\nabla} \cdot \boldsymbol{E} &= \rho \\ \boldsymbol{\nabla} \times \boldsymbol{B} - \frac{\partial \boldsymbol{E}}{\partial t} &= \boldsymbol{j} \\ \boldsymbol{\nabla} \cdot \boldsymbol{B} &= 0 \\ \boldsymbol{\nabla} \times \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} &= 0 \end{aligned} \tag{7.1}$$

You may never have seen them in this simple form. It arises because: (1) We use Heaviside-Lorentz units. (See the appendix in Jackson regarding units.) In this system, all the factors of 4π are absent from Maxwell's equation. (They retaliate by showing up in Coulomb's law, but we don't care about that.) There are no dielectrics or magnetic materials, so $\epsilon = \mu = 1$. (3) We set $c = 1$ as usual. The electromagnetic current is conserved.

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{j} = 0 \tag{7.2}$$

The vector fields can be derived from potentials.

$$\boldsymbol{E} = -\boldsymbol{\nabla} A^0 - \frac{\partial \boldsymbol{A}}{\partial t} \qquad \boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A} \tag{7.3}$$

This can all be put in covariant notation by defining the following tensors:

$$\begin{aligned} A^\mu &= (A^0, \boldsymbol{A}) \\ j^\mu &= (\rho, \boldsymbol{j}) \\ F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu \end{aligned} \tag{7.4}$$

Then current conservation is just

$$\partial_\mu j^\mu = 0 \tag{7.5}$$

and Maxwell's equations can be summarized as:

$$\begin{aligned} \partial_\mu F^{\mu\nu} &= j^\nu \\ \partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} &= 0 \end{aligned} \tag{7.6}$$

The fact that these can be written in covariant form indicates that relativity is "built in" to Maxwell's equations.

We need a Lagrangian to do field theory, Jackson has some obscure things to say about Lagrangians in general, but if you know the answer ahead of time, it's really very easy.

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - j_\mu A^\mu \tag{7.7}$$

To show that this does indeed generate the right equations of motion, simplify the expression as follows:

$$\begin{aligned} -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} &= -\frac{1}{4} (\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) \\ &= -\frac{1}{2} [(\partial_\mu A_\nu)(\partial^\mu A^\nu) - (\partial_\mu A_\nu)(\partial^\nu A^\mu)] \\ \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} &= -\partial^\mu A^\nu + \partial^\nu A^\mu = -F^{\mu\nu} \end{aligned} \tag{7.8}$$

The Euler-Lagrange equations are

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \right) - \frac{\partial \mathcal{L}}{\partial A_\nu} = -\partial_\mu F^{\mu\nu} + j^\nu = 0,$$

which is exactly (7.6).

The potential contains some unphysical information in the sense that Maxwell's equations are invariant under gauge transformations. The familiar form for gauge transformation is

$$\mathbf{A} \to \mathbf{A} + \nabla \Lambda$$

$$\phi \to \phi - \frac{\partial \Lambda}{\partial t}$$

where $\Lambda(x)$ is any arbitrary function of space and time. We can write this in covariant form as follows.

$$A^\mu(x) \to A^\mu(x) - \partial^\mu \Lambda(x) \tag{7.9}$$

This leaves the field tensor invariant $F^{\mu\nu}(x) \to F^{\mu\nu}(x)$, and hence preserves Maxwell's equations. We can use this symmetry to make A^μ satisfy certain relations.[1] For example, we can choose Λ so that $A^0(x) = 0$. Imposing constraints like this is called "gauge fixing." This particular constraint defines what is called temporal gauge. It is possible to impose an additional constraint

$$\boldsymbol{\nabla} \cdot \boldsymbol{A} = 0 \tag{7.10}$$

This is known as Coulomb gauge or radiation gauge. Another choice is Lorentz gauge.

$$\partial_\mu A^\mu = 0 \tag{7.11}$$

This has the advantage of being covariant, but it does not fully fix the gauge. This poses a dilemma: we can't fully fix the gauge without making the equations non-covariant. We can't make the theory covariant without building into the theory unphysical degrees of freedom known as Fadeev-Popov ghosts.[2] The subject is so difficult and convoluted that one distinguished text[3] simply refuses to discuss it!

Components of A^μ that can be set equal to zero by a gauge transformation cannot have any physical significance. It's a well-known result from classical field theory that any vector can be decomposed into a longitudinal and a transverse part.

$$\boldsymbol{A} = \boldsymbol{A}_T + \boldsymbol{A}_L$$

where by definition

$$\boldsymbol{\nabla} \times \boldsymbol{A}_L = 0 \qquad \boldsymbol{\nabla} \cdot \boldsymbol{A}_T = 0$$

Furthermore, the decomposition is unique up to an additive constant. Since F^{ij} is basically the curl of A, it cannot depend on the longitudinal components. Only the transverse components are meaningful. There are only two

[1] This subject is discussed extensively in Jackson and our Ph632 class. Here I am quoting the main results without proof.

[2] Where is Buffy now that we need her?

[3] Peskin and Schroder

physical degrees of freedom, which correspond to the two polarization states of an electromagnetic wave.

Perhaps the easiest way out of this mess is to quantize in Coulomb gauge, though even this is not straightforward. This is the approach taken in the previous generation of texts.[4] Modern texts do this with path integrals. It's ironic that the photon propagator, which is all you need to do perturbation theory in QED, is very simple. This is my excuse for spending minimal effort on this vexatious subject.

7.2 Quantization in the Coulomb Gauge

We have two goals.[5] First, to find an expansion of the photon field in terms of creation and annihilation operators, and second, find the photon propagator. The Lagrangian density for the free fields is

$$\mathcal{L}(x) = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} = \frac{1}{2}(\boldsymbol{E}^2 - \boldsymbol{B}^2). \tag{7.12}$$

The conjugate momenta to A^μ are,

$$\pi_0(x) = \frac{\partial\mathcal{L}}{\partial(\partial_0 A^0)} = 0$$

$$\pi_i(x) = \frac{\partial\mathcal{L}}{\partial(\partial_0 A^i)} = -E_i \tag{7.13}$$

and the Hamiltonian density is

$$\mathcal{H} = \pi_\mu \dot{A}^\mu - \mathcal{L} = \frac{1}{2}(\boldsymbol{E}^2 + \boldsymbol{B}^2) + \boldsymbol{E}\cdot\boldsymbol{\nabla}A^0 \tag{7.14}$$

The next step is to impose the equal-time commutation relations. (Compare with (73) in the previous chapter.)

$$[A_\alpha(\boldsymbol{x},t), \pi_\beta(\boldsymbol{y},t)] = i\delta_{\alpha\beta}\delta(\boldsymbol{x}-\boldsymbol{y}) \tag{7.15}$$

$$[A_\alpha(\boldsymbol{x},t), A_\beta(\boldsymbol{y},t)] = [\pi_\alpha(\boldsymbol{x},t), \pi_\beta(\boldsymbol{y},t)] = 0$$

We are missing one relation, since $\pi_0 = 0$. This is an invitation to duck into coulomb gauge and set $A^0 = \boldsymbol{\nabla}\cdot\boldsymbol{A} = 0$. We are not out of the woods yet,

[4] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics and Relativistic Quantum Fields* is the Gold Standard.

