PHY 488F/1488F Lecture Notes

Michael Luke

Fall, 1999

These notes are perpetually under construction. Please let me know of any typos or errors. The notes are in large part an abridged and revised version of Sidney Coleman's field theory lectures from Harvard, written up by Brian Hill.

Contents

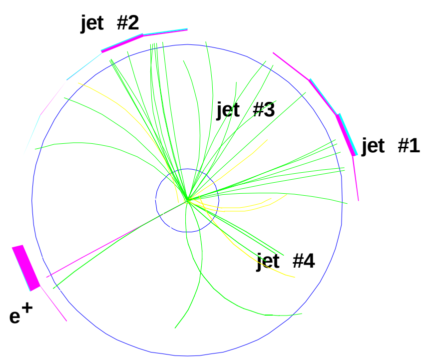

Figure 1.1: The results of a proton-antiproton collision at the Tevatron. The proton and antiproton beams travel perpendicular to the page, colliding at the origin of the tracks. Each of the curved tracks indicates a charged particle in the final state. The tracks are curved because the detector is placed in a magnetic field; the radius of the curvature of the path of a particle provides a means to determine its mass, and therefore identify it.

1 Introduction

1.1 Relativistic Quantum Mechanics

Usually, additional symmetries simplify physical problems. For example, in non-relativistic quantum mechanics (NRQM) rotational invariance greatly simplifies scattering problems. Why does the addition of Lorentz invariance complicate quantum mechanics? The answer is very simple: in relativistic systems, the number of particles is not conserved. In a quantum system, this has profound implications.

Consider, for example, scattering a particle in potential. At low energies, $E \ll mc^2$ where relativity is unimportant, NRQM provides a perfectly adequate description. The incident particle is in some initial state, and one can fairly simply calculate the amplitude for it to scatter into any final state. There is only one particle, before and after the scattering process. At higher energies where relativity is important things gets more complicated, because if $E \sim mc^2$ there is enough energy to pop additional particles out of the vacuum (we will discuss how this works at length in the course). For example, in p-p (proton-proton) scattering with a centre of mass energy $E > m_\pi c^2$ (where $m_\pi \sim 140\,\mathrm{MeV}$ is the mass of the neutral pion) the process

$$p + p \to p + p + \pi^0$$

is possible. At higher energies, $E > 2m_p c^2$, one can produce an additional proton-antiproton pair:

$$p + p \to p + p + p + \overline{p}$$

and so on. Therefore, what started out as a simple two-body scattering process has turned into a many-body problem, and it is necessary to calculate the amplitude to produce a variety of many-body final states. The most energetic accelerator today is the Tevatron at Fermilab, outside Chicago, which collides protons and antiprotons with energies greater than $1\,\mathrm{TeV}$, or about $10^3 m_p c^2$, so typical collisions produce a huge slew of particles (see Fig. 1.1).

Clearly we will have to construct a many-particle quantum theory to describe such a process. However, the problems with NRQM run much deeper, as a brief contemplation of the uncertainty principle indicates. Consider the familiar problem of a particle in a box. In the nonrelativistic description, we can localize the particle in an arbitrarily small region, as long as we accept an arbitrarily large uncertainty in its momentum. But relativity tells us that this description must break down if the box gets too small. Consider a particle of mass μ trapped in a container with reflecting walls of side L. The uncertainty in the particle's momentum is therefore of order $\hbar/L$. In the relativistic regime, this translates to an uncertainty of order $\hbar c/L$ in the particle's energy. For L small enough, $L \lesssim \hbar/\mu c$ (where $\hbar/\mu c \equiv \lambda_c$, the *Compton wavelength* of the particle), the uncertainty in the energy of the system is large enough for particle creation to occur - particle anti-particle pairs can pop out of the vacuum, making the number of particles in the container uncertain! The physical state of the system is a quantum-mechanical superposition of states with different particle number. Even the vacuum state - which in an interacting quantum theory is not the zero-particle state, but rather the state of lowest energy - is complicated. The smaller the distance scale you look at it, the more complex its structure.

There is therefore no sense in which it is possible to localize a particle in a region smaller than its Compton wavelength. In atomic physics, where NRQM works very well, this does not introduce any problems. The Compton wavelength of an electron (mass $\mu = 0.511\,\mathrm{MeV}/c^2$), is $1/0.511\,\mathrm{MeV} \times 197\,\mathrm{MeV\,fm} \sim 4 \times 10^{-11}\,\mathrm{cm}$, or about 10^{-3} Bohr radii. So there is no problem localizing an electron on atomic scales, and the relativistic corrections due to multi-particle states are small. On the other hand, the up and down quarks which make up the proton have masses of order 10 MeV ($\lambda_c \simeq 20$ fm) and are confined to a region the size of a proton, or about 1 fm. Clearly the internal structure of the proton is much more complex than a simple three quark system, and relativistic effects will be huge.

Thus, there is no such thing in relativistic quantum mechanics as the two, one, or even zero body problem! In principle, one is always dealing with the infinite

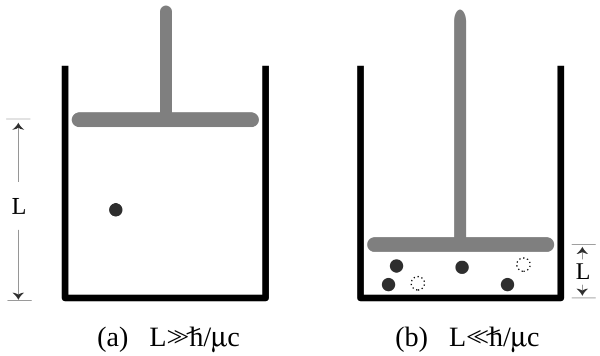

Figure 1.2: A particle of mass μ cannot be localized in a region smaller than its Compton wavelength, $\lambda_c = \hbar/\mu$. At smaller scales, the uncertainty in the energy of the system allows particle production to occur; the number of particles in the box is therefore indeterminate.

body problem. Thus, except in very simple toy models (typically in one spatial dimension), it is impossible to solve any relativistic quantum system exactly. Even the nature of the vacuum state in the real world, a horribly complex sea of quark-antiquark pairs, gluons, electron-positron pairs as well as more exotic beasts like Higgs condensates and gravitons, is totally intractable analytically. Nevertheless, as we shall see in this course, even incomplete (usually perturbative) solutions will give us a great deal of understanding and predictive power.

As a general conclusion, *you cannot have a consistent, relativistic, single particle quantum theory.* So we will have to set up a formalism to handle many-particle systems. Furthermore, it should be clear from this discussion that our old friend the position operator $\vec{X}$ from NRQM does not make sense in a relativistic theory: the $\{|\vec{x}\rangle\}$ basis of NRQM simply does not exist, since particles cannot be localized to arbitrarily small regions. The first casualty of relativistic QM is the position operator, and it will not arise in the formalism which we will develop.

There is a second, intimately related problem which arises in a relativistic quantum theory, which is that of causality. In both relativistic and nonrelativistic quantum mechanics observables correspond to Hermitian operators. In NRQM, however, observables are not attached to space-time points - one simply talks about the position operator, the momentum operator, and so on. However, in a relativistic theory we have to be more careful, because making a measurement forces the system into an eigenstate of the corresponding operator. Unless we are careful about

only defining observables *locally* (i.e. having different observables at each space-time point) we will run into trouble with causality, because observables separated by spacelike separations will be able to interfere with one another.

Consider applying the NRQM approach to observables to a situation with two observers, One and Two, at space-time points x_1 and x_2, which are separated by a spacelike interval. Observer One could be here and Observer Two in the An-

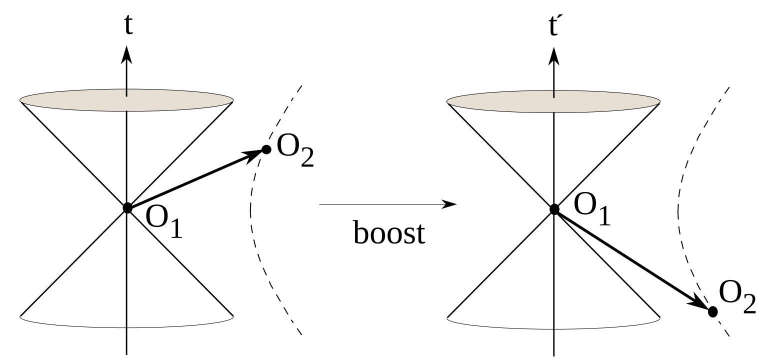

Figure 1.3: Observers O_1 and O_2 are separated by a spacelike interval. A Lorentz boost will move the observer O_2 along the hyperboloid $(\Delta t)^2 - |\Delta\vec{x}|^2 =$ constant, so the time ordering is frame dependent, and they are not in causal contact. Therefore, measurements made at the two points cannot interfere, so observables at point 1 must commute with all observables at point 2.

dromeda galaxy. Now suppose that these two observers both decide to measure non-commuting observables. In this case their measurements can interfere with one another. So imagine that Observer One has an electron and measures the $x-$component of its spin, forcing it into an eigenstate of the spin operator σ_x. If Observer Two measures a non-commuting observable such as σ_y, the next time Observer One measures σ_x it has a 50% chance of being in the opposite spin state, and so she can immediately tell that Observer Two has made a measurement. They have communicated at faster than the speed of light. This of course violates causality, since there are reference frames in which Observer One's second measurement *preceded* Observer Two's measurement (recall that the time-ordering of spacelike separated events depends on the frame of reference), and leads to all sorts of paradoxes (maybe Observer Two then changes his mind and doesn't make the measurement ...).

The problem with NRQM in this context is that it has action at a distance built in: observables are universal, and don't refer to particular space-time points. Classical physics got away from action at a distance by introducing electromagnetic and gravitational fields. The fields are defined at all spacetime points, and the dynamics of the fields are purely local - the dynamics of the field at a point x^μ are determined entirely by the physical quantities (the various fields and their derivatives, as well as the charge density) at that point. In relativistic quantum mechanics, therefore, we can get away from action at a distance by promoting all of our operators to *quantum fields*: operator-valued functions of space-time whose dynamics is purely local. Hence, relativistic quantum mechanics is usually known as "Quantum Field Theory." The requirement that causality be respected then simply translates into a requirement that spacelike separated observables commute: as our example demonstrates, if $O_1(x_1)$ and $O_2(x_2)$ are observables which are defined at the space-time points x_1 and x_2, we must have

$$[O_1(x_1), O_2(x_2)] = 0 \text{ for } (x_1 - x_2)^2 < 0. \tag{1.1}$$

Spacelike separated measurements cannot interfere with one another.

1.2 Conventions and Notation

Before delving into QFT, we will set a few conventions for the notation we will be using in this course.

1.2.1 Units

We will choose the "natural" system of units to simplify formulas and calculations, in which $\hbar = c = 1$ (we do this by choosing units such that one unit of velocity is c and one unit of action is $\hbar$.) This makes life much simpler. For example, by setting $\hbar = 1$, we no longer have to distinguish between wavenumber $\vec{k}$ and momentum $\vec{P} = \hbar\vec{k}$, or between frequency ω and energy $E = \hbar\omega$. The only unit left is the unit of mass, and since the unit of mass is the same as the unit of energy, we may take the unit of mass to be MeV. From the fact that velocity $[L/T]$ and action $[ML^2/T]$ are dimensionless we find that length and time have units of MeV^{-1}.

Consider the fine structure constant which is a fundamental dimensionless number characterizing the strength of the electromagnetic interaction to a single charged particle. In the old units it is

$$\alpha = \frac{e^2}{4\pi\hbar c} = \frac{1}{137.04}.$$

In the new units it is

$$\alpha = \frac{e^2}{4\pi} = \frac{1}{137.04}.$$

Thus the charge e has units of $(\hbar c)^{1/2}$ in the old units, but it is dimensionless in the new units.

It is easy to convert a physical quantity back to conventional units by using the following.

$$\begin{aligned} \hbar &= 6.58 \times 10^{-22}\ \mathrm{MeV\ sec} \\ \hbar c &= 1.97 \times 10^{-11}\ \mathrm{MeV\ cm} \end{aligned} \tag{1.2}$$

By multiplying or dividing by these factors you can convert factors of MeV into sec or cm. A useful conversion is

$$\hbar c = 197\ \mathrm{MeV\ fm} \tag{1.3}$$

where 1 fm (femtometer, or "fermi")$= 10^{-13}$ cm is a typical nuclear scale. Some particle masses in natural units are:

particle	mass
e^- (electron)	511 keV
μ^- (muon)	105.7 MeV
π^0 (pion)	134 MeV
p (proton)	938.3 MeV
n (neutron)	939.6 MeV
B (B meson)	5.279 GeV
W^+ (W boson)	80.2 GeV
Z^0 (Z boson)	91.17 GeV

1.2.2 Relativistic Notation

When dealing with non-orthogonal coordinates, it is of crucial importance to distinguish between *contravariant* coordinates x^μ and *covariant* coordinates x_μ. Just to remind you of the distinction, consider the set of two-dimensional non-orthogonal coordinates on the plane shown in Fig. (1.4).

Now consider the coordinates of a point x in the $(1,2)$ basis. In terms of the unit vectors $\hat{e}_1$ and $\hat{e}_2$ (where the $(1,2)$ subscripts are *labels*, not indices: $\hat{e}_1$ and $\hat{e}_2$ are vectors, not coordinates), we can write

$$\hat{x} = x^1\hat{e}_1 + x^2\hat{e}_2 \tag{1.4}$$

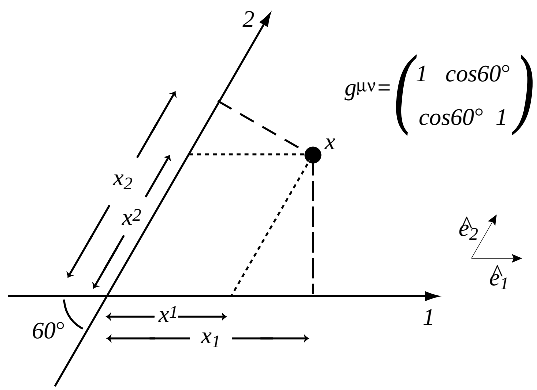

Figure 1.4: Non-orthogonal coordinates on the plane.

which defines the contravariant coordinates x^1 and x^2; these distances are marked on the diagram.

The *covariant* coordinates (x_1, x_2) are defined by

$$x_{1,2} \equiv \vec{x} \cdot \hat{e}_{1,2} \tag{1.5}$$

which are also shown on the figure. Note that for orthogonal axes in flat (Euclidean) space there is no distinction between covariant and contravariant coordinates, which is how you made it this far without worrying about the distinction. However, away from Euclidean space (in particular, in Minkowski space-time) the distinction is crucial.

Given the two sets of coordinates, it is simple to take the scalar product of two vectors. From the definitions above, we have

$$\begin{aligned} \vec{x} \cdot \vec{y} &= \vec{x} \cdot (y^1 \hat{e}_1 + y^2 \hat{e}_2) \\ &= y^1 \vec{x} \cdot \hat{e}_1 + y^2 \vec{x} \cdot \hat{e}_2 \\ &= y^1 x_1 + y^2 x_2 = y_1 x^1 + y_2 x^2 \end{aligned} \tag{1.6}$$

so scalar products are always obtained by pairing upper with lower indices. The relation between contravariant and covariant coordinates is straightforward to derive:

$$\begin{aligned} x_i &= (x^1 \hat{e}_1 + x^2 \hat{e}_2) \cdot \hat{e}_i \\ &= x^j (\hat{e}_i \cdot \hat{e}_j) \\ &\equiv g_{ij} x^j \end{aligned} \tag{1.7}$$

where we have defined the *metric tensor*

$$g_{ij} \equiv \hat{e}_i \cdot \hat{e}_j. \tag{1.8}$$

Note that we are also using the Einstein summation convention: repeated indices (always paired - upper and lower) are implicitly summed over.

One can also define the metric tensor with raised and mixed indices via the relations

$$g_{ij} \equiv g_{ik} g_j^k \equiv g_{ik} g_{jl} g^{kl} \tag{1.9}$$

(note that $g_i^j = \delta_i^j$, the Kronecker delta). The metric tensor g^{ij} raises indices in the natural way,

$$x^i = g^{ij} x_j. \tag{1.10}$$

Minkowskian space is a simple situation in which we use non-orthogonal basis vectors, because time and space look different. The contravariant components of the four-vector x^μ are $(t, \mathbf{r}) = (t, x, y, z)$ where $\mu = 0, 1, 2, 3$. The flat Minkowski space metric is

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{1.11}$$

This is used to raise and lower indices: $x_\mu = g_{\mu\nu} x^\nu = (t, -\mathbf{r})$. The scalar product of two four-vectors is written as

$$a_\mu b^\mu = a^\mu b_\mu = a^\mu g_{\mu\nu} b^\nu = a^0 b^0 - \mathbf{a} \cdot \mathbf{b}. \tag{1.12}$$

It easily follows that this is Lorentz invariant, $a'_\mu b'^\mu = a_\mu b^\mu$.

Note that as before, repeated indices are summed over, and upper indices are always paired with lower indices (see Fig. (1.5)). This ensures that the result of the contraction is a Lorentz scalar. If you get an expression like $a^\mu b^\mu$ (this isn't a scalar because the upper and lower indices aren't paired) or (worse) $a^\mu b^\mu c_\mu d_\mu$ (which indices are paired with which?) you've probably made a mistake. If in doubt, it's sometimes helpful to include explicit summations until you get the hang of it. Remember, this notation was designed to make your life easier!

Under a Lorentz transformation a four-vector transforms according to matrix multiplication:

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu. \tag{1.13}$$

Figure 1.5: Be careful with indices.

where the 4×4 matrix $\Lambda^\mu{}_\nu$ defines the Lorentz transformation. Special cases of $\Lambda^\mu{}_\nu$ include space rotations and "boosts", which look as follows:

$$\begin{aligned} \Lambda^\mu{}_\nu(\text{rotation about } z\text{−axis}) &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ \Lambda^\mu{}_\nu(\text{boost in } x \text{ direction}) &= \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned} \tag{1.14}$$

with $\gamma = (1-v^2)^{-1/2}$. The set of all Lorentz transformations may be defined as those transformations which leave $g_{\mu\nu}$ invariant:

$$g_{\mu\nu} = g_{\alpha\beta}\Lambda^\alpha{}_\mu\Lambda^\beta{}_\nu. \tag{1.15}$$

To see how derivatives transform under Lorentz transformations, we note that the variation

$$\delta\phi = \frac{\partial\phi}{\partial x^\mu}\delta x^\mu \tag{1.16}$$

is a scalar and we would therefore like to write it as $\delta\phi = \partial_\mu\phi\delta x^\mu$. Thus we define

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \tag{1.17}$$

and

$$\partial^\mu \equiv \frac{\partial}{\partial x_\mu} = \left(\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}\right). \tag{1.18}$$

Thus, $\partial/\partial x^\mu$ transforms as a covariant (lower indices) four-vector. Note that

$$\partial_\mu A^\mu = \partial_0 A^0 + \partial_j A^j \tag{1.19}$$

and

$$\partial^\mu\partial_\mu = \frac{\partial^2}{\partial t^2} - \nabla^2 = \Box. \tag{1.20}$$

The energy and momentum of a particle together form the components of its 4-momentum $P^\mu = (E, \vec{p})$.

Finally, we will make use (particularly in the section of Dirac fields) of the completely antisymmetric tensor $\epsilon^{\mu\nu\alpha\beta}$ (often known as the *Levi-Civita* tensor). It is defined by

$$\epsilon^{\mu\nu\alpha\beta} = \begin{cases} 1 & \text{if } (\mu,\nu,\alpha,\beta) \text{ is an even permutation of (0,1,2,3),} \\ -1 & \text{if } (\mu,\nu,\alpha,\beta) \text{ is an odd permutation of (0,1,2,3),} \\ 0 & \text{if } (\mu,\nu,\alpha,\beta) \text{ is not a permutation of (0,1,2,3).} \end{cases} \tag{1.21}$$

Note that you must be careful with raised or lowered indices, since $\epsilon^{0123} = -\epsilon_{0123} = 1$. You should verify that (like the metric tensor $g^{\mu\nu}$) $\epsilon^{\mu\nu\alpha\beta}$ is a relativistically invariant tensor; that is, that under a Lorentz transformation the properties (1.21) still hold.

1.2.3 Fourier Transforms

We will frequently need to go back and forth between the position (x) and momentum (or wavenumber) (p or k) space descriptions of a function, via the Fourier transform. As you should recall, the Fourier transform $\tilde{f}(k)$ allows any function $f(x)$ to be expanded on a continuous basis of plane waves. In quantum mechanics, plane waves correspond to eigenstates of momentum, so Fourier transforming a field will allow us to write it as a sum of modes with definite momentum, which is frequently a very useful thing to do. In n dimensions we therefore write

$$f(x) = \int \frac{d^n k}{(2\pi)^n}\tilde{f}(k)e^{ik\cdot x}. \tag{1.22}$$

It is simple to show that $\tilde{f}(k)$ is therefore given by

$$\tilde{f}(k) = \int d^n x f(x) e^{-ik\cdot x} \tag{1.23}$$

We have introduced two conventions here which we shall stick to in the rest of the course, the sign of the exponentials (we could just as easily have reversed the signs of the exponentials in Eqs. (1.22) and (1.23)) and the placement of the factors of 2π. The latter convention will prove to be convenient because it allows us to easily keep track of powers of 2π - every time you see a $d^n k$ it comes with a factor of $(2\pi)^{-n}$, while $d^n x$'s have no such factors. Also remember that in Minkowski space, $k \cdot x = Et - \vec{k} \cdot \vec{x}$, where $E = k_0$ and $t = x_0$.

1.2.4 The Dirac Delta "Function"

We will frequently be making use in this course of the Dirac delta function $\delta(x)$, which satisfies

$$\int_{-\infty}^{\infty} dx\, \delta(x) = 1 \tag{1.24}$$

and

$$\delta(x) = 0,\ x \neq 0. \tag{1.25}$$

Similarly, in n dimensions we may define the n dimensional delta function

$$\delta^{(n)}(x) \equiv \delta(x_0)\delta(x_1)...\delta(x_n) \tag{1.26}$$

which satisfies

$$\int d^n x\, \delta^{(n)}(x) = 1. \tag{1.27}$$

The δ function can be written as the Fourier transform of a constant,

$$\delta^{(n)}(x) = \frac{1}{(2\pi)^n} \int d^n p\, e^{ip\cdot x}. \tag{1.28}$$

We will also make use of the (one-dimensional) step function

$$\theta(x) = \begin{cases} 1, & x > 0 \\ 0, & x < 0 \end{cases} \tag{1.29}$$

which satisfies

$$\frac{d\theta(x)}{dx} = \delta(x). \tag{1.30}$$

Note that the symbol x will sometimes denote an n-dimensional vector with components x^μ, as in Eq. (1.26), and sometimes a single coordinate, as in Eq. (1.29) - it should be clear from context. For clarity, however, we will usually distinguish three-vectors ($\vec{x}$) from four-vectors (x or x^μ).

1.3 A Naïve Relativistic Theory

Having dispensed with the formalities, in this section we will illustrate with a simple example the somewhat abstract worries about causality we had in the previous section. We will construct a relativistic quantum theory as an obvious relativistic generalization of NRQM, and discover that the theory violates causality: a single free particle will have a nonzero amplitude to be found to have travelled faster than the speed of light.

Consider a free, spinless particle of mass μ. The state of the particle is completely determined by its three-momentum $\vec{k}$ (that is, the components of momentum form a complete set of commuting observables). We may choose as a set of basis states the set of momentum eigenstates $\{|\,\vec{k}\rangle\}$:

$$\vec{P}|\,\vec{k}\rangle = \vec{k}|\,\vec{k}\rangle \tag{1.31}$$

where $\vec{P}$ is the momentum operator. (Note that in our notation, $\vec{P}$ is an operator on the Hilbert space, while the components of $\vec{k}$ are just numbers.) These states are normalized

$$\langle \vec{k}\,|\vec{k}'\rangle = \delta^{(3)}(\vec{k} - \vec{k}') \tag{1.32}$$

and satisfy the completeness relation

$$\int d^3k\,|\,\vec{k}\rangle\langle\vec{k}\,| = 1. \tag{1.33}$$

An arbitrary state $|\,\psi\rangle$ is a linear combination of momentum eigenstates

$$|\,\psi\rangle = \int d^3k\,\psi(\vec{k})\,|\,\vec{k}\rangle \tag{1.34}$$

$$\psi(\vec{k}) \equiv \langle\vec{k}\,|\psi\rangle. \tag{1.35}$$

The time evolution of the system is determined by the Schrödinger equation

$$i\frac{\partial}{\partial t}|\,\psi(t)\rangle = H|\,\psi(t)\rangle. \tag{1.36}$$

where the operator H is the Hamiltonian of the system. The solution to Eq. (1.36) is

$$|\,\psi(t')\rangle = e^{-iH(t'-t)}|\,\psi(t)\rangle. \tag{1.37}$$

In NRQM, for a free particle of mass μ,

$$H|\,\vec{k}\rangle = \frac{|\vec{k}|^2}{2\mu}|\,\vec{k}\rangle. \tag{1.38}$$

If we rashly neglect the warnings of the first section about the perils of single-particle relativistic theories, it appears that we can make this theory relativistic simply by replacing the Hamiltonian in Eq. (1.38) by the relativistic Hamiltonian

$$H_{\rm rel} = \sqrt{|\vec{P}|^2 + \mu^2}. \tag{1.39}$$

The basis states now satisfy

$$H_{\rm rel}|\,\vec{k}\rangle = \omega_k|\,\vec{k}\rangle \tag{1.40}$$

where

$$\omega_k \equiv \sqrt{|\vec{k}|^2 + \mu^2}. \tag{1.41}$$

is the energy of the particle.

This theory looks innocuous enough. We have already argued on general grounds that it cannot be consistent with causality. Nevertheless, it is instructive to show this explicitly. We will find that, if we prepare a particle localized at one position, there is a non-zero probability of finding it outside of its forward light cone at some later time.

To measure the position of a particle, we introduce the position operator, $\vec{X}$, satisfying

$$[X_i, P_j] = i\delta_{ij} \tag{1.42}$$

(remember, we are setting $\hbar = 1$ in everything that follows). In the $\{|\,\vec{k}\rangle\}$ basis, matrix elements of $\vec{X}$ are given by

$$\langle\vec{k}\,|X_i|\,\psi\rangle = i\frac{\partial}{\partial k_i}\psi(\vec{k}) \tag{1.43}$$

and position eigenstates by

$$\langle\vec{k}\,|\vec{x}\rangle = \frac{1}{(2\pi)^{3/2}}e^{-i\vec{k}\cdot\vec{x}}. \tag{1.44}$$

Now let us imagine that at $t = 0$ we have localized a particle at the origin:

$$|\,\psi(0)\rangle = |\,\vec{x} = 0\rangle. \tag{1.45}$$

After a time t we can calculate the amplitude to find the particle at the position $\vec{x}$. This is just

$$\langle\vec{x}\,|\psi(t)\rangle = \langle\vec{x}\,|e^{-iHt}|\,\vec{x} = 0\rangle. \tag{1.46}$$

Inserting the completeness relation Eq. (1.33) and using Eqs. (1.44) and (1.40) we can express this as

$$\begin{aligned}\langle\vec{x}\,|\psi(t)\rangle &= \int d^3k\langle\vec{x}\,|\vec{k}\rangle\langle\vec{k}\,|e^{-iHt}|\,\vec{x} = 0\rangle \\ &= \int d^3k\frac{1}{(2\pi)^3}e^{i\vec{k}\cdot\vec{x}}e^{-i\omega_k t} \\ &= \int_0^\infty \frac{k^2dk}{(2\pi)^3}\int_0^\pi d\theta\,\sin\theta\int_0^{2\pi}d\phi\,e^{ikr\cos\theta}e^{-i\omega_k t}\end{aligned} \tag{1.47}$$

where we have defined $k \equiv |\vec{k}|$ and $r \equiv |\vec{x}|$. The angular integrals are straightforward, giving

$$\langle\vec{x}\,|\psi(t)\rangle = -\frac{i}{(2\pi)^2 r}\int_{-\infty}^{\infty} k\,dk\,e^{ikr}e^{-i\omega_k t}. \tag{1.48}$$

For $r > t$, *i.e.* for a point outside the particle's forward light cone, we can prove using contour integration that this integral is non-zero.

Consider the integral Eq. (1.48) defined in the complex k plane. The integral is along the real axis, and the integrand is analytic everywhere in the plane except for branch cuts at $k = \pm i\mu$, arising from the square root in ω_k. The contour integral can be deformed as shown in Fig. (1.6). For $r > t$, the integrand vanishes

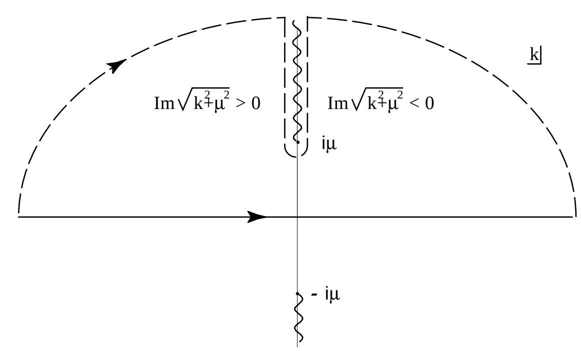

Figure 1.6: Contour integral for evaluating the integral in Eq. (1.48). The original path of integration is along the real axis; it is deformed to the dashed path (where the radius of the semicircle is infinite). The only contribution to the integral comes from integrating along the branch cut.

exponentially on the circle at infinity in the upper half plane, so the integral may be rewritten as an integral along the branch cut. Changing variables to $z = -ik$,

$$\begin{aligned}\langle \vec{x}\,|\psi(t)\rangle &= -\frac{i}{(2\pi)^2 r}\int_\mu^\infty (iz)d(iz)e^{-zr}\left(e^{\sqrt{z^2-\mu^2}t} - e^{-\sqrt{z^2-\mu^2}t}\right)\\ &= \frac{i}{2\pi^2 r}e^{-\mu r}\int_\mu^\infty dz\, ze^{-(z-\mu)r}\sinh\left(\sqrt{z^2-\mu^2}\,t\right). \qquad (1.49)\end{aligned}$$

The integrand is positive definite, so the integral is non-zero. The particle has a small but non-zero probability to be found outside of its forward light-cone, so the theory is acausal. Note the exponential envelope, $e^{-\mu r}$ in Eq. (1.49) means that for distances $r \gg 1/\mu$ there is a negligible chance to find the particle outside the light-cone, so at distances much greater than the Compton wavelength of a particle, the single-particle theory will not lead to measurable violations of causality. This is in accordance with our earlier arguments based on the uncertainty principle: multi-particle effects become important when you are working at distance scales of order the Compton wavelength of a particle.

How does the multi-particle element of quantum field theory save us from these difficulties? It turns out to do this in a quite miraculous way. We will see in a few lectures that one of the most striking predictions of QFT is the existence of *antiparticles* with the same mass as, but opposite quantum numbers of, the corresponding particle. Now, since the time ordering of two spacelike-separated events at points x and y is frame-dependent, there is no Lorentz invariant distinction between emitting a particle at x and absorbing it at y, and emitting an antiparticle at y and absorbing it at x: in Fig. (1.3), what appears to be a particle travelling from O_1 to O_2 in the frame on the left looks like an antiparticle travelling from O_2 to O_1 in the frame on the right. In a Lorentz invariant theory, both processes must occur, and they are indistinguishable. Therefore, if we wish to determine whether or not a measurement at x can influence a measurement at y, we must add the *amplitudes* for these two processes. As it turns out, the amplitudes exactly cancel, so causality is preserved.