[5] I am following B. Hatfield, *Quantum Field Theory of Point Particles and Strings.*

however. Eq. (7.3) gives $\boldsymbol{\nabla} \cdot \boldsymbol{E} = 0$. This is Gauss's law; so far so good. If we take the divergence of (7.15)

$$[A_i(\boldsymbol{x}, t), E_j(\boldsymbol{y}, t)] = -i\delta_{ij}\delta(\boldsymbol{x} - \boldsymbol{y}), \tag{7.16}$$

we get the odd result

$$0 = -i\partial_i\delta(\boldsymbol{x} - \boldsymbol{y})$$

This is time for some creative thinking. Suppose we modify (7.15) slightly so that we avoid this contradiction. Lets replace the δ-function with a function $f(\boldsymbol{x} - \boldsymbol{y})$ such that $\partial_i f(\boldsymbol{x} - \boldsymbol{y}) = 0$. The usual choice is the transverse δ-function,

$$\delta^{\rm tr}_{ij}(\boldsymbol{x} - \boldsymbol{y}) \equiv \int \frac{d^4k}{(2\pi)^3} e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{y})} \left(\delta_{ij} - \frac{k_i k_j}{\boldsymbol{k}^2}\right). \tag{7.17}$$

$\delta^{\rm tr}_{ij}(\boldsymbol{x})$ has the desired property: $\partial_i\delta^{\rm tr}_{ij}(\boldsymbol{x}) = 0$. We modify (16) to read

$$[A_i(\boldsymbol{x}, t), E_j(\boldsymbol{y}, t)] = -i\delta^{\rm tr}_{ij}\delta(\boldsymbol{x} - \boldsymbol{y}), \tag{7.18}$$

The following comments are in order:

- $\delta^{\rm tr}_{ij}(\boldsymbol{x}) \neq 0$ so A_i and E_i don't commute at space-like separations. This in itself is OK since $\boldsymbol{A}$ is not measurable.
- Electric and magnetic fields calculated using (7.18) do commute at space-like separations, and in fact, we get the same commutators for $\boldsymbol{E}$ and $\boldsymbol{B}$ regardless whether we use the transverse or ordinary δ-function in (7.18).
- The underlying idea here is that there are only two degrees of freedom, but (7.16) tries to quantize three. The transverse δ-function remedies this problem.

Now we are home safe. It remains only to expand $\boldsymbol{A}$ in plane waves and calculate the propagator.

$$\boldsymbol{A}(x) = \int \frac{d^3k}{\sqrt{(2\pi)^3 2k_0}} \sum_{\lambda=1}^{2} \boldsymbol{\varepsilon}^\lambda(k) \left[a^\lambda(k)e^{-ikx} + a^{\lambda\dagger}(k)e^{ikx}\right] \tag{7.19}$$

Since there is no antiphoton (or more correctly, the photon is its own antiparticle) there is only one set of creation and destruction operators. Since we are working in transverse gauge, $\boldsymbol{A}$ must be perpendicular to the direction

of propagation. This is accomplished by defining the photon polarization vectors $\boldsymbol{\varepsilon}^\lambda$ such that

$$\boldsymbol{\varepsilon}^\lambda \cdot \boldsymbol{k} = 0 \tag{7.20}$$

We also require that they be orthonormal.

$$\boldsymbol{\varepsilon}^\lambda(k) \cdot \boldsymbol{\varepsilon}^{\lambda'}(k) = \delta^{\lambda\lambda'} \tag{7.21}$$

Equation (7.18) will be satisfied if

$$[a^\lambda(k), a^{\lambda\dagger}(k')] = \delta^{\lambda\lambda'}\delta^{(3)}(\boldsymbol{k} - \boldsymbol{k'}) \tag{7.22}$$

Our last mission is to calculate the propagator. The process should be quite familiar by now.

$$\begin{aligned} iD_F(x'-x)_{\mu\nu} &= \langle 0|T[A_\mu(x')A_\nu(x)]|0\rangle \\ &= i\int \frac{d^4k}{(2\pi)^4}\frac{e^{-ik(x'-x)}}{k^2+i\epsilon}\sum_{\lambda=1}^{2}\boldsymbol{\varepsilon}_\mu^\lambda(k)\boldsymbol{\varepsilon}_\nu^\lambda(k) \end{aligned} \tag{7.23}$$

The polarization tensors are not in a convenient form, and since they refer to a particular frame of reference, they are not covariant. That is the price of working in Coulomb gauge. Here is an argument that addresses both issues. We construct a set of orthonormal vectors based on the coordinate system in which the polarization vectors are defined. In this frame they are perpendicular to $\boldsymbol{k}$, and since they have no time component $\varepsilon_\mu k^\mu = 0$ (regardless of the time component of k). We introduce a time-like unit vector $\eta = (\ 1 \quad 0 \quad 0 \quad 0\ )$, $\eta^\mu\varepsilon_\mu = 0$, but $\eta_\mu k^\mu \neq 0$. We complete our set of four unit vectors by projecting out the part of k^μ along η^μ.

$$\hat{k}^\mu = \frac{k^\mu - (k\cdot\eta)\eta^\mu}{\sqrt{(k\cdot\eta)^2 - k^2}} \tag{7.24}$$

The four unit vectors then are η, ε^1, ε^2, and $\hat{k}$. If you think of them as column matrices,

$$\eta\eta^\dagger - \varepsilon^1\varepsilon^{1\dagger} - \varepsilon^2\varepsilon^{2\dagger} - \hat{k}\hat{k}^\dagger = g,$$

where g is the usual metric tensor. Writing this as a tensor equation gives,

$$\sum_{\lambda=1}^{2}\varepsilon_\mu^\lambda(k)\varepsilon_\nu^\lambda(k) = -g_{\mu\nu} + \eta_\mu\eta_\nu - \hat{k}_\mu\hat{k}_\nu. \tag{7.25}$$

The usual four-vector argument goes like this: equation (7.25) is valid in one particular reference frame. $g_{\mu\nu}$ and $\hat{k}_\mu\hat{k}_\nu$ are legitimate four-tensors. *If* $\eta_\mu\eta_\nu$

were also, then (24) would be valid in all reference frames. Unfortunately, it isn't, so it's not. Let's go ahead and substitute (7.24) into (7.25) anyhow.

$$\sum_{\lambda=1}^{2} \varepsilon_\mu^\lambda(k)\varepsilon_\nu^\lambda(k) = -g_{\mu\nu} - \frac{k_\mu k_\nu}{(k\cdot\eta)^2 - k^2} + \frac{(k\cdot\eta)(k_\mu\eta_\nu + k_\nu\eta_\mu)}{(k\cdot\eta)^2 - k^2} - \frac{k^2\eta_\mu\eta_\nu}{(k\cdot\eta)^2 - k^2} \tag{7.26}$$

This is not covariant, but when the photon propagator is used in a actual calculation it is always coupled to a conserved current. Eq. (7.5) written in momentum space is $k_\mu j^\mu = 0$. All the terms above with uncontracted k's vanish! The last term is something of an embarrassment. When substituted into (7.23) it gives the term

$$-\eta_\mu\eta_\nu \frac{\delta(t-t')}{4\pi|\boldsymbol{x}-\boldsymbol{x}'|}$$

This is the instantaneous Coulomb interaction. It is a well-known artifact of Coulomb gauge in classical E&M.[6] It can't be right, of course. Even the Coulomb interaction doesn't propagate instantaneously through space. In this case it is cancelled by another spurious term in the Hamiltonian.[7] The remaining term in (7.26) gives the correct propagator,

$$D_F(x'-x)_{\mu\nu} = -g_{\mu\nu}\int \frac{d^4k}{(2\pi)^4}\frac{e^{-ik(x'-x)}}{k^2+i\epsilon}. \tag{7.27}$$

In momentum space, this is simply,

$$D_F(k)_{\mu\nu} = -\frac{g_{\mu\nu}}{k^2+i\epsilon} \tag{7.28}$$

This argument guarantees that so long as we use (7.27) or (7.28) in our calculations, the resulting S matrix elements and Green's functions will have the correct covariant behavior, even though the quantization was done in Coulomb gauge.

[6] See Jackson, Sec. 6.5 and references therein.

[7] See M. Kaku, *Quantum Field Theory, Section 4.3*

Chapter 8

Quantum Electrodynamics

8.1 Gauge Invariance

Before we can do any realistic calculations, we must figure out how the electromagnetic field couples to electrons. We could guess this as follows: we know that the classical electromagnetic Lagrangian has the term $-j^\mu A_\mu$, so we assume that the interaction Hamiltonian must be of the form $j^\mu A_\mu$. Furthermore, the electron probability current is $j^\mu = \bar{\psi}\gamma^\mu\psi$, so the charge current must be $e\bar{\psi}\gamma^\mu\psi$. We conclude that $\mathcal{H}_{\text{int}} = e\bar{\psi}\gamma^\mu\psi A_\mu$. This is true, but there is a profound principle here that we should linger over.

The free-particle Dirac Lagrangian is $\mathcal{L} = \bar{\psi}(i\not{\partial} - m)\psi$. If we make the transformation $\psi \to e^{iq\theta}\psi$, then $\mathcal{L} \to \mathcal{L}$. This is called a "global gauge transformation." It's not what we usually think of as a gauge transformation in E&M, but they are related as you will see. Try an infinitesimal transformation $\delta\psi = iq\delta\theta\psi$. If the Lagrangian is to remain invariant under this transformation we must have

$$\delta\mathcal{L} = 0 = \frac{\partial\mathcal{L}}{\partial\psi}\delta\psi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\delta(\partial_\mu\psi) \tag{8.1}$$

$$= iq\delta\theta\left[\frac{\partial\mathcal{L}}{\partial\psi}\psi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\partial_\mu\psi\right]$$

Here θ is an infinitesimal and q is something like a charge. The Euler-Lagrange equation of motion is

$$\frac{\partial\mathcal{L}}{\partial\psi} = \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)} \tag{8.2}$$

Substituting (8.2) into (8.1) gives

$$iq\delta\theta\partial_\mu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\psi\right]=0 \tag{8.3}$$

Evidentally, the quantity in brackets is a conserved current.

$$j^\mu\propto\frac{\delta\mathcal{L}}{\delta(\partial_\mu\psi)}\psi \tag{8.4}$$

and $\partial_\mu j^\mu=0$. Applying this formula to the Dirac Lagrangian gives $j_\mu=\bar{\psi}\gamma^\mu\psi$, which we knew before.

There is a stronger statement to be made if we make a different transformation at each point in space, *i.e.* if we make a local transformation.

$$\psi'=e^{iq\theta(x)}\psi. \tag{8.5}$$

The function $\theta(x)$ is arbitrary. Now

$$\partial_\mu\psi'=e^{iq\theta}\partial_\mu\psi+iq(\partial_\mu\theta)e^{iq\theta}\psi$$

Now $\mathcal{L}$ is not invariant by itself, but it can be made invariant by incorporating an additional field with special transformation properties. Let's insist that we replace $\partial_\mu\psi$ everywhere with $(\partial_\mu-ieqA_\mu)\psi$. If the following statement is true, the electron Lagrangian will be invariant:

$$(\partial_\mu-ieqA'_\mu)\psi'=e^{iq\theta}(\partial_\mu-ieqA_\mu)\psi,$$

but

$$(\partial_\mu-ieqA'_\mu)\psi'=e^{iq\theta}\partial_\mu\psi+iq(\partial_\mu\theta)e^{iq\theta}\psi-ieqA'_\mu e^{iq\theta}\psi$$

We conclude that

$$A'_\mu=\frac{1}{e}\partial_\mu\theta+A_\mu. \tag{8.6}$$

In other words, the electron Lagrangian will be invariant under the transformation (8.5) if our new fields transform according to (8.6), but these new fields must also appear in our Lagrangian, and this Lagrangian must be invariant under the transformation

$$A_\mu\rightarrow A_\mu+\frac{1}{e}\,\partial_\mu\theta(x). \tag{8.7}$$

Equation (8.7) is of course, the usual statement of gauge invariance in classical E&M, but now we see it in a larger context; quantum gauge invariance, (8.5), *requires the existence on an additional vector field.*

We know that the Lagrangian for classical E&M (excluding sources) is

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

where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

Transformation (8.7) does leave this invariant, but if this were the Lagrangian for a massive field, it would contain an additional term

$$m^2 A^{\dagger\mu} A_\mu$$

which would not be invariant. We conclude that local gauge invariance requires a *massless* vector field.

You will notice that the "charge" q cancels in the above calculation. As a consequence the sign and magnitude of e are undetermined. As a convention we usually write

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - e\gamma^\mu A^\mu - m)\psi \tag{8.8}$$

so that

$$\mathcal{H}_{\text{int}} = -\mathcal{L}_{\text{int}} = e\bar{\psi}\gamma^\mu \psi A_\mu = j^\mu A_\mu \tag{8.9}$$

8.2 Noether's Theorem

Zee calls Noether's theorem, "one of the most profound observations in theoretical physics." The theorem simply states that every continuous symmetry transformation is connected with a conserved current and hence, with some sort of conserved "charge." The converse is also true, for every conserved current, there is a corresponding symmetry group. The proof is a simple generalization to that given above. Given the Lagrangian density composed of N fields φ_i and their derivatives, $\mathcal{L} = \mathcal{L}(\varphi_i, \partial_\mu \varphi_i)$. Consider an infinitesimal change, $\delta\varphi_i$, brought about by some small parameter or parameters $\delta\epsilon_j$, and calculate the action.

$$\delta S = \sum_{i=1}^{N} \int d^4x \left[\frac{\partial \mathcal{L}}{\partial \varphi_i} \, \delta\varphi_i + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_i)} \partial_\mu \delta\varphi_i \right] \tag{8.10}$$

(The repeated μ indices are summed.) We can use the N Euler-Lagrange equations of motion

$$0 = \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_i)} \right] - \frac{\partial \mathcal{L}}{\partial \varphi_i}$$

to combine the two terms in (8.10).

$$\delta S = \sum_{i,j=1}^{N} \int d^4x \; \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_i)} \frac{\partial \varphi_i}{\partial \epsilon_j} \delta \epsilon_j \right] \tag{8.11}$$

If the action is unchanged by this variation, we can define a current

$$J_j^\mu = \sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_i)} \frac{\partial \varphi_i}{\partial \epsilon_j} \tag{8.12}$$

Equation (8.11) says that $\partial_\mu J_j^\mu = 0$. From this conserved current, we can also create a conserved charge by integrating over the time component of the current:

$$Q_j \equiv \int d^3x \; J_j^0 \tag{8.13}$$

Now integrate the conservation equation:

$$0 = \int d^3x \; \partial_\mu J_j^\mu = \int d^3x \; \partial_0 J_j^0 + \int d^3x \; \nabla \cdot \mathbf{J}_j$$

$$= \frac{d}{dt} Q_j + \int_s d\mathbf{s} \cdot \mathbf{J}_j$$

In the last term I have used the divergence theorem to convert the volume integral into a surface integral. Assume that the fields vanish at infinity so the last term can be neglected. Then:

$$\frac{d}{dt} Q_j(t) = 0 \tag{8.14}$$

In summary: a continuous symmetry implies a conserved current, which in turn implies a conserved charge. It's important to understand that the "charge" so defined usually has no connection with electrical charge. Most of the conserved quantities in particle physics, such as strangeness, charm, lepton number, *etc.* come about in this way. Many theories in particle physics came about because the conserved quantities were first discovered experimentally. Theorists then reasoned "backwards" from this discovery to construct the underlying conserved currents.