2 Constructing Quantum Field Theory

2.1 Multi-particle Basis States

2.1.1 Fock Space

Having killed the idea of a single particle, relativistic, causal quantum theory, we now proceed to set up the formalism for a consistent theory. The first thing we need to do is define the states of the system. The basis for our Hilbert space in relativistic quantum mechanics consists of any number of spinless mesons (the space is called "Fock Space".) However, we saw in the last section that a consistent relativistic theory has no position operator. In QFT, position is no longer an *observable*, but instead is simply a *parameter*, like the time t. In other words, the unphysical question "where is the particle at time t" is replaced by physical questions such as "what is the expectation value of the observable O (the electric field, the energy density, etc.) at the space-time point $(t, \vec{x})$." Therefore, we can't use position eigenstates as our basis states. The momentum operator is fine; momentum is a conserved quantity and can be measured in an arbitrarily small volume element. Therefore, we choose as our single particle basis states the same states as before,

$$\{|\,\vec{k}\rangle\}, \qquad (2.1)$$

but now this is only a piece of the Hilbert space. The basis of two-particle states is

$$\{|\,\vec{k}_1, \vec{k}_2\rangle\}. \qquad (2.2)$$

Because the particles are bosons, these states are even under particle interchange[1]

$$|\,\vec{k}_1, \vec{k}_2\rangle = |\,\vec{k}_2, \vec{k}_1\rangle. \qquad (2.3)$$

They also satisfy

$$\begin{aligned}\langle \vec{k}_1, \vec{k}_2\,|\vec{k}_1', \vec{k}_2'\rangle &= \delta^{(3)}(\vec{k}_1 - \vec{k}_1')\delta^{(3)}(\vec{k}_2 - \vec{k}_2') + \delta^{(3)}(\vec{k}_1 - \vec{k}_2')\delta^{(3)}(\vec{k}_2 - \vec{k}_1')\\ H|\,\vec{k}_1, \vec{k}_2\rangle &= (\omega_{k_1} + \omega_{k_2})|\,\vec{k}_1, \vec{k}_2\rangle\\ \vec{P}|\,\vec{k}_1, \vec{k}_2\rangle &= (\vec{k}_1 + \vec{k}_2)|\,\vec{k}_1, \vec{k}_2\rangle. \qquad (2.4)\end{aligned}$$

States with 2,3,4, ... particles are defined analogously. There is also a zero-particle state, the vacuum $|\,0\rangle$:

$$\begin{aligned}\langle 0\,|0\rangle &= 1\\ H|\,0\rangle &= 0, \qquad \vec{P}|\,0\rangle = 0 \qquad (2.5)\end{aligned}$$

[1] We will postpone the study of fermions until later on, when we discuss spinor fi elds.

and the completeness relation for the Hilbert space is

$$1 = |\,0\rangle\langle 0\,| + \int d^3k |\,\vec{k}\rangle\langle\vec{k}\,| + \frac{1}{2!}\int d^3k_1 d^3k_2 |\,\vec{k_1},\vec{k_2}\rangle\langle\vec{k_1},\vec{k_2}\,| + ... \tag{2.6}$$

(The factor of 1/2! is there to avoid double-counting the two-particle states.) This is starting to look unwieldy. An arbitrary state will have a wave function over the single-particle basis which is a function of 3 variables ($k_x,\ k_y,\ k_z$), a wave function over the two-particle basis which is a function of 6 variables, and so forth. An interaction term in the Hamiltonian which creates a particle will connect the single-particle wave-function to the two-particle wave-function, the two-particle to the three-particle, ... This will be a mess. We need a better description, preferably one which has no explicit multi-particle wave-functions.

As a pedagogical device, it will often be convenient in this course to consider systems confined to a periodic box of side L. This is nice because the wavefunctions in the box are normalizable, and the allowed values of $\vec{k}$ are discrete. Since translation by L must leave the system unchanged, the allowed momenta must be of the form

$$\vec{k} = \left(\frac{2\pi n_x}{L}, \frac{2\pi n_y}{L}, \frac{2\pi n_z}{L}\right) \tag{2.7}$$

for n_x, n_y, n_z integers.

We can then write our states in the *occupation number representation*,

$$|\,...n(\vec{k}), n(\vec{k}'), ...\rangle \tag{2.8}$$

where the $n(\vec{k})$'s give the number of particles of each momentum in the state. Sometimes the state (2.8) is written

$$|\,n(\cdot)\rangle$$

where the $(\cdot)$ indicates that the state depends on the function n for all $\vec{k}$'s, not any single $\vec{k}$. The number operator $N(\vec{k})$ counts the occupation number for a given $\vec{k}$,

$$N(\vec{k})|\,n(\cdot)\rangle = n(\vec{k})|\,n(\cdot)\rangle. \tag{2.9}$$

In terms of $N(\vec{k})$ the Hamiltonian and momentum operator are

$$H = \sum_{\vec{k}} \omega_k N(\vec{k}) \quad \vec{P} = \sum_{\vec{k}} \vec{k} N(\vec{k}). \tag{2.10}$$

This is bears a striking resemblance to a system we have seen before, the simple harmonic oscillator. For a single oscillator, $H_{\rm SHO} = \omega(N+\frac{1}{2})$, where N is the excitation level of the oscillator. Fock space is in a 1-1 correspondence with the space of an infinite system of independent harmonic oscillators, and up to an (irrelevant) overall constant, the Hamiltonians for the two theories look the same. We can make use of that correspondence to define a compact notation for our multiparticle theory.

2.1.2 Review of the Simple Harmonic Oscillator

The Hamiltonian for the one dimensional S.H.O. is

$$H_{\rm SHO} = \frac{P^2}{2\mu} + \frac{1}{2}\omega^2\mu X^2. \tag{2.11}$$

We can write this in a simpler form by performing the canonical transformation

$$P \to p = \frac{P}{\sqrt{\mu\omega}},\ \ X \to q = \sqrt{\mu\omega}X \tag{2.12}$$

(the transformation is canonical because it preserves the commutation relation $[P,X] = [p,q] = -i$). In terms of p and q the Hamiltonian (2.11) is

$$H_{\rm SHO} = \frac{\omega}{2}(p^2+q^2). \tag{2.13}$$

The raising and lowering operators a and $a^\dagger$ are defined as

$$a = \frac{q+ip}{\sqrt{2}},\ \ a^\dagger = \frac{q-ip}{\sqrt{2}} \tag{2.14}$$

and satisfy the commutation relations

$$[a,a^\dagger] = 1,\ \ [H,a^\dagger] = \omega a^\dagger,\ \ [H,a] = -\omega a \tag{2.15}$$

where $H = \omega(a^\dagger a + 1/2) \equiv \omega(N+1/2)$. If $H|\,E\rangle = E|\,E\rangle$, it follows from (2.15) that

$$\begin{aligned} Ha^\dagger|\,E\rangle &= (E+\omega)a^\dagger|\,E\rangle \\ Ha|\,E\rangle &= (E-\omega)a^\dagger|\,E\rangle. \end{aligned} \tag{2.16}$$

so there is a ladder of states with energies ..., $E-\omega, E, E+\omega, E+2\omega$, ... Since $\langle\psi\,|a^\dagger a|\,\psi\rangle = |a|\,\psi\rangle|^2 \geq 0$, there is a lowest weight state $|\,0\rangle$ satisfying $N|\,0\rangle = 0$ and $a|\,0\rangle = 0$. The higher states are made by repeated applications of $a^\dagger$,

$$|\,n\rangle = c_n(a^\dagger)^n|\,0\rangle,\ \ N|\,n\rangle = n|\,n\rangle. \tag{2.17}$$

Since $\langle n\,|aa^\dagger|\,n\rangle = n+1$, it is easy to show that the constant of proportionality $c_n = 1/\sqrt{n!}$.

2.1.3 An Operator Formalism for Fock Space

Now we can apply this formalism to Fock space. Define *creation* and *annihilation* operators a_k and $a_k^\dagger$ for each momentum $\vec{k}$ (remember, we are still working in a box so the allowed momenta are discrete). These obey the commutation relations

$$[a_k, a_{k'}^\dagger] = \delta_{kk'}, \quad [a_k, a_{k'}] = [a_k^\dagger, a_{k'}^\dagger] = 0. \tag{2.18}$$

The single particle states are

$$|\,\vec{k}\rangle = a_k^\dagger |\,0\rangle, \tag{2.19}$$

the two-particle states are

$$|\,\vec{k}, \vec{k}'\rangle = a_k^\dagger a_{k'}^\dagger |\,0\rangle \tag{2.20}$$

and so on. The vacuum state, $|\,0\rangle$, satisfies

$$a_k |\,0\rangle = 0 \tag{2.21}$$

and the Hamiltonian is

$$H = \sum_{\vec{k}} \omega_k \, a_k^\dagger a_k. \tag{2.22}$$

At this point we can remove the box and, with the obvious substitutions, define creation and annihilation operators in the continuum. Taking

$$[a_k, a_{k'}^\dagger] = \delta^{(3)}(\vec{k} - \vec{k}'), \quad [a_k, a_{k'}] = [a_k^\dagger, a_{k'}^\dagger] = 0 \tag{2.23}$$

it is easy to check that we recover the normalization condition $\langle \vec{k}' |\,\vec{k}\rangle = \delta^{(3)}(\vec{k} - \vec{k}')$ and that $H|\,\vec{k}\rangle = \omega_k |\,\vec{k}\rangle$, $\vec{P}|\,\vec{k}\rangle = \vec{k}|\,\vec{k}\rangle$.

We have seen explicitly that the energy and momentum operators may be written in terms of creation and annihilation operators. In fact, *any* observable may be written in terms of creation and annihilation operators, which is what makes them so useful.

2.1.4 Relativistically Normalized States

The states $\{|\,0\rangle, |\,\vec{k}\rangle, |\,\vec{k}_1, \vec{k}_2\rangle, ...\}$ form a perfectly good basis for Fock Space, but will sometimes be awkward in a relativistic theory because they don't transform simply under Lorentz transformations. This is not unexpected, since the normalization and completeness relations clearly treat spatial components of k^μ differently from the time component. Since multi-particle states are just tensor products of single-particle states, we can see how our basis states transform under Lorentz transformations by just looking at the single-particle states.

Let $O(\Lambda)$ be the operator acting on the Hilbert space which corresponds to the Lorentz transformation $x'^\mu = \Lambda^\mu{}_\nu x^\nu$. The components of the four-vector $k^\mu = (\omega_k, \vec{k})$ transform according to

$$k'^\mu = \Lambda^\mu{}_\nu k^\nu. \tag{2.24}$$

Therefore, under a Lorentz transformation, a state with three momentum $\vec{k}$ is obviously transformed into one with three momentum $\vec{k}'$. But this tells us nothing about the normalization of the transformed state; it only tells us that

$$O(\Lambda)|\,\vec{k}\rangle = \lambda(\vec{k}, \vec{k}')|\,\vec{k}'\rangle \tag{2.25}$$

where $\vec{k}'$ is given by Eq. (2.24), and λ is a proportionality constant to be determined. Of course, for states which have a nice relativistic normalization, λ would be one. Unfortunately, our states don't have a nice relativistic normalization. This is easy to see from the completeness relation, Eq. (1.33), because d^3k is not a Lorentz invariant measure. As we will show in a moment, under the Lorentz transformation (2.24) the volume element d^3k transforms as

$$d^3k \rightarrow d^3k' = \frac{\omega_{k'}}{\omega_k} d^3k. \tag{2.26}$$

Since the completeness relation, Eq. (1.33), holds for both primed and unprimed states,

$$\int d^3k \,|\,\vec{k}\rangle\langle\vec{k}\,| = \int d^3k' \,|\,\vec{k}'\rangle\langle\vec{k}'\,| = 1 \tag{2.27}$$

we must have

$$O(\Lambda)|\,\vec{k}\rangle = \sqrt{\frac{\omega_{k'}}{\omega_k}}|\,\vec{k}'\rangle \tag{2.28}$$

which is not a simple transformation law. Therefore we will often make use of the relativistically normalized states

$$|\,k\rangle \equiv \sqrt{(2\pi)^3}\sqrt{2\omega_k}\,|\,\vec{k}\rangle \tag{2.29}$$

(The factor of $(2\pi)^{3/2}$ is there by convention - it will make factors of 2π come out right in the Feynman rules we derive later on.) The states $|\,k\rangle$ now transform simply under Lorentz transformations:

$$O(\Lambda)|\,k\rangle = |\,k'\rangle. \tag{2.30}$$

The convention I will attempt to adhere to from this point on is states with three-vectors, such as $|\,\vec{k}\rangle$, are non-relativistically normalized, whereas states with four-vectors, such as $|\,k\rangle$, are relativistically normalized.

The easiest way to derive Eq. (2.26) is simply to note that d^3k is not a Lorentz invariant measure, but the four-volume element d^4k is. Since the free-particle states satisfy $k^2 = \mu^2$, we can restrict k^μ to the hyperboloid $k^2 = \mu^2$ by multiplying the measure by a Lorentz invariant function:

$$\begin{aligned} & d^4k\,\delta(k^2-\mu^2)\,\theta(k^0) \\ &= d^4k\,\delta((k^0)^2-|\vec{k}|^2-\mu^2)\,\theta(k^0) \\ &= \frac{d^4k}{2k^0}\,\delta(k^0-\omega_k)\,\theta(k^0). \end{aligned} \tag{2.31}$$

(Note that the θ function restricts us to positive energy states. Since a proper Lorentz transformation doesn't change the direction of time, this term is also invariant under a proper L.T.) Performing the k^0 integral with the δ function yields the measure

$$\frac{d^3k}{2\omega_k}. \tag{2.32}$$

Under a Lorentz boost our measure is now invariant:

$$\frac{d^3k}{\omega_k} = \frac{d^3k'}{\omega_{k'}} \tag{2.33}$$

which immediately gives Eq. (2.26).

Finally, whereas the nonrelativistically normalized states obeyed the orthogonality condition

$$\langle \vec{k}'\,|\vec{k}\rangle = \delta^{(3)}(\vec{k}-\vec{k}') \tag{2.34}$$

the relativistically normalized states obey

$$\langle k'\,|k\rangle = (2\pi)^3 2\omega_k \delta^{(3)}(\vec{k}-\vec{k}'). \tag{2.35}$$

The factor of ω_k compensates for the fact that the δ function is not relativistically invariant.

2.2 Canonical Quantization

Having now set up a slick operator formalism for a multiparticle theory based on the SHO, we now have to construct a theory which determines the dynamics of observables. As we argued in the last section, we expect that causality will require us to define observables at each point in space-time, which suggests that the fundamental degrees of freedom in our theory should be fields, $\phi^a(x)$. In the quantum theory they will be operator valued functions of space-time. For the theory to be causal, we must have $[\phi(x), \phi(y)] = 0$ for $(x-y)^2 < 0$ (that is, for x and y spacelike separated). To see how to achieve this, let us recall how we got quantum mechanics from classical mechanics.

2.2.1 Classical Particle Mechanics

In CPM, the state of a system is defined by generalized coordinates $q_a(t)$ (for example $\{x, y, z\}$ or $\{r, \theta, \phi\}$), and the dynamics are determined by the Lagrangian, a function of the q_a's, their time derivatives $\dot{q}_a$ and the time t: $L(q_1, q_2, ...q_n, \dot{q}_1, ..., \dot{q}_n, t) = T - V$, where T is the kinetic energy and V the potential energy. We will restrict ourselves to systems where L has no explicit dependence on t (we will not consider time-dependent external potentials). The action, S, is defined by

$$S \equiv \int_{t_1}^{t_2} L(t)dt. \tag{2.36}$$

Hamilton's Principle then determines the equations of motion: under the variation $q_a(t) \to q_a(t) + \delta q_a(t),\ \delta q_a(t_1) = \delta q_a(t_2) = 0$ the action is stationary, $\delta S = 0$.

Explicitly, this gives

$$\delta S = \int_{t_1}^{t_2} dt \sum_a \left[\frac{\partial L}{\partial q_a}\delta q_a + \frac{\partial L}{\partial \dot{q}_a}\delta \dot{q}_a\right]. \tag{2.37}$$

Define the canonical momentum conjugate to q_a by

$$p_a \equiv \frac{\partial L}{\partial \dot{q}_a}. \tag{2.38}$$

Integrating the second term in Eq. (2.37) by parts, we get

$$\delta S = \int_{t_1}^{t_2} dt \sum_a \left[\frac{\partial L}{\partial q_a} - \dot{p}_a\right]\delta q_a + p_a \delta q_a\Big|_{t_1}^{t_2}. \tag{2.39}$$

Since we are only considering variations which vanish at t_1 and t_2, the last term vanishes. Since the δq_a's are arbitrary, Eq. (2.37) gives the Euler-Lagrange equations

$$\frac{\partial L}{\partial q_a} = \dot{p}_a. \tag{2.40}$$

An equivalent formalism is the Hamiltonian formulation of particle mechanics. Define the Hamiltonian

$$H(q_1, ..., q_n, p_1, ..., p_n) = \sum_a p_a \dot{q}_a - L. \tag{2.41}$$

Note that H is a function of the p's and q's, not the $\dot{q}$'s. Varying the p's and q's we find

$$\begin{aligned} dH &= \sum_a dp_a \dot{q}_a + p_a d\dot{q}_a - \frac{\partial L}{\partial q_a}dq_a - \frac{\partial L}{\partial \dot{q}_a}d\dot{q}_a \\ &= \sum_a dp_a \dot{q}_a - \dot{p}_a dq_a \end{aligned} \tag{2.42}$$

where we have used the Euler-Lagrange equations and the definition of the canonical momentum. Varying p and q separately, Eq. (2.42) gives Hamilton's equations

$$\frac{\partial H}{\partial p_a} = \dot{q}_a, \quad \frac{\partial H}{\partial q_a} = -\dot{p}_a. \tag{2.43}$$

Note that when L does not explicitly depend on time (that is, its time dependence arises solely from its dependence on the $q_a(t)$'s and $\dot{q}_a(t)$'s) we have

$$\begin{aligned} \frac{dH}{dt} &= \sum_a \frac{\partial H}{\partial p_a}\dot{p}_a + \frac{\partial H}{\partial q_a}\dot{q}_a \\ &= \sum_a \dot{q}_a \dot{p}_a - \dot{p}_a \dot{q}_a = 0 \end{aligned} \tag{2.44}$$

so H is conserved. In fact, H is the energy of the system (we shall show this later on when we discuss symmetries and conservation laws.)

2.2.2 Quantum Particle Mechanics

Given a classical system with generalized coordinates q_a and conjugate momenta p_a, we obtain the quantum theory by replacing the functions $q_a(t)$ and $p_a(t)$ by operator valued functions $\hat{q}_a(t),\ \hat{p}_a(t)$, with the commutation relations

$$\begin{aligned} [\hat{q}_a(t), \hat{q}_b(t)] &= [\hat{p}_a(t), \hat{p}_b(t)] = 0 \\ [\hat{p}_a(t), \hat{q}_b(t)] &= -i\delta_{ab} \end{aligned} \tag{2.45}$$

(recall we have set $\hbar = 1$). At this point let's drop the ˆ's on the operators - it should be obvious by context whether we are talking about quantum operators or classical coordinates and momenta. Note that we have included explicit time dependence in the operators $q_a(t)$ and $p_a(t)$. This is because we are going to work in the Heisenberg picture[2],in which states are time-independent and operators carry the time dependence, rather than the more familiar Schrödinger picture, in which the states carry the time dependence. (In both cases, we are considering operators with no explicit time dependence in their definition). (See Mandl & Shaw, Appendix to Chapter 1).

You are probably used to doing quantum mechanics in the "Schrödinger picture" (SP). In the SP, operators with no explicit time dependence in their definition are time independent. The time dependence of the system is carried by the states through the Schrödinger equation

$$i\frac{d}{dt}|\,\psi(t)\rangle_S = H|\,\psi(t)\rangle_S \Longrightarrow |\,\psi(t)\rangle_S = e^{-iH(t-t_0)}|\,\psi(t_0)\rangle_S. \tag{2.46}$$

[2]Actually, we will later be working in the"interaction picture", but for free fi elds this is equivalent to the Heisenberg picture. We will discuss this in a few lectures.

However, there are many equivalent ways to define quantum mechanics which give the same physics. This is simply because we never measure states directly; all we measure are the matrix elements of Hermitian operators between various states. Therefore, any formalism which differs from the SP by a transformation on both the states and the operators which leaves matrix elements invariant will leave the physics unchanged. One such formalism is the Heisenberg picture (HP). In the HP states are time independent

$$|\,\psi(t)\rangle_H = |\,\psi(t_0)\rangle_H. \tag{2.47}$$

Thus, Heisenberg states are related to the Schrödinger states via the unitary transformation

$$|\,\psi(t)\rangle_H = e^{iH(t-t_0)}|\,\psi(t)\rangle_S. \tag{2.48}$$

Since physical matrix elements must be the same in the two pictures,

$${}_S\langle\psi(t)\,|O_S|\,\psi(t)\rangle_S = {}_S\langle\psi(0)\,|e^{iHt}O_S e^{-iHt}|\,\psi(0)\rangle_S = {}_H\langle\psi(t)\,|O_H(t)|\,\psi(t)\rangle_H, \tag{2.49}$$

from Eq. (2.48) we see that in the HP it is the operators, not the states, which carry the time dependence:

$$O_H(t) = e^{iHT}O_S e^{-iHt} = e^{iHT}O_H(0)e^{-iHt} \tag{2.50}$$

(since at $t = 0$ the two descriptions coincide, $O_S = O_H(0)$). This is the solution of the Heisenberg equation of motion

$$i\frac{d}{dt}O_H(t) = [O_H(t), H]. \tag{2.51}$$

Since we are setting up an operator formalism for our quantum theory (recall that we showed in the first section that it was much more convenient to talk about creation and annihilation operators rather than wave-functions in a multi-particle theory), the HP will turn out to be much more convenient than the SP.

Notice that Eq. (2.51) gives

$$\frac{dq_a(t)}{dt} = i[H, q_a(t)]. \tag{2.52}$$

A useful property of commutators is that $[q_a, F(q,p)] = i\partial F/\partial p_a$ where F is a function of the p's and q's. Therefore $[q_a, H] = i\partial H/\partial p_a$ and we recover the first of Hamilton's equations,

$$\frac{dq_a}{dt} = \frac{\partial H}{\partial p_a}. \tag{2.53}$$

Similarly, it is easy to show that $\dot{p}_a = -\partial H/\partial q_a$. Thus, the Heisenberg picture has the nice property that the equations of motion are the same in the quantum theory and the classical theory. Of course, this does not mean the quantum and classical mechanics are the same thing - observables are constructed differently in classical and quantum mechanics. But it does mean that the quantum theory has the correspondence principle built in. That is, for states which look classical, so that expectation values of q_a and p_a may be interpreted as the actual coordinates and conjugate momenta of the state, the classical equations of motion will be reproduced by the quantum theory.

2.2.3 Classical Field Theory

In this quantum theory, observables are constructed out of the q's and p's. In a classical field theory, such as classical electrodynamics, observables (in this case the electric and magnetic field, or equivalently the vector and scalar potentials) are defined at each point in space-time. The generalized coordinates of the system are just the components of the field at each point x. We could label them just as before, $q_{x,a}$ where the index x is continuous and a is discrete, but instead we'll call our generalized coordinates $\phi_a(x)$. Note that x is *not* a generalized coordinate, but rather a label on the field, describing its position in spacetime. It is like t in particle mechanics. The subscript a labels the field; for fields which aren't scalars under Lorentz transformations (such as the electromagnetic field) it will also denote the various Lorentz components of the field.

We will be rather cavalier about going to a continuous index from a discrete index on our observables. Everything we said before about classical particle mechanics will go through just as before with the obvious replacements

$$\begin{aligned} \sum_a &\rightarrow \int d^3x \sum_a \\ \delta_{ab} &\rightarrow \delta_{ab}\delta^{(3)}(\vec{x}-\vec{x}'). \end{aligned} \tag{2.54}$$

Since the Lagrangian for particle mechanics can couple coordinates with different labels a, the most general Lagrangian we could write down for the fields could couple fields at different coordinates x. However, since we are trying to make a causal theory, we don't want to introduce action at a distance - the dynamics of the field should be local in space (as well as time). Furthermore, since we are attempting to construct a Lorentz invariant theory and the Lagrangian only depends on first derivatives with respect to time, we will only include terms with first derivatives with respect to spatial indices. We can write a Lagrangian of this form as

$$L(t) = \sum_a \int d^3x \mathcal{L}(\phi_a(x), \partial_\mu\phi_a(x)) \tag{2.55}$$

where the action is given by

$$S = \int_{t_1}^{t_2} dt L(t) = \int d^4x \mathcal{L}(t,\vec{x}). \tag{2.56}$$

The function $\mathcal{L}(t,\vec{x})$ is called the "Lagrange density"; however, we will usually be sloppy and follow the rest of the world in calling it the Lagrangian. Note that both $\mathcal{L}$ and S are Lorentz invariant, while L is not.

Once again we can vary the fields $\phi_a \rightarrow \phi_a + \delta\phi_a$ to obtain the Euler-Lagrange equations:

$$\begin{aligned} 0 &= \delta S \\ &= \sum_a \int d^4x \left(\frac{\partial\mathcal{L}}{\partial\phi_a}\delta\phi_a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\partial_\mu\phi_a\right) \\ &= \sum_a \int d^4x \left(\left[\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\Pi^\mu_a\right]\delta\phi_a + \partial_\mu\left[\Pi^\mu_a\delta\phi_a\right]\right) \\ &= \sum_a \int d^4x \left(\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\Pi^\mu_a\right)\delta\phi_a \end{aligned} \tag{2.57}$$

where we have defined

$$\Pi^\mu_a \equiv \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} \tag{2.58}$$

and the integral of the total derivative in Eq. (2.57) vanishes since the $\delta\phi_a$'s vanish on the boundaries of integration. Thus we derive the equations of motion for a classical field,

$$\frac{\partial\mathcal{L}}{\partial\phi_a} = \partial_\mu\Pi^\mu_a. \tag{2.59}$$

The analogue of the conjugate momentum p_a is the time component of Π^μ_a, Π^0_a, and we will often abbreviate it as Π_a. The Hamiltonian of the system is

$$H = \sum_a \int d^3x \left(\Pi^0_a\partial_0\phi^0_a - \mathcal{L}\right) \equiv \int d^3x\, \mathcal{H}(x) \tag{2.60}$$

where $\mathcal{H}(x)$ is the Hamiltonian density.

Now let's construct a simple Lorentz invariant Lagrangian with a single scalar field. The simplest thing we can write down that is quadratic in ϕ and $\partial_\mu\phi$ is

$$\mathcal{L} = \tfrac{1}{2}a\left[\partial_\mu\phi\partial^\mu\phi + b\phi^2\right]. \tag{2.61}$$

The parameter a is really irrelevant here; we can easily get rid of it by rescaling our fields $\phi \rightarrow \phi/\sqrt{a}$. So let's take instead

$$\mathcal{L} = \pm\tfrac{1}{2}\left[\partial_\mu\phi\partial^\mu\phi + b\phi^2\right]. \tag{2.62}$$

What does this describe? Well, the conjugate momenta are

$$\Pi^\mu = \pm\partial^\mu\phi \tag{2.63}$$

so the Hamiltonian is

$$H = \pm\tfrac{1}{2}\int d^3x \left[\Pi^2 + (\nabla\phi)^2 - b\phi^2\right]. \tag{2.64}$$

For the theory to be physically sensible, there must be a state of lowest energy. H must be bounded below. Since there are field configurations for which each of the terms in Eq. (2.65) may be made arbitrarily large, the overall sign of H must be $+$, and we must have $b < 0$. Defining $b = -\mu^2$, we have the Hamiltonian

$$H = \tfrac{1}{2}\int d^3x \left[\Pi^2 + (\nabla\phi)^2 + \mu^2\phi^2\right]. \tag{2.65}$$

Each term in H is positive definite: the first corresponds to the energy required for the field to change in time, the second to the energy corresponding to spatial variations, and the last to the energy required just to have the field around in the first place. The equation of motion for this theory is

$$\left(\partial_\mu\partial^\mu + \mu^2\right)\phi(x) = 0. \tag{2.66}$$

This looks promising. In fact, this equation is called the *Klein-Gordon* equation. It was actually first written down by Schrödinger, at the same time he wrote down

$$i\frac{\partial}{\partial t}\psi(x) = -\frac{1}{2\mu}\nabla^2\psi(x). \tag{2.67}$$

In quantum mechanics for a wave $e^{i(\vec{k}\cdot\vec{x}-\omega t)}$, we know $E = \omega$, $\vec{p} = \vec{k}$, so this equation is just $E = \vec{p}^2/2\mu$. Of course, Schrödinger knew about relativity, so from $E^2 = \vec{p}^2 + \mu^2$ he also got

$$\left[-\frac{\partial^2}{\partial t^2} + \nabla^2 - \mu^2\right]\psi = 0 \tag{2.68}$$

or, in our notation,

$$\left(\partial_\mu\partial^\mu + \mu^2\right)\psi(x) = 0. \tag{2.69}$$

Unfortunately, this is a disaster if we want to interpret $\psi(x)$ as a wavefunction as in the Schrödinger Equation: this equation has both positive and negative energy solutions, $E = \pm\sqrt{\vec{p}^2 + \mu^2}$. The energy is unbounded below and the theory has

no ground state. This should not be such a surprise, since we already know that single particle relativistic quantum mechanics is inconsistent.

In Eq. (2.66), though, $\phi(x)$ is *not* a wavefunction. It is a classical field, and we just showed that the Hamiltonian is positive definite. Soon it will be a quantum field which is also not a wavefunction; it is a Hermitian operator. It will turn out that the positive energy solutions to Eq. (2.66) correspond to the creation of a particle of mass μ by the field operator, and the negative energy solutions correspond to the annihilation of a particle of the same mass by the field operator. (It took eight years after the discovery of quantum mechanics before the negative energy solutions of the Klein-Gordon equation were correctly interpreted by Pauli and Weisskopf.) The Hamiltonian will still be positive definite. So let's quantize our classical field theory and construct the quantum field. Then we'll try and figure out what we've created.

2.2.4 Quantum Field Theory

To quantize our classical field theory we do exactly what we did to quantize CPM, with little more than a change of notation. Replace $\phi(x)$ and $\Pi^\mu(x)$ by operator-valued functions satisfying the commutation relations

$$\begin{aligned} [\phi_a(\vec{x},t), \phi_b(\vec{y},t)] &= [\Pi^0(\vec{x},t), \Pi^0(\vec{y},t)] = 0 \\ [\phi_a(\vec{x},t), \Pi^0_b(\vec{y},t)] &= i\delta_{ab}\,\delta^{(3)}(\vec{x}-\vec{y}). \end{aligned} \tag{2.70}$$

As before, $\phi^a(\vec{x},t)$ and $\Pi_a(\vec{y},t)$ are Heisenberg operators, satisfying

$$\frac{d\phi_a(x)}{dt} = i[H, \phi_a(x)], \quad \frac{d\Pi_a(x)}{dt} = i[H, \Pi_a(x)]. \tag{2.71}$$

For the Klein-Gordon field it is easy to show using the explicit form of the Hamiltonian Eq. (2.65) that the operators satisfy

$$\dot{\phi}_a(x) = \Pi(x), \quad \dot{\Pi}(x) = \nabla^2\phi - \mu^2\phi \tag{2.72}$$

and so the quantum fields also obey the Klein-Gordon equation.