8.3 Feynman's Rules for QED

The derivation of Feynman's rules for electrons and photons proceeds along the same lines as the derivation for scalar particles in Chapter 3. Because

fermion fields anticommute, there are some additional minus signs to keep track of. The time-ordered product was defined in (6.103)

$$T[\psi_\alpha(x_1)\overline{\psi}_\beta(x_2)] = \psi_\alpha(x_1)\overline{\psi}_\beta(x_2)\theta(x_1^0 - x_2^0) - \overline{\psi}(x_2)_\beta\psi(x_1)_\alpha\theta(x_2^0 - x_1^0) \tag{8.15}$$

Decompose the field into positive- and negative-energy parts. $\psi = \psi^+ + \psi^-$. Remember that ψ^+ contains the electron annihilation operator $\hat{b}_k$ and ψ^- contains the positron creation operator $\hat{d}_k^\dagger$. Rewrite (8.15) in such a way that all creation operators stand to the right of all annihilation operators.

$$\begin{aligned} &T[\psi_\alpha(x_1)\overline{\psi}_\beta(x_2)] \\ &= \{\psi_\alpha^+(x_1), \overline{\psi}_\beta^+(x_2)\}\theta(x_1^0 - x_2^0) - \{\overline{\psi}_\beta^-(x_2), \psi_\alpha^-(x_1)\}\theta(x_2^0 - x_1^0) \\ &-\overline{\psi}_\beta^+(x_2)\psi_\alpha^+(x_2) + \psi_\alpha^+(x_1)\overline{\psi}_\beta^-(x_2) + \psi_\alpha^-(x_1)\overline{\psi}_\beta^+(x_2) + \psi_\alpha^-(x_1)\overline{\psi}_\beta^-(x_2) \end{aligned} \tag{8.16}$$

You see that the second line of this equation is a c number and the third, a q number. We call the c-number part the "contraction" of the two fields. The q-number part is "normal ordered" in the sense that all creation operators stand to the left of all annihilation operators, but because of the additional minus sign we must amend the definition of normal ordering somewhat.

Definition: The fields $:\psi(x_1)\psi(x_2)\cdots\psi(x_n):$ are normal ordered if in each term, all creation operators stand to the left of all annihilation operators. Each term is positive unless it required on odd number of permutations to bring the various creation and annihilation operators into that order.

In a way analogous to (3.36), we introduce the notion of the contraction of two fields.

$$\overbrace{\psi_\alpha(x_1)\overline{\psi}_\beta(x_2)} \equiv T[\psi_\alpha(x_1)\overline{\psi}_\beta(x_2)] - :\psi_\alpha(x_1)\overline{\psi}_\beta(x_2): \tag{8.17}$$

Since the vacuum expectation value of normal-ordered fields vanishes, we have

$$\overbrace{\psi_\alpha(x_1)\overline{\psi}_\beta(x_2)} = \langle 0|T[\psi_\alpha(x_1)\overline{\psi}_\beta(x_2)]|0\rangle = i(S_F)_{\alpha\beta}(x_1 - x_2), \tag{8.18}$$

where I have borrowed the definition of the propagator from (6.104). It's clear that

$$\overbrace{\psi(x_1)\psi(x_2)} = \overbrace{\overline{\psi}(x_1)\overline{\psi}(x_2)} = 0 \tag{8.19}$$

An additional complication arises when there are more than two fields to time order. In this case we must include a minus sign for each operator

interchange. For example, (compare with (3.33))

$$\begin{aligned}
&T[\psi(x_1)\psi(x_2)\overline{\psi}(x_3)\overline{\psi}(x_4)] =: \psi(x_1)\psi(x_2)\overline{\psi}(x_3)\overline{\psi}(x_4) : \\
&-\overbrace{\psi(x_1)\overline{\psi}(x_3)} : \psi(x_2)\overline{\psi}(x_4) : + \overbrace{\psi(x_1)\overline{\psi}(x_4)} : \psi(x_2)\overline{\psi}(x_3) : \\
&+\overbrace{\psi(x_2)\overline{\psi}(x_3)} : \psi(x_1)\overline{\psi}(x_4) : - \overbrace{\psi(x_2)\overline{\psi}(x_4)} : \psi(x_1)\overline{\psi}(x_3) : \\
&-\overbrace{\psi(x_1)\overline{\psi}(x_3)}\,\overbrace{\psi(x_2)\overline{\psi}(x_4)} + \overbrace{\psi(x_1)\overline{\psi}(x_4)}\,\overbrace{\psi(x_2)\overline{\psi}(x_3)}
\end{aligned} \tag{8.20}$$

Wick's theorem enables us to calculate the vacuum expectation value of any product of boson and fermion fields. The next step toward calculating the S matrix requires a fermion version of the LSZ reduction formulas developed in Section 7.3. The derivation is similar to that leading to (3.67). There are no new ideas, just a lot of niggling details regarding minus signs. I will simply state the results and refer the reader to the classic exposition in Bjorken and Drell.[1]

In analogy with (3.54) we define the fermion in-field

$$\lim_{t\to-\infty} \psi(x) = \sqrt{Z_2}\ \psi_{\text{in}}(x) \tag{8.21}$$

The constant Z_2 is called the electron wave function renormalization constant. There are three such constants in QED. One for the electron wave function, one for the photon, and one for the electric charge. If it were not for the fact that all Feynman diagrams with closed loops give divergent results, we would be able to calculate them. As it is, we must be sure that they do not appear in any final results. As it turns out, the photon renormalization constant can be set equal to unity and the other two can be gathered into the definition of the physical mass and charge of the electron. This is to say that QED is *renormalizable.* One of the key ideas to come out of this is that renormalizability is a very special property possessed by only a few theories, and non-renormalizable theories are wrong!

Define the plane-wave solutions of Dirac's equation as follows.

$$\begin{aligned}
U_{ks}(x) &= \sqrt{\frac{m}{(2\pi)^3 E_k}}\ u(k,s)e^{-ikx} \\
V_{ks}(x) &= \sqrt{\frac{m}{(2\pi)^3 E_k}}\ v(k,s)e^{ikx}
\end{aligned} \tag{8.22}$$

[1] *Relativistic Quantum Fields*, J. D. Bjorken and S. D. Drell, McGraw-Hill 1965, Section 16.9

In this notation the second-quantized Dirac field is

$$\psi_{\text{in}}(x) = \int d^3k \sum_s [b_{\text{in}}(k,s)U_{ks}(x) + d^\dagger_{\text{in}}(k,s)V_{ks}] \tag{8.23}$$

There are several formulas analogous to (3.65) depending on whether we remove a particle or antiparticle from the in-state or out-state. If we remove a particle from the in-state (refer to Section 7.3 for the signification of the notation)

$$\begin{aligned} S_{\beta\alpha} &= \langle \beta \text{ out}|\alpha \text{ in}\rangle \\ &\sim -\frac{i}{\sqrt{Z_2}} \int d^4x \ \langle \beta \text{ out}|\overline{\psi}(x)|\alpha - k \text{ in}\rangle \overleftarrow{(-i \not{\partial} - m)} U_{ks}(x) \end{aligned} \tag{8.24}$$

Removing an antiparticle from the in-state leads to

$$\sim \frac{i}{\sqrt{Z_2}} \int d^4x \ \overline{V}_{ks} \overrightarrow{(i \not{\partial} - m)} \langle \beta \text{ out}|\overline{\psi}(x)|\alpha - k \text{ in}\rangle \tag{8.25}$$

Removing a particle from the out-state leads to

$$\sim -\frac{i}{\sqrt{Z_2}} \int d^4x \ \overline{U}_{ks} \overrightarrow{(i \not{\partial} - m)} \langle \beta - k \text{ out}|\overline{\psi}(x)|\alpha \text{ in}\rangle U_{ks}(x) \tag{8.26}$$

Removing an antiparticle from the out-state gives

$$\sim \frac{i}{\sqrt{Z_2}} \int d^4x \ \langle \beta - k \text{ out}|\overline{\psi}(x)|\alpha \text{ in}\rangle \overleftarrow{(-i \not{\partial} - m)} V_{ks}(x) \tag{8.27}$$

After two or more particles have been "reduced," the fields inside the bra-ket combination are time ordered.

Finally, we need to be able to reduce photon fields. In this case the asymptotic condition is

$$\lim_{t \to -\infty} \mathbf{A}(\mathbf{x}, t) = \sqrt{Z_3} \mathbf{A}_{\text{in}}(\mathbf{x}, t) \tag{8.28}$$

The constant Z_3 will turn out to renormalize the electron charge. The wave function is

$$\mathbf{A}_{k\lambda} = \frac{1}{\sqrt{(2\pi)^3 2\omega}} e^{-ikx} \epsilon(k, \lambda) \tag{8.29}$$

so that the Fourier expansion of $\mathbf{A}_{\text{in}}(x)$ is

$$\mathbf{A}_{\text{in}} = \int d^3k \sum_\lambda [a_{\text{in}}(k,\lambda)\mathbf{A}_{k\lambda}(x) + a^\dagger_{\text{in}}(k,\lambda)\mathbf{A}^*_{k\lambda}(x)] \tag{8.30}$$

The reduction formulas are virtually identical to (3.67) (with $m = 0$ of course) except for one minus sign.

$$\begin{aligned} S_{\beta\alpha} &= \langle \beta \text{ out}|\alpha \text{ in}\rangle \\ &\sim \frac{-i}{\sqrt{Z_3}} \int d^4x \, \langle \beta \text{ out}|A_\mu(x)|\alpha - (k,\lambda)\rangle \overleftarrow{\Box} A^{\mu *}_{k\lambda}(x) \end{aligned} \tag{8.31}$$

The additional minus sign in (8.31) comes from the space-like nature of the polarization unit vector

$$\epsilon_\mu \epsilon^\mu = -\boldsymbol{\epsilon} \cdot \boldsymbol{\epsilon} = -1$$

This would be a good point to summarize the results so far. According to (5.5) the S-matrix can be rewritten

$$S_{fi} = \delta_{fi} - i(2\pi)^4 \delta^{(4)}(P_f - P - i)\mathcal{T}_{fi} \tag{8.32}$$

The δ_{fi} stands for all those terms appearing in the reduction procedure in which the two initial particles don't scatter.[2] The $(2\pi)^4\delta(P_f - P_i)$ appears when we do the last integral over coordinate space. The interesting part is $\mathcal{T}_{fi}$, which is the crucial ingredient in (5.21) and (5.23) to calculate cross sections and decay rates. To calculate the S-matrix, we start with the basic QED Hamiltonian (8.9) and substitute it into the formula for the S-operator (**??**). This is an infinite series. We keep only those terms corresponding to the desired degree of approximation. These terms are substituted into (**??**)

$$S_{\beta\alpha} = \langle \beta \text{ out}|S|\alpha \text{ in}\rangle. \tag{8.33}$$

In this context, α stands for the specific initial state you have set up in your experiment, and β represents the particular final state whose probability you wish to calculate. We then use (8.24), (8.25), (8.26), (8.27), and (8.31) to reduce out the various electron and photon fields. Finally, do all the implied differentiation and integration.

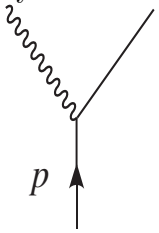
This is a "cookbook" procedure. No problem-solving ability is required. You should be able to do it while in a coma. You have to have patience though, even a modest calculation can involve, implicitly at least, tens of thousands of terms. After you have done a few such calculations, you will realize that you are doing the same few things over and over. These things can be boiled down to a few simple rules known as Feynman's rules.

[2] In a real experiment, the vast majority of incident beam particles don't scatter. All this business about calculating cross sections has to do with the infinitesimal fraction that does.

1. Draw all possible connected, topologically distinct diagrams, including loops, with n external legs. Ignore vacuum-to-vacuum graphs. Each vertex must have a continuous electron or positron line and a photon line terminating on the fermion.

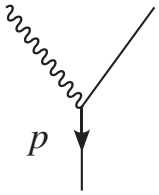
2. There is a factor given by the following rules for each external line.

Incoming fermion:



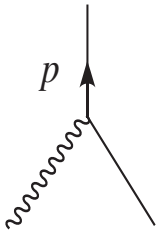
$$= \frac{u(p,s)}{\sqrt{Z_2(2\pi)^3 2E_p}}$$

Incoming antifermion:



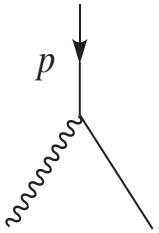
$$= \frac{\overline{v}(p,s)}{\sqrt{Z_2(2\pi)^3 2E_p}}$$

Outgoing fermion:



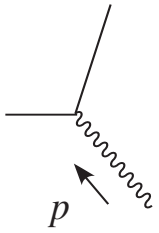
$$= \frac{\overline{u}(p,s)}{\sqrt{Z_2(2\pi)^3 2E_p}}$$

Outgoing antifermion:

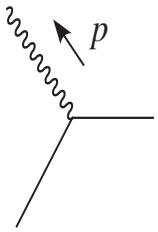


$$= \frac{v(p,s)}{\sqrt{Z_2(2\pi)^3 2E_p}}$$

Incoming photon:



$$= \frac{\epsilon^{\mu}(p)}{\sqrt{Z_3(2\pi)^3 2E_p}}$$

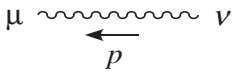
Outgoing photon: 

$$= \frac{\epsilon^{\mu *}(p)}{\sqrt{Z_3 (2\pi)^3 2E_p}}$$

3. For each internal line there is a propagator

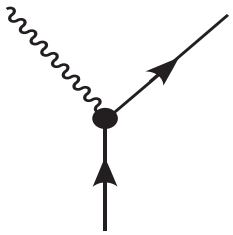
   Electron propagator: 

$$= \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon}$$

   Photon propagator: 

$$= \frac{-ig_{\mu\nu}}{p^2 + i\epsilon}$$

4. At each vertex, place a factor of $-ie\gamma^\mu$.

   Electron-photon vertex: 
$$= -ie\gamma^\mu$$

5. Insert an additional factor of -1 for each closed fermion loop.

6. For each internal loop, integrate over:

$$\int \frac{d^4 q}{(2\pi)^4} \tag{8.34}$$

7. A relative factor -1 appears between graphs that differ from each other by an interchange of two identical external fermion lines.

8. Internal fermion lines appear with arrows in both clockwise and counterclockwise directions. However. diagrams that are topologically equivalent are counted only once.

Remember that the electron-photon coupling has the form $e\overline{\psi}\gamma^\mu\psi A_\mu$. As a consequence, the spinors associated with any vertex must appear in the order $\overline{u}\gamma^\mu u$ (with v replacing u as needed).

With this in mind, let's do some examples.

8.4 The Reaction $e^- + e^+ \to \mu^- + \mu^+$

Consider the reaction $e^- + e^+ \to \mu^- + \mu^+$. In terms of momentum, that is $p + p' \to k + k'$, and in terms of spin indices we have s and $s' \to r$ and r'. The momentum of the exchanged photon is $q = p + p' = k + k'$. The muons are almost identical to electrons, *i.e.* they have the same electromagnetic interactions, but they are heavier by a factor of 200. This simplifies the problem in several ways. For one thing, we are usually entitled to neglect the mass of the electron. For another, there are no identical particles in this example. Feyman's rules give

$$iM = \frac{ie^2}{q^2} \left(\overline{v}_e(p')\gamma^\mu u_e(p)\right) \left(\overline{u}_\mu(k)\gamma_\mu v_\mu(k')\right) \tag{8.35}$$

I have added subscripts e and μ to the spinors to indicate that they contain different masses. The spin indices are implicit. I will get back to spins in a minute.