Let's try and get some feeling for $\phi(x)$ by expanding it in a plane wave basis. (Since ϕ is a solution to the KG equation this is completely general.) The plane wave solutions to Eq. (2.66) are exponentials $e^{ik\cdot x}$ where $k^2 = \mu^2$. We can therefore write $\phi(x)$ as

$$\phi(x) = \int d^3k \left[\alpha_k e^{-ik\cdot x} + \alpha_k^\dagger e^{ik\cdot x}\right] \tag{2.73}$$

where the α_k's and $\alpha_k^\dagger$'s are operators. Since $\phi(x)$ is going to be an observable, it must be Hermitian, which is why we have to have the $\alpha_k^\dagger$ term. We can solve for α_k and $\alpha_k^\dagger$. First of all,

$$\begin{aligned}\phi(\vec{x},0) &= \int d^3k \left[\alpha_k e^{i\vec{k}\cdot\vec{x}} + \alpha_k^\dagger e^{-i\vec{k}\cdot\vec{x}}\right] \\ \partial_0\phi(\vec{x},0) &= \int d^3k(-i\omega_k)\left[\alpha_k e^{i\vec{k}\cdot\vec{x}} - \alpha_k^\dagger e^{-i\vec{k}\cdot\vec{x}}\right]. \end{aligned} \tag{2.74}$$

Recalling that the Fourier transform of $e^{-i\vec{k}\cdot\vec{x}}$ is a delta function:

$$\int \frac{d^3x}{(2\pi)^3} e^{-i(\vec{k}-\vec{k}')\cdot\vec{x}} = \delta^{(3)}(\vec{k}-\vec{k}') \tag{2.75}$$

we get

$$\begin{aligned}\int \frac{d^3x}{(2\pi)^3}\phi(\vec{x},0)e^{-i\vec{k}\cdot\vec{x}} &= \alpha_k + \alpha_{-k}^\dagger \\ \int \frac{d^3x}{(2\pi)^3}\dot{\phi}(\vec{x},0)e^{-i\vec{k}\cdot\vec{x}} &= (-i\omega_k)(\alpha_k - \alpha_{-k}^\dagger)\end{aligned} \tag{2.76}$$

and so

$$\begin{aligned}\alpha_k &= \tfrac{1}{2}\int \frac{d^3x}{(2\pi)^3}\left[\phi(\vec{x},0) + \frac{i}{\omega_k}\partial_0\phi(\vec{x},0)\right] e^{-i\vec{k}\cdot\vec{x}} \\ \alpha_k^\dagger &= \tfrac{1}{2}\int \frac{d^3x}{(2\pi)^3}\left[\phi(\vec{x},0) - \frac{i}{\omega_k}\partial_0\phi(\vec{x},0)\right] e^{i\vec{k}\cdot\vec{x}}.\end{aligned} \tag{2.77}$$

Using the equal time commutation relations Eq. (2.70), we can calculate $[\alpha_k, \alpha_{k'}^\dagger]$:

$$\begin{aligned}[\alpha_k, \alpha_{k'}^\dagger] &= -\frac{1}{4}\int \frac{d^3x d^3y}{(2\pi)^6}\left[\frac{i}{\omega_{k'}}[\phi(\vec{x},0),\dot{\phi}(\vec{y},0)] + \frac{i}{\omega_k}[\phi(\vec{y},0),\dot{\phi}(\vec{x},0)]\right] e^{-i\vec{k}\cdot\vec{x}+i\vec{k}'\cdot\vec{y}} \\ &= -\frac{1}{4}\int \frac{d^3x d^3y}{(2\pi)^6}\left[\frac{i}{\omega_{k'}}[i\delta^{(3)}(\vec{x}-\vec{y})] + \frac{i}{\omega_k}[i\delta^{(3)}(\vec{x}-\vec{y})]\right] e^{-i\vec{k}\cdot\vec{x}+i\vec{k}'\cdot\vec{y}} \\ &= \frac{1}{4}\int \frac{d^3x}{(2\pi)^6}\left[\frac{1}{\omega_{k'}} + \frac{1}{\omega_k}\right] e^{-i(\vec{k}-\vec{k}')\cdot\vec{x}} \\ &= \frac{1}{(2\pi)^3 2\omega_k}\delta^{(3)}(\vec{k}-\vec{k}').\end{aligned} \tag{2.78}$$

This is starting to look familiar. If we define $a_k \equiv \alpha_k/(2\pi)^{3/2}\sqrt{2\omega_k}$, then

$$[a_k, a_{k'}^\dagger] = \delta^{(3)}(\vec{k}-\vec{k}'). \tag{2.79}$$

These are just the commutation relations for creation and annihilation operators. So the quantum field $\phi(x)$ is a sum over all momenta of creation and annihilation operators:

$$\phi(x) = \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}}\left[a_k e^{-ik\cdot x} + a_k^\dagger e^{ik\cdot x}\right]. \tag{2.80}$$

Actually, if we are to interpret a_k and $a_k^\dagger$ as our old annihilation and creation operators, they had better have the right commutation relations with the Hamiltonian

$$[H, a_k^\dagger] = \omega_k a_k^\dagger, \quad [H, a_k] = -\omega_k a_k \tag{2.81}$$

so that they really do create and annihilate mesons. From the explicit form of the Hamiltonian (Eq. (2.65)), we can substitute the expression for the fields in terms of $a_k^\dagger$ and a_k and the commutation relation Eq. (2.79) to obtain an expression for the Hamiltonian in terms of the $a_k^\dagger$'s and a_k's. After some algebra (do it!), we obtain

$$\begin{aligned}H = \tfrac{1}{2}\int \frac{d^3k}{2\omega_k} &\left[a_k a_{-k} e^{-2i\omega_k t}(-\omega_k^2 + \vec{k}^2 + \mu^2)\right. \\ &+ a_k^\dagger a_k(\omega_k^2 + \vec{k}^2 + \mu^2) \\ &+ a_k a_k^\dagger(\omega_k^2 + \vec{k}^2 + \mu^2) \\ &\left.+ a_k^\dagger a_{-k}^\dagger e^{2i\omega_k t}(-\omega_k^2 + \vec{k}^2 + \mu^2)\right].\end{aligned} \tag{2.82}$$

Since $\omega_k^2 = \vec{k}^2 + \mu^2$, the time-dependent terms drop out and we get

$$H = \tfrac{1}{2}\int d^3k\,\omega_k\left[a_k a_k^\dagger + a_k^\dagger a_k\right]. \tag{2.83}$$

This is almost, but not quite, what we had before,

$$H = \int d^3k\,\omega_k\, a_k^\dagger a_k. \tag{2.84}$$

Commuting the a_k and $a_k^\dagger$ in Eq. (2.83) we get

$$H = \int d^3k\,\omega_k\left[a_k^\dagger a_k + \tfrac{1}{2}\delta^{(3)}(0)\right]. \tag{2.85}$$

$\delta^{(3)}(0)$? That doesn't look right. Let's go back to our box normalization for a moment. Then

$$H = \tfrac{1}{2}\sum_{\vec{k}}\omega_k\left[a_k a_k^\dagger + a_k^\dagger a_k\right] = \sum_{\vec{k}}\left[a_k^\dagger a_k + \tfrac{1}{2}\right] \tag{2.86}$$

so the $\delta^{(3)}(0)$ is just the infinite sum of the zero point energies of all the modes. The energy of each mode starts at $\frac{1}{2}\omega_k$, not zero, and since there are an infinite number of modes we got an infinite energy in the ground state.

This is no big deal. It's just an overall energy shift, and it doesn't matter where we define our zero of energy. Only energy differences have any physical meaning, and these are finite. However, since the infinity gets in the way, let's use this opportunity to banish it forever. We can do this by noticing that the zero point energy of the SHO is really the result of an *ordering ambiguity*. For example, when quantizing the simple harmonic oscillator we could have just as well written down the classical Hamiltonian

$$H_{SHO} = \frac{\omega}{2}(q - ip)(q + ip). \tag{2.87}$$

When p and q are numbers, this is the same as the usual Hamiltonian $\frac{\omega}{2}(p^2 + q^2)$. But when p and q are operators, this becomes

$$H_{SHO} = \omega\, a^\dagger a \tag{2.88}$$

instead of the usual $\omega(a^\dagger a + 1/2)$. So by a judicious choice of ordering, we should be able to eliminate the (unphysical) infinite zero-point energy. For a set of free fields $\phi_1(x_1),\ \phi_2(x_2), ..., \phi_n(x_n)$, define the *normal-ordered product*[3]

$$:\phi_1(x_1)...\phi_n(x_n): \tag{2.89}$$

as the usual product, but with all the creation operators on the left and all the annihilation operators on the right. Since creation operators commute with one another, as do annihilation operators, this uniquely specifies the ordering. So instead of H, we can use $:H:$ and the infinite energy of the ground state goes away:

$$:H:= \int d^3k\, \omega_k\, a_k^\dagger a_k. \tag{2.90}$$

That was easy. But there is a lesson to be learned here, which is that if you ask a silly question in quantum field theory, you will get a silly answer. Asking about absolute energies is a silly question[4]. In general in quantum field theory, if you ask an unphysical question (and it may not be at all obvious that it's unphysical) you

[3]Mandl & Shaw use the notation $N[\phi_1(x_1)...\phi_n(x_n)]$.

[4]except if you want to worry about gravity. In general relativity the curvature couples to the absolute energy, and so it is a physical quantity. In fact, for reasons nobody understands, the observed absolute energy of the universe appears to be almost precisely zero (the famous cosmological constant problem - the energy density is at least 56 orders of magnitude smaller than dimensional analysis would suggest). We won't worry about gravity in this course.

will get infinity for your answer. Taming these infinities is a major headache in QFT.

At this point it's worth stepping back and thinking about what we have done. The classical theory of a scalar field that we wrote down has nothing to do with particles; it simply had as solutions to its equations of motion travelling waves satisfying the energy-momentum relation of a particle of mass μ. The canonical commutation relations we imposed on the fields ensured that the Heisenberg equation of motion for the operators in the quantum theory reproduced the classical equations of motion, thus building the correspondence principle into the theory. However, these commutation relations also ensured that the Hamiltonian had a discrete particle spectrum, and from the energy-momentum relation we saw that the parameter μ in the Lagrangian corresponded to the mass of the particle. Hence, quantizing the classical field theory immediately forced upon us a particle interpretation of the field: these are generally referred to as the *quanta* of the field. For the scalar field, these are spinless bosons (such as pions, kaons, or the Higgs boson of the Standard Model). As we will see later on, the quanta of the electromagnetic (vector) field are photons, while fermions like the electron are the quanta of the corresponding fermi field. In this latter case, however, there is not such a simple correspondence to a classical field: the Pauli exclusion principle means that you can't make a coherent state of fermions, so there is no classical equivalent of an electron field.

At this stage, the field operator ϕ may still seem a bit abstract - an operator-valued function of space-time from which observables are built. To get a better feeling for it, let us consider the interpretation of the state $\phi(\vec{x},0)|\,0\rangle$. From the field expansion Eq. (2.80), we have

$$\phi(\vec{x},0)|\,0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{-i\vec{k}\cdot\vec{x}}|\,k\rangle. \tag{2.91}$$

Thus, when the field operator acts on the vacuum, it pops out a linear combination of momentum eigenstates. (Think of the field operator as a hammer which hits the vacuum and shakes quanta out of it.) Taking the inner product of this state with a momentum eigenstate $|\,p\rangle$, we find

$$\begin{aligned} \langle p\,|\phi(\vec{x},0)|\,0\rangle &= \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}e^{-i\vec{k}\cdot\vec{x}}\langle p\,|k\rangle \\ &= e^{-i\vec{p}\cdot\vec{x}}. \end{aligned} \tag{2.92}$$

Recalling the nonrelativistic relation between momentum and position eigenstates,

$$\langle\vec{p}\,|\vec{x}\rangle = e^{-i\vec{p}\cdot\vec{x}} \tag{2.93}$$

we see that we can interpret $\phi(\vec{x}, 0)$ as an operator which, acting on the vacuum, creates a particle at position $\vec{x}$. Since it contains both creation and annihilation operators, when it acts on an n particle state it has an amplitude to produce both an $n+1$ and an $n-1$ particle state.

Free field theory is only so interesting, of course, since there are no interactions. Particles just move freely. There is no scattering, and in fact, no way to measure anything. A more general theory would have a potential in the Hamiltonian as well, for example

$$\mathcal{L} = \mathcal{L}_0 - \lambda\phi(x)^4 \tag{2.94}$$

where $\mathcal{L}_0$ is the free Hamiltonian. The field now has self-interactions, so the dynamics are nontrivial. To see how such a potential affects the dynamics of the field quanta, consider the potential as a small perturbation (so that we can still expand the fields in terms of solutions to the free-field Hamiltonian). Writing $\phi(x)$ in terms of $a_k^\dagger$'s and a_k's, we see that the interaction term has pieces with n creation operators and $4-n$ annihilation operators. For example, there will be a piece which looks like $a_{k_1}^\dagger a_{k_2}^\dagger a_{k_3} a_{k_4}$, containing two annihilation and two creation operators. This will contribute to $2 \to 2$ scattering when acting on an incoming 2 meson state, and the amplitude for the scattering process will be proportional to λ. At second order in perturbation theory we can get $2 \to 4$ scattering, or pair production, occurring with an amplitude proportional to λ^2. At higher order more complicated processes can occur. This is where we are aiming. But before we set up perturbation theory and scattering theory, we are going to derive some more exact results from field theory which will prove useful.

2.2.5 Covariant Commutators

As a final, slightly technical, aside, I should point out that it is not obvious that we have constructed a Lorentz invariant theory. $\mathcal{L}$ is certainly L.I., but the canonical commutation relations Eq. (2.70) are not - they single out equal times. On the other hand, since they are local we shouldn't run into trouble with causality, which was our motivation in the first place. Let's make sure we can write things in a Lorentz invariant manner. Define

$$\begin{aligned} \phi^+(x) &= \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}} a_k e^{-ik\cdot x}, \\ \phi^-(x) &= \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}} a_k^\dagger e^{ik\cdot x}, \\ \phi(x) &= \phi^+(x) + \phi^-(x) \end{aligned} \tag{2.95}$$

(the $\pm$ convention is opposite to what you might expect, but such is life). Then we have

$$\begin{aligned} [\phi^+(x), \phi^-(y)] &= \int \frac{d^3k\, d^3k'}{(2\pi)^3 2\sqrt{\omega_k \omega_{k'}}} [a_k, a_{k'}^\dagger] e^{-ik\cdot x + ik'\cdot y} \\ &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{-ik\cdot(x-y)} \equiv i\Delta_+(x-y). \end{aligned} \tag{2.96}$$

$\Delta_+(x-y)$ is manifestly Lorentz invariant because d^3k/ω_k is a Lorentz invariant measure. Clearly $[\phi^+(x), \phi^+(y)] = [\phi^-(x), \phi^-(y)] = 0$, so we have

$$\begin{aligned} [\phi(x), \phi(y)] &= [\phi^+(x), \phi^-(y)] + [\phi^-(x), \phi^+(y)] \\ &= i[\Delta_+(x-y) - \Delta_+(y-x)]. \end{aligned} \tag{2.97}$$

Again, this is manifestly Lorentz invariant, and so can only depend on $(x-y)^2$. But we already know that $[\phi(\vec{x}, t), \phi(\vec{y}, t)] = 0$ for any $\vec{x}$ and $\vec{y}$, and hence for any value of $(x-y)^2 < 0$. Therefore for all $(x-y)^2 < 0$, we must have $[\phi(x), \phi(y)] = 0$.

3 Symmetries and Conservation Laws

The dynamics of interacting field theories, such as ϕ^4 theory in Eq. (2.94), are extremely complex. The resulting equations of motion are not analytically soluble. In fact, free field theory (with the optional addition of a source term, as we will discuss) is the only field theory in four dimensions which has an analytic solution. Nevertheless, in more complicated interacting theories it is often possible to discover many important features about the solution simply by examining the symmetries of the theory. In this chapter we will look at this question in detail and develop some techniques which will allow us to extract dynamical information from the symmetries of a theory.

3.1 Classical Mechanics

Let's return to classical mechanics for a moment, where the Lagrangian is $L = T - V$. As a simple example, consider two particles in one dimension in a potential

$$L = \tfrac{1}{2}m_1\dot{q}_1^2 + \tfrac{1}{2}m_2\dot{q}_2^2 - V(q_1, q_2). \tag{3.1}$$

The momenta conjugate to the q_i's are $p_i = m_i\dot{q}_i$, and from the Euler-Lagrange equations

$$\dot{p}_i = -\frac{\partial V}{\partial q_i}, \quad \dot{P} \equiv \dot{p}_1 + \dot{p}_2 = -\left(\frac{\partial V}{\partial q_1} + \frac{\partial V}{\partial q_2}\right). \tag{3.2}$$

If V depends only on $q_1 - q_2$ (that is, the particles aren't attached to springs or anything else which defines a fixed reference frame) then the system is invariant under the shift $q_i \to q_i + \alpha$, and $\partial V/\partial q_1 = -\partial V/\partial q_2$, so $\dot{P} = 0$. The total momentum of the system is conserved. A symmetry ($L(q_i + \alpha, \dot{q}_i) = L(q_i, \dot{q}_i)$) has resulted in a conservation law.

We also saw earlier that when $\partial L/\partial t = 0$ (that is, L depends on t only through the coordinates q_i and their derivatives), then $dH/dt = 0$. H (the energy) is therefore a conserved quantity when the system is invariant under time translation.

This is a very general result which goes under the name of Noether's theorem: for every symmetry, there is a corresponding conserved quantity. It is useful because it allows you to make exact statements about the solutions of a theory without solving it explicitly. Since in quantum field theory we won't be able to solve anything exactly, symmetry arguments will be extremely important.

To prove Noether's theorem, we first need to define "symmetry." Given some general transformation $q_a(t) \to q_a(t, \lambda)$, where $q_a(t, 0) = q_a(t)$, define

$$Dq_a \equiv \frac{\partial q_a}{\partial \lambda}\Big|_{\lambda=0} \tag{3.3}$$

For example, for the transformation $\vec{r} \to \vec{r} + \lambda\hat{e}$ (translation in the $\hat{e}$ direction), $D\vec{r} = \hat{e}$. For time translation, $q_a(t) \to q_a(t + \lambda) = q_a(t) + \lambda dq_a/dt + \mathcal{O}(\lambda^2)$, $Dq_a = dq_a/dt$.

You might imagine that a symmetry is defined to be a transformation which leaves the Lagrangian invariant, $DL = 0$. Actually, this is too restrictive. Time translation, for example, doesn't satisfy this requirement: if L has no explicit t dependence,

$$L(t, \lambda) = L(q_a(t + \lambda), \dot{q}_a(t + \lambda)) = L(0) + \lambda\frac{dL}{dt} + ... \tag{3.4}$$

so $DL = dL/dt$. So more generally, a transformation is a symmetry iff $DL = dF/dt$ for some function $F(q_a, \dot{q}_a, t)$. Why is this a good definition? Consider the variation of the action S:

$$DS = \int_{t_1}^{t_2} dtDL = \int_{t_1}^{t_2} dt\frac{dF}{dt} = F(q_a(t_2), \dot{q}_a(t_2), t_2) - F(q_a(t_1), \dot{q}_a(t_1), t_1). \tag{3.5}$$

Recall that when we derived the equations of motion, we didn't vary the q_a's and $\dot{q}_a$'s at the endpoints, $\delta q_a(t_1) = \delta q_a(t_2) = 0$. Therefore the additional term doesn't contribute to δS and therefore doesn't affect the equations of motion.

It is now easy to prove Noether's theorem by calculating DL in two ways. First of all,

$$\begin{aligned} DL &= \sum_a \frac{\partial L}{\partial q_a}Dq_a + \frac{\partial L}{\partial \dot{q}_a}D\dot{q}_a \\ &= \sum_a \dot{p}_a Dq_a + p_a D\dot{q}_a \\ &= \frac{d}{dt}\sum_a p_a Dq_a \end{aligned} \tag{3.6}$$

where we have used the equations of motion and the equality of mixed partials ($D\dot{q}_a = d(Dq_a)/dt$). But by the definition of a symmetry, $DL = dF/dt$. So

$$\frac{d}{dt}\left(\sum_a p_a Dq_a - F\right) = 0. \tag{3.7}$$

So the quantity $\sum_a p_a Dq_a - F$ is conserved.

Let's apply this to our two previous examples.

1. Space translation: $q_i \to q_i + \alpha$. Then $DL = 0$, $p_i = m_i\dot{q}_i$ and $Dq_i = 1$, so $p_1 + p_2 = m_1\dot{q}_1 + m_2\dot{q}_2$ is conserved. We will call any conserved quantity associated with spatial translation invariance momentum, even if the system looks nothing like particle mechanics.

2. Time translation: $t \to t + \lambda$. Then $Dq_a = dq_a/dt$, $DL = dL/dt$, $F = L$ and so the conserved quantity is $\sum_a (p_a \dot{q}_a) - L$. This is the Hamiltonian, justifying our previous assertion that the Hamiltonian is the energy of the system. Again, we will call the conserved quantity associated with time translation invariance the energy of the system.

This works for classical particle mechanics. Since the canonical commutation relations are set up to reproduce the E-L equations of motion for the operators, it will work for quantum particle mechanics as well.

3.2 Symmetries in Field Theory

Since field theory is just the continuum limit of classical particle mechanics, the same arguments must go through as well. In fact, stronger statements may be made in field theory, because not only are conserved quantities globally conserved, they must be locally conserved as well. For example, in a theory which conserves electric charge we can't have two separated opposite charges simultaneously wink out of existence. This conserves charge globally, but not locally. Recall from electromagnetism that the charge density satisfies

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{\jmath} = 0. \tag{3.8}$$

This just expresses current conservation. Integrating over some volume V, and defining $Q_V = \int_V d^3x \rho(x)$, we have

$$\frac{dQ_V}{dt} = -\int_v \nabla \cdot \vec{\jmath} = -\int_S dS \cdot \vec{\jmath} \tag{3.9}$$

where S is the surface of V. This means that the rate of change of charge inside some region is given by the flux through the surface. Taking the surface to infinity, we find that the total charge Q is conserved. However, we have the stronger statement of current conservation, Eq. (3.8). Therefore, in field theory conservation laws will be of the form $\partial_\mu \mathcal{J}^\mu = 0$ for some four-current $\mathcal{J}^\mu$.

As before, we consider the transformations $\phi_a(x) \to \phi_a(x, \lambda)$, $\phi_a(x, 0) = \phi_a(x)$, and define

$$D\phi_a = \frac{\partial \phi_a}{\partial \lambda}\Big|_{\lambda=0}. \tag{3.10}$$

A transformation is a symmetry iff $D\mathcal{L} = \partial_\mu F^\mu$ for some $F^\mu(\phi_a, \dot{\phi}_a, x)$. I will leave it to you to show that, just as in particle mechanics, a transformation of this

form doesn't affect the equations of motion. We now have

$$\begin{aligned} D\mathcal{L} &= \sum_a \frac{\partial L}{\partial \phi_a} D\phi_a + \Pi^\mu_a D(\partial_\mu \phi_a) \\ &= \sum_a \partial_\mu \Pi^\mu_a D\phi_a + \Pi^\mu_a \partial_\mu D\phi_a \\ &= \partial_\mu \sum_a (\Pi^\mu_a D\phi_a) = \partial_\mu F^\mu \end{aligned} \tag{3.11}$$

so the four components of

$$J^\mu = \sum_a \Pi^\mu_a D\phi_a - F^\mu \tag{3.12}$$

satisfy $\partial_\mu J^\mu = 0$. If we integrate over all space, so that no charge can flow out through the boundaries, this gives the global conservation law

$$\frac{dQ}{dt} \equiv \frac{d}{dt} \int d^3x J^0(x) = 0. \tag{3.13}$$

3.2.1 Space-Time Translations and the Energy-Momentum Tensor

We can use the techniques from the previous section to calculate the conserved current and charge in field theory corresponding to a space or time translation. Under a shift $x \to x + \lambda e$, where e is some fixed four-vector, we have

$$\begin{aligned} \phi_a(x) &\to \phi_a(x + \lambda e) \\ &= \phi_a(x) + \lambda e_\mu \partial^\mu \phi_a(x) + ... \end{aligned} \tag{3.14}$$

so

$$D\phi_a(x) = e_\mu \partial^\mu \phi_a(x). \tag{3.15}$$

Similarly, since $\mathcal{L}$ contains no explicit dependence on x but only depends on it through the fields ϕ^a, we have $D\mathcal{L} = \partial_\mu(e^\mu \mathcal{L})$, so $F = e^\mu \mathcal{L}$. The conserved current is therefore

$$\begin{aligned} J^\mu &= \sum_a \Pi^\mu_a D\phi - F \\ &= \sum_a \Pi^\mu_a e_\nu \partial^\nu \phi_a - e^\mu \mathcal{L} \\ &= e_\nu \left[\sum_a \Pi^\mu_a \partial^\nu \phi_a - g^{\mu\nu} \mathcal{L} \right] \\ &\equiv e_\nu T^{\mu\nu} \end{aligned} \tag{3.16}$$

where $T^{\mu\nu} = \sum_a \Pi_a^\mu \partial^\nu \phi_a - g^{\mu\nu}\mathcal{L}$ is called the *energy-momentum tensor.* Since $\partial_\mu J^\mu = 0 = \partial_\mu T^{\mu\nu} e_\nu$ for arbitrary e, we also have

$$\partial_\mu T^{\mu\nu} = 0. \tag{3.17}$$

For time translation, $e_\nu = (1, \vec{0})$. $T^{\mu 0}$ is therefore the "energy current", and the corresponding conserved quantity is

$$Q = \int d^3x J^0 = \int d^3x T^{00} = \int d^3x \sum_a \left(\Pi_a^0 \partial_0 \phi_a - \mathcal{L}\right) = \int d^3x \mathcal{H} \tag{3.18}$$

where $\mathcal{H}$ is the Hamiltonian density we had before. So the Hamiltonian, as we had claimed, really is the energy of the system (that is, it corresponds to the conserved quantity associated with time translation invariance.)

Similarly, if we choose $e^\mu = (0, \hat{x})$ then we will find the conserved charge to be the x-component of momentum. For the Klein-Gordon field, a straightforward substitution of the expansion of the fields in terms of creation and annihilation operators into the expression for $\int d^3x T^{01}$ gives the expression we obtained earlier for the momentum operator,

$$: \vec{P} := \int d^3k \, \vec{k} \, a_k^\dagger a_k \tag{3.19}$$

where again we have normal-ordered the expression to remove spurious infinities.

Note that the *physical* momentum $\vec{P}$, the conserved charge associated with space translation, has nothing to do with the *conjugate* momentum Π_a of the field ϕ_a. It is important not to confuse these two uses of the term "momentum."

3.2.2 Lorentz Transformations

Under a Lorentz transformation

$$x^\mu \to \Lambda^\mu{}_\nu x^\nu \tag{3.20}$$

a four-vector transforms as

$$a^\mu \to \Lambda^\mu{}_\nu a^\nu \tag{3.21}$$

as discussed in the first section. Since a scalar field by definition does not transform under Lorentz transformations, it has the simple transformation law

$$\phi(x) \to \phi(\Lambda^{-1}x). \tag{3.22}$$

This simply states that the field itself does not transform at all; the value of the field at the coordinate x in the new frame is the same as the field at that same point in the old frame. Fields with spin have more complicated transformation laws, since the various components of the fields rotate into one another under Lorentz transformations. For example, a vector field (spin 1) A_μ transforms as

$$A^\mu(x) \to \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x). \tag{3.23}$$

As usual, we will restrict ourselves to scalar fields at this stage in the course.

To use the machinery of the previous section, let us consider a one parameter subgroup of Lorentz transformations parameterized by λ. This could be rotations about a specified axis by an angle λ, or boosts in some specified direction with $\gamma = \lambda$. This will define a family of Lorentz transformations $\Lambda(\lambda)^\mu{}_\nu$, from which we wish to get $D\phi = \partial\phi/\partial\lambda|_{\lambda=0}$. Let us define

$$\epsilon^\mu{}_\nu \equiv D\Lambda^\mu{}_\nu. \tag{3.24}$$

Then under a Lorentz transformation $a^\mu \to \Lambda^\mu{}_\nu a^\nu$, we have

$$Da^\mu = \epsilon^\mu{}_\nu a^\nu. \tag{3.25}$$

It is straightforward to show that $\epsilon_{\mu\nu}$ is antisymmetric. From the fact that $a^\mu b_\mu$ is Lorentz invariant, we have

$$\begin{aligned} 0 &= D(a^\mu b_\mu) = (Da^\mu)b_\mu + a^\mu(Db_\mu) \\ &= \epsilon^\mu{}_\nu a^\nu b_\mu + a^\mu \epsilon_\mu{}^\nu b_\nu \\ &= \epsilon_{\mu\nu} a^\nu b^\mu + \epsilon_{\nu\mu} a^\nu b^\mu \\ &= (\epsilon_{\mu\nu} + \epsilon_{\nu\mu})\, a^\nu b^\mu \end{aligned} \tag{3.26}$$

where in the third line we have relabelled the dummy indices. Since this holds for arbitrary four vectors a^μ and b^ν, we must have

$$\epsilon_{\mu\nu} = -\epsilon_{\nu\mu}. \tag{3.27}$$

The indices μ and ν range from 0 to 3, which means there are $4(4-1)/2 = 6$ independent components of ϵ. This is good because there are six independent Lorentz transformations - three rotations (one about each axis) and three boosts (one in each direction).

Let's take a moment and do a couple of examples to demystify this. Take $\epsilon_{12} = -\epsilon_{21}$ and all the other components zero. Then we have

$$\begin{aligned} Da^1 &= \epsilon^1_2 a^2 = -\epsilon_{12} a^2 = -a^2 \\ Da^2 &= \epsilon^2_1 a^1 = -\epsilon_{21} a^2 = +a^1. \end{aligned} \tag{3.28}$$

This just corresponds to the a rotation about the z axis,

$$\begin{pmatrix} a^1 \\ a^2 \end{pmatrix} \rightarrow \begin{pmatrix} \cos\lambda & -\sin\lambda \\ \sin\lambda & \cos\lambda \end{pmatrix} \begin{pmatrix} a^1 \\ a^2 \end{pmatrix}. \tag{3.29}$$

On the other hand, taking $\epsilon_{01} = -\epsilon_{10} = +1$ and all other components zero, we get

$$\begin{aligned} Da^0 &= \epsilon^0_1 a^1 = \epsilon_{01} a^1 = +a^1 \\ Da^1 &= \epsilon^1_0 a^1 = -\epsilon_{10} a^2 = +a^0. \end{aligned} \tag{3.30}$$

Note that the signs are different because lowering a 0 index doesn't bring in a factor of -1. This is just the infinitesimal version of

$$\begin{pmatrix} a^0 \\ a^1 \end{pmatrix} \rightarrow \begin{pmatrix} \cosh\lambda & \sinh\lambda \\ \sinh\lambda & \cosh\lambda \end{pmatrix} \begin{pmatrix} a^0 \\ a^1 \end{pmatrix}. \tag{3.31}$$

which corresponds to a boost along the x axis.

Now we're set to construct the six conserved currents corresponding to the six different Lorentz transformations. Using the chain rule, we find

$$\begin{aligned} D\phi(x) &= \frac{\partial}{\partial\lambda}\, \phi(\Lambda^{-1}(\lambda)^\mu{}_\nu x^\nu)\Big|_{\lambda=0} \\ &= \partial_\alpha\phi(x)\, \frac{\partial}{\partial\lambda}\, (\Lambda^{-1}(\lambda)(x)^\alpha\Big|_{\lambda=0} \\ &= \partial_\alpha\phi(x)\, D\left(\Lambda^{-1}(\lambda)^\alpha{}_\beta x^\beta\right) \\ &= \partial_\alpha\phi(x)\left(-\epsilon^\alpha_\beta\right) x^\beta \\ &= -\epsilon_{\alpha\beta} x^\beta \partial^\alpha \phi(x). \end{aligned} \tag{3.32}$$

Since $\mathcal{L}$ is a scalar, it depends on x only through its dependence on the field and its derivatives. Therefore we have

$$\begin{aligned} D\mathcal{L} &= \epsilon_{\alpha\beta} x^\alpha \partial^\beta \mathcal{L} \\ &= \partial_\mu \left(\epsilon_{\alpha\beta} x^\beta g^{\mu\alpha}\right) \mathcal{L} \end{aligned} \tag{3.33}$$

and so the conserved current J^μ is

$$\begin{aligned} J^\mu &= \sum_a \left(\Pi^\mu \epsilon_{\alpha\beta} x^\alpha \partial^\beta \phi - \epsilon_{\alpha\beta} x^\alpha g^{\mu\beta} \mathcal{L}\right) \\ &= \epsilon_{\alpha\beta}\left(\Pi^\mu x^\alpha \partial^\beta \phi - x^\alpha g^{\mu\beta}\mathcal{L}\right). \end{aligned} \tag{3.34}$$

Since the current must be conserved for all six antisymmetric matrices $\epsilon_{\alpha\beta}$, the part of the quantity in the parentheses that is antisymmetric in α and β must be conserved. That is,

$$\partial_\mu M^{\mu\alpha\beta} = 0 \tag{3.35}$$

where

$$\begin{aligned} M^{\mu\alpha\beta} &= \Pi^\mu x^\alpha \partial^\beta \phi - x^\alpha g^{\mu\beta}\mathcal{L} - \alpha \leftrightarrow \beta \\ &= x^\alpha\left(\Pi^\mu \partial^\beta \phi - g^{\mu\beta}\mathcal{L}\right) - \alpha \leftrightarrow \beta \\ &= x^\alpha T^{\mu\beta} - x^\beta T^{\mu\alpha} \end{aligned} \tag{3.36}$$

where $T^{\mu\nu}$ is the energy-momentum tensor defined in Eq. (3.16). The six conserved charges are given by the six independent components of

$$J^{\alpha\beta} = \int d^3x\, M^{0\alpha\beta} = \int d^3x\, \left(x^\alpha T^{0\beta} - x^\beta T^{0\alpha}\right). \tag{3.37}$$

Just as we called the conserved quantity corresponding to space translation the momentum, we will call the conserved quantity corresponding to rotations the angular momentum. So for example J^{12}, the conserved quantity coming from invariance under rotations about the 3 axis, is

$$J^{12} = \int d^3x\, \left(x^1 T^{02} - x^2 T^{01}\right). \tag{3.38}$$

This is the field theoretic analogue of angular momentum. We can see that this definition matches our previous definition of angular momentum in the case of a point particle with position $\vec{r}(t)$. In this case, the energy momentum tensor is

$$T^{0i}(\vec{x}, t) = p^i \delta^{(3)}(\vec{x} - \vec{r}(t)) \tag{3.39}$$

which gives

$$J^{12} = x^1 p^2 - x^2 p^1 = (\vec{r} \times \vec{p})_3 \tag{3.40}$$

which is the familiar expression for the third component of the angular momentum. Note that this is only for scalar particles. Particles with spin carry intrinsic angular momentum which is not included in this expression - this is only the orbital angular momentum. Particles with spin are described by fields with tensorial character, which is reflected by additional terms in the J^{ij}.

That takes care of three of the invariants corresponding to Lorentz transformations. Together with energy and linear momentum, they make up the conserved quantities you learned about in first year physics. What about boosts? There must be three more conserved quantities. What are they? Consider

$$J^{0i} = \int d^3x\, \left(x^0 T^{0i} - x^i T^{00}\right). \tag{3.41}$$

This has an explicit reference to x^0, the time, which is something we haven't seen before in a conservation law. But there's nothing in principle wrong with this. The x^0 may be pulled out of the spatial integral, and the conservation law gives

$$\begin{aligned} 0 &= \frac{d}{dt}J^{0i} = \frac{d}{dt}\left[t\int d^3x\, T^{0i} - \int d^3x\, x^i T^{00}\right] \\ &= t\frac{d}{dt}\int d^3x\, T^{0i} + \int d^3x\, T^{0i} - \frac{d}{dt}\int d^3x\, x^i T^{00} \\ &= t\frac{d}{dt}p^i + p^i - \frac{d}{dt}\int d^3x\, x^i T^{00}. \end{aligned} \quad (3.42)$$

The first term is zero by momentum conservation, and the second term, p^i, is a constant. Therefore we get

$$p^i = \frac{d}{dt}\int d^3x\, x^i T^{00} = \text{constant}. \quad (3.43)$$

This is just the field theoretic and relativistic generalization of the statement that the centre of mass moves with a constant velocity. The centre of mass is replaced by the "centre of energy." Although you are not used to seeing this presented as a separate conservation law from conservation of momentum, we see that in field theory the relation between the T^{0i}'s and the first moment of T^{00} is the result of Lorentz invariance. The three conserved quantities $\int d^3x\, x^i T^{00}(x)$ are the Lorentz partners of the angular momentum.

3.3 Internal Symmetries

Energy, momentum and angular momentum conservation are clearly properties of any Lorentz invariant field theory. We could write down an expression for the energy-momentum tensor $T^{\mu\nu}$ without knowing the explicit form of $\mathcal{L}$. However, there are a number of other quantities which are experimentally known to be conserved, such as electric charge, baryon number and lepton number which are not automatically conserved in any field theory. By Noether's theorem, these must also be related to continuous symmetries. Experimental observation of these conservation laws in nature is crucial in helping us to figure out the Lagrangian of the real world, since they require $\mathcal{L}$ to have the appropriate symmetry and so tend to greatly restrict the form of $\mathcal{L}$. We will call these transformations which don't correspond to space-time transformations internal symmetries.

3.3.1 $U(1)$ Invariance and Antiparticles

Here is a theory with an internal symmetry:

$$\mathcal{L} = \tfrac{1}{2}\sum_{a=1}^{2}\partial_\mu\phi_a\partial^\mu\phi_a - \mu^2\phi_a\phi_a - g\left(\sum_a(\phi_a)^2\right)^2. \quad (3.44)$$

It is a theory of two scalar fields, ϕ_1 and ϕ_2, with a common mass μ and a potential $g\left(\sum_a(\phi_a)^2\right)^2 = g\left((\phi_1)^2+(\phi_2)^2\right)^2$. This Lagrangian is invariant under the transformation

$$\begin{aligned} \phi_1 &\to \phi_1\cos\lambda + \phi_2\sin\lambda \\ \phi_2 &\to -\phi_1\sin\lambda + \phi_2\cos\lambda. \end{aligned} \quad (3.45)$$

This is just a rotation of ϕ_1 into ϕ_2 in field space. It leaves $\mathcal{L}$ invariant (try it) because $\mathcal{L}$ depends only on $\phi_1^2+\phi_2^2$ and $(\partial_\mu\phi_1)^2+(\partial_\mu\phi_2)^2$, and just as $r^2 = x^2+y^2$ is invariant under real rotations, these are invariant under the transformation (3.45).

We can write this in matrix form:

$$\begin{pmatrix}\phi_1'\\ \phi_2'\end{pmatrix} = \begin{pmatrix}\cos\lambda & \sin\lambda\\ -\sin\lambda & \cos\lambda\end{pmatrix}\begin{pmatrix}\phi_1\\ \phi_2\end{pmatrix}. \quad (3.46)$$

In the language of group theory, this is known as an $SO(2)$ transformation. The S stands for "special", meaning that the transformation matrix has unit determinant, the O for "orthogonal" and the 2 because it's a 2×2 matrix. We say that $\mathcal{L}$ has an $SO(2)$ symmetry.

Once again we can calculate the conserved charge:

$$\begin{aligned} D\phi_1 &= \phi_2 \\ D\phi_2 &= -\phi_1 \\ D\mathcal{L} &= 0 \to F^\mu = \text{constant}. \end{aligned} \quad (3.47)$$

Since F^μ is a constant, we can just forget about it (if J^μ is a conserved current, so is J^μ plus any constant). So the conserved current is

$$J^\mu = \Pi_1^\mu D\phi_1 + \Pi_2^\mu D\phi_2 = (\partial^\mu\phi_1)\phi_2 - (\partial^\mu\phi_2)\phi_1 \quad (3.48)$$

and the conserved charge is

$$Q = \int d^3x J^0 = \int d^3x(\dot\phi_1\phi_2 - \dot\phi_2\phi_1). \quad (3.49)$$

This isn't very illuminating at this stage. At the level of classical field theory, this symmetry isn't terribly interesting. But in the quantized theory it has a very nice interpretation in terms of particles and antiparticles. So let's consider quantizing the theory by imposing the usual equal time commutation relations. At this stage, let's also forget about the potential term in Eq. (3.44). Then we have a theory of two free fields and we can expand the fields in terms of creation and annihilation operators. We will denote the corresponding creation and annihilation operators by $a^\dagger_{ki}$ and a_{ki}, where $i = 1, 2$. They create and destroy two different types of meson, which we denote by

$$a^\dagger_{k1}|\,0\rangle = |\,k,1\rangle, \qquad a^\dagger_{k2}|\,0\rangle = |\,k,2\rangle. \tag{3.50}$$

Substituting the expansion

$$\phi_i = \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}} \left[a_{ki}e^{-ik\cdot x} + a^\dagger_{ki}e^{ik\cdot x}\right] \tag{3.51}$$

into Eq. (3.49) gives, after some algebra,

$$Q = i\int d^3k(a^\dagger_{k1}a_{k2} - a^\dagger_{k2}a_{k1}). \tag{3.52}$$

We are almost there. This looks like the expression for the number operator, except for the fact that the terms are off-diagonal. Let's fix that by defining new creation and annihilation operators which are a linear combination of the old ones:

$$\begin{aligned} b_k &\equiv \frac{a_{k1} + ia_{k2}}{\sqrt{2}}, \;\; b^\dagger_k \equiv \frac{a^\dagger_{k1} - ia^\dagger_{k2}}{\sqrt{2}} \\ c_k &\equiv \frac{a_{k1} - ia_{k2}}{\sqrt{2}}, \;\; c^\dagger_k \equiv \frac{a^\dagger_{k1} + ia^\dagger_{k2}}{\sqrt{2}}. \end{aligned} \tag{3.53}$$

It is easy to verify that the b_k's and c_k's also have the right commutation relations to be creation and annihilation operators. They create linear combinations of states with type 1 and type 2 mesons,

$$b^\dagger_k|\,0\rangle = \frac{1}{\sqrt{2}}(|\,k,1\rangle - i|\,k,2\rangle). \tag{3.54}$$

Linear combinations of states are perfectly good states, so let's work with these as our basis states. We can call them particles of type b and type c

$$b^\dagger_k|\,0\rangle = |\,k,b\rangle, \;\; c^\dagger_k|\,0\rangle = |\,k,c\rangle. \tag{3.55}$$

In terms of our new operators, it is easy to show that

$$\begin{aligned} Q &= i\int d^3k\,(a^\dagger_{k1}a_{k2} - a^\dagger_{k2}a_{k1}) \\ &= \int d^3k\,(b^\dagger_k b_k - c^\dagger_k c_k) \\ &= N_b - N_c \end{aligned} \tag{3.56}$$

where $N_i = \int d^k a^\dagger_{ki}a_{ki}$ is the number operator for a field of type i. The total charge is therefore the number of b's minus the number of c's, so we clearly have b-type particles with charge $+1$ and c-type particles with charge -1. We say that c and b are one another's *antiparticle*: they are the same in all respects except that they carry the opposite conserved charge. Note that we couldn't have a theory with b particles and not c particles: they both came out of the Lagrangian Eq. (3.44). The existence of antiparticles for all particles carrying a conserved charge is a generic prediction of QFT.

Now, that was all a bit involved since we had to rotate bases in midstream to interpret the conserved charge. With the benefit of hindsight we can go back to our original Lagrangian and write it in terms of the *complex* fields

$$\begin{aligned} \psi &\equiv \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \\ \psi^\dagger &\equiv \frac{1}{\sqrt{2}}(\phi_1 - i\phi_2). \end{aligned} \tag{3.57}$$

In terms of ψ and $\psi^\dagger$, $\mathcal{L}$ is

$$\mathcal{L} = \partial_\mu\psi^\dagger\partial^\mu\psi - \mu^2\psi^\dagger\psi \tag{3.58}$$

(note that there is no factor of $\frac{1}{2}$ in front). In terms of creation and annihilation operators, ψ and $\psi^\dagger$ have the expansions

$$\begin{aligned} \psi &= \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}} \left(b_k e^{-ik\cdot x} + c^\dagger_k e^{ik\cdot x}\right) \\ \psi^\dagger &= \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}} \left(c_k e^{-ik\cdot x} + b^\dagger_k e^{ik\cdot x}\right) \end{aligned} \tag{3.59}$$

so ψ creates c-type particles and annihilates their antiparticle b, whereas $\psi^\dagger$ creates b-type particles and annihilates c's. Thus ψ always changes the Q of a state by -1 (by creating a c or annihilating a b in the state) whereas $\psi^\dagger$ acting on a state increases the charge by one. We can also see this from the commutator $[Q,\psi]$: from the expression for the conserved charge Eq. (3.56) it is easy to show that

$$[Q,\psi] = -\psi, \;\; [Q,\psi^\dagger] = \psi^\dagger. \tag{3.60}$$

If we have a state $|\,q\rangle$ with charge q (that is, $|\,q\rangle$ is an eigenstate of the charge operator Q with eigenvalue q), then

$$Q(\psi|\,q\rangle) = [Q,\psi]|\,q\rangle + \psi Q|\,q\rangle = (-1+q)\psi|\,q\rangle \tag{3.61}$$

so $\psi|\,q\rangle$ has charge $q-1$, as we asserted.

The transformation Eq. (3.45) may be written as

$$\psi' = e^{-i\lambda}\psi. \tag{3.62}$$

This is called a $U(1)$ transformation or a phase transformation (the "U" stands for "unitary".) Clearly a $U(1)$ transformation on complex fields is equivalent to an $SO(2)$ transformation on real fields, and is somewhat simpler to work with. In fact, we can now work from our ψ fields right from the start. In terms of the classical fields, start with the Lagrangian

$$\mathcal{L} = \partial_\mu\psi^*\partial^\mu\psi - \mu^2\psi^*\psi \tag{3.63}$$

(these are classical fields, not operators, so the complex conjugate of ψ is ψ^*, not $\psi^\dagger$.) We can quantize the theory correctly and obtain the equations of motion if we follow the same rules as before, but treat ψ and ψ^* as *independent* fields. That is, we vary them independently and assign a conjugate momentum to each:

$$\Pi^\mu_\psi = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)},\ \ \Pi^\mu_{\psi^*} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi^*)}. \tag{3.64}$$

Therefore we have

$$\Pi^\mu_\psi = \partial^\mu\psi^*,\ \ \Pi^\mu_{\psi^*} = \partial^\mu\psi \tag{3.65}$$

which leads to the Euler-Lagrange equations

$$\partial_\mu\Pi^\mu_\psi = \frac{\partial\mathcal{L}}{\partial\psi} \to (\Box+\mu^2)\psi^* = 0. \tag{3.66}$$

Similarly, we find $(\Box+\mu^2)\psi = 0$. Adding and subtracting these equations, we clearly recover the equations of motion for ϕ_1 and ϕ_2.

We can similarly canonically quantize the theory by imposing the appropriate commutation relations

$$[\psi(\vec{x},t),\Pi^0_\psi(\vec{y},t)] = i\delta^{(3)}(\vec{x}-\vec{y}),\ \ [\psi^\dagger(\vec{x},t),\Pi^0_{\psi^\dagger}(\vec{y},t)] = i\delta^{(3)}(\vec{x}-\vec{y}),.... \tag{3.67}$$

We will leave it as an exercise to show that this reproduces the correct commutation relations for the ϕ fields and their conjugate momenta.

Clearly ψ and ψ^* are not independent. Still, this rule of thumb works because there are two real degrees of freedom in ϕ_1 and ϕ_2, and two real degrees of freedom in ψ, which may be independently varied. We can see how this works to give us the correct equations of motion. Consider the Euler-Lagrange equations for a general theory of a complex field ψ. For a variation in the fields $\delta\psi$ and $\delta\psi^*$, we find an expression for the variation in the action of the form

$$\delta S = \int d^4x(A\delta\psi + A^*\delta\psi^*) = 0 \tag{3.68}$$

where A is some function of the fields. The correct way to obtain the equations of motion is to first perform a variation $\delta\psi$ which is purely real, $\delta\psi = \delta\psi^*$. This gives the Euler-Lagrange equation

$$A + A^* = 0. \tag{3.69}$$

Then performing a variation $\delta\psi$ which is purely imaginary, $\delta\psi = -\delta\psi^*$, gives $A - A^* = 0$. Combining the two, we get $A = A^* = 0$.

If we instead apply our rule of thumb, we imagine that ψ and ψ^* are unrelated, so we can vary them independently. We first take $\delta\psi^* = 0$ and from Eq. (3.68) get $A = 0$. Then taking $\delta\psi = 0$ we get $A^* = 0$. So we get the same equations of motion, $A = A^* = 0$.

We will refer to complex fields as "charged" fields from now on. Note that since we haven't yet introduced electromagnetism into the theory the fields aren't charged in the usual electromagnetic sense; "charged" only indicates that they carry a conserved $U(1)$ quantum number. A better analogue of the "charge" in this theory is baryon or lepton number. Later on we will show that the only consistent way to couple a matter field to the electromagnetic field is for the interaction to couple a conserved $U(1)$ charge to the photon field, at which point the $U(1)$ charge will correspond to electric charge.

3.3.2 Non-Abelian Internal Symmetries

A theory with a more complicated group of internal symmetries is

$$\mathcal{L} = \tfrac{1}{2}\sum_{a=1}^{n}\Big(\partial_\mu\phi_a\partial^\mu\phi_a - \mu^2\phi_a\phi_a\Big) - g\left(\sum_{a=1}^{n}(\phi_a)^2\right)^2. \tag{3.70}$$

This is the same as the previous example except that we have n fields instead of just two. Just as in the first example the Lagrangian was invariant under rotations mixing up ϕ_1 and ϕ_2, this Lagrangian is invariant under rotations mixing up $\phi_1...\phi_n$, since it only depends on the "length" of $(\phi_1,\phi_2,...,\phi_n)$. Therefore the internal symmetry group is the group of rotations in n dimensions,

$$\phi_a \to \sum_b R_{ab}\phi_b \tag{3.71}$$

where R_{ab} is an $n\times n$ rotation matrix. There are $n(n-1)/2$ independent planes in n dimensions, and we can rotate in each of them, so there are $n(n-1)/2$ conserved currents and associated charges. This example is quite different from the first one because the various rotations don't in general commute - the group of rotations in $n>2$ dimensions is nonabelian. The group of rotation matrices in n dimensions is called $SO(n)$ (Special, Orthogonal, n dimensions), and this theory has an $SO(n)$ symmetry. A new feature of nonabelian symmetries is that, just as the rotations don't in general commute, neither do the currents or charges in the quantum theory. For example, for a theory with $SO(3)$ invariance, the currents are

$$\begin{aligned} J^{\mu}_{[1,2]} &= (\partial^{\mu}\phi_1\phi_2) - (\partial^{\mu}\phi_2\phi_1) \\ J^{\mu}_{[1,3]} &= (\partial^{\mu}\phi_1\phi_3) - (\partial^{\mu}\phi_3\phi_1) \\ J^{\mu}_{[2,3]} &= (\partial^{\mu}\phi_2\phi_3) - (\partial^{\mu}\phi_3\phi_2) \end{aligned} \tag{3.72}$$

and in the quantum theory the (appropriately normalized) charges obey the commutation relations

$$\begin{aligned} &[Q_{[2,3]}, Q_{[1,3]}] = iQ_{[1,2]} \\ &[Q_{[1,3]}, Q_{[1,2]}] = iQ_{[2,3]} \\ &[Q_{[2,3]}, Q_{[1,2]}] = iQ_{[1,3]} \end{aligned} \tag{3.73}$$

This means that it not possible to simultaneously measure more than one of the $SO(3)$ charges of a state: the charges are non-commuting observables.

For n complex fields with a common mass,

$$\mathcal{L} = \sum_{a=1}^{n}\left(\partial_{\mu}\psi_a^*\partial^{\mu}\psi_a - \mu^2\psi_a^*\psi_a\right) - g\left(\sum_{a=1}^{n}|\psi_a|^2\right)^2 \tag{3.74}$$

the theory is invariant under the group of transformations

$$\psi_a \to \sum_b U_{ab}\psi_b \tag{3.75}$$

where U_{ab} is any unitary $n\times n$ matrix. We can write this as a product of a $U(1)$ symmetry, which is just multiplication of each of the fields by a common phase, and an $n\times n$ unitary matrix with unit determinant, a so-called $SU(n)$ matrix. The symmetry group of the theory is the direct product of these transformations, or $SU(n)\times U(1)$.

We won't be discussing non-Abelian symmetries much in the course, but we just note here that there are a number of non-Abelian symmetries of importance in particle physics. The familiar isospin symmetry of the strong interactions is

an $SU(2)$ symmetry, and the charges of the strong interactions correspond to an $SU(3)$ symmetry of the quarks (as compared to the $U(1)$ charge of electromagnetism). The charges of the electroweak theory correspond to those of an $SU(2)\times U(1)$ symmetry group. "Grand Unified Theories" attempt to embed the observed strong, electromagnetic and weak charges into a single symmetry group such as $SU(5)$ or $SO(10)$. We could proceed much further here into group theory and representations, but then we'd never get to calculate a cross section. So we won't delve deeper into non-Abelian symmetries at this stage.

3.4 Example: Non-Relativistic Quantum Mechanics ("Second Quantization")

To put some flesh on the formalism we have developed so far, let's pause and work through an example. The following problem was used as a midterm test the first time I taught this course (with rather bleak results ...). In the following years I gave it as a problem set. I suggest you work through it before looking at the solution.

The Problem

Consider a theory of a complex scalar field ψ

$$\mathcal{L}_0 = i\psi^*\partial_0\psi + b\vec{\nabla}\psi^* \cdot \vec{\nabla}\psi,$$

where b is some real number (this Lagrange density is not real, but that's all right: the action integral is real). As the investigation proceeds, you should recognize this theory as good old non-relativistic quantum mechanics. Treating the theory in this manner is called *second quantization*, and is a useful formalism for studying multi-particle quantum mechanics.

1. Consider $\mathcal{L}_0$ as defining a classical field theory. Find the Euler-Lagrange equations. Find the plane-wave solutions, those for which $\psi = e^{i(\vec{k}\cdot\vec{x}-\omega t)}$, and find ω as a function of k. Although this theory is not Lorentz-invariant, it is invariant under space-time translations and an internal $U(1)$ symmetry transformation. Thus it possesses a conserved energy, a conserved linear momentum and a conserved charge associated with the internal symmetry. Find these quantities as integrals of the fields and their derivatives. Fix the sign of b by demanding the energy be bounded below. (As explained in class, in dealing with complex fields, you just turn the crank, ignoring the fact that ψ and ψ^* are complex conjugates. Everything should turn out all right in the end: the equation of motion for ψ will be the complex conjugate of that for ψ^*, and the conserved quantities will all be real.) (WARNING: Even though this is a non-relativistic problem, our formalism is set up with relativistic conventions; don't miss minus signs associated with raising and lowering spatial indices.)

2. Canonically quantize the theory. (HINT: You may be bothered by the fact that the momentum conjugate to ψ^* vanishes. Don't be. Because the equations of motion are first-order in time, a complete and independent set of initial-value data consists of ψ and its conjugate momentum alone. It is only on these that you need to impose the canonical quantization conditions.) Identify appropriately normalized coefficients in the expansion of the fields in terms of plane wave solutions with annihilation and/or creation operators, and write the energy, linear momentum and internal-symmetry charge in terms of these operators. (Normal-order freely.) Find the equation of motion for the single particle state $|\vec{k}\rangle$ and the two particle state $|\vec{k}_1, \vec{k}_2\rangle$ in the Schrödinger Picture. What physical quantities do b and the internal symmetry charge correspond to?

Solution

1. The Euler Lagrange equations are

$$\partial_\mu \Pi^\mu_a = \frac{\partial \mathcal{L}}{\partial \phi_a} \tag{3.76}$$

so we first need the momenta conjugate to the fields. Treating ψ and ψ^* as independent fields, we find

$$\Pi^0_\psi = \frac{\partial \mathcal{L}}{\partial(\partial_0\psi)} = i\psi^*, \qquad \Pi^i_\psi = \frac{\partial \mathcal{L}}{\partial(\partial_i\psi)} = -b\partial^i\psi^*$$
$$\Pi^0_{\psi^*} = \frac{\partial \mathcal{L}}{\partial(\partial_i\psi^*)} = 0, \qquad \Pi^i_{\psi^*} = \frac{\partial \mathcal{L}}{\partial(\partial_i\psi^*)} = b\partial^i\psi. \tag{3.77}$$

Thus, the equations of motion for the two fields are

$$\begin{aligned} i\partial_0\psi &= b\nabla^2\psi \\ i\partial_0\psi^* &= -b\nabla^2\psi^* \end{aligned} \tag{3.78}$$

(note that, as required, the equations of motion are conjugates of each other. This is actually ensured by the fact that the action is real.) This is a wave equations for ψ; expanding in normal modes

$$\psi = e^{i(\vec{k}\cdot\vec{x}-\omega_k t)}, \tag{3.79}$$

the equations of motion gives the dispersion relation

$$\omega_k = -b|\vec{k}|^2. \tag{3.80}$$

The internal $U(1)$ symmetry is (of course)

$$\psi \to e^{-i\lambda}\psi, \quad \psi^* \to e^{i\lambda}\psi^*. \tag{3.81}$$

Recall that the conserved current is given in general by

$$J^\mu = \sum_a \Pi^\mu_a D\psi_a - F^\mu. \tag{3.82}$$

In our case, $F^\mu = 0$ (or equivalently a constant), since $D\mathcal{L} = 0$. We also have $D\psi = -i\psi$. Hence, the conserved charge density is the time component of J^μ

$$J^0 = \psi^*\psi \tag{3.83}$$

and the conserved charge Q is the integral of this quantity over all space,

$$Q = \int d^3x\, J^0 = \int d^3x\, \psi^*\psi. \tag{3.84}$$

For the invariance under space–time translations $\psi(x) \to \psi(x + \lambda_\mu a^\mu)$, where a^μ is and arbitrary four vector (unit vector), we find

$$D\psi = a^\mu \partial_\mu \psi, \qquad F^\mu = a^\mu \mathcal{L}. \tag{3.85}$$

Therefore

$$\begin{aligned} J^0 &= \Pi^0 D\psi - a^0\mathcal{L} \\ &= i\psi^* a^\mu \partial_\mu \psi - a^0 i\psi^*\partial_0\psi - a^0 b\vec{\nabla}\psi^* \cdot \vec{\nabla}\psi \\ &= -i\psi^*(\vec{a}\cdot\vec{\nabla})\psi - a^0 b\vec{\nabla}\psi^*\cdot\vec{\nabla}\psi. \end{aligned}$$

For a time translation: $a^0 = 1$, $a^i = 0$,

$$H = E = \int d^3x[-b\vec{\nabla}\psi^*\cdot\vec{\nabla}\psi] = -b\int d^3x\, |\nabla\psi|^2\,.$$

For this energy to be bounded from below, we need $b < 0$.

For a space translation: $a^0 = 0$, $a^i = \vec{x}$,

$$P^i = \int d^3x[i\psi^*\partial^i\psi]$$

It's easy to see that both the energy and momentum are Hermitian.

2. Since the momentum conjugate to ψ^* vanishes, the only surviving equal time commutation relation to impose is on ψ and its conjugate, $i\psi^*$. Cancelling the i's, we get

$$[\psi(\vec{x},t), \psi^*(\vec{x},t)] = \delta(\vec{x}-\vec{y}),$$

and

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

Now expand the fields in the plane wave solutions given in part (1) to get

$$\begin{aligned} \psi(\vec{x},t) &= \int d^3k A_k e^{i(\vec{k}\cdot\vec{x}-\omega t)} \\ \psi^*(\vec{y},t) &= \int d^3k B_k e^{-i(\vec{k}\cdot\vec{y}-\omega t)} \end{aligned}$$

Assume that $A_k = \alpha a_k$ and $B_k = \alpha a_k^\dagger$, then

$$\begin{aligned} [\psi(\vec{x},t), \psi^*(\vec{y},t)] &= \int d^3k \int d^3k' e^{i(\vec{k}\cdot\vec{x}-\omega t)} e^{-i(\vec{k}'\cdot\vec{y}-\omega t)}\, [A_k, B_k] \\ &= \int d^3k \int d^3k' e^{i(\vec{k}\cdot\vec{x}-\omega t)} e^{-i(\vec{k}'\cdot\vec{y}-\omega t)} \alpha^2\delta(\vec{k}-\vec{k}') \\ &= \alpha^2 \int d^3k e^{i\vec{k}\cdot(\vec{x}-\vec{y})} \\ &= (2\pi)^3\alpha^2\delta(\vec{x}-\vec{y}) \end{aligned}$$

This means that $\alpha = \frac{1}{(2\pi)^{3/2}}$ and $A_k = \frac{1}{(2\pi)^{3/2}}a_k$ is an annihilation operator, whereas $B_k = \frac{1}{(2\pi)^{3/2}}a_k^\dagger$ is a creation operator. The field ψ therefore only annihilates a particle and ψ^* only creates particles.

Now we can go ahead and write the energy, the momentum and the internal symmetry charge in terms of these creation and annihilation operators. We find

$$\begin{aligned} E &= \int d^3x[-b\vec{\nabla}\psi^*\cdot\vec{\nabla}\psi] \\ &= -b\frac{1}{(2\pi^3)}\int d^3x \int d^3k\, d^3k'\, a_k^\dagger e^{-i(\vec{k}\vec{x}-\omega t)} a_{k'} e^{i(\vec{k}'\cdot\vec{x}-\omega' t)}\vec{k}\cdot\vec{k}' \\ &= -b\frac{1}{(2\pi^3)}\int d^3k\, d^3k'\, a_k^\dagger a_{k'} e^{-i\omega t} e^{i\omega' t}\vec{k}\cdot\vec{k}'(2\pi)^3\delta(\vec{k}-\vec{k}') \\ &= -b\int d^3k\, a_k^\dagger a_k |\vec{k}|^2 \end{aligned}$$

Similarly we find

$$\begin{aligned} Q &= \int d^3k\, a_k^\dagger a_k \\ P^i &= \int d^3k\, k^i a_k^\dagger a_k \end{aligned}$$

This form for the momentum operator is to be expected, since $a_k^\dagger a_k$ is the usual number operator. The Hamiltonian acting on a one-particle state is therefore

$$: H : |\vec{k}\rangle = -b|\vec{k}|^2|\vec{k}\rangle$$

and on the two-particle state is

$$: H : |\vec{k}_1, \vec{k}_2\rangle = -b\left(|\vec{k}_1|^2 + |\vec{k}_2|^2\right)|\vec{k}_1, \vec{k}_2\rangle$$

(this is straightforward to show using the commutation relations of the creation and annihilation operators in the usual way). This clearly corresponds to the usual energy of one- and two-particle states if $b = -1/2m$. The equations of motion for these states in the Schrödinger picture are therefore

$$i\frac{\partial}{\partial t}|\vec{k}\rangle = \frac{|\vec{k}|^2}{2m}|\vec{k}\rangle$$

and

$$i\frac{\partial}{\partial t}|\vec{k}_1, \vec{k}_2\rangle = \frac{1}{2m}\left(|\vec{k}_1|^2 + |\vec{k}_2|^2\right)|\vec{k}_1, \vec{k}_2\rangle.$$

This is just the usual EOM for one- and two-particle states in NRQCD.

The conserved charge

$$Q = \int d^3k\, a_k^\dagger a_k$$

is just the number operator. This is a conserved quantity in a nonrelativistic theory, since particle creation is a relativistic effect.

4 Interacting Fields

In this section we will put the formalism we have spend the past few lectures deriving to work. Although we have been talking about symmetries of general (possibly very complicated) Lagrangians, the only equation of motion we have solved is the Klein-Gordon equation, which is just a theory of free fields. For a real field, ϕ, we had

$$\mathcal{L}_\phi = \tfrac{1}{2}(\partial_\mu\phi\partial^\mu\phi - \mu^2\phi^2) \tag{4.1}$$

and for a complex field ψ we had

$$\mathcal{L}_\psi = \partial_\mu\psi^\dagger\partial^\mu\psi - m^2\psi^\dagger\psi. \tag{4.2}$$

Because we could solve the Klein-Gordon equation, we could expand the fields as sums of plane waves multiplied by creation and annihilation operators,

$$\begin{aligned}
\phi(x) &= \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}}\left[a_k e^{-ik\cdot x} + a_k^\dagger e^{ik\cdot x}\right] \\
\psi(x) &= \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}}\left[b_k e^{-ik\cdot x} + c_k^\dagger e^{ik\cdot x}\right] \\
\psi^\dagger(x) &= \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}}\left[c_k e^{-ik\cdot x} + b_k^\dagger e^{ik\cdot x}\right].
\end{aligned} \tag{4.3}$$

We have expressions for the energy, momentum and $U(1)$ charge in our theory, but it is incredibly dull because nothing happens. We just have plane waves propagating. In the quantum theory, as we have seen, this corresponds to a theory of noninteracting, spinless bosons. $\mathcal{L} = \mathcal{L}_\phi + \mathcal{L}_\psi$ is a theory of ϕ particles and ψ particles, but they never interact because the two Lagrangians are decoupled. We can make things more interesting by adding interaction terms to the Lagrangian.