In order to obtain a differential cross section from M we must do three difficult things: square it, sum over final spins, and average over initial spins. In order to make the notation more compact, I will write (8.35) as

$$iM = \frac{ie^2}{q^2} a^\mu b_\mu \tag{8.36}$$

In order to calculate $|M|^2$, we will need to figure out

$$|ab|^2 = a^\mu a^{\nu\dagger} b_\mu b_\nu^\dagger \tag{8.37}$$

Notice that a and b are 1×1 matrices in spin space and four-vectors in Minkowski space. There are two useful consequences; they all commute and complex conjugation and Hermitian conjugation have the same effect. This allows us to write

$$a^{\nu\dagger} = u^\dagger \gamma^{\nu\dagger} \gamma^{0\dagger} v = u^\dagger \gamma^0 \gamma^\nu v = \overline{u}\gamma^\nu v \tag{8.38}$$

With that insight (8.37) becomes

$$|M|^2 = \frac{e^4}{q^4} \left(\overline{v}(p')\gamma^\mu u(p)\overline{u}(p)\gamma^\nu v(p')\right) \left(\overline{u}(k)\gamma_\mu v(k')\overline{v}(k')\gamma_\nu u(k)\right) \tag{8.39}$$

Now make use of the spin information. The differential cross section will be proportional to

$$\frac{1}{2}\sum_s \frac{1}{2}\sum_{s'} \sum_r \sum_{r'} |M(s, s', r, r')|^2 \tag{8.40}$$

Remember from a previous chapter

$$\sum_s u(p,s)\overline{u}(p,s) = \frac{\not{p}+m}{2m} \tag{8.41}$$

$$\sum_s v(p,s)\overline{v}(p,s) = \frac{\not{p}-m}{2m}$$

Let's do the spin sums on the first term of (8.39)

$$\sum_{s,s'} \overline{v}(p',s')\gamma^\mu u(p,s)\overline{u}(p,s)\gamma^\nu v(p',s') \tag{8.42}$$

$$= \frac{1}{2m}\sum_{s'} \overline{v}_a(p',s')\gamma^\mu_{ab}(\not{p}+m)_{bc}\gamma^\nu_{cd}v_d(p',s')$$

I apologize for the notational overload. An object like γ^μ is a 4×4 matrix in spin space. I have turned it in to a tensor by adding the matrix indices a and b. Tensors commute as you know, so (8.42) can be rewritten

$$= \frac{1}{2m}\sum_{s'} v_d(p',s')\overline{v}_a(p',s')\gamma^\mu_{ab}(\not{p}+m)_{bc}\gamma^\nu_{cd}$$

Now v and $\overline{v}$ are in the right relation to one another so that we can use (8.41)

$$= \left(\frac{1}{2m}\right)^2 (\not{p}'-m)_{da}\gamma^\mu_{ab}(\not{p}+m)_{bc}\gamma^\nu_{cd}$$

Now this is recognizable as the trace of a product of matrices.

$$= \left(\frac{1}{2m}\right)^2 \mathrm{Tr}\left[(\not{p}'-m)\gamma^\mu(\not{p}+m)\gamma^\nu\right]$$

Finally combining (8.39), (8.40), and (8.42), we get

$$\begin{aligned}\frac{1}{4}\sum_{\text{spins}}|M|^2 = \frac{1}{4}\left(\frac{e}{2mq}\right)^4 &\mathrm{Tr}\left[(\not{p}'-m_e)\gamma^\mu(\not{p}+m_e)\gamma^\nu\right] \\ &\times \mathrm{Tr}\left[(\not{p}'-m_\mu)\gamma_\mu(\not{p}+m_\mu)\gamma_\nu\right]\end{aligned} \tag{8.43}$$

There you have it. QED calculations eventually come down to taking traces.

8.4.1 Trace Technology

The things we are taking traces of are all 4×4 matrices, even when they don't look like that. This results in some paradoxical formulas like Tr $1 = 4$ and Tr $m = 4m$. You can check from the representation of the gamma matrices that $\mathrm{Tr}\gamma^\mu = 0$. Traces of an even number of γ's are generally not zero. Here is a simple result.

$$\mathrm{Tr}(\not{a}\ \not{b}) = a_\mu b_\nu \mathrm{Tr}(\gamma^\mu\gamma^\nu) = \frac{1}{2}a_\mu b_\nu \mathrm{Tr}\{\gamma^\mu, \gamma^\nu\} = a_\mu b_\nu \mathrm{Tr}(g^{\mu\nu}) = 4a \cdot b$$

I have used the important fact that the trace of a product of matrices such as $\mathrm{Tr}(abc\cdots z)$ is invariant under any cyclic permutation of the matrices. Thus in the equation above $\mathrm{Tr}(\gamma^\mu\gamma^\nu) = \mathrm{Tr}(\gamma^\nu\gamma^\mu)$. Our result can be summarized

$$\mathrm{Tr}(\gamma^\mu\gamma^\nu) = 4g^{\mu\nu} \tag{8.44}$$

The trace of an odd number of γ's always vanishes. The proof requires a trick. Remember the γ^5 matrix. It has the properties

$$\mathrm{Tr}(\gamma^5) = 0 \qquad (\gamma^5)^2 = 1 \qquad \{\gamma^5, \gamma^\mu\} = 0$$

Now the line of argument goes

$$\mathrm{Tr}(\not{a}_1\ \not{a}_2 \cdots \not{a}_n) = \mathrm{Tr}(\not{a}_1\ \not{a}_2 \cdots \not{a}_n\gamma^5\gamma^5) = \mathrm{Tr}(\gamma^5\ \not{a}_1\ \not{a}_2 \cdots \not{a}_n\gamma^5)$$
$$= (-1)^n \mathrm{Tr}(\not{a}_1\ \not{a}_2 \cdots \not{a}_n\gamma^5\gamma^5) = (-1)^n \mathrm{Tr}(\not{a}_1\ \not{a}_2 \cdots \not{a}_n)$$

The trace is zero for n odd.

There are many other identities that are easily proved. In a moment I will use

$$\mathrm{Tr}(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma) = 4(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma} + g^{\mu\sigma}g^{\nu\rho}) \tag{8.45}$$

which is obtained from repeated application of (8.44). Others can be found in any standard text. At the moment I am looking at Peskin and Schroder, *Introduction to Quantum Field Theory* Appendix A3.

8.4.2 Kinematics

A short calculation using (8.45) yields

$$\mathrm{Tr}\left[(\not{p}' - m)\gamma^\mu(\not{p} + m)\gamma^\nu\right] = 4[p'^\mu p^\nu + p'^\nu p^\mu - g^{\mu\nu}(p \cdot p' + m_e^2)] \tag{8.46}$$

We can set $m_e^2 = 0$. Then (8.43) can be evaluated by using (8.46) twice.

$$\frac{1}{4}\sum_{\text{spins}} |M|^2 = \frac{8e^4}{q^4}[(p \cdot k)(p' \cdot k') + (p \cdot k')(p' \cdot k) + m_\mu^2(p \cdot p')] \tag{8.47}$$

Equation (8.47) is valid in any Lorentz frame. This is fine as far as it goes, but in order to compare the result with an experiment, we need to specialize to a definite frame of reference. Such an experiment is typically done with a colliding beam machine for which the laboratory is the CM frame. Let us assume then that $k = (E, \boldsymbol{k})$, $k' = (E, -\boldsymbol{k})$, $p = (E, E\hat{\boldsymbol{z}})$ and $p' = (E, -E\hat{\boldsymbol{z}})$. The angle between $\boldsymbol{k}$ and the z axis is θ. The following kinematic identities follow with $m_e = 0$.

$$q^2 = (p + p')^2 = 4E^2 \qquad p \cdot p' = 2E^2$$

$$p \cdot k = p' \cdot k' = E^2 - E|\mathbf{k}|\cos\theta \qquad p' \cdot k = p \cdot k' = E^2 + E|\mathbf{k}|\cos\theta$$

Eqn. (8.47) can be rewritten

$$\frac{1}{4}\sum_{\text{spins}} |M|^2 = e^4 \left[\left(1 + \frac{m_\mu^2}{E^2}\right) + \left(1 - \frac{m_\mu^2}{E^2}\right)\cos^2\theta \right] \tag{8.48}$$

Back in Chapter 8 I derived the relationship between M and the differential cross section. Since then we have introduced the Dirac spinors, which must be normalized differently from the Kline-Gordon fields. I will therefore state (without proof) the appropriate version of ?? is

$$\left(\frac{d\sigma}{d\Omega}\right)_{CM} = \frac{1}{2E_A E_B |\mathbf{v}_A - \mathbf{v}_B|} \, \frac{|\mathbf{p}_1|}{(2\pi)^2 4E_{CM}} \, |M(p_A + p_B \to p_1 + p_2)|^2 \tag{8.49}$$

This holds in any coordinate system in which generic particles A and B scatter into a final state consisting of two particles with momenta p_1 and p_2. In our case particles A and B are collinear and assumed massless. In this case

$$|\mathbf{v}_A - \mathbf{v}_B| = \left| \frac{\mathbf{k}_A}{E_A} - \frac{\mathbf{k}_B}{E_B} \right| = 2 \tag{8.50}$$

In our example the M in (8.49) is replaced by the spin-averaged M of (8.48).