4.1 Particle Creation by a Classical Source[5]

The simplest type of interaction we can introduce into the theory is to couple the ϕ field to a classical source:

$$\mathcal{L} = \mathcal{L}_\phi - \rho(x)\phi(x) \tag{4.4}$$

where $\rho(x)$ is some fixed, known function of space and time which is only nonzero for a finite time interval. This leads to the equation of motion

$$\partial_\mu\partial^\mu\phi + \mu^2\phi = -\rho(x). \tag{4.5}$$

[5]See Peskin & Schroeder, pp. 32–33.

To realize why $\rho(x)$ is a source term, recall from classical electromagnetism that in the presence of a charge distribution $\varrho(\vec{x},t)$ and a current $\vec{\jmath}(\vec{x},t)$ the potentials obey the inhomogeneous wave equations

$$\begin{aligned} \nabla^2\varphi - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} &= -4\pi\varrho \\ \nabla^2\vec{A} - \frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} &= -\frac{4\pi}{c}\vec{\jmath}. \end{aligned} \tag{4.6}$$

φ and $\vec{A}$ form the components of the four-vector $A^\mu = (\varphi, \vec{A})$, so in four-vector notation we may write this as

$$\partial_\mu\partial^\mu A^\nu = 4\pi J^\nu \tag{4.7}$$

where $J^\mu = (\varrho, \vec{\jmath})$ is the 4-current. Except for the fact that ϕ is massive, has no vector index and is a quantum field, these two theories look quite similar, so we may interpret $\rho(x)$ as a source for the ϕ field, just as a charge distribution is a source for electric field.

This theory is actually simple enough that we can solve it exactly. If we start in the vacuum state, what will we find at some time in the far future, after the source $\rho(x)$ has been turned on and off again? We can answer this by solving the field equations directly.

Before $\rho(x)$ is turned on, the theory is free, and $\phi_0(x)$ may be expanded in terms of creation and annihilation operators, as in Eq. (4.3). After the source has turned on, we can construct the solution to the equation of motion as follows:

$$\phi(x) = \phi_0(x) + i\int d^4y\, D_R(x-y)\,\rho(y) \tag{4.8}$$

where $D_R(x-y)$ is the *retarded Green function*, satisfying

$$\begin{aligned} (\partial_\mu\partial^\mu + \mu^2)D_R(x-y) &= -i\delta^{(4)}(x-y) \\ D_R(x-y) &= 0, \;\; x^0 < y^0. \end{aligned} \tag{4.9}$$

The second requirement, that D_R be the retarded Green function, is required so that the boundary condition $\phi(x) \to \phi_0(x)$ as $x_0 \to -\infty$ is satisfied.

The simplest way to find the Green function is to rewrite Eq. (4.9) in momentum space. Writing

$$D_R(x-y) = \int \frac{d^4k}{(2\pi)^4} e^{-ik\cdot(x-y)}\,\tilde{D}_R(k) \tag{4.10}$$

we find the algebraic equation for $\tilde{D}_R(k)$,

$$(-k^2+\mu^2)\tilde{D}_R(k) = -i \tag{4.11}$$

which immediately gives us

$$D_R(x-y) = \int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2-\mu^2}\,e^{-ik\cdot(x-y)}. \tag{4.12}$$

This doesn't quite define D_R: the k^0 integrand in Eq. (4.12) has poles at $k^0 = \pm\omega_k$. In order to define the integral, we must choose a path of integration around the poles. Let us choose a path of integration which passes above both poles. Then

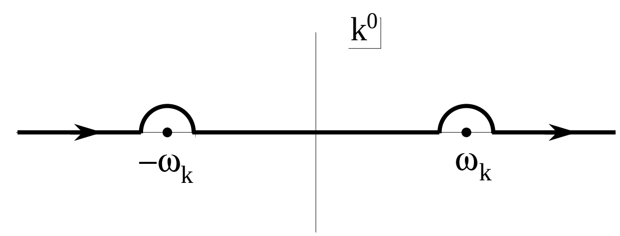

Figure 4.1: The contour defining $D_R(x-y)$.

for $y_0 > x_0$ we can close the contour in the upper half plane, giving zero for the integral since the path of integration doesn't enclose any singularities. Thus, the Green function vanishes for $y_0 > x_0$, making this the appropriate contour for $D_R(x-y)$. For $x_0 > y_0$, we can close the contour in the bottom half plane, obtaining for the integral

$$\begin{aligned} D_R(x-y) &\overset{x_0\geq y_0}{=} \int\frac{d^3k}{(2\pi)^3}\left[\frac{1}{2\omega_k}e^{-ik\cdot(x-y)}\Big|_{k^0=\omega_k} + \frac{1}{-2\omega_k}e^{-ik\cdot(x-y)}\Big|_{k^0=-\omega_k}\right] \\ &= \int\frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}\left(e^{-ik\cdot(x-y)} - e^{ik\cdot(x-y)}\right) \\ &= i\left[\Delta_+(x-y) - \Delta_+(y-x)\right] \\ &= [\phi(x),\phi(y)] \end{aligned} \tag{4.13}$$

where the functions $\Delta_+(x)$ and $\Delta_-(x)$ were introduced in Section 2. The retarded Green function $D_R(x-y)$ is therefore related to the commutator of two fields, or equivalently (since the commutator is a c-number, not an operator), the vacuum expecation value of the commutator:

$$\begin{aligned} D_R(x-y) &= \theta(x_0-y_0)[\phi(x),\phi(y)] \\ &= \theta(x_0-y_0)\langle 0\,|[\phi(x),\phi(y)]|\,0\rangle. \end{aligned} \tag{4.14}$$

For our present purposes, we only need the second line in Eq. (4.13). Inserting this expression into Eq. (4.8) gives

$$\begin{aligned} \phi(x) &= \phi_0(x) + i\int d^4y \int \frac{d^3k}{(2\pi)^3 2\omega_k}\theta(x_0-y_0)\left(e^{-ik\cdot(x-y)} - e^{-ik\cdot(x-y)}\right)\rho(y) \\ &\overset{x_0\to\infty}{=} \phi_0(x) + i\int \frac{d^3k}{(2\pi)^3 2\omega_k}\int d^4y \left(e^{-ik\cdot(x-y)} - e^{ik\cdot(x-y)}\right)\rho(y) \\ &= \phi_0(x) + i\int \frac{d^3k}{(2\pi)^3 2\omega_k}\left(e^{-ik\cdot x}\tilde\rho(k) - e^{-ik\cdot x}\tilde\rho(-k)\right) \end{aligned} \tag{4.15}$$

where in the second line we have used the fact that if we wait until all of $\rho(x)$ is in the past, the theta function equals one over the whole domain of integration and may be dropped. We have also defined the Fourier transform

$$\tilde\rho(k) = \int d^4y\, e^{ik\cdot y}\rho(y). \tag{4.16}$$

Thus we find, after the source has been turned off,

$$\phi(x) = \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}}\left\{\left(a_k + \frac{i}{(2\pi)^{3/2}\sqrt{2\omega_k}}\tilde\rho(k)\right)e^{-ik\cdot x} + \text{h.c.}\right\}. \tag{4.17}$$

Since all observables are built out of the fields, we have solved the theory. The Hamiltonian in the far future is now

$$H = \int d^3k\,\omega_k\left(a_k^\dagger - \frac{i}{(2\pi)^{3/2}\sqrt{2\omega_k}}\tilde\rho^*(k)\right)\left(a_k + \frac{i}{(2\pi)^{3/2}\sqrt{2\omega_k}}\tilde\rho(k)\right) \tag{4.18}$$

(this is obvious if you go back to the original derivation of H in terms of $\phi(x)$) and so the expectation value of the energy of the system in the far future is

$$\langle 0\,|H|\,0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2}|\tilde\rho(k)|^2. \tag{4.19}$$

Note that because we are in the Heisenberg representation, we are still in the ground state of the free theory – the state hasn't evolved. The time evolution of the system is all contained in the evolution of the fields. Now, since in the far future we

have free field theory again, the spectrum of the Hamiltonian is just free particles, which means that the expectation value of the total number of particles created with momentum k is

$$dN(\vec k) = \frac{|\tilde\rho(k)|^2}{(2\pi)^3 2\omega_k} \tag{4.20}$$

and so each Fourier component of ρ produces particles with the corresponding four-momentum with a probability proportional to $|\tilde\rho(k)|^2$. The expectation value of the total number of particles produced is

$$\int dN = \int \frac{d^3k}{(2\pi)^3 2\omega_k}|\tilde\rho(k)|^2. \tag{4.21}$$

4.2 More on Green Functions

Since Green functions are of central importance to scattering theory, let's pause for a moment and study the expression (4.12) a bit more. The retarded Green function $D_R(x-y)$ was obtain by choosing the path of integration shown in Fig. (4.1). Other paths of integration give Green functions which are useful for solving problems with different boundary conditions. Choosing a path of integration which passes below both poles would give the advanced Green function, obeying $G_A(x-y)=0$ for $x_0 > y_0$. This would be useful if we knew the value of the field in the far future and were interested in its value before the source was turned on. Another possibility is a path which goes below the pole at $-\omega_k$ and above the pole at ω_k. In this case, when $x_0 > y_0$ we perform the k^0 integral by closing the contour below, obtaining the result $i\Delta_+(x-y)$ for the integral. When $x_0 < y_0$ we close the

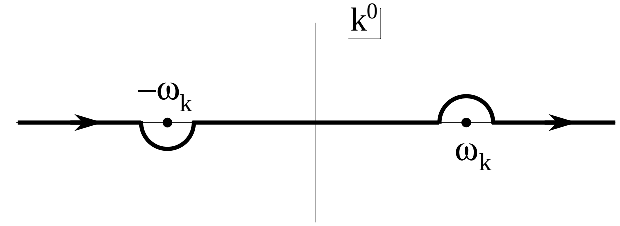

Figure 4.2: The contour defining $D_F(x-y)$.

contour above, obtaining the same expression but with x and y interchanged. This defines the Green function

$$D_F(x-y) = \begin{cases} i\Delta_+(x-y), & x_0 > y_0; \\ i\Delta_+(y-x), & x_0 < y_0. \end{cases}$$

$$= \theta(x^0 - y^0)\langle 0\,|\phi(x)\phi(y)|\,0\rangle + \theta(y_0 - x_0)\langle 0\,|\phi(y)\phi(x)|\,0\rangle$$
$$\equiv \langle 0\,|T\phi(x)\phi(y)|\,0\rangle \tag{4.22}$$

where the last line defines the *time ordering* symbol T, which instructs us to place the operators that follow in order with the latest on the left. This Green function, called the *Feynman propagator*, will be of central importance to scattering theory, and we shall return to it shortly. It is convenient to write the Feynman propagator as

$$D_F(x-y) = \int \frac{d^4k}{(2\pi)^4}\frac{i}{k^2-\mu^2+i\epsilon}e^{-ik\cdot(x-y)} \tag{4.23}$$

where the limit $\epsilon \to 0^+$ is understood and the path of integration in the k^0 plane is now along the real axis, since the poles are then at $k^0 = \pm(\omega_k - i\epsilon)$ and are displaced properly above and below the real axis. Note that the sign of the $i\epsilon$ term is crucial: if ϵ were negative, the contours would enclose the opposite poles, and the time ordering would come out reversed.

4.3 Mesons Coupled to a Dynamical Source

The Lagrangian Eq. (4.4) is analogous to electromagnetism coupled to a current which is unaffected by the dynamics of the field. While this is in many cases a good approximation, in the real world the current itself interacts with the electromagnetic field, and the resulting dynamics are quite complicated. For scalar field theory, the analogous situation is described by a potential which couples the two fields ψ and ϕ:

$$\mathcal{L} = \mathcal{L}_\phi + \mathcal{L}_\psi - g\psi^\dagger\psi\phi. \tag{4.24}$$

Note that the potential only depends on ψ and $\psi^\dagger$ in the combination $\psi^\dagger\psi$, so the interaction term doesn't break the $U(1)$ symmetry. We are therefore guaranteed that the interacting theory will also conserve charge. Furthermore, the interaction depends only on the fields, not their derivatives, so the conjugate momenta are the same as they were in the free theory. This equations of motion are

$$\partial_\mu\partial^\mu\phi + \mu^2\phi = -g\psi^\dagger\psi,$$
$$\partial_\mu\partial^\mu\psi + m^2\psi = -g\psi\phi. \tag{4.25}$$

The field equations are now coupled, so the fields interact. In fact, comparing this with Eq. (4.4), we see that $\psi^\dagger\psi$ is a current density, a source for the ϕ field, just like $\rho(x)$. This model is much more complicated that the previous one, however, because there is a back-reaction: the current $\psi^\dagger\psi$ in turn depends on the field ϕ. The source is now not a prescribed function of space-time, as it was in the previous case, but a full dynamical variable, so solving this theory is going to be

much harder. In general we cannot solve this system of coupled nonlinear partial differential equations exactly. Instead, we will have to solve them perturbatively: that is, if g is small we can treat the interaction term as a small perturbation of free field theory. We will be able to solve the equations of motion as a power series in g.[6] In fact, most of the rest of this course will be concerned with applying perturbation theory to an assortment of different theories. Much of what is known about quantum field theory comes from perturbation theory.

The theory defined in Eq. (4.24) describes the interactions of two types of meson, one of which carries a conserved charge. This doesn't look anything like the particles we see in the real world, but we will use it in this section as a toy model to illustrate our perturbative approach to scattering theory. However, we have seen that the equations of motion look quite similar to the equations of motion of an electric field coupled to a current. If the ψ fields were spin $1/2$ fermions instead of spin 0 bosons we would have a theory of the strong interactions between nucleons, where the force is transmitted through the exchange of ϕ mesons. We'll take advantage of this analogy and refer to the ψ particles as "nucleons" (in quotation marks) and the ϕ's as mesons. We'll call this our "nucleon"-meson theory.

4.4 The Interaction Picture

How do we set this problem up? First of all, we would like to make use of some of our previous results for free field theories. In particular, we would like to be able to write our fields in terms of creation and annihilation operators, because in this form we know exactly how the fields act on the states of the theory. Unfortunately, the solution to the Heisenberg equations of motion are no longer plane waves but instead something awful. We can fix this with a clever trick called the *interaction picture*.

We have already discussed the Schrödinger and Heisenberg pictures. The interaction picture combines elements of each. All three pictures will coincide at $t = 0$:

$$|\,\psi(0)\rangle_S = |\,\psi(0)\rangle_H = |\,\psi(0)\rangle_I$$
$$O_S(0) = O_H(0) = O_I(0) \tag{4.26}$$

where the subscript I refers to the interaction picture, and O represents a generic operator with no explicit t dependence.

Recall that in the Schrödinger picture, the operators don't evolve with time,

[6]Since g has dimensions of mass, the power series will actually be a series in g/M, where M is some typical mass or energy in the problem.

and the t dependence is carried entirely by the states

$$\begin{aligned} O_S(t) &= O_S(0) \\ i\frac{d}{dt}|\,\psi(t)\rangle_S &= H|\,\psi(t)\rangle_S, \end{aligned} \tag{4.27}$$

while in the Heisenberg picture the states are independent of time and the operators (and in particular, the fields) carry the time dependence

$$\begin{aligned} |\,\psi(t)\rangle_H &= |\,\psi(0)\rangle_H \\ i\frac{d}{dt}|\,O_H(t)\rangle &= [O_H(t), H]. \end{aligned} \tag{4.28}$$

We showed earlier that matrix elements are the same in the two pictures.

In the interaction picture (IP) we split the Hamiltonian up into two pieces,

$$H = H_0 + H_I \tag{4.29}$$

where H_0 is the free Hamiltonian (that is, the Hamiltonian corresponding to the free-field Lagrangian), and H_I contains the interaction term. Since

$$H = \int d^3x \sum_a \Pi_a^0 \dot{\phi}_a - L = \int d^3x \sum_a \Pi_a^0 \dot{\phi}_a - L_0 - L_I, \tag{4.30}$$

if $\mathcal{L}_I$ contains no derivatives of the fields (so it doesn't change the conjugate momenta from the free theory), we see immediately that

$$H_I = -L_I. \tag{4.31}$$

In our example, $\mathcal{H}_I = -\mathcal{L}_I = g\psi^\dagger\psi\phi$. States in the I.P. are defined by

$$|\,\psi(t)\rangle_I \equiv e^{iH_0 t}|\,\psi(t)\rangle_S. \tag{4.32}$$

If we were dealing with a free field theory, $H_I = 0$, this would immediately give $|\,\psi(t)\rangle_I = |\,\psi(t)\rangle_H = |\,\psi(0)\rangle_H$ and the states would be independent of time, just as in the Heisenberg picture.

Demanding that matrix elements be identical in all three pictures, we find

$${}_S\langle\psi(t)\,|O_S|\,\psi(t)\rangle_S = {}_I\langle\psi(t)\,|O_I(t)|\,\psi(t)\rangle_I = {}_S\langle\psi(t)\,|e^{-iH_0 t}O_I(t)e^{iH_0 t}|\,\psi(t)\rangle_S \tag{4.33}$$

and so in the I.P. the operators evolve according to the free Hamiltonian:

$$O_I(t) = e^{iH_0 t}O_I(0)e^{-iH_0 t} \tag{4.34}$$

(where we have used Eq. (4.26)). This is the solution of the equation of motion

$$i\frac{d}{dt}O_I(t) = [O_I(t), H_0]. \tag{4.35}$$

This is useful because fields in an interacting theory in the I.P. will evolve just like free fields in the Heisenberg picture, so we can continue to use all of our results for free fields. All of the complications have been relegated to the equation of motion for the states. From the equations of motion of the Schrödinger field, Eq. (4.27), we have

$$\begin{aligned} & i\frac{d}{dt}e^{-iH_0 t}|\,\psi(t)\rangle_I = H_S e^{-iH_0 t}|\,\psi(t)\rangle_I \\ \Rightarrow\quad & H_0 e^{-iH_0 t}|\,\psi(t)\rangle_I + e^{-iH_0 t} i\frac{d}{dt}|\,\psi(t)\rangle_I = (H_0(0) + H_I(0))e^{-iH_0 t}|\,\psi(t)\rangle_I \\ \Rightarrow\quad & i\frac{d}{dt}|\,\psi(t)\rangle_I = e^{iH_0 t}H_I(0)e^{-iH_0 t}|\,\psi(t)\rangle_I = H_I(t)|\,\psi(t)\rangle_I \end{aligned} \tag{4.36}$$

where $H_I(t) = e^{iH_0 t}H_I(0)e^{-iH_0 t}$, as expected from Eq. (4.34). Again we see explicitly that when $H_I = 0$ the fields in the I.P. are independent of time.

4.5 Dyson's Formula

We can already get an idea of how perturbation theory is going to work in the interaction picture. The time dependence of the operators is trivial, simply given by the free field equations. On the other hand, the time dependence of the states is going to be taken into account perturbatively, order by order in the interaction Hamiltonian. Since the Hamiltonian generates time evolution, at first order in perturbation theory the Hamiltonian can act once on the states. The interaction term $\psi^\dagger\psi\phi$ contains a collection of creation and annihilation operators, such as

$$c^\dagger b^\dagger a,\ c^\dagger ca,\ bb^\dagger a^\dagger,\ bca^\dagger, \ldots. \tag{4.37}$$

These interactions don't conserve particle number, and can contribute to a number of processes. In the first term, the Hamiltonian acts on the initial state, annihilates a ϕ particle and creates a ψ particle and antiparticle: this corresponds to the decay process $\phi \to \overline{\psi}\psi$. The second corresponds to the absorption $\psi + \phi \to \psi$, and so on. At second order in perturbation theory the Hamiltonian can acts twice on the state, producing more complicated processes like $\overline{\psi} + \psi \to \phi \to \overline{\psi} + \psi$ (ψ anti-ψ scattering through the creation of an intermediate ϕ). In this section we will set up a formalism to apply perturbation theory to scattering processes.

Scattering processes are particularly convenient to study because in many cases *the initial and final states look like systems of noninteracting particles*. What do we mean by this? In a scattering process, we start out with some initial state $|\, i\rangle$ consisting of a number of isolated particles. Since the particles are widely separated, we don't expect them to feel the effects of the potential in Eq. (4.24), and so they will look like free plane wave states (that is, eigenstates of the free Hamiltonian H_0.) In particular, we expect them to be eigenstates of particle number, even though N will not in general commute with the interaction Hamiltonian H_I. We say we are colliding two electrons, or two protons, or whatever, with some particular momentum. The initial state looks simple.

As the particles approach one another, they begin to feel the potential, and the states start to evolve according to Eq. (4.36) in a complicated and non-linear way. At this intermediate stage, the system will look extremely complicated when expressed in terms of our basis of free particles. Particles will be created and destroyed, since H_I in general doesn't commute with N. We no longer have, for example, just two colliding protons, but a complicated mess of protons, pions, photons, and so forth.

We can imagine several results of the scattering process. Several initial particles could collide and form a bound state, such as $p+p \to {}_2D$ (two protons fusing to form a deuterium nucleus). In this case, no matter how long we wait after the scattering process has occurred the final state will never look like an eigenstate of the free Hamiltonian, because the interaction is responsible for the bound state. If we turn the interaction off, the bound state will fly apart. The formalism we are going to develop for scattering theory will not be very useful in this situation.

Instead, we could have a process in which no bound states are formed. Then some long time after the interaction the system will consist of a bunch of widely separated particles, perhaps three protons, an antiproton and fourteen pions. The system will again look like a collection of noninteracting particles. Again it will look simple. This is the type of process we will be considering.

Before we go any further, I should tell you that this is a bit of a fake. In fact, no matter how far you go into the past or future from a scattering process you *never* end up with a collection of free particles. We already know this from electromagnetism: long after the collision process, an electron still carries its electromagnetic field along with it. When we quantize electromagnetism, we will see that this corresponds to a cloud of photons around the electron. Similarly, the "nucleons" in our toy model will always have a cloud of mesons around them. If we turn off the interaction, the states will change, so our simple picture is not quite right. Despite this, our quick and dirty scattering theory will still work. You can see that this might be the case by imagining that instead of Eq. (4.24), our theory is defined by the Lagrangian

$$\mathcal{L} = \mathcal{L}_\phi + \mathcal{L}_\psi - g f(t)\psi^\dagger\psi\phi \tag{4.38}$$

where $f(t) = 0$ for large $|t|$ and $f(t) = 1$ for t near 0, as shown in Figure 4.3. For processes where bound states occur, $f(t)$ clearly drastically changes the states

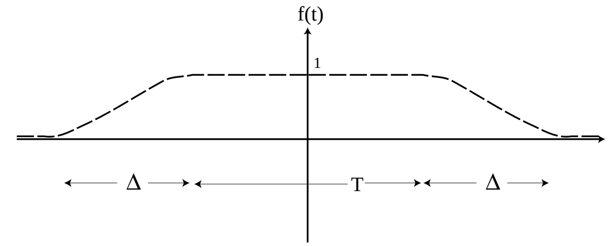

Figure 4.3: The "turning on and off" function $f(t)$ in Eq. (4.38). In the limit $\Delta \to \infty$, $T \to \infty$, $\Delta/T \to 0$ we expect to recover the results of the original theory Eq. (4.24). The scattering process occurs near $t = 0$.

in the far future, since when $f(t) \to 0$ the interaction turns off and the states will fly apart. But in cases where there are no bound states formed, you might imagine that adding $f(t)$ to the interaction won't change the scattering amplitude at all. In particular, if we imagine that a long time $T/2$ after the scattering process occurs, we turn the interaction off very slowly (adiabatically) over a time period Δ, we expect that *the simple states in the real theory will slowly turn into the eigenstates of the free Hamiltonian with unit probability.* In other words, there must be a $1-1$ correspondence between the asymptotic (simple) eigenstates of the full Hamiltonian and the eigenstates of the free Hamiltonian. This means that we can't consider bound states, which are not eigenstates of the free Hamiltonian. In the limit $T \to \infty$, $\Delta \to \infty$ and $\Delta/T \to 0$ (the last requirement ensures that edge effects vanish) we should recover the full theory.

This description is really meant as a hand-waving way of justifying our approach in which the initial and final states are taken to be eigenstates of the free Hamiltonian. It is possible to justify this approach (more) rigorously, but this would take us into technical details which we don't have time for in this course. The hand-waving approach will have to suffice at this stage. [7]

So we want to solve

$$i\frac{d}{dt}|\, \psi(t)\rangle = H_I(t)|\, \psi(t)\rangle \tag{4.39}$$

(we will drop the subscript I on the states, since we will always be working in the

[7] See Peskin and Schroeder, Section 7.2, for the proper treatment of this problem.

I.P. from now on) with the boundary condition

$$|\,\psi(-\infty)\rangle = |\,i\rangle. \tag{4.40}$$

We want to connect the simple description in the far past with the simple description in the far future, long after the collision has taken place. If we define the scattering operator S

$$|\,\psi(\infty)\rangle = S|\,\psi(-\infty)\rangle = S|\,i\rangle \tag{4.41}$$

then the amplitude to find the system in some given state $|\,f\rangle$ in the far future is

$$\langle f\,|S|\,i\rangle \equiv S_{fi}. \tag{4.42}$$

This is conventionally known as the "S-matrix element". We can solve for S iteratively: integrating both sides of Eq. (4.39) from $t_1 = -\infty$ to t, we find

$$|\,\psi(t)\rangle = |\,i\rangle + (-i)\int_{-\infty}^{t} dt_1\, H_I(t_1)|\,\psi(t_1)\rangle. \tag{4.43}$$

Iterating this gives

$$\begin{aligned} |\,\psi(t)\rangle = |\,i\rangle \;&+\; (-i)\int_{-\infty}^{t} dt_1\, H_I(t_1)|\,i\rangle \\ &+\; (-i)^2\int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2\, H_I(t_1)H_I(t_2)|\,\psi(t_2)\rangle. \end{aligned} \tag{4.44}$$

Repeating this procedure indefinitely and taking $t \to \infty$, we obtain the following expansion for S:

$$S = \sum_{n=0}^{\infty}(-i)^n \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 \ldots \int_{-\infty}^{t_{n-1}} dt_n\, H_I(t_1)\ldots H_I(t_n). \tag{4.45}$$

There is a more symmetric way to write this. Look at the $n = 2$ term, for example:

$$\int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 H_I(t_1)H_I(t_2). \tag{4.46}$$

This corresponds to integrating over the region $-\infty < t_2 < t_1 < \infty$ shown in part (a) of the figure. We can reverse the order of integration, and noting that this is the same region of integration as in part (b) of the figure, we can write the term as

$$\begin{aligned} &\int_{-\infty}^{\infty} dt_2 \int_{t_2}^{\infty} dt_1 H_I(t_1)H_I(t_2) \\ =\; &\int_{-\infty}^{\infty} dt_1 \int_{t_1}^{\infty} dt_2 H_I(t_2)H_I(t_1), \end{aligned} \tag{4.47}$$

so we can write the second term of the expansion as

$$\frac{1}{2!}\left[\int_{-\infty}^{\infty} dt_1 \int_{t_1}^{\infty} dt_2\, H_I(t_2)H_I(t_1) + \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2\, H_I(t_1)H_I(t_2)\right]. \tag{4.48}$$

Notice that in the first term $t_2 < t_1$, while in the second $t_1 < t_2$. So the H_I's

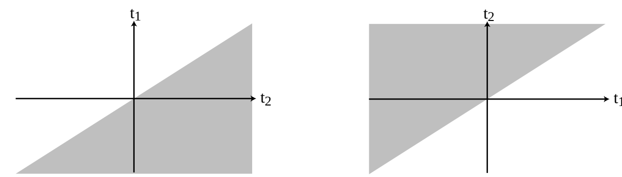

Figure 4.4: The shaded regions correspond to the region of integration in (a) Eq. (4.46) and (b) Eq. (4.47).

are always ordered with the earlier one on the right. As before, we define the *time-ordered product* $T(O_1O_2)$ of two operators $O_1(x_2)$ and $O_2(x_2)$ by

$$T(O_1(x_1)O_2(x_2)) = \begin{cases} O_1(x_1)O_2(x_2), & t_1 > t_2; \\ O_2(x_2)O_1(x_1), & t_1 < t_2. \end{cases} \tag{4.49}$$

In terms of the time-ordered product, we can write the second term in the expansion of S as

$$\frac{1}{2!}\int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2\, T(H_I(t_1)H_I(t_2)). \tag{4.50}$$

Similarly, for n operators we define the time ordered product (or T-product) such that the operators are ordered chronologically, the earliest on the right and the latest on the left. H_I commutes with itself at equal times, so there is no ambiguity in this definition. The n'th term in the expansion of S may then be written as

$$\frac{1}{n!}\int_{-\infty}^{\infty} dt_1 \ldots \int_{-\infty}^{\infty} dt_n\, T(H_I(t_1)\ldots H_I(t_n)) \tag{4.51}$$

and the expansion for S is then

$$\begin{aligned} S \;&=\; \sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\int_{-\infty}^{\infty} dt_1 \ldots \int_{-\infty}^{\infty} dt_n\, T(H_I(t_1)\ldots H_I(t_n)) \\ &=\; \sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\int d^4x_1 \ldots \int d^4x_n\, T(\mathcal{H}_I(x_1)\ldots \mathcal{H}_I(x_n)). \end{aligned} \tag{4.52}$$

We can even be slick and write this series as a time-ordered exponential,

$$S = Te^{-i\int d^4x \mathcal{H}_I(x)}, \tag{4.53}$$

where the time-ordering acts on each term in the series expansion. This is *Dyson's formula*.

4.6 Wick's Theorem

To evaluate the individual terms in Dyson's formula we will have to calculate matrix elements of time ordered products of fields between the initial and final scattering states. For example, in our meson-"nucleon" theory at second order in g we have to evaluate matrix elements of the form

$$\langle f\,|T(\mathcal{H}_I(x_1)\mathcal{H}_I(x_2))|\,i\rangle = \langle f\,|T(\psi^\dagger(x_1)\psi(x_1)\phi(x_1)\psi^\dagger(x_2)\psi(x_2)\phi(x_2))|\,i\rangle. \tag{4.54}$$

For the scattering process $N+N \to N+N$ (elastic scattering of two "nucleons"), we have $|\,i\rangle = |\,\vec{k}_1(N);\vec{k}_2(N)\rangle$, $|\,f\rangle = |\,\vec{k}_3(N);\vec{k}_4(N)\rangle$, where $\vec{k}_4 = \vec{k}_1+\vec{k}_2-\vec{k}_3$ since our theory conserves momentum. Since we know how the fields act on the states in the I.P., this matrix element is straightforward to calculate. However, in this form it's still rather messy, because the T-product contains 16 arrangements of "nucleon" creation and annihilation operators. It would be much simpler if we could normal-order this expression, because then the only ordering which would contribute to this process would be ones with two "nucleon" annihilation operators on the right and two "nucleon" creation operators on the left. In fact, there is a relation between time-ordered and normal-ordered products, which goes by the name of *Wick's theorem*. To state Wick's theorem, we define the *contraction* of two fields,

$$\overbrace{A(x)B}(y) \equiv T(A(x)B(y)) - :A(x)B(y): \tag{4.55}$$

It is easy to see that $\overbrace{A(x)B}(y)$ is a number, not an operator. Consider first the case $x^0 > y^0$. Then

$$T(A(x)B(y)) = (A^{(+)}+A^{(-)})(B^{(+)}+B^{(-)}) =: AB : +[A^{(+)},B^{(-)}] \tag{4.56}$$

so $\overbrace{A(x)B}(y)$ is a number (given by the canonical commutation relations). Similarly, it is a number when $x^0 < y^0$, so we can sandwich both sides of Eq. (4.55) between vacuum states to find that

$$\begin{aligned}\overbrace{A(x)B}(y) &= \langle 0\,|\overbrace{A(x)B}(y)|\,0\rangle \\ &= \langle 0\,|T(A(x)B(y))|\,0\rangle - \langle 0\,|:A(x)B(y):|\,0\rangle \\ &= \langle 0\,|T(A(x)B(y))|\,0\rangle \end{aligned} \tag{4.57}$$

since the vacuum expectation value of a normal ordered product of fields vanishes (the annihilation operators on the right annihilate the vacuum). So we have found that the contraction of two fields is just the vacuum expectation value of the time ordered product of the fields. We have already seen this object before - it is the Feynman propagator for the field,

$$\overbrace{\phi(x)\phi}(y) = D_F(x-y) = \langle 0\,|T(\phi(x)\phi(y))|\,0\rangle = \int \frac{d^4k}{(2\pi)^4}e^{ik\cdot(x-y)}\frac{i}{k^2-\mu^2+i\epsilon}. \tag{4.58}$$

where the $\lim_{\epsilon\to 0^+}$ is implicit in this expression.

For the charged fields, it is straightforward to show that the propagator is

$$\overbrace{\psi(x)\psi}^\dagger(y) = \overbrace{\psi^\dagger(x)\psi}(y) = \int \frac{d^4k}{(2\pi)^4}e^{ik\cdot(x-y)}\frac{i}{k^2-m^2+i\epsilon} \tag{4.59}$$

while other contractions vanish:

$$\overbrace{\psi(x)\psi}(y) = \overbrace{\psi^\dagger(x)\psi}^\dagger(y) = 0. \tag{4.60}$$

(The last equation is true because ψ only creates c-type particles and annihilates b-type particles, therefore $\langle 0\,|T(\psi(x)\psi(y))|\,0\rangle = 0$.)

Having defined the propagator of a field, we can now state Wick's theorem. For any collection of fields $\phi_1 \equiv \phi_{a_1}(x_1),\ \phi_2 \equiv \phi_{a_2}(x_2), ...$ the T-product of the fields has the following expansion

$$\begin{aligned}T(\phi_1...\phi_n) &= \ :\phi_1...\phi_n: \\ &+ \ :\overbrace{\phi_1\phi_2}\phi_3...\phi_n: + :\overbrace{\phi_1\phi_2\phi_3}...\phi_n: + ... + :\phi_1\phi_2...\overbrace{\phi_{n-1}\phi_n}: \\ &+ \ :\overbrace{\phi_1\phi_2}\overbrace{\phi_3\phi_4}\phi_5...\phi_n: + ... + \phi_1\phi_2...:\overbrace{\phi_{n-3}\phi_{n-2}}\overbrace{\phi_{n-1}\phi_n}: \\ &+ \ ...\end{aligned} \tag{4.61}$$

On the right-hand side of the equation we have all possible terms with all possible contractions of two fields. We are also using the notation

$$:A(x)\overbrace{B(y)C}(z)D(w): \equiv :A(x)C(z):\overbrace{B(y)D}(w) \tag{4.62}$$

Wick's theorem is true by definition for $n=2$. The proof that this is true for all n is by induction, and so not terribly illuminating, so we won't repeat it here.

Wick's theorem has unravelled the messy combinatorics of the T-product, leaving us with an expression in terms of propagators and normal-ordered products, whose matrix elements are easy to take without worrying about commutation relations. In its general form, Eq. (4.61), it looks rather daunting, so let's get a feeling for it by applying it to the expression for S at $\mathcal{O}(g^2)$ in our model:

$$\frac{(-ig)^2}{2!}\int d^4x_1\int d^4x_2 T(\psi^\dagger(x_1)\psi(x_1)\phi(x_1)\psi^\dagger(x_2)\psi(x_2)\phi(x_2)). \tag{4.63}$$

Wick's theorem relates this to a number of normal-ordered products. One of these terms is

$$\frac{(-ig)^2}{2!}\int d^4x_1\int d^4x_2 :\psi^\dagger(x_1)\psi(x_1)\overbrace{\phi(x_1)\psi^\dagger(x_2)\psi(x_2)\phi}(x_2): \tag{4.64}$$

This term can contribute to a variety of physical processes. The ψ field contains operators which annihilate a "nucleon" and create an "anti-nucleon." The $\psi^\dagger$ field contains operators which annihilate an "anti-nucleon" and create a "nucleon." So the operator

$$:\psi^\dagger(x_1)\psi(x_1)\overbrace{\phi(x_1)\psi^\dagger(x_2)\psi(x_2)\phi}(x_2) :\equiv: \psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2): \overbrace{\phi(x_1)\phi}(x_2) \tag{4.65}$$

can contribute to elastic NN scattering, $N+N\to N+N$. That is to say, the matrix element

$$\langle\vec{k_3}(N);\vec{k_4}(N)\,|:\psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2):|\,\vec{k_1}(N);\vec{k_2}(N)\rangle \tag{4.66}$$

is nonzero, because there are terms in the two ψ fields that can annihilate the two nucleons in the initial state and terms in the two $\psi^\dagger$ fields that can create two nucleons, to give a nonzero matrix element. Other combinations of annihilation and creation operators in this term can also contribute to $\overline{N}+\overline{N}\to\overline{N}+\overline{N}$ and $N+\overline{N}\to N+\overline{N}$. You can also see that there is no combination of creation and annihilation operators that will contribute to $N+N\to\overline{N}+\overline{N}$. The ψ fields would have to annihilate the nucleons, and the $\psi^\dagger$ fields can't create anti-nucelons. We already knew this had to be the case, because the theory has a conserved $U(1)$ charge which wouldn't be conserved in this process. It is reassuring to see that this actually works in practice.

Another term in the expansion of the T-product is

$$\frac{(-ig)^2}{2!}\int d^4x_1\int d^4x_2 :\psi^\dagger(x_1)\overbrace{\psi(x_1)\phi(x_1)\psi^\dagger}(x_2)\psi(x_2)\phi(x_2): \tag{4.67}$$

This term can contribute to the following $2\to 2$ scattering processes (you should verify this):

$$N+\phi\to N+\phi,\;\; \overline{N}+\phi\to\overline{N}+\phi,\;\; N+\overline{N}\to\phi+\phi,\;\; \phi+\phi\to N+\overline{N}.$$

A single term is able to contribute to a variety of processes like this because each field can either destroy or create particles.

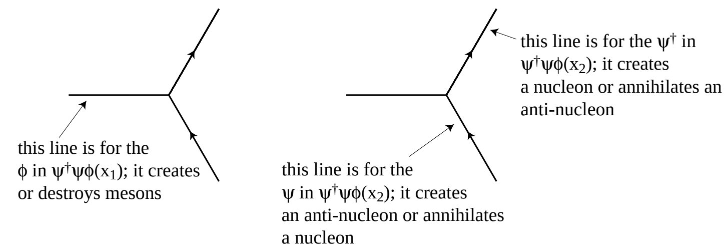

Figure 4.5: Vertices at second order in the Wick expansion.

4.7 Diagrammatic Perturbation Theory

In practice, nobody ever bothers thinking about Dyson's formula or Wick's theorem when calculating scattering amplitudes. This is because there is a very simple diagrammatic shorthand which has all of this formalism built into it. At n'th order in perturbation theory, you start by writing down n interaction vertices. Represent each field at the vertex by a line. To distinguish ψ's from $\psi^\dagger$'s, we can draw an arrow on the corresponding line as shown in Figure 4.5. Contractions are represented by connecting the lines. Any time there is a contraction, join the lines of the contracted fields. The arrows will always line up, because the contractions for which they don't, $\overbrace{\psi(x)\psi}(y)$ and $\overbrace{\psi^\dagger(x)\psi^\dagger}(y)$, are zero. An unarrowed line will never be connected to an arrowed line because $\overbrace{\psi(x)\phi}(y)$ is clearly zero as well. So the term in Eq. (4.64) corresponds to the diagram in Fig. 4.6, while the term in

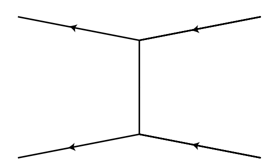

Figure 4.6: Wick diagram corresponding to Eq. (4.64).

Eq. (4.67) corresponds to the diagram in Fig. 4.7. (Since the arrows always line up, we have only drawn one arrow on the contracted nucleon lines). *These diagrams are in one to one correspondence with the terms in the Wick expansion.*

At the moment, our diagrams correspond to operators, individual terms in the Wick expansion. However, for a given process we are interested not in having an

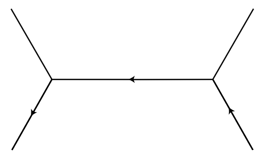

Figure 4.7: Wick diagram corresponding to Eq. (4.67).

expression for the operator S, but instead for the matrix element

$$\langle f\,|(S-1)|\,i\rangle. \tag{4.68}$$

We really want $S-1$, not S, because we aren't interested in processes in which no scattering at all occurs, which corresponds to the leading order term of the Wick expansion. For $NN \to NN$ scattering we want the matrix element

$$\langle p_1'(N), p_2'(N)\,|(S-1)|\,p_1(N), p_2(N)\rangle. \tag{4.69}$$

Note that there are no arrows over the momenta in the states. We are now doing relativistic field theory in earnest and so we are going to use our relativistically normalized states from the first lecture,

$$|\,k\rangle = (2\pi)^{3/2}\sqrt{2\omega_k}|\,\vec{k}\rangle. \tag{4.70}$$

We can write these states as

$$|\,k\rangle = a^\dagger(k)|\,0\rangle \tag{4.71}$$

where the relativistically normalized creation operator $a^\dagger(k)$ is defined as

$$a^\dagger(k) \equiv (2\pi)^{3/2}\sqrt{2\omega_k}\,a_k^\dagger \tag{4.72}$$

and the scalar field ϕ has the expansion

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k}\left[a(k)e^{-ik\cdot x} + a^\dagger(k)e^{ik\cdot x}\right]. \tag{4.73}$$

From Eqs. (4.70) and (4.71), we also find

$$\begin{aligned} a(k')|\,k\rangle &= a(k')a^\dagger(k)|\,0\rangle \\ &= [a(k'), a^\dagger(k)]|\,0\rangle \\ &= (2\pi)^3 2\omega_k\,\delta^{(3)}(\vec{k}-\vec{k}')|\,0\rangle \end{aligned} \tag{4.74}$$

and so

$$\int \frac{d^3k'}{(2\pi)^3 2\omega_k} a(k')|\,k\rangle = |\,0\rangle. \tag{4.75}$$

Similar relations holds for the relativistically normalized "nucleon" and "anti-nucleon" creation and annihilation operators, so a relativistically normalized incoming two nucleon state is

$$|\,p_1(N); p_2(N)\rangle = b^\dagger(p_1)b^\dagger(p_2)|\,0\rangle. \tag{4.76}$$

Now, to evaluate Eq. (4.69) at second order in the Wick expansion we need the matrix element

$$\langle p_1'; p_2'\,|:\psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2):|\,p_1; p_2\rangle \tag{4.77}$$

(since we only have nucleons in the initial and final states, I'm going to suppress the "N" label on the states). The only way to get a nonzero matrix element is by using the nucleon annihilation terms in $\psi(x_1)$ and $\psi(x_2)$ to annihilate the two incoming nucleons, and using the nucleon creation terms in $\psi^\dagger(x_1)$ and $\psi^\dagger(x_2)$ to create the two nucleons in the final state. Any other combination of creation and annihilation operators will give zero inner product. So in equations,

$$\begin{aligned} &\langle p_1'; p_2'\,|:\psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2):|\,p_1; p_2\rangle = \\ &\qquad \langle p_1'; p_2'\,|\psi^\dagger(x_1)\psi^\dagger(x_2)|\,0\rangle\langle 0\,|\psi(x_1)\psi(x_2)|\,p_1; p_2\rangle. \end{aligned} \tag{4.78}$$

From the explicit expansion of ψ in terms of $b^\dagger(k)$ and $c(k)$ and Eq. (4.75), you can easily show that

$$\langle 0\,|\psi(x_1)\psi(x_2)|\,p_1; p_2\rangle = e^{-ip_1\cdot x_1 - ip_2\cdot x_2} + e^{-ip_1\cdot x_2 - ip_2\cdot x_1}. \tag{4.79}$$

Using this and its complex conjugate, we find four terms contributing to the matrix element

$$\begin{aligned} &\langle p_1'; p_2'\,|:\psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2):|\,p_1; p_2\rangle = \\ &\quad (e^{ip_1'\cdot x_1 + ip_2'\cdot x_2} + e^{ip_1'\cdot x_2 + ip_2'\cdot x_1})(e^{-ip_1\cdot x_1 - ip_2\cdot x_2} + e^{-ip_1\cdot x_2 - ip_2\cdot x_1}) \\ &\quad = e^{ip_1'\cdot x_1 + ip_2'\cdot x_2 - ip_1\cdot x_1 - ip_2\cdot x_2} + e^{ip_1'\cdot x_2 + ip_2'\cdot x_1 - ip_1\cdot x_2 - ip_2\cdot x_1} \\ &\quad + e^{ip_1'\cdot x_2 + ip_2'\cdot x_1 - ip_1\cdot x_1 - ip_2\cdot x_2} + e^{ip_1'\cdot x_1 + ip_2'\cdot x_2 - ip_1\cdot x_2 - ip_2\cdot x_1} \end{aligned} \tag{4.80}$$

Notice that the first two terms on the first line of the final answer differs by the interchange $x_1 \leftrightarrow x_2$. The same is true for the last two terms. Since we are integrating over x_1 and x_2 symmetrically, and since $\overbrace{\phi(x_1)\phi}(x_2)$ is symmetric under $x_1 \leftrightarrow x_2$, these terms must give identical contributions to the matrix element. This factor of 2 cancels the $1/2!$ in Dyson's formula. Using our expression for the ϕ

propagator, Eq. (4.58), we obtain the following expression for the second order contribution to NN scattering

$$
\begin{aligned}
& (-ig)^2 \int d^4x_1\, d^4x_2 \overbrace{\phi(x_1)\phi}(x_2) \\
& \quad \times \left(e^{ip_1'\cdot x_1 + ip_2'\cdot x_2 - ip_1\cdot x_1 - ip_2\cdot x_2} + e^{ip_1'\cdot x_2 + ip_2'\cdot x_1 - ip_1\cdot x_2 - ip_2\cdot x_1} \right) \\
= & (-ig)^2 \int d^4x_1\, d^4x_2 \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2-\mu^2+i\epsilon} \qquad (4.81) \\
& \quad \times \left(e^{i(p_1'-p_1+k)\cdot x_1 + i(p_2'-p_2-k)\cdot x_2} + e^{i(p_2'-p_1+k)\cdot x_1 + i(p_1'-p_2-k)\cdot x_2} \right).
\end{aligned}
$$

The x_1 and x_2 integrations are easy to do – they just give us δ functions, so this becomes

$$
\begin{aligned}
(-ig)^2 \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2-\mu^2+i\epsilon} & \Big[(2\pi)^4\delta^{(4)}(p_1'-p_1+k)(2\pi)^4\delta^{(4)}(p_2'-p_2-k) \\
& + (2\pi)^4\delta^{(4)}(p_2'-p_1+k)(2\pi)^4\delta^{(4)}(p_1'-p_2-k) \Big]. \qquad (4.82)
\end{aligned}
$$

Finally, we can do the k integration using the δ functions, and we get

$$
(-ig)^2(2\pi)^4\delta^{(4)}(p_1'+p_2'-p_1-p_2) \left(\frac{i}{(p_1'-p_1)^2-\mu^2+i\epsilon} + \frac{i}{(p_2'-p_1)^2-\mu^2+i\epsilon} \right). \qquad (4.83)
$$

We can now incorporate this into our diagram for NN scattering so that each term in the matrix element Eq. (4.82) corresponds to a diagram by implementing the following rules:

(a) Identify each uncontracted line in the diagram with an external particle, and label it with the corresponding momentum. Draw a separate diagram for each distinct labelling of external legs. (Note that there are different conventions about the direction in which these diagrams are read. I prefer to read them right to left, with the external lines on the right corresponding to incoming particles, and those on the left corresponding to outgoing particles. This is convenient because this is the same order as in the Dirac notation for the matrix element $\langle f\,|(S-1)|\,i\rangle$.)

For nucleons, the direction of the arrow indicates the direction of flow of the $U(1)$ charge. An incoming arrow in the initial state corresponds to a nucleon being annihilated; an incoming arrow in the final state corresponds to an anti-nucleon being created. Similarly, an outgoing arrow in the initial state corresponds to an anti-nucleon and an outgoing arrow in the final state corresponds to an outgoing nucleon.

(b) At each vertex, write down a factor of

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

where $\sum_{k_i}$ is the sum of all momenta flowing *into* (or out of, if you like, as long as you're consistent) the vertex.

(c) For each internal line with momentum k flowing through it, write down a factor

$$
\int \frac{d^4k}{(2\pi)^4} D(k^2)
$$

where $D(k^2)$ is the propagator for the appropriate field:

$$
D(k^2) = \frac{i}{k^2-\mu^2+i\epsilon}
$$

for a meson, and

$$
D(k^2) = \frac{i}{k^2-m^2+i\epsilon}
$$

for a "nucleon".

That's it. Note that there is no excuse for not getting the factors of $(2\pi)^4$ right. Every factor of d^4k always comes along with a factor of $(2\pi)^{-4}$, and every $\delta^{(4)}$ function always comes with a factor of $(2\pi)^4$.

For NN scattering, there are two distinct labellings of the external momenta, so we can write down two diagrams: Following the rules (a)–(c), we can imme-

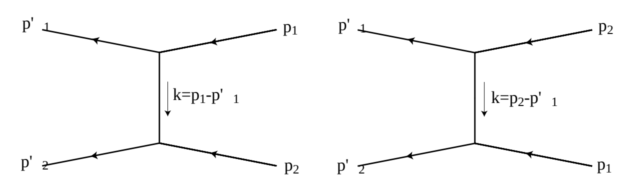

Figure 4.8: Feynman diagrams contributing to NN scattering at order g^2.

diately read off the relevant matrix element, Eq. (4.82). These graphs are called *Feynman diagrams*, and the set of rules (a)–(c) are the *Feynman rules* for our toy model. Note that you can shortcut the trivial integrations over δ functions just by assigning internal momenta flowing through propagators so as to conserve energy and momentum whenever an internal momentum is determined by the other momenta at a vertex.

These diagrams have a very simple physical interpretation. For the first diagram for NN scattering, you can say that a nucleon with momentum p_1 comes in and interacts, scattering into a nucleon with momentum p_1' and a meson with momentum $k = p_1 - p_1'$. Energy and momentum are conserved in this process, but the virtual meson doesn't satisfy $k^2 = \mu^2$. In terms of the uncertainty principle, the meson must not live long enough for its energy to be measured to great enough accuracy to measure this discrepancy. It therefore can't exist as a real particle, but must be reabsorbed after a short time. To distinguish it from a physical particle, it is referred to as a "virtual" meson, and it is reabsorbed by a nucleon with momentum p_2, scattering it into a nucleon with momentum p_2'. (Note that although we are writing this as though there is a definite ordering to these events, the graph has no time-ordering in it. We could just as well say that the meson is emitted from the second nucleon and then absorbed by the first.)

The second diagram must be there because of Bose statistics. Since the two incoming nuclei are identical, it is in principle impossible to say which of the incident nuclei carries p_1 and which carries p_2. The processes occuring in the two graphs are indistinguishable, and so the amplitude must sum over both of them. Note that Bose statistics are automatically built into our creation and annihilation operator formalism.

Notice that performing the final integral over δ functions leaves us with a factor of

$$(2\pi)^4\delta^{(4)}(p_f - p_i)$$

where p_f is the sum of all final momenta, and p_i is the sum of initial momenta. This just enforces energy-momentum conservation on the graph as a whole. Since it is always there in any diagram, it is traditional to define the invariant Feynman amplitude $\mathcal{A}_{fi}$ (Mandl & Shaw call this $\mathcal{M}$) by

$$\langle f\,|(S-1)|\,i\rangle = i\mathcal{A}_{fi}(2\pi)^4\delta^{(4)}(p_f - p_i). \tag{4.84}$$

The factor of i is there by convention; it reproduces the phase conventions of NRQM.

4.8 More on Feynman Diagrams

In the previous section we introduced Feynman diagrams as a convenient way to calculate matrix elements of the individual terms in Dyson's formula. Each vertex and line in the diagram was associated with a factor given by the corresponding Feynman rule. We also noted that there was always an overall energy-momentum

conserving δ function, which we factored out of the matrix element to define the invariant Feynman amplitude $\mathcal{A}$:

$$\langle f\,|(S-1)|\,i\rangle = i\mathcal{A}_{fi}(2\pi)^4\delta^{(4)}(p_F - p_I) \tag{4.85}$$

where p_I and p_F are the sums of the total initial and final momenta, respectively. We also noted that we could shortcut some of the trivial delta functions and integrations by simply imposing energy-momentum conservation on each vertex. We can incorporate these simplifications into our Feynman rules for $i\mathcal{A}_{fi}$:

(a) Draw all possible diagrams at each order which can contribute to the process. Assign a momentum to each line (internal and external) and enforce energy-momentum conservation at each vertex. Draw a separate diagram for each distinct labelling of the momenta of the external legs.

(b) At each vertex, write down a factor of $(-ig)$.

(c) For each contracted line, write down a factor of the propagator for that field.

This is fine for graphs like the ones we have been considering. However, there are also diagrams with closed loops for which energy-momentum conservation at the vertices is not sufficient to fix all the internal momenta. For example, the diagram in Fig. 4.9 corresponds to matrix elements obtained from the contraction

$$\langle p\,|:\overbrace{\psi^\dagger(x_1)\psi}(x_2)\psi(x_1)\psi^\dagger(x_2)\overbrace{\phi(x_1)\phi}(x_2):|\,p\rangle. \tag{4.86}$$

Enforcing energy-momentum conservation at each vertex is still not sufficient to

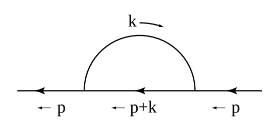

Figure 4.9: Feynman diagram corresponding to matrix element (4.86).

fix the momentum k flowing through the loop, and so we must keep the factor of

$$\int \frac{d^4k}{(2\pi)^4} \tag{4.87}$$

Similarly, the fully contracted term

$$\langle 0\,|:\overbrace{\psi^\dagger(x_1)\psi}(x_2)\overbrace{\psi(x_1)\psi^\dagger}(x_2)\overbrace{\phi(x_1)\phi}(x_2):|\,0\rangle \tag{4.88}$$

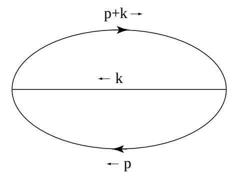

Figure 4.10: Feynman diagram corresponding to matrix element (4.88).

corresponds to the two-loop graph in Fig. (4.10). In this diagram, neither p nor k is constrained, so we must integrate over both momenta. Thus we add an additional Feynman rule for $i\mathcal{A}$:

(d) For each internal loop with momentum k unconstrained by energy-momentum conservation, write down a factor of

$$\frac{d^4k}{(2\pi)^4}.$$

Now, for our nucleon-nucleon scattering process, we found the following expression for $i\mathcal{A}$:

$$i\mathcal{A} = (-ig)^2 \left[\frac{i}{(p_1 - p_1')^2 - \mu^2 + i\epsilon} + \frac{i}{(p_1 - p_2')^2 - \mu^2 + i\epsilon} \right]. \tag{4.89}$$

In the centre of mass frame, we can write the momenta as

$$\begin{aligned} p_1 &= (\sqrt{p^2 + m^2}, p\hat{e}) \\ p_2 &= (\sqrt{p^2 + m^2}, -p\hat{e}) \\ p_1' &= (\sqrt{p^2 + m^2}, p\hat{e}') \\ p_2' &= (\sqrt{p^2 + m^2}, -p\hat{e}') \end{aligned} \tag{4.90}$$

where $\hat{e} \cdot \hat{e}' = \cos\theta$, and θ is the scattering angle. This immediately gives

$$(p_1 - p_1')^2 = -2p^2(1 - \cos\theta), \quad (p_1 - p_2')^2 = -2p^2(1 + \cos\theta) \tag{4.91}$$

and so

$$i\mathcal{A} = ig^2 \left[\frac{1}{2p^2(1 - \cos\theta) + \mu^2} + \frac{1}{2p^2(1 + \cos\theta) + \mu^2} \right]. \tag{4.92}$$

Here we've dropped the $i\epsilon$ because the denominator never vanishes. In fact, the $i\epsilon$ can always be dropped for calculations with no closed loops. However, for diagrams with closed loops it is required to make the integration over the loop momentum, Eq. (4.87), well-defined.

Notice that we can see again that the two terms in Eq. (4.92) are required because of Bose statistics. Scattering into two identical particles at an angle θ is indistinguishable from scattering at an angle $\pi - \theta$, and so the probability must be symmetrical under the interchange of the two processes. Since these particles are bosons, the amplitude must also be symmetric. We can now write down some more Feynman diagrams which contribute to scattering at $\mathcal{O}(g^2)$:

- $N(p_1) + \overline{N}(p_2) \to N(p_1') + \overline{N}(p_2')$: There are two Feynman graphs contributing to this process, shown in Fig.(4.11). Applying our Feynman rules

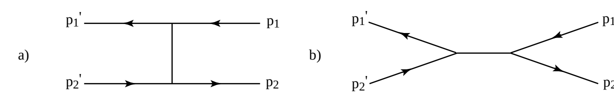

Figure 4.11: Feynman Diagrams contributing to $\overline{N}N \to \overline{N}N$

to these diagrams gives

$$i\mathcal{A} = (-ig)^2 \left[\frac{i}{(p_1 - p_1')^2 - \mu^2} + \frac{i}{(p_1 + p_2)^2 - \mu^2} \right]. \tag{4.93}$$

It is important to be able to recognize which diagrams are and aren't distinct. Since the diagrams are simply a shorthand for matrix elements of operators in the Wick expansion, the orientation of the lines inside the graphs have absolutely no significance. We could just as well have drawn diagrams (a) and (b) as shown in Fig. (4.12). Both figures (a) are really the same diagram,

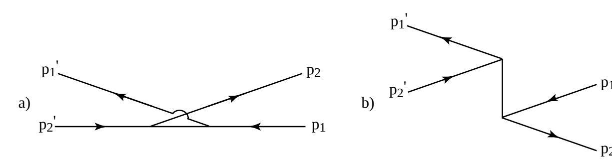

Figure 4.12: Alternate drawing of the Feynman diagrams in Fig. (4.11).

because they have the same arrangement of lines and vertices: the vertices are $N(p_1) - N(p_1') - \phi$ and $N(p_2) - N(p_2') - \phi$ in both diagrams, with the two ϕ's contracted. Similarly, both diagrams labelled (b) are identical. We could even be perverse and draw diagram (b) as shown in Fig. (4.13).

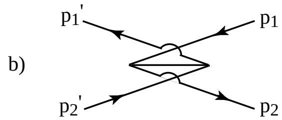

Figure 4.13: Alternate drawing of diagram (b).

- $N(p_1) + \overline{N}(p_2) \to \phi(p_1')\phi(p_2')$, or nucleon-nucleon annihilation into two mesons. These are given by the diagrams in Fig. (4.14), which gives

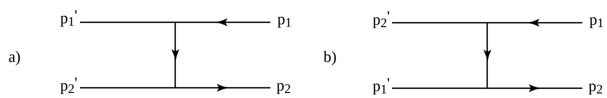

Figure 4.14: Diagrams contributing to $N\overline{N} \to \phi\phi$.

$$i\mathcal{A} = (-ig)^2 \left[\frac{i}{(p_1 - p_1')^2 - m^2} + \frac{i}{(p_1 - p_2')^2 - m^2} \right]. \tag{4.94}$$

In this case we have virtual nucleons in the intermediate state, instead of virtual mesons. Once again, Bose statistics are taken into account by the two diagrams, which differ only by the exchange of the identical particles in the final state.

- $N(p_1) + \phi(p_2) \to N(p_1') + \phi(p_2')$, or nucleon-meson scattering. From the two diagrams in Fig. (4.15) we obtain

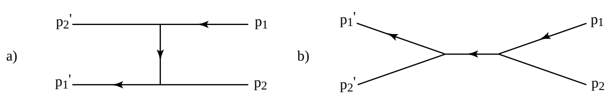

Figure 4.15: Diagrams contributing to $N\phi \to N\phi$.

$$i\mathcal{A} = (-ig)^2 \left[\frac{i}{(p_1 - p_2')^2 - m^2} + \frac{i}{(p_1 + p_2)^2 - m^2} \right]. \tag{4.95}$$

Once again, we could have drawn the diagram (a) as shown in Fig. (4.16)

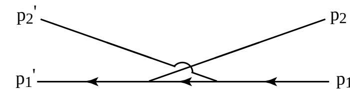

Figure 4.16: Alternate drawing of diagram (a).

This completes the list of interesting scattering processes at $\mathcal{O}(g^2)$. Note that there are processes such as $\overline{N}\overline{N} \to \overline{N}\overline{N}$ and $\overline{N}\phi \to \overline{N}\phi$ which we didn't write down; clearly these are simply related to the analogous process with particles instead of antiparticles. That the amplitudes are identical is related to an additional invariance of the theory which we have not yet discussed, C (charge-conjugation) invariance. We will discuss this in more detail later on in these notes.

In some cases there are additional combinatoric factors which must be incorporated into the Feynman rules. Consider the Lagrangian of a self-coupled scalar field

$$\mathcal{L} = \mathcal{L}_0 - \frac{\lambda}{4!}\phi^4. \tag{4.96}$$

The reason for the factor of $4!$ in the definition of the coupling is made immediately clear by examining the perturbative expansion of the theory. This theory has a single interaction vertex, shown in Fig. (4.17). . At $\mathcal{O}(\lambda)$ in perturbation theory, the

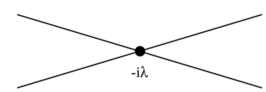

Figure 4.17: Interaction vertex for ϕ^4 interaction.

only term which contributes to $\phi\phi \to \phi\phi$ scattering is the completely uncontracted term

$$-\frac{\lambda}{4!}\langle k_1', k_2'| : \phi(x)\phi(x)\phi(x)\phi(x) : |k_1, k_2\rangle. \tag{4.97}$$

Now, any one of the ϕ fields can annihilate the first meson; any one of the remaining three can annihilate the second, leaving either of the remaining fields to create either of the final mesons, giving a total of $4!$ different combinations. For the Feynman rule for this vertex, the factors of $4!$ cancel and we are left simply with $-i\lambda$. At higher orders in perturbation theory, there are more complicated diagrams contributing to these scattering processes. For example, for $NN \to NN$ scattering in our meson-"nucleon" theory, at $\mathcal{O}(g^4)$ we have diagrams like the two shown in Fig. (4.18). Diagram (a) arises from Wick contractions of the form

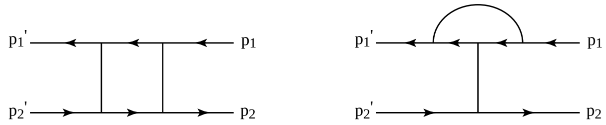

Figure 4.18: Two representative graphs which contribute to $NN \to NN$ scattering at $\mathcal{O}(g^4)$.

$$: \psi^\dagger(x_1)\psi(x_2)\psi(x_3)\psi^\dagger(x_4)\overbrace{\psi(x_1)\psi^\dagger}(x_3)\overbrace{\psi^\dagger(x_2)\psi}(x_4)\overbrace{\phi(x_1)\phi}(x_2)\overbrace{\phi(x_3)\phi}(x_4) : \quad (4.98)$$

whereas diagram (b) arises from Wick contractions of the form

$$: \psi^\dagger(x_1)\psi(x_2)\psi(x_4)\psi^\dagger(x_4)\overbrace{\psi(x_1)\psi^\dagger}(x_3)\overbrace{\psi^\dagger(x_2)\psi}(x_3)\overbrace{\phi(x_1)\phi}(x_2)\overbrace{\phi(x_3)\phi}(x_4) : \quad (4.99)$$

At $\mathcal{O}(g^4)$ we also get a new process, $\phi\phi \to \phi\phi$ scattering, from the graph in Fig. (4.19). The momenta flowing through the internal lines in this figure have been

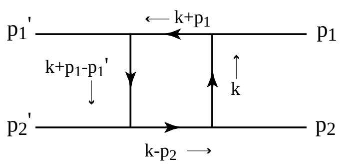

Figure 4.19: Diagram contributing to $\phi\phi \to \phi\phi$ scattering.

explicitly shown. Because of the overall energy-momentum conserving δ function, it does not matter whether we label, for example, the bottom line by $k - p_2$ or $k + p_1 - p_1' - p_2'$. We can also see explicitly that energy-momentum conservation at the vertices leaves one unconstrained momentum k which must be integrated over. According to our Feynman rules, this last graph is

$$\begin{aligned} i\mathcal{A} &= (-ig)^4 \int \frac{d^4k}{(2\pi)^4} \frac{i^4}{(k^2 - m^2 + i\epsilon)((k+p_1)^2 - m^2 + i\epsilon)} \\ &\times \frac{1}{((k + p_1 - p_1')^2 - m^2 + i\epsilon)((k - p_2)^2 - m^2 + i\epsilon)}. \end{aligned} \quad (4.100)$$

The evaluation of integrals of this type is a delicate procedure, and we won't discuss it in this course. Note, however, that for large k^μ the integral behaves as

$$\int \frac{d^4k}{k^8}$$

and so is convergent. This is not generally the case: in many situations loop integrals diverge, giving infinite coefficients at each order in perturbation theory. This was a serious problem in the early years of quantum field theory. However, it turns out that these infinities are similar in spirit to the infinity we faced when we found a divergent vacuum energy. By a sufficiently shrewd redefinition of the parameters in the Lagrangian, all infinities in observable quantities may be eliminated. There is a well-defined procedure known as *renormalization* which accomplishes this feat.