8.5 Introduction to Renormalization

Every Feynman diagram with one or more closed loops yields a divergent integral.[3] One of the most important discoveries of twentieth-century physics is that these divergent terms contain real physical content that is part of

[3] There are some exceptions in the case of spin-zero particles, but it's universally true in qed.

the "right answer" to any higher-order calculation. The program of extracting useful information from meaningless infinities is called *renormalization.* This is an extremely technical and complicated subject, but I would like to expose you to some of the basic ideas. To do this, I will concentrate on a subset of divergent diagrams that go under the heading of electron self-energy.

Consider the quantum process in which a single photon of momentum k is emitted and reabsorbed by a virtual electron with initial and final momentum p. In the interim between emitting and reabsorbing the photon the electron has momentum $q = p - k$. According to Feynman's rules, Section 11.3, the amplitude for this process is

$$\begin{aligned} & iS_F(p) \int \frac{d^4k}{(2\pi)^4} (-ie\gamma^\mu) \left(\frac{-ig_{\mu\nu}}{k^2 + i\epsilon} \right) \frac{i(\not{p} - \not{k} + m)}{(p-k)^2 - m^2 + i\epsilon} (-ie\gamma^\nu) iS_F(p) \\ & = iS_F(p) \left[-ie^2 \int \frac{d^4k}{(2\pi)^4} \frac{\gamma^\mu (m + \not{p} - \not{k}) \gamma_\mu}{(-k^2 - i\epsilon)(m^2 - (p-k)^2 - i\epsilon)} \right] S_F(p) \\ & \equiv iS_F(p) \Sigma(p) S_F(p) \end{aligned} \tag{8.51}$$

The last equality implicitly defines the divergent quantity Σ, know as a "self-energy insertion." It is clearly divergent, since there are five powers of k in the numerator and only four in the denominator. We will eventually evaluate it, but for the time being I would like to investigate some of its general properties. Such as term would arise in perturbation theory as a second-order correction to the electron propagator. We could define a "more nearly exact" propagator

$$iS'_F(p) \approx iS_F(p) + iS_F(p)\Sigma(p)S_F(p)$$

I say "more nearly exact" meaning that $iS'_F(p)$ contains all the second-order corrections that would arise in the usual perturbation expansion of the propagator. We could make an even better approximation by including another self-energy insertion.

$$iS'_F \approx iS_F + iS_F\Sigma S_F + iS_F\Sigma S_F\Sigma S_F$$

This is not quite correct, even to fourth order, since I have ignored the diagram in which the electron emits another photon before reabsorbing the first. Diagrams of this sort are called "overlapping divergences." Such diagrams present a huge additional complication, but "sufficient unto the day is the evil thereof." I will ignore them. My conclusions will still be qualitatively correct.

I can go on adding diagrams like this *ad infinitum*. The sum would look like this.

$$iS[1 + \Sigma S + \Sigma S \Sigma S + (\Sigma S)^3 + \cdots] \tag{8.52}$$

(I have suppressed the F subscript and well as the (p) to make the equations easier to read.) This looks like the familiar geometric series $(1-x)^{-1} = 1 + x^2 + x^3 + \cdots$, so I am tempted to write

$$iS' = iS\frac{1}{1 - \Sigma S} \tag{8.53}$$

The question is, what does it mean to divide by a 4×4 matrix? You can regard (8.53) as notational shorthand for (8.52). Here are some further observations.

- It's clear from (8.52) that Σ has the form

  $$\Sigma = mA(p^2) + \not{p}B(p^2) \tag{8.54}$$

  It's true that A and B are infinite, but let's ignore that for the time being. Furthermore

  $$S = \frac{\not{p} + m}{p^2 - m^2}$$

  (I will ignore the $i\epsilon$'s for the time being also. I don't want to distract you with details. I promise I haven't forgotten them and will put them back when necessary.) It follows that $[S, \Sigma] = 0$. Therefore the following forms are equivalent.

  $$S' = S\frac{1}{1 - \Sigma S} = \frac{1}{1 - \Sigma S}S = \frac{S}{1 - \Sigma S} \tag{8.55}$$

- Since $\not{p}\,\not{p} = p^2$ it makes sense to write in the same spirit

  $$S = \frac{\not{p} + m}{p^2 - m^2} = \frac{\not{p} + m}{(\not{p} - m)(\not{p} + m)} = \frac{1}{\not{p} - m}, \tag{8.56}$$

  and finally

  $$S^{-1} = \not{p} - m. \tag{8.57}$$

  The propagator can now be rewritten

  $$iS' = \frac{iS}{1 - \Sigma S}\frac{S^{-1}}{S^{-1}} = \frac{i}{\not{p} - m - \Sigma} \tag{8.58}$$

Leaving aside the fact that Σ is infinite, there is still a strange pathology in (8.57). Every propagator we have encountered so far has had a pole at the mass of the particle. It is almost axiomatic that a propagator is the probability amplitude for the particle to propagate from x to y *such that* $p^2 = m^2$ on the mass shell. It must be that the series of self-energy terms we have summed *has modified the mass of the particle.* We have to confess that at least in the context of perturbation theory, the m that appears in the Hamiltonian is not the physical mass of the particle. We call the m in the Hamiltonian the "bare" mass and the mass that marks the pole of the exact propagator, the "dressed" or physical mass, which I will call $\overline{m}$. It seems that given the bare mass, we should be able to calculate the physical mass or vice versa. The fact that Σ is infinite makes this permanently impossible. We have not faced up to this problem so far, because we have always worked at the lowest non-trivial order of perturbation theory. To that order we can assume the the bare and physical masses are identical, and everything works out OK. As soon as we include the self-energy terms, we must keep track of both m and $\overline{m}$. Worse yet, the same sort of thing happens when we calculate higher-order corrections to the electron-photon vertex. There we find that the charge e put into the Hamiltonian is not the physical charge of the electron but some "bare" charge, and again we must keep track of the e's and $\overline{e}$'s.

The agenda of keeping track of these corrections goes by the name "renormalization." It is a remarkable fact that all the infinite corrections brought about by all loop diagrams to all orders of perturbation theory can swept under the rug of these two constants. A theory for which this is possible is said to be "renormalizable." It is an article of faith that no non-renormalizable theory can be completely right, however useful it might be phenomenologically. Fermi's theory of weak interactions, which we will study in the next chapter, is a good example. It explains nuclear beta decay quite nicely, but it would require an infinite number of infinite constants to take care of all the loops! It was eventually modified by the addition of the intermediate vector bosons. The resulting theory is called the "standard model." It is renormalizable and right! It is also possible to do quantum field theory with gravitons. The theory at least predicts Newton's law of gravitation. To this extent it is correct. It is non-renormalizable, however, and we have no idea what the true theory might be.

It is necessary to rewrite (8.58) in such a way that the physical mass appears in the denominator. To this end we expand the self energy term as follows.

$$\Sigma(\not{p}) = \Sigma(\overline{m}) + (\not{p} - \overline{m})\Sigma'(\overline{m}) + (\not{p} - \overline{m})^2 R(p^2) \tag{8.59}$$

This is the key equation, and it deserves some comment.

- It looks like a power series expansion about the point $\not{p} = m$, but in fact, $\not{p}$ can never equal m, since m is diagonal and $\not{p}$ never is. Rather, (8.59) defines implicitly what is meant by $R(p^2)$. I will sometimes write it
$$\Sigma(\not{p}) = \Sigma(\overline{m}) + (\not{p} - \overline{m})\Sigma'(\overline{m}) + \Sigma_R(\not{p}) \tag{8.60}$$
In which case it defines Σ_R. By construction $\Sigma_R(\overline{m}) = 0$.

- It appears that in the term Σ' we have differentiated with respect to a matrix! That is not so strange as it seems. The point is that $p^2 = \not{p}^2$. For example, the $A(p^2)$ term in (8.54) can be differentiated as follows.
$$A'(\overline{m}) = \left.\frac{dA(p^2)}{d\ \not{p}}\right|_{\not{p}=\overline{m}} = \left.\frac{dA(p^2)}{dp^2}\frac{d\ \not{p}^2}{d\ \not{p}}\right|_{\not{p}=\overline{m}} = 2\overline{m}\left.\frac{dA(p^2)}{dp^2}\right|_{p^2=\overline{m}^2} \tag{8.61}$$

Strictly speaking, $\Sigma(\overline{m})$ and $\Sigma'(\overline{m})$ are both infinite, so the operations in (8.60) and (8.61) are not well defined. Our strategy for giving them a precise meaning works as follows. I will introduce an additional parameter ϵ into (8.51) in such a way the the integrals converge for $\epsilon > 0$. It will turn out that $A(p^2)$ and $B(p^2)$ in (8.54) can each be split into two parts. One part will depend on ϵ. It becomes infinite as $\epsilon \to 0$, but for $\epsilon \neq 0$ it is a well-behaved mathematical expression. The other part will depend on p^2 but not on ϵ. It is finite, calculable, and important. I am about to show that all the epsilon-dependent terms can be lumped together with the bare mass to give the physical mass $\overline{m}$ and a new epsilon-dependent term Z_2 called the wave function renormalization factor, which cancels out in any complete calculation of a Feynman diagram. The point is that all these calculations are done with finite (though epsilon-dependent) terms, and when the time comes to calculate some physical quantity such as a cross section or reaction rate, *there are no epsilon-dependent terms remaining.*

Now substitute (8.60) into (8.58)

$$iS' = \frac{-i}{m + \Sigma(\overline{m}) - \not{p} + (\not{p} - \overline{m})\Sigma'(\overline{m}) + \Sigma_R} \tag{8.62}$$

I looks like the "real" mass is $\overline{m} = m + \Sigma(\overline{m})$. We need one further constant.[4]

$$Z_2^{-1} \equiv 1 - \Sigma'(\overline{m}) \tag{8.63}$$

[4]This is standard notation. There are other Z's, but since we are only dealing with the electron self-energy part, these other constants don't arise.

Equation (8.62) can be rewritten

$$iS' = \frac{i(\not{p} + \overline{m})}{p^2 - \overline{m}^2} \frac{Z_2}{1 + (\overline{m} - \not{p})Z_2 R(p^2)} \tag{8.64}$$

The looks like the "bare" propagator iS (with m replaced by $\overline{m}$) multiplied by a factor, which, if Z_2 were not epsilon-dependent, would be finite and calculable. Before we set $\epsilon = 0$, however, Z_2 is finite, and we can show why it does not appear in the final result of any complete calculation of a Feynman diagram. First consider the Z_2 in the denominator of (8.64). You can see from (8.51) that Σ is proportional to e^2. We must confess again the the e that appears in the Hamiltonian is not the measured charge of the electron, but rather a "bare" charge, which is modified by all the loop diagrams to all orders of perturbation theory. Part of this modification is due to Z_2. Let's tentatively define a "dressed" charge, $\overline{e} = \sqrt{Z_2}e$. Then Z_2 disappears from the denominator because it is absorbed into Σ. The Z_2 disappears from the numerator when we realize that our virtual electron will not be observable until it couples to a photon at each end. (Draw some Feynman diagrams to convince yourself this is true.) Each electron-photon vertex comes with a factor of e, so the entire diagram is proportional to $e^2 Z_2 = \overline{e}^2$.

Now I will evaluate the very difficult integral (8.51). Notice that the denominator is the product of two factors that can be labelled

$$A = m^2 - (p - k)^2 - i\epsilon \qquad B = -k^2 - i\epsilon$$

Feyman realized that a four-dimensional integral can be simplified by making it five-dimensional. The trick is

$$\frac{1}{AB} = \int_0^1 \frac{dz}{[Az + B(1 - z)]^2} \equiv \int_0^1 \frac{dz}{D^2} \tag{8.65}$$

The term D can be simplified by completing the square. Define $k^\mu \equiv k'^\mu + zp^\mu$. Then

$$D = C^2 - k'^2 \qquad \text{where} \qquad C^2 \equiv z[m^2 - p^2(1 - z)]$$

With these substitutions, the numerator of (8.51) becomes

$$N(p, z) = \gamma^\mu [m + \not{p}(1 - z) - \not{k}']\gamma_\mu$$

I can delete the $\not{k}'$ term using the familiar argument that an odd function integrated over even limits yields zero. Using the fact that $\gamma^\mu \gamma_\mu = 4$, gives us the final form of the numerator function.

$$N = 2[2m - \not{p}(1 - z)] \tag{8.66}$$

With all these definitions the self-energy term becomes

$$\Sigma(p) = -i\overline{e}^2 \int_0^1 dz N(p,z) \int \frac{d^4k}{(2\pi)^4} \frac{1}{(C^2 - k^2 - i\epsilon)^2} \tag{8.67}$$

The integral can be done for integer dimension d using the remarkable formula[5]

$$\int \frac{d^dk}{(2\pi)^d} \frac{1}{(C^2 - k^2 - i\epsilon)^n} = \frac{i}{(4\pi)^{d/2}} \frac{\Gamma(n - d/2)}{\Gamma(n)} \left(\frac{1}{C^2}\right)^{n-d/2} \tag{8.68}$$

Here Γ is the Euler gamma function defined by

$$\Gamma(\alpha) \equiv \int_0^\infty dt\, t^{\alpha-1} e^{-t} \tag{8.69}$$

If α is an integer, say n, $\Gamma(n) = (n-1)!$, and in general $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$. It's useful to get Maple to plot this function for you. You will see that Γ interpolates smoothly between the integer values of α. You will also see that $\Gamma(0) = \infty$. We expected that. The integral (8.67) has to diverge. Now of course, the integral on the left side of (8.68) only makes sense when d is an integer, but the expression on the right *interpolates smoothly between integer values of d.* Even if d is only infinitesimally less than four, it is still finite. Now define $\epsilon \equiv 4 - d$. For our purposes, $n = 2$.

$$\frac{\Gamma(n - d/2)}{\Gamma(n)} = \frac{\Gamma(3 - d/2)}{2 - d/2} = \frac{2}{\epsilon}$$

$$\Sigma(p) = \frac{2\overline{e}^2}{(4\pi)^2} \int_0^1 dz N \frac{C^{-\epsilon}}{\epsilon} \tag{8.70}$$

Insert the expansion

$$C^{-\epsilon} = 1 - \epsilon \ln C + O(\epsilon^2)$$

$$\Sigma(p) = \frac{2\overline{e}^2}{(4\pi)^2} \int_0^1 dz N(\not{p}, z) \left(\frac{1}{\epsilon} - \ln C(p^2, z)\right) + O(\epsilon) \tag{8.71}$$

The integrals in (8.71) are easy enough to do, although the resulting expressions are not very illuminating. The important point is that $\Sigma(\not{p})$ can be divided up as follows.

$$\Sigma(\not{p}) = (mB_1 + \not{p}B_2)/\epsilon + \tilde{\Sigma}(\not{p}) \tag{8.72}$$

[5]This is proved in Gross, *Relativistic Quantum Mechanics and Field Theory*, page 344

where $\tilde{\Sigma}(\not{p})$ does not contain ϵ and so remains finite in the limit $\epsilon \to 0$. B_1 and B_2 are simple constants. When we expand $\Sigma(\not{p})$ using (8.60), $\Sigma_R(\not{p})$ is as promised, finite. Thus we can absorb all the infinite terms into the two constants $\overline{m}$ and Z_2, and calculate $\Sigma_R(\not{p})$and $R(p^2)$ which are, again as promised, finite and important.