4.9 The Yukawa Potential

People were scattering nucleons off nucleons long before quantum field theory was around, and at low energies they could describe scattering processes adequately with non-relativistic quantum mechanics. Let's look at the nonrelativistic limit of the "nucleon-nucleon" scattering amplitude and try to understand it in terms of NRQM.

First of all, recall[8] the *Born approximation* from NRQM: at first order in perturbation theory, the amplitude for an incoming state with momentum $\vec{k}$ to scatter off a potential $U(\vec{r})$ into an outgoing state with momentum $\vec{k}'$ is proportional to the Fourier transform of the potential,

$$\begin{aligned} i\mathcal{A}_{\rm NR}(\vec{k} \to \vec{k}') &\propto \int d^3r\, e^{-i(\vec{k}'-\vec{k})\cdot\vec{r}} U(\vec{r}) \\ &\equiv \tilde{U}(\vec{k}' - \vec{k}). \end{aligned} \quad (4.101)$$

In the centre of mass frame, two-body scattering is simplified to the problem of scattering off a potential, both classicially and quantum mechanically. Now, in the nonrelativistic limit, the meson propagator from the first diagram in Fig. (4.8) is

$$\frac{i}{(p_1 - p_2)^2 - \mu^2} \simeq \frac{-i}{|\vec{p}_1 - \vec{p}_2|^2 + \mu^2} \quad (4.102)$$

where we have taken the nonrelativistic limit, $p_1^0 = p_2^0 = \mu$. Therefore, to explain the first term in the scattering amplitude, Eq. (4.83), we must have

$$\tilde{U}(\vec{q}) \propto \frac{g^2}{|\vec{q}|^2 + \mu^2}. \quad (4.103)$$

[8]See, for example, Cohen-Tannoudji, Diu and Lalo¨e, *Quantum Mechanics*, Vol. II, Chapter VIII, especially section B. 4

It is a simple matter to invert the Fourier transform to find $U(\vec{r})$:

$$\begin{aligned} U(\vec{r}) &\propto \int \frac{d^3q}{(2\pi)^3}\frac{-g^2}{|\vec{q}|^2+\mu^2}e^{i\vec{q}\cdot r} \\ &= -\frac{g^2}{4\pi^2}\int_0^\infty dq\, q^2\frac{e^{iqr}-e^{-iqr}}{iqr}\frac{1}{q^2+\mu^2} \\ &= -\frac{g^2}{4\pi r}\int_{-\infty}^{\infty} dq\,\frac{qe^{iqr}}{q^2+\mu^2}. \end{aligned} \tag{4.104}$$

Closing the contour of the integral in the upper half complex plane, we pick up the residue of the single pole at $q=+i\mu$. Thus we find

$$U(\vec{r}) \propto \frac{g^2}{r}e^{-\mu|\vec{r}|}. \tag{4.105}$$

If we had been careful about the signs, we would have found that the sign of the potential is negative; thus, the first diagram corresponds to a nucleon scattering off an *attractive* "Yukawa potential." The range of the potential is $1/\mu$, the Compton wavelength of the ϕ meson. This was the potential written down by Yukawa to describe the nuclear force. He worked backward from the range of the force (about 1 fm) to predict the mass (about 200 MeV) of the required boson, the pion.

What about the second term? This corresponds to something in NRQM called the "exchange potential." It arises due to the indistinguishability of the scattering pions, and as we have seen, its presence is required by Bose statistics.

Note that the sign of the Yukawa term in the amplitude is the same in antinucleon-antinucleon scattering, nucleon-nucleon scattering, and nucleon-antinucleon scattering. Thus we conclude that the potential due to scalar boson exchange is universally attractive, in contrast to the electrostatic potential.

5 Decay Widths, Cross Sections and Phase Space

At this stage we are now able to calculate amplitudes for a variety of processes by evaluating Feynman diagrams,

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

but we have yet to make contact with anything measurable. In order to calculate *probabilities*, we must square the amplitudes and sum over all observed final states. But it looks like the probability is going to be proportional to

$$|S_{fi}|^2 \sim |\delta^{(4)}(p_F-p_I)|^2.$$

$|\delta^{(4)}(p_F-p_I)|^2$?? Squaring a delta function makes no sense. What happened?

The problem is that we are not working with "square-integrable" states. Instead, our states are normalized to $\delta^{(3)}$ functions. They are not normalizable because they are plane waves, existing at every point in space-time. Thus the scattering process is in fact occurring at every point in space, for all time. No wonder we got divergent nonsense. This is clearly not what we wanted.

The proper way to solve this problem is to take our plane wave states and build up localized wave packets, which are normalizable and for which the scattering process really is restricted to some finite region of space-time. Another approach, which is simpler and will give the right answer, is to return to our old crutch and put the system in a box of volume V, and turn the interaction on for only a finite time T. This will solve the normalization problem because plane wave states in the box are square-integrable, the states being normalized to

$$\langle\vec{k}\,|\vec{k}'\rangle = \delta_{\vec{k}\vec{k}'} \tag{5.1}$$

instead of $\delta^{(3)}(\vec{k}-\vec{k}')$. Furthermore, if we divide our answer by T, we will get the transition probability/unit time, which is really what we are interested in. Finally, we can take the limit $T, V\to\infty$ and hope it makes sense (it will).

As we discussed earlier in the course, in a box measuring L on each side with periodic boundary conditions, the allowed values of momenta must be of the form

$$k_x=\frac{2\pi n_x}{L},\;\; k_y=\frac{2\pi n_y}{L},\;\; k_z=\frac{2\pi n_z}{L} \tag{5.2}$$

where $n_x,\ n_y$ and n_z are integers, as shown in the k_x-k_y plane in Fig. (5.1). The integrals over momentum for the expansion of the fields therefore becomes a sum over discrete momenta, and the scalar field ϕ has the expansion

$$\phi(x)=\sum_{\vec{k}}\left[\frac{a_k e^{-ik\cdot x}}{\sqrt{2\omega_k}\sqrt{V}}+\frac{a_k^\dagger e^{ik\cdot x}}{\sqrt{2\omega_k}\sqrt{V}}\right] \tag{5.3}$$

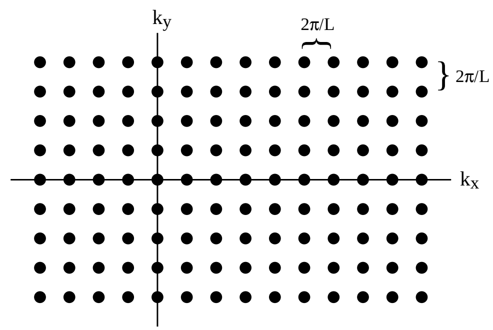

Figure 5.1: Allowed values of k_x and k_y in a box of measuring L on each side.

(you can check that this is the right expansion by seeing that the commutation relations for $a_k^\dagger$ and a_k reproduce the correct canonical commutation relations for the fields). Switching back to our non-relativistic normalization for our states, we see that each time a field creates or annihilates a state it will bring in an additional factor of

$$\frac{e^{\pm ik\cdot x}}{\sqrt{2\omega_k}\sqrt{V}} \tag{5.4}$$

(in contrast to the factor of $e^{\pm ik\cdot x}$ we had in the last set of lecture notes, when we were working with relativistically normalized states in infinite volume). Thus, we have for finite $V = L^3$ and T,

$$\langle f\,|(S-1)|\,i\rangle_{VT} = i\mathcal{A}_{fi}^{VT}(2\pi)^4\delta_{VT}^{(4)}(p_F - p_I) \times \prod_f \frac{1}{\sqrt{2\omega_k}\sqrt{V}} \prod_i \frac{1}{\sqrt{2\omega_k}\sqrt{V}} \tag{5.5}$$

where the products are over final (f) and initial (i) particles, and the notation VT indicates finite volume and time. The function

$$\delta_{VT}^{(4)}(p) \equiv \frac{1}{(2\pi)^4}\int_{-T/2}^{T/2} dt \int_V d^3\vec{x}\, e^{ip\cdot x} \tag{5.6}$$

approaches a δ function in the $V, T \to \infty$ limit.

Each quantity in Eq. (5.5) is finite, so squaring it is now sensible. However, since we want to make contact with the real world, we note that no experimentalist can measure the cross section for the scattering process $N(p_1)+N(p_2) \to N(p_1')+N(p_2')$ for any particular values of the momenta since it is impossible to resolve a single state. It is only possible to measure all states about some small region Δk in

momentum space. From the figure, it is clear that in a region of size $\Delta k_x \Delta k_y \Delta k_z$, there are

$$\frac{L}{2\pi}\Delta k_x \frac{L}{2\pi}\Delta k_y \frac{L}{2\pi}\Delta k_z = \frac{V}{(2\pi)^3}\Delta k_x \Delta k_y \Delta k_z \tag{5.7}$$

states. If there are N particles in the final state, in the infinitesimal region of size $d^3p_1 d^3p_2 ... d^3p_N$ there will be

$$\prod_{f=1}^{N} \frac{V}{(2\pi)^3} d^3p_f \tag{5.8}$$

states, which must be summed over. Squaring our expression for the amplitude, summing over all final states and dividing by the total time T, we find the following expression for the differential transition probability per unit time w_{VT}/T:

$$\frac{w_{VT}}{T} = \frac{1}{T}|\mathcal{A}_{fi}^{VT}|^2 (2\pi)^8 \left|\delta_{VT}^{(4)}(p_F - p_I)\right|^2 \times \prod_f \frac{d^3p_f}{(2\pi)^3 2\omega_p} \prod_i \frac{1}{2\omega_i V}. \tag{5.9}$$

Note that the factors of V cancel in the product over final particles. We will find that for decay rates and cross sections the V in the product over initial particles also cancel. The only tricky part of taking the limit $V, T \to \infty$ is the $|\delta_{VT}^{(4)}(p)|^2$ function. This will approach a function which is infinitely peaked at the origin, so we might anticipate it will be proportional to a delta function. So let's look at

$$\int d^4p \left|\delta_{VT}^{(4)}(p)\right|^2 = \frac{1}{(2\pi)^8}\int d^4p \int_{-T/2}^{T/2} dt \int_{-T/2}^{T/2} dt' \int_V d^3\vec{x} \int_V d^3\vec{x}'\, e^{ip\cdot x - ip\cdot x'}. \tag{5.10}$$

Performing the integral $\int d^4p$, the exponential factors just give us $(2\pi)^4\delta^{(4)}(x - x')$, so we can trivially do the integrals over t' and x', and we find

$$\int d^4p \left|\delta_{VT}^{(4)}(p)\right|^2 = \frac{1}{(2\pi)^4}\int_{-T/2}^{T/2} dx \int_V d^3\vec{x} = \frac{VT}{(2\pi)^4}. \tag{5.11}$$

So indeed $\left|\delta^{(4)}(p)\right|^2$ is proportional to a δ function, with a coefficient which diverges in the limit $T, V \to \infty$:

$$\lim_{V,T\to\infty} \left|(2\pi)^4 \delta_{VT}^{(4)}(p)\right|^2 = VT(2\pi)^4\delta^{(4)}(p). \tag{5.12}$$

Substituting this into Eq. (5.9) and taking the limit $V, T \to \infty$, we find

$$\begin{aligned} \frac{w}{T} &= |\mathcal{A}_{fi}|^2 V (2\pi)^4 \delta_{VT}^{(4)}(p_F - p_I) \times \prod_{\text{final particles } f} \frac{d^3p_f}{(2\pi)^3 2E_f} \prod_{\text{initial particles } i} \frac{1}{2E_i V} \\ &\equiv |\mathcal{A}_{fi}|^2 V D \prod_{\text{initial particles } i} \frac{1}{2E_i V}, \end{aligned} \tag{5.13}$$

where we are using E and ω interchangeably for the energies of the particles, and we have defined the factor D by

$$D \equiv (2\pi)^4 \delta^{(4)}(p_F - p_I) \prod_{\text{final particles } f} \frac{d^3 p_f}{(2\pi)^3 2E_f}. \tag{5.14}$$

Note that D is manifestly Lorentz invariant, since the measure $d^3p_f/(2\pi)^3 2E_f$ is the invariant measure we derived earlier on. Also note that just as in the case of our Feynman rules, each $\delta^{(n)}$ function comes with a factor of $(2\pi)^n$, and each integration $d^n k$ comes with a factor of $(2\pi)^{-n}$, so there is no excuse for getting the factors of 2π wrong.

Now, in fact we are only really interested in processes with one or two particles in the initial state (but still an arbitrary number of particles in the final state), corresponding to decays and $2 \to N$ particle scattering. The relevant physical quantities we wish to calculate are lifetimes and cross sections. So let's examine each of these in turn.

5.1 Decays

For a decay process there is a single particle in the initial state, so

$$\frac{w}{T} = \frac{1}{2E} |\mathcal{A}_{fi}|^2 D. \tag{5.15}$$

Note that the factors of V have cancelled, as they must in order to have a sensible $V \to \infty$ limit. In the particle's rest frame, we will define the quantity $d\Gamma$ as the differential decay probability/unit time:

$$d\Gamma \equiv \frac{1}{2M} |\mathcal{A}_{fi}|^2 D. \tag{5.16}$$

Then the total decay probability/unit time, Γ, is

$$\Gamma = \frac{1}{2M} \int_{\text{all final states}} |\mathcal{A}_{fi}|^2 D. \tag{5.17}$$

Since the probability of the particle decaying/unit time is Γ, after a time t the probability that the particle has not decayed is just $e^{-\Gamma t}$. Therefore, $\Gamma = 1/\tau$, where τ is the particle's lifetime (in natural units). Γ is called the "decay width." If we consider the uncertainty principle, we see that it does in fact correspond to a width. Since the particle exists for a time τ, any measurement of its energy (or mass, in its rest frame) must be uncertain by $\sim 1/\tau = \Gamma$. Thus, a series of measurements of the particle's mass will have a characteristic spread of order Γ, as indicated in Fig. (5.2).

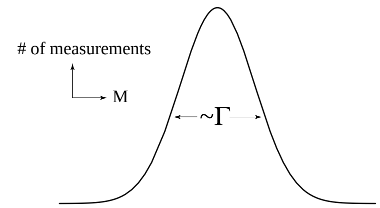

Figure 5.2: The result of a series of measurements of the mass of a particle with lifetime $\tau = 1/\Gamma$. The width of the distribution is proportional to Γ.

5.2 Cross Sections

In a physical scattering experiment, a beam of particles is collided with a target (or another beam of particles coming in the opposite direction), and a measurement is made of the number of particles incident on a detector. So for an incident flux F =# of particles/unit time/unit area, an infinitesimal detector element will record some number dN scatterings/unit time

$$dN = F d\sigma \tag{5.18}$$

where $d\sigma$ is called the differential cross section. The total number of scatterings per unit time is then $N = F\sigma$, where σ is the total cross section. With this definition, we have

$$\begin{aligned} d\sigma &= \frac{\text{differential probability}}{\text{unit time} \times \text{unit flux}} \\ &= \frac{\mathcal{A}_{fi}^2}{4E_1E_2V} D \times \frac{1}{\text{flux}} \\ &= \frac{\mathcal{A}_{fi}^2}{4E_1E_2} \frac{1}{|\vec{v_1} - \vec{v_2}|} D \end{aligned} \tag{5.19}$$

where $\vec{v}_1$ and $\vec{v}_2$ are the 3-velocities of the colliding particles, in terms of which the flux is $|\vec{v_1} - \vec{v_2}|/V$. This is easy to see. Consider first a beam of particles moving perpendicular to a plane of area A and moving with 3-velocity $\vec{v}$. If the density of particles is d, then after a time t, the total number of particles passing through the plane is

$$N = |\vec{v}| A t d. \tag{5.20}$$

Therefore the flux is $N/At = |\vec{v}|d$. With our normalization, there is one particle in the box of volume V, so $d = 1/V$, and the flux is $|\vec{v}|/V$. In the case of two beams colliding, the probability of finding either particle in a unit volume is $1/V$, but since the collision can occur anywhere in the box the total flux is $|\vec{v}_1 - \vec{v}_2|/V^2 \times V = |\vec{v}_1 - \vec{v}_2|/V$.

From Eq. (5.19), the total cross section is

$$\sigma = \frac{1}{4E_1E_2}\frac{1}{|\vec{v}_1 - \vec{v}_2|}\int_{\text{all final states}} |\mathcal{A}_{fi}|^2 D. \tag{5.21}$$

Once again, the factors of V cancel and the result is well-behaved in the limit $T, V \to \infty$.

5.3 D for Two Body Final States

These formulas for the decay widths and the cross sections are true for arbitrary numbers of particles in the final states. For two particles, there are six integrals to do ($d^3\vec{p}_1 d^3\vec{p}_2$), but four of the variables are constrained by the energy-momentum conserving δ function $\delta^{(4)}(p_1 + p_2 - p_I)$, leaving only two independent variables to integrate over. Thus, we can write D in a simpler form. For a two-body final state,

$$D = \int \frac{d^3\vec{p}_1}{(2\pi)^3 2E_1}\frac{d^3\vec{p}_2}{(2\pi)^3 2E_2}(2\pi)^4\delta^{(4)}(p_1 + p_2 - p_I). \tag{5.22}$$

In the centre of mass frame, $\vec{p}_I = 0$, and $E_i \equiv E_T$, the total energy available in the process. Therefore

$$\begin{aligned} D &= \frac{d^3\vec{p}_1}{(2\pi)^3 2E_1}\frac{d^3\vec{p}_2}{(2\pi)^3 2E_2}(2\pi)^3\delta^{(3)}(\vec{p}_1 + \vec{p}_2)(2\pi)\delta(E_1 + E_2 - E_T) \\ &\Rightarrow \frac{d^3\vec{p}_1}{(2\pi)^3 4E_1E_2}(2\pi)\delta(E_1 + E_2 - E_T) \\ &= \frac{1}{(2\pi)^3 4E_1E_2}p_1^2 dp_1 d\Omega_1 (2\pi)\delta(E_1 + E_2 - E_T) \end{aligned} \tag{5.23}$$

where we have performed the integral over $\vec{p}_2$, and $\vec{p}_2 = -\vec{p}_1$ is now implicit. We have also written $d^3\vec{p}_1 = p_1^2 dp_1 d\cos\theta_1 d\phi_1 \equiv p_1^2 dp_1 d\Omega_1$, where θ and ϕ are the polar angles of $\vec{p}_1$.

To eliminate the last dependent variable, the δ function of energy must be converted to a δ function of p_1. Using the general formula

$$\delta[f(x)] = \sum_{x_0 \in \text{zeroes of } f} \frac{1}{|f'(x_0)|}\delta(x - x_0) \tag{5.24}$$

to change variables from E_1 to p_1, we must include a factor of

$$\left|\frac{\partial(E_1 + E_2)}{\partial p_1}\right|^{-1}.$$

Since $E_1^2 = p_1^2 + m^2$, and $E_2^2 = p_2^2 + m^2 = p_1^2 + m^2$ (from $\vec{p}_2 = -\vec{p}_1$),

$$\frac{\partial E_1}{\partial p_1} = \frac{p_1}{E_1}, \quad \frac{\partial E_2}{\partial p_2} = \frac{p_1}{E_2} \tag{5.25}$$

and so

$$\left|\frac{\partial(E_1 + E_2)}{\partial p_1}\right| = \frac{p_1(E_1 + E_2)}{E_1E_2} = \frac{p_1 E_T}{E_1E_2}. \tag{5.26}$$

The desired result for a two body final state *in the centre of mass frame* is therefore

$$D = \frac{1}{16\pi^2}\frac{p_1\, d\Omega_1}{E_T}. \tag{5.27}$$

In this derivation, we have assumed that the particles A and B in the final state are *distinguishable*, because we treated the final states $|A(\vec{p}_1), B(\vec{p}_2)\rangle$ and $|A(\vec{p}_2), B(\vec{p}_1)\rangle$ as distinct. In fact, if the particles are identical then these states are in fact identical, and so we've double-counted by a factor of 2!. In general, for n identical particles in the final state, we must multiply D by a factor of $1/n!$.

Now let's apply this to a couple of examples. Going back to our QMD theory, suppose $\mu^2 > 4m^2$, so that the decay $\phi \to \overline{N}N$ is kinematically allowed. There is only one diagram contributing to this decay at leading order in perturbation theory, shown in Fig. (5.3), so $i\mathcal{A} = -ig$ (simple!) and the decay width of the ϕ is

$$\begin{aligned} \Gamma &= \frac{g^2}{2\mu}\frac{p_1}{16\pi^2\mu}\int d\Omega_1 \\ &= \frac{g^2 p_1}{8\pi\mu^2} \end{aligned} \tag{5.28}$$

since $\int d\Omega = 4\pi$. p_1 is straightforward to compute from energy-momentum con-

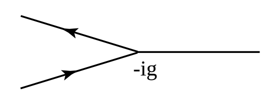

Figure 5.3: Leading contribution to $\mu \to \overline{N}N$.

servation. The initial four-momentum is $(\mu, \vec{0})$ and the final momenta of the nucleons are $P_1 = (\sqrt{p_1^2 + m^2}, \vec{p}_1)$, $P_2 = (\sqrt{p_1^2 + m^2}, -\vec{p}_1)$, so $p_1 = \sqrt{\mu^2 - 4m^2}/2$.

As a second example, we consider $2 \to 2$ particle scattering in the centre of mass frame. Since the results which follow are just kinematics and don't depend on the amplitude $\mathcal{A}$, they are valid in *any* theory. The 3-velocities are $\vec{v}_1 =$

$\vec{p}_1/(\gamma m_1) = \vec{p}_1/E_1$, and $\vec{v}_2 = \vec{p}_2/E_2 = -\vec{p}_1/E_2$, so

$$|\vec{v_1} - \vec{v_2}| = p_1 \left(\frac{1}{E_1} + \frac{1}{E_2}\right) = p_1 \frac{E_2 + E_1}{E_1 E_2} = \frac{p_1 E_T}{E_1 E_2}. \tag{5.29}$$

This leads to

$$d\sigma = \frac{1}{4p_i E_T} \frac{1}{16\pi^2} \frac{p_f d\Omega_1}{E_T} |\mathcal{A}_{fi}|^2 \tag{5.30}$$

and so

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2 E_T^2} \frac{p_f}{p_i} |\mathcal{A}_{fi}|^2 \tag{5.31}$$

where p_i and p_f are the magnitudes of the three-momenta of the incoming and outgoing particles, respectively.

5.4 D for 3 Body Final States

For three body final states, there are nine integrals to do and four constraints from the δ function, leaving five independent variables. The derivation is straightforward but more lengthy than for the 2 body final state, so we will just quote the result here. If the outgoing particles have energies E_1, E_2 and E_3, then we will choose the independent variables to be $E_1,\ E_2,\ \theta_1, \phi_1$ and ϕ_{12}, where ϕ_{12} is the angle between particles 1 and 2. In terms of these variables,

$$D = \frac{1}{256\pi^5} dE_1\, dE_2\, d\Omega_1\, d\phi_{12} \tag{5.32}$$

in the centre of mass frame. In some cases (such as the decay of a spinless meson), the amplitude is independent of Ω_1 and ϕ_{12}; in this case, we can integrate over those three variables ($\int d\Omega_1 d\phi_{12} = 8\pi^2$) to obtain

$$D = \frac{1}{32\pi^3} dE_1 dE_2. \tag{5.33}$$

6 Discrete Symmetries: C, P and T

6.1 Charge Conjugation, C

In our "nucleon"-meson theory, we noticed that the amplitude for $N(p_1)+\phi(p_2) \to N(p_1')+\phi(p_2')$ was the same as for $\overline{N}(p_1)+\phi(p_2) \to \overline{N}(p_1')+\phi(p_2')$. The same is true for other processes in this theory which differ only by the exchange of particles for antiparticles, such as $NN \to NN$ and $\overline{NN} \to \overline{NN}$. It is also true to all orders in perturbation theory, not only at $\mathcal{O}(g^2)$, and arises because the theory has an additional symmetry which we have neglected until now, called charge conjugation and denoted C. This is a symmetry transformation which interchanges particles with their antiparticles. Unlike the previous symmetries we have seen, it is not a continuous symmetry parameterized by some continuously varying parameter, and it does not correspond to a conserved current. Instead, it is a *discrete* symmetry.

Given an arbitrary state $|N(\vec{k}_1), \overline{N}(\vec{k}_2), ..., N(\vec{k}_n)\rangle$ composed of nucleons and anti-nucleons we can define a unitary operator U_c which effects this discrete transformation. Clearly,

$$U_c|N(\vec{k}_1), \overline{N}(\vec{k}_2), ..., N(\vec{k}_n)\rangle = |\overline{N}(\vec{k}_1), N(\vec{k}_2), ..., \overline{N}(\vec{k}_n)\rangle. \tag{6.1}$$

We also see that with this definition $U_c^2 = 1$, so $U_c^{-1} = U_c^\dagger = U_c$. We can now see how the fields transform under C. Consider some general state $|\,\psi\rangle$ and its charge conjugate $|\,\overline{\psi}\rangle$. Then $b_k^\dagger|\,\psi\rangle \equiv |N(k), \psi\rangle$, and

$$U_c b_k^\dagger|\,\psi\rangle = U_c|N(k), \psi\rangle = |\overline{N}(k), \overline{\psi}\rangle = c_k^\dagger|\,\overline{\psi}\rangle = c_k^\dagger U_c|\,\psi\rangle. \tag{6.2}$$

Since this is true for arbitrary states $|\,\psi\rangle$, we must have $U_c b_k^\dagger = c_k^\dagger U_c$, or

$$b_k^\dagger \to b_k'^\dagger = U_c b_k^\dagger U_c^\dagger = c_k^\dagger,\ \ c_k^\dagger \to c_k'^\dagger = U_c c_k^\dagger U_c^\dagger = b_k^\dagger. \tag{6.3}$$

A similar equation is true for annihilation operators, which is easily seen by taking the complex conjugate of both equations. As expected, the transformation exchanges particle creation operators for anti-particle creation operators, and vice-versa. Expanding the fields in terms of creation and annihilation operators, we immediately see that

$$\psi \to \psi' = U_c \psi U_c^\dagger = \psi^\dagger,\ \ \psi^\dagger \to \psi'^\dagger = U_c \psi^\dagger U_c^\dagger = \psi. \tag{6.4}$$

It is clearly a symmetry of the theory, $U_c L U_c^\dagger = L$, since ψ and $\psi^\dagger$ always occur together in each term of the Lagrangian. It is straightforward to show that scattering matrix elements are therefore unchanged by charge conjugation. Denoting the initial and final states by $|\,i\rangle$ and $\langle f\,|$ and their charge conjugates by $|\,\bar{\imath}\rangle$ and $\langle\overline{f}\,|$,

$$\begin{aligned}\langle f\,|S|\,i\rangle &= \langle f\,|U_c^\dagger U_c S U_c^\dagger U_c|\,i\rangle \\ &= \langle\overline{f}\,|U_c S U_c^\dagger|\,\bar{\imath}\rangle \\ &= \langle\overline{f}\,|S|\,\bar{\imath}\rangle \end{aligned} \tag{6.5}$$

since $S = T\exp(i\int d^4x\mathcal{H}_I)$, and so $U_c\mathcal{H}_I U_c^\dagger = \mathcal{H}_I \Rightarrow U_c S U_c^\dagger = S$.

While we're at it, there are two other discrete symmetries of $\mathcal{L}_I$ which will be useful in other contexts (it is perhaps worth pointing out here that none of these three symmetries is particularly interesting in this simple theory we are studying. However, they will be much more interesting later on, when we study theories with spin 1/2 and spin 1 fields). These are *parity* (P) and *time reversal* (T). We will look at these in turn.

6.2 Parity, P

A parity transformation corresponds to a reflection of the axes through the origin, $\vec{x} \to -\vec{x}$. Similarly, momenta are reflected, so

$$U_p|\,\vec{k}\rangle = |\,-\vec{k}\rangle \tag{6.6}$$

where U_p is the unitary operator effecting the parity transformation. Clearly we will also have

$$U_p \left\{ \begin{array}{c} a_{\vec{k}} \\ a^\dagger_{\vec{k}} \end{array} \right\} U_p^\dagger = \left\{ \begin{array}{c} a_{-\vec{k}} \\ a^\dagger_{-\vec{k}} \end{array} \right\} \tag{6.7}$$

and so under a parity transformation the fields have the transformation

$$\begin{aligned}\phi(\vec{x},t) &\to U_p\phi(\vec{x},t)U_p^\dagger \\ &= U_p\int\frac{d^3k}{\sqrt{2\omega_k}(2\pi)^{3/2}}\left[a_k e^{i\vec{k}\cdot\vec{x}-i\omega_k t} + a_k^\dagger e^{-i\vec{k}\cdot\vec{x}+i\omega_k t}\right]U_p^\dagger \\ &= \int\frac{d^3k}{\sqrt{2\omega_k}(2\pi)^{3/2}}\left[a_{-\vec{k}} e^{i\vec{k}\cdot\vec{x}-i\omega_k t} + a_{-\vec{k}}^\dagger e^{-i\vec{k}\cdot\vec{x}+i\omega_k t}\right] \\ &= \int\frac{d^3k}{\sqrt{2\omega_k}(2\pi)^{3/2}}\left[a_k e^{-i\vec{k}\cdot\vec{x}-i\omega_k t} + a_k^\dagger e^{i\vec{k}\cdot\vec{x}+i\omega_k t}\right] \\ &= \phi(-\vec{x},t) \end{aligned} \tag{6.8}$$

where we have changed variables $\vec{k} \to -\vec{k}$ in the integration. Just as before, $U_p L U_p^\dagger = L$, so our theory conserves parity.

Actually, this transformation $\phi(\vec{x},t) \to \phi(-\vec{x},t)$ is not unique. Suppose we had a theory with an additional discrete symmetry $\phi \to -\phi$. This is not true for our nucleon-meson theory, since the interaction term changes sign under this transformation, but it is a symmetry of the Lagrangian $\mathcal{L} = \mathcal{L}_0 - \lambda\phi^4/4!$ which we looked at briefly in the last section. In this case, we could equally well have defined the fields to transform under parity as

$$\phi(\vec{x},t) \to -\phi(-\vec{x},t), \tag{6.9}$$

since that is also a symmetry of $\mathcal{L}$. In fact, to be completely general, if we had a theory of n identical fields $\phi_1...\phi_n$, we could define a parity transformation to be of the form

$$\phi_a(\vec{x},t) \to \phi'_a(\vec{x},t) = R_{ab}\phi_b(-\vec{x},t) \tag{6.10}$$

for some $n\times n$ matrix R_{ab}. So long as this transformation is a symmetry of $\mathcal{L}$ it is a perfectly decent definition of parity. The point is, if you have a number of discrete symmetries of a theory there is always some ambiguity in how you define P (or C, or T, for that matter). But this is just a question of terminology. The important thing is to recognize the symmetries of the theory.

In our meson-nucleon theory, $\phi \to -\phi$ is not a symmetry, so the only sensible definition of parity is Eq. (6.8). When ϕ does not change sign under a parity transformation, we call it a *scalar*. In other situations, Eq. (6.8) is not a symmetry of the theory, but Eq. (6.9) is. In this case, we call ϕ a *pseudoscalar*. When there are only spin-0 particles in the theory, theories with pseudoscalars look a little contrived. The simplest example I've seen is

$$\mathcal{L} = \tfrac{1}{2}\sum_{a=1}^{4}\left(\partial^\mu\phi_a\partial_\mu\phi_a - m_a^2\phi_a^2\right) - i\epsilon^{\mu\nu\alpha\beta}\partial_\mu\phi_1\partial_\mu\phi_2\partial_\mu\phi_3\partial_\mu\phi_4 \tag{6.11}$$

where $\epsilon^{\mu\nu\alpha\beta}$ is a completely antisymmetric four-index tensor, and $\epsilon^{0123} = 1$. Under parity, if $\phi_a(\vec{x},t) \to \pm\phi_a(-\vec{x},t)$, then

$$\begin{aligned}\partial_0\phi_a(\vec{x},t) &\to \pm\partial_0\phi_a(-\vec{x},t) \\ \partial_i\phi_a(\vec{x},t) &\to \mp\partial_i\phi_a(-\vec{x},t)\end{aligned} \tag{6.12}$$

where $i = 1,2,3$, since parity reverses the sign of $\vec{x}$ but leaves t unchanged. Now, the interaction term in Eq. (6.11) always contains three spatial derivatives and one time derivative because $\epsilon^{\mu\nu\alpha\beta} = 0$ unless all four indices are different. Therefore in order for parity to be a symmetry of this Lagrangian, an odd number of the fields ϕ_a (it doesn't matter which ones) must also change sign under a parity transformation. Thus, three of the fields will be scalars, and one pseudoscalar, or else three must be pseudoscalars and one scalar. It doesn't matter which.

6.3 Time Reversal, T

The last discrete symmetry we will look at is time reversal, T, in which $t \to -t$. A more symmetric transformation is PT in which all four components of x^μ flip sign: $x^\mu \to -x^\mu$. However, time reversal is a little more complicated than P and T because it cannot be represented by a unitary, linear transformation.

We can see why this is the case by going back to particle mechanics and quantizing the Lagrangian

$$L = \frac{1}{2}\dot{q}^2. \tag{6.13}$$

Suppose the unitary operator U_T corresponds to T. Then

$$\begin{aligned} U_T q(t) U_T^\dagger &= q(-t) \\ U_T p(t) U_T^\dagger &= U_T \frac{dq(t)}{dt} U_T^\dagger = -\dot{q}(-t) = -p(-t) \end{aligned} \tag{6.14}$$

and so

$$U_T[q(t), p(t)]U_T^\dagger = U_T i U_T^\dagger = i = -[q(-t), p(-t)] \tag{6.15}$$

and so we cannot consistently apply the canonical commutation relations for all time! Clearly U_T can't be a unitary operator. We need something else.

What we need, in fact, is an operator which is *anti*-linear. Under an antilinear operator Ω,

$$a|\,\psi\rangle \to \Omega\,[a|\,\psi\rangle] = a^*\Omega|\,\psi\rangle. \tag{6.16}$$

That is, numbers are complex conjugated under an antilinear transformation. Since Dirac notation is set up to deal with linear operators, it is somewhat awkward to express antilinear operators in this notation.

The simplest anti-linear operator is just complex conjugation,

$$a|\,\psi\rangle \to \Omega\,[a|\,\psi\rangle] = a^*|\,\psi\rangle \tag{6.17}$$

and in fact this is precisely what we need. First of all, it doesn't screw up the commutation relations because $\Omega i \Omega^{-1} = -i$, so there is an extra minus sign in Eq. (6.15) and there is no contradiction:

$$\Omega_T[aq(t)]\Omega_T^{-1} = a^* q(-t) \tag{6.18}$$

so

$$\Omega_t[q(t), p(t)]\Omega_T^{-1} = i^* = -i = -[q(-t), p(-t)] \tag{6.19}$$

as required. In field theory, complex conjugation corresponds to the operator PT. It has no effect on the creation and annihilation operators,

$$\Omega_{PT} \left\{ \begin{matrix} a_k \\ a_k^\dagger \end{matrix} \right\} \Omega_{PT}^{-1} = \left\{ \begin{matrix} a_k \\ a_k^\dagger \end{matrix} \right\} \tag{6.20}$$

or on the states

$$\Omega_{PT}|\vec{k}_1, ...\vec{k}_n\rangle = |\vec{k}_1, ...\vec{k}_n\rangle \tag{6.21}$$

(this is to be expected, since time reversal flips the direction of all the momenta, and a parity transformation flips them back). The only thing it acts on is the i in the exponents occurring in the expansion of the fields

$$\begin{aligned} \phi(\vec{x}, t) &\to \Omega_{PT}\phi(\vec{x}, t)\Omega_{PT}^\dagger \\ &= \Omega_{PT} \int \frac{d^3k}{\sqrt{2\omega_k}(2\pi)^{3/2}} \left[a_k e^{i\vec{k}\cdot\vec{x} - i\omega_k t} + a_k^\dagger e^{-i\vec{k}\cdot\vec{x} + i\omega_k t}\right] \Omega_{PT}^{-1} \\ &= \int \frac{d^3k}{\sqrt{2\omega_k}(2\pi)^{3/2}} \left[a_k e^{-i\vec{k}\cdot\vec{x} + i\omega_k t} + a_k^\dagger e^{i\vec{k}\cdot\vec{x} - i\omega_k t}\right] \\ &= \phi(-\vec{x}, -t). \end{aligned} \tag{6.22}$$

Hence this is exactly what is required for a PT transformation.

Now, our meson-nucleon Lagrangian was very dull in that it was invariant under each of these three symmetries separately. For example, the amplitude for

$$N(p_1) + N(p_2) \to N(p_1') + N(p_2') \tag{6.23}$$

was related by C to

$$\overline{N}(p_1) + \overline{N}(p_2) \to \overline{N}(p_1') + \overline{N}(p_2'). \tag{6.24}$$

Under P, this becomes

$$\overline{N}(\omega_1, -\vec{p}_1) + \overline{N}(\omega_2, -\vec{p}_2) \to \overline{N}(\omega_1', -\vec{p}_1') + \overline{N}(\omega_2', -\vec{p}_2'). \tag{6.25}$$

Under T, the incoming and outgoing states are reversed, and the signs of the 3-momenta change sign, so under a T transformation this becomes

$$\overline{N}(p_1') + \overline{N}(p_2') \to \overline{N}(p_1) + \overline{N}(p_2). \tag{6.26}$$

In our theory, the amplitudes for all these processes are identical by the symmetries. In a more general theory, any of C, P or T may be broken. However, it is a *general* property of any local, relativistic field theory that the amplitude must be invariant under the combined action of CPT (this is called the CPT theorem). Hence, while the amplitudes Eq. (6.24) and Eq. (6.25) need not be equal to Eq. (6.23) in some more complicated theory, the amplitudes for Eq. (6.23) and the CPT transformed process Eq. (6.26) will always be the same. Diagrammatically, we can see that this ought to be the case. Consider an arbitrary Feynman diagram with four external nucleon lines, indicated in Fig. (6.1), and define the momenta p_1–p_4 to be

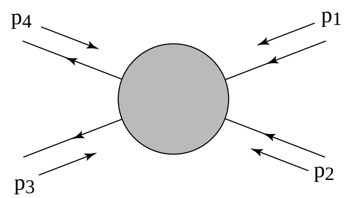

Figure 6.1: A Feynman diagram contributing to the processes Eq. (6.23)–Eq. (6.26). The blob represents an arbitrarily complicated diagram.

directed inward. The blob will be some function $F(p_1, p_2, p_3, p_4)$ of the external momenta. The amplitude for the process (6.23) is then $F(p_1, p_2, -p_1', -p_2')$ while the amplitude for the CPT transformed process, Eq. (6.26) is $F(-p_1, -p_2, p_1', p_2')$ (the diagram doesn't change if we invert it on the page). CPT invariance then simply means that for any diagram,

$$F(p_1, p_2, p_3, p_4) = F(-p_1, -p_2, -p_3, -p_4). \tag{6.27}$$

In a scalar theory, this clearly must be the case. Since F is a Lorentz scalar, it can only depend on the scalar quantities $p_i \cdot p_j$ or (in a more complicated theory) the $\epsilon^{\mu\nu\alpha\beta}$ tensor contracted with the external momenta, $\epsilon^{\mu\nu\alpha\beta} p_\mu p_\nu p_\alpha p_\beta$. All such combinations are invariant under the transformation $p_i \to -p_i$.

The CPT theorem is also true in more general theories with spin, which we will now discuss.

7 Spin 1/2 Fields

7.1 Transformation Properties

So far we have only looked at the theory of an interacting *scalar* field $\phi(x)$. Recall that since ϕ is a scalar, under a Lorentz transformation $x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu$, ϕ transforms according to

$$\phi(x) \to \phi'(x) = \phi(\Lambda^{-1}x). \tag{7.1}$$

This simply states that the field itself does not transform at all; the value of the field at the coordinate x in the new frame is the same as the field at that same point in the old frame. In general, ϕ could have a more complicated transformation law. For example, we could have four fields ϕ^μ, $\mu = 1..4$, which make up the components of a 4-vector. In this case, ϕ^μ will transform under a Lorentz transformation as

$$\phi^\mu(x) \to \phi'^\mu(x) = \Lambda^\mu{}_\nu \phi^\nu(\Lambda^{-1}x). \tag{7.2}$$

In general, a field will transform in some well-defined way under the Lorentz group,

$$\phi_a(x) \to D_{ab}(\Lambda)\phi_b(\Lambda^{-1}x). \tag{7.3}$$

If ϕ_a has n components, $D_{ab}(\Lambda)$ is an $n \times n$ matrix. The matrices $D(\Lambda)$ form an n-dimensional *representation* of the Lorentz group: if Λ_1 and Λ_2 define two Lorentz transformations,

$$\sum_b D_{ab}(\Lambda_1) D_{bc}(\Lambda_2) = D_{ac}(\Lambda_1\Lambda_2). \tag{7.4}$$

In addition, $D(\Lambda^{-1}) = D(\Lambda)^{-1}$, and $D(1) = I$, the identity matrix.

We are interested in describing particles of spin $1/2$. From our previous experience in quantum mechanics, we already know how such objects transform under rotations, a subgroup of the Lorentz group. A spin $1/2$ state $|\,\psi\rangle$ has two components:

$$|\,\psi\rangle = \begin{pmatrix} |\,\psi_\uparrow\rangle \\ |\,\psi_\downarrow\rangle \end{pmatrix}. \tag{7.5}$$

The spin operators S_x, S_y and S_z are given by

$$S_x = \tfrac{1}{2}\sigma_x, \;\; S_y = \tfrac{1}{2}\sigma_y, \;\; S_z = \tfrac{1}{2}\sigma_z \tag{7.6}$$

where σ_x, σ_y and σ_z are the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \;\; \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \;\; \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{7.7}$$

For a particle with no orbital angular momentum, the total angular momentum J is just given by the spin operators. In general, rotations are generated by the angular momentum operator J: [9] a general state $|\,\psi\rangle$ transforms under a rotation about the $\hat{e}$ axis by an angle θ as

$$|\,\psi\rangle \rightarrow U_R(\hat{e},\theta)|\,\psi\rangle \tag{7.8}$$

where the rotation operator

$$U_R(\hat{e},\theta) = e^{-iJ\cdot\hat{e}\theta} \tag{7.9}$$

is unitary. Therefore, a state with total angular momentum $1/2$ transforms under this rotation as[10]

$$|\,\psi\rangle \rightarrow e^{-i\vec{\sigma}\cdot\hat{e}\theta/2}|\,\psi\rangle \tag{7.10}$$

so the matrices

$$e^{-i\vec{\sigma}\cdot\hat{e}\theta/2}$$

form the *spinor* representation of the rotation group. Hence, a quantum field u which creates and annihilates spin $1/2$ particles will transform under rotations according to

$$u'(x) = U_R^\dagger u(x) U_R = e^{-i\vec{\sigma}\cdot\hat{e}\theta/2} u(R^{-1}x) \tag{7.11}$$

where R is the rotation matrix for vectors, and we are suppressing spinor indices.

In a relativistic theory, we must determine the transformation properties of spinors under the full Lorentz group, not just the rotation group. The most satisfying way to do this would be to pause for a moment from field theory and develop the theory of representations of the Lorentz group. We could find all possible representations of the group, and then finally restrict ourselves to the spinor representation. But since we have only a small amount of time, and since we are really not interested in any representations of the Lorentz group beside the scalar, vector (both of which we already understand) and spinor representations, corresponding to particles of spin 0, 1 and $1/2$ respectively, we will simply introduce a nice trick for obtaining the spinor representation from the vector representation. This is not generalizable to other representations, but it will serve our purposes.

Let us return to the rotation subgroup of the Lorentz group. The rotations are the group of transformations $(x,y,z) \rightarrow (x',y',z')$ which leave $r^2 = x^2+y^2+z^2$ invariant (and which retain the handedness of the coordinates, so we do not include reflections). We can assemble the components of a 3-vector into a two by two traceless Hermitian matrix

$$X = \begin{pmatrix} z & x-iy \\ x+iy & -z \end{pmatrix} = x_i\sigma_i. \tag{7.12}$$

A two by two complex matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has eight independent components (two for each complex entry). Hermiticity requires $\mathrm{Re}\,a = \mathrm{Re}\,d$, $\mathrm{Im}\,a = \mathrm{Im}\,d = 0$, $\mathrm{Re}\,b = \mathrm{Re}\,c$ and $\mathrm{Im}\,b = -\mathrm{Im}\,c$, reducing the number of independent components to four, and tracelessness reduces this to three, so Eq. (7.12) is the most general form for a two by two traceless Hermitian matrix. Consider the transformation $X' = UXU^\dagger$, where U is a two by two unitary matrix with unit determinant. Then X' is also a traceless, Hermitian matrix

$$\begin{aligned} \mathrm{Tr}X' &= \mathrm{Tr}UXU^\dagger = \mathrm{Tr}U^\dagger UX = \mathrm{Tr}X \\ X'^\dagger &= \left(UXU^\dagger\right)^\dagger = X' \end{aligned} \tag{7.13}$$

so in general we can write it as

$$X' = \begin{pmatrix} z' & x'-iy' \\ x'+iy' & -z' \end{pmatrix}. \tag{7.14}$$

Since

$$\det X' = \det U \det X \det U^\dagger = \det X \tag{7.15}$$

we have

$$x'^2 + y'^2 + z'^2 = x^2 + y^2 + z^2 \tag{7.16}$$

so this is another way of writing the transformation law of a vector under rotations. It is easy to show by direct matrix multiplication that

$$U = e^{-i\vec{\sigma}\cdot\hat{e}\theta/2} \tag{7.17}$$

is the appropriate transformation matrix for a rotation about the $\hat{e}$ axis.

The U's by themselves form a two-dimensional represenation of the rotation group, and a spinor is defined to be a two-component column vector which transforms under rotations through multiplication by U:

$$u' = Uu. \tag{7.18}$$

This agrees with our previous assertion, Eq. (7.10).

[9] See, for example, Cohen-Tannoudji, Diu and Lalo¨e, *Quantum Mechanics*, Volume 1, Chapter VI (especially Complement $\mathrm{B_{VI}}$).

[10] Recall that the exponential of a matrix M is defi ned by the power series $e^M = 1+M+M^2/2!+M^3/3!+\ldots$. This is only equal to the exponential of the entries in the matrix if M is diagonal.

We can extend this construction to the whole connected Lorentz group. Removing the tracelessness condition on X increases the number of free parameters by one, so it now takes the general form

$$X = \begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix}. \tag{7.19}$$

Now consider the transformation

$$X' = QXQ^\dagger \tag{7.20}$$

where Q is no longer required to be unitary, but still $\det Q = 1$. Then under the transformation Eq. (7.20), $\det X = \det X'$, so

$$t'^2 - x'^2 - y'^2 - z'^2 = t^2 - x^2 - y^2 - z^2 \tag{7.21}$$

and so the transformation corresponds to a (proper[11]) Lorentz transformation. We note that the matrix Q has six independent parameters, which is the same as a proper Lorentz transformation (three independent rotations and three independent boosts).

Consider the transformation of the 4-vector $(1, \vec{0})$ under a boost in the $\hat{z}$ direction,

$$(1, \vec{0}) \to (\gamma, -\sqrt{\gamma^2 - 1}\hat{z}). \tag{7.22}$$

It is convenient to introduce the parameter ϕ, defined by

$$\cosh\phi = \gamma, \quad \sinh\phi = -\sqrt{\gamma^2 - 1} \tag{7.23}$$

where $\gamma = \sqrt{v^2 - 1}$ parameterizes the boost. Then the vector transforms as

$$(1, \vec{0}) \to (\cosh\phi, \sinh\phi). \tag{7.24}$$

It is straightforward to verify that in our matrix representation, this boost corresponds to the transformation matrix $Q_z = \exp(\sigma_z\phi/2)$ (note that there is no i in the exponential; Q_z is Hermitian, not unitary, so $Q_z^\dagger = Q_z$):

$$\begin{aligned} X' &= e^{\phi\sigma_z/2} X e^{\phi\sigma_z/2} \\ &= \begin{pmatrix} e^{\phi/2} & 0 \\ 0 & e^{-\phi/2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} e^{\phi/2} & 0 \\ 0 & e^{-\phi/2} \end{pmatrix} \\ &= \begin{pmatrix} e^{\phi} & 0 \\ 0 & e^{-\phi} \end{pmatrix} \\ &= \begin{pmatrix} \cosh\phi + \sinh\phi & 0 \\ 0 & \cosh\phi - \sinh\phi \end{pmatrix} \end{aligned} \tag{7.25}$$

[11]The *proper* or *connected* Lorentz transformations do not include reflections or time reversal. Any proper Lorentz transformation may be written as a product of a rotation and a boost.

and so $t' = \cosh\phi$ and $z' = \sinh\phi$, as required. In general, you can verify that a boost in the $\hat{e}$ direction is given by

$$Q = e^{\vec{\sigma}\cdot\hat{e}\phi/2}. \tag{7.26}$$

The Q's, the group of unitary two by two matrices with unit determinant (including the rotation matrices U) form a representation of the connected Lorentz group. Under a boost, a spinor transforms as $u \to Qu$.

This construction is not unique. If we have a set of matrices $Q(\Lambda)$ which form a representation of a group, so do the matrices $Q^*(\Lambda)$, as do the matrices $SQ^*(\Lambda)S^\dagger$ for some unitary matrix S. This is easy to verify; for example, the new representation preserves the group multiplication rule:

$$SQ^*(\Lambda_1)S^\dagger SQ^*(\Lambda_2)S^\dagger = S\left[Q(\Lambda_1)Q(\Lambda_2)\right]^* S^\dagger = SQ^*(\Lambda_1\Lambda_2)S^\dagger \tag{7.27}$$

where we have used the fact that the Q's form a representation. However there may or may not be any physical difference between the two representations. Two representations Q and $\tilde{Q}$ are said to be *equivalent* if there is some unitary matrix S such that

$$Q(\Lambda) = S\tilde{Q}(\Lambda)S^\dagger \tag{7.28}$$

for all transformations Λ. This is physically sensible because if two representations are equivalent, I can always transform an object u transforming under one representation to one transforming under the other representation by performing the change of basis

$$u \to Su. \tag{7.29}$$

There is no physics in a change of basis, so a set of fields transforming under Q are physically equivalent to a set of fields transforming under $\tilde{Q}$. On the other hand, if no such matrix S exists, then the two representations are *inequivalent*, and there is a physical difference between fields transforming according the two representations. We will see a practical illustration of this shortly.

For the Lorentz group, the two representations Q and Q^* are, in fact, inequivalent. Therefore there are two different types of spinor fields we can define; those which transform according to Q and those which transform according to Q^*. However, for the rotation subgroup, U and U^* can be shown to be equivalent representations, with $S = i\sigma_2$:

$$U(R) = i\sigma_2 U^*(R)(i\sigma_2)^\dagger = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} U^*(R) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{7.30}$$

for all rotation matrices $U(R) = \exp(-i\vec{\sigma}\cdot\hat{e}\theta/2)$. This is why we never encountered two different types of spinors when dealing only with rotations. However, for boosts, it can be shown that

$$i\sigma_2 \left(e^{-\vec{\sigma}\cdot\hat{e}\phi/2}\right)^* (i\sigma_2)^\dagger = e^{\vec{\sigma}\cdot\hat{e}\phi/2} \neq e^{-\vec{\sigma}\cdot\hat{e}\phi/2} \tag{7.31}$$

and so the two representations Q and $SQ^*S^\dagger$ of the Lorentz group are not equivalent.

Thus we can define two types of spinors which we shall denote by u_+ and u_-. They transform in the same way under rotations,

$$u_\pm \to e^{-i\vec{\sigma}\cdot\hat{e}\theta/2} u_\pm \tag{7.32}$$

but differently under boosts

$$u_\pm \to e^{\pm\vec{\sigma}\cdot\hat{e}\phi/2} u_\pm. \tag{7.33}$$

In group theory jargon, u_+ is said to transform according to the $\mathcal{D}^{(0,1/2)}$ representation of the Lorentz group and u_- transforms according to the $\mathcal{D}^{(1/2,0)}$ representation.

In order to construct Lorentz invariant Lagrangians which are bilinear in the fields, we shall need to know how terms bilinear in the u's transform. Not surprisingly, since in some sense the spinors were the "square roots" of the vectors, we can construct four-vectors from pairs of spinors. First consider the bilinear $u_+^\dagger u_+$. Under a Lorentz transformation,

$$u_+^\dagger u_+ \to u_+^\dagger Q^\dagger Q u_+. \tag{7.34}$$

If Q is purely a rotation, $Q^\dagger = Q^{-1}$ (Q is unitary) so $u_+^\dagger u_+ \to u_+^\dagger u_+$. Therefore it is a scalar under rotations (but not under Lorentz boosts). The three components of

$$u_+^\dagger \vec{\sigma} u_+ \equiv (u_+^\dagger \sigma_x u_+, u_+^\dagger \sigma_y u_+, u_+^\dagger \sigma_z u_+) \tag{7.35}$$

form a three-vector under rotations; for example, under a rotation about the z axis

$$u_+^\dagger \sigma_x u_+ \to u_+^\dagger e^{i\sigma_z\theta/2} \sigma_x e^{-i\sigma_z\theta/2} u_+ = u_+^\dagger \left(\sigma_x \cos\theta - \sigma_y \sin\theta\right) u_+. \tag{7.36}$$

This is the correct transformation of a three vector. Putting these together, the four components of

$$V^\mu = (u_+^\dagger u_+, u_+^\dagger \vec{\sigma} u_+) \tag{7.37}$$

form a four-vector. Again, we can check this explictly for a boost in the z direction:

$$\begin{aligned} V^0 &= u_+^\dagger u_+ \to u_+^\dagger e^{\sigma_z\phi} u_+ = u_+^\dagger(\cosh\phi + \sigma_z \sinh\phi) u_+ \\ &\qquad = \cosh\phi\, u_+^\dagger u_+ + \sinh\phi\, u_+^\dagger \sigma_z u_+ = \cosh\phi V^0 + \sinh\phi V^3 \\ V^3 &= u_+^\dagger \sigma_z u_+ \to u_+^\dagger e^{\sigma_z\phi/2} \sigma_z e^{\sigma_z\phi/2} u_+ = u_+^\dagger(\sigma_z \cosh\phi + \sinh\phi) u_+ \\ &\qquad = \cosh\phi\, u_+^\dagger \sigma_z u_+ + \sinh\phi\, u_+^\dagger u_+ \\ &= \cosh\phi V^3 + \sinh\phi V^0 \end{aligned} \tag{7.38}$$

as expected for the components of a four vector. Similarly, V_x and V_y may be shown to have the correct transformation properties under a Lorentz transformation.

A similar construction shows that the components of

$$W^\mu = (u_-^\dagger u_-, -u_-^\dagger \vec{\sigma} u_-) \tag{7.39}$$

transform as a four vector under a proper Lorentz transformation.

7.2 The Weyl Lagrangian

We will now promote our spinors to fields, that is spinor functions of space and time, transforming under Lorentz transformations according to

$$u_+(x) \to U_\Lambda(\Lambda)^\dagger u_+(x) U_\Lambda(\Lambda) = D(\Lambda) u_+(\Lambda^{-1}x) \tag{7.40}$$

and similarly for u_-, where $U_\Lambda(\Lambda)$ is the unitary operator corresponding to Lorentz transformations,

$$U_\Lambda(\Lambda)|\, k\rangle = |\, \Lambda k\rangle. \tag{7.41}$$

Note that the u's are complex fields. We can now construct a Lagrangian for u_+, keeping in mind the following restrictions:

1. The action S should be real. This is because we want just as many field equations as there are fields. By breaking up any complex fields into their real and imaginary parts, we can always think of S as being a function only of a number of *real* fields, say N of them. If S were complex, with independent real and imaginary parts, then the real and imaginary parts of the resulting Euler-Lagrange equations would yield $2N$ field equations for N fields, too many to be satisfied except in special cases (see Weinberg, *The Quantum Theory of Fields*, Vol. I, pg. 300).

2. $\mathcal{L}$ should be bilinear in the fields, to produce a linear equation of motion. In the absence of interactions, we want u_+ to be a free field with plane wave solutions, which requires linear equations of motion.

3. $\mathcal{L}$ should be invariant under the internal symmetry transformation $u_+ \to e^{-i\lambda} u_+,\ u_+^\dagger \to e^{i\lambda} u_+^\dagger$. This is because we want to have some contact with the real world, and all observed fermions carry some conserved charge (like baryon number or lepton number).

We've already seen that bilinears in u_+ and $u_+^\dagger$ form the components of a four-vector; to make this a scalar we have to contract with another vector. The only other vector we have at our disposal is the derivative ∂_μ. Hence, the simplest Lagrangian we can write down satisfying the above requirements is

$$\mathcal{L} = i\left(u_+^\dagger \partial_0 u_+ + u_+^\dagger \vec{\sigma}\cdot\vec{\nabla} u_+\right). \tag{7.42}$$

The i in front is required for the action to be real, which you can verify by integrating by parts. The sign of $\mathcal{L}$ is not fixed; we will take it at this point to be $+$. We will see later on that this theory has problems with positivity of the energy no matter what sign we choose, so we will defer the discussion to a later section.

The Lagrangian Eq. (7.42) is called the *Weyl* Lagrangian. We can get the equation of motion by varying with respect to $u_+^\dagger$:

$$\Pi^\mu_{u_+^\dagger} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu u_+^\dagger)} = 0 \tag{7.43}$$

so the equation of motion is

$$\frac{\partial\mathcal{L}}{\partial u_+^\dagger} = 0 \Rightarrow (\partial_0 + \vec{\sigma}\cdot\vec{\nabla})u_+ = 0. \tag{7.44}$$

Multiplying this equation by $\partial_0 - \vec{\sigma}\cdot\vec{\nabla}$ and using the relation

$$\sigma_i\sigma_j = i\sum_k \epsilon_{ijk}\sigma_k + \delta_{ij} \tag{7.45}$$

gives us $(\vec{\sigma}\cdot\vec{\nabla})^2 = \vec{\nabla}^2$ and so

$$\left(\partial_0^2 - \vec{\nabla}^2\right)u_+ = \Box u_+ = 0. \tag{7.46}$$

Remember that u_+ is a column vector, so both components of u_+ obey the Klein-Gordon equation for a massless field.

Defining the energy to be positive,

$$k_0 = \sqrt{|\vec{k}|^2} \tag{7.47}$$

there are two solutions for $u_+(x)$:

$$u_+(x) = u_+ e^{-ik\cdot x},\;\; u_+(x) = v_+ e^{ik\cdot x} \tag{7.48}$$

where u_+ and v_+ are constant 2 component spinors. Based on our previous experience with complex fields, we expect that when we quantize the theory, u_+ will

multiply an annihilation operator for a particle and v_+ will multiply a creation operator for an antiparticle. Substituting the positive energy solution into the Weyl equation gives

$$(\partial_0 + \vec{\sigma}\cdot\vec{\nabla})u_+(x) = (-ik_0 + i\vec{\sigma}\cdot\vec{k})u_+(x) = 0 \tag{7.49}$$

and so

$$(k_0 - \vec{\sigma}\cdot\vec{k})u_+ = 0. \tag{7.50}$$

Consider $\vec{k}$ to be in the $\hat{z}$ direction, $\vec{k} = k_0\hat{z}$. Then we have $(1-\sigma_z)u_+ = 0$, or

$$u_+ \propto \begin{pmatrix}1\\0\end{pmatrix}. \tag{7.51}$$

What does this tell us about the states of the quantum theory? Well, in the quantum theory we expect that u_+ will multiply an annihilation operator. Consider a state $|\,k\rangle$ moving in the positive z direction, $\vec{k} = (0,0,k_z)$, $k_0 > 0$. Then we expect that

$$\langle 0\,|u_+(x)|\,k\rangle \propto e^{-ik\cdot x}\begin{pmatrix}1\\0\end{pmatrix} \tag{7.52}$$

or

$$\langle 0\,|u_+(0)|\,k\rangle \propto \begin{pmatrix}1\\0\end{pmatrix}. \tag{7.53}$$

It will turn out that this state is in an eigenstate of the z component of angular momentum, J_z:

$$J_z|\,k\rangle = \lambda|\,k\rangle. \tag{7.54}$$

It is straightforward to find λ. Since u is a spinor field, we know how it transforms under rotations about the z axis by an angle θ,

$$U_R^\dagger(\hat{z},\theta)u_+(0)U_R(\hat{z},\theta) = e^{-i\sigma_z\theta/2}u_+(0) \tag{7.55}$$

and therefore

$$\begin{aligned}\langle 0\,|U_R^\dagger u_+(0)U_R|\,k\rangle &= \langle 0\,|e^{-i\sigma_z\theta/2}u_+(0)|\,k\rangle \\ &\propto e^{-i\sigma_z\theta/2}\begin{pmatrix}1\\0\end{pmatrix} = e^{-i\theta/2}\begin{pmatrix}1\\0\end{pmatrix}.\end{aligned} \tag{7.56}$$

But since

$$\begin{aligned}U_R|\,k\rangle &= e^{-i\lambda\theta}|\,k\rangle \\ U_R|\,0\rangle &= |\,0\rangle\end{aligned} \tag{7.57}$$

we also have

$$\langle 0 | U_R^\dagger(\hat{z},\theta) u_+(0) U_R(\hat{z},\theta) | k\rangle = e^{-i\lambda\theta}\langle 0 | u_+(0) | k\rangle \propto e^{-i\lambda\theta}\begin{pmatrix}1\\0\end{pmatrix} \tag{7.58}$$

and so

$$\lambda = 1/2. \tag{7.59}$$

Therefore in the quantum theory the annihilation operator multiplying u_+ will annihilate states with angular momentum 1/2 along the direction of motion. Similarly, we can show that

$$v_+ \propto \begin{pmatrix}1\\0\end{pmatrix},$$

and that v_+ will multiply a creation operator that creates states with angular momentum -1/2 along the direction of motion.

Therefore, the quanta of this theory consist of particles carrying spin $+1/2$ in the direction of motion and antiparticles carrying spin $-1/2$ in the direction of motion, while there are no corresponding states with particles (antiparticles) carrying spin antiparallel (parallel) to the direction of motion (since there is no corresponding solution to the equations of motion for the fields). These states don't look much like electrons, since electrons (or any other massive spin 1/2 particle) can have spin either parallel or antiparallel to the direction of motion. In fact, it is not consistent for a massive particle to have only one spin state, since for a massive particle it is always possible to boost to a frame going faster than the particle. In this frame, the particle's 3-momentum is in the opposite direction but its spin is unchanged. Thus, if the spin was parallel to the direction of motion in one frame, it is antiparallel in the other.

Thus, spin along the direction of motion is only a good quantum number for massless particles, and is usually called the *helicity* of the particle. *Spin* is usually reserved for massive particles to describe their angular momentum in the rest frame. A particle with positive helicity (along the direction of motion) is referred to as "right-handed," while if the helicity is negative (antiparallel to the direction of motion) it is "left-handed." Thus, in the quantum theory we expect that $u_+(x)$ will annihilate right-handed particles. Since the field operator therefore changes the helicity of the state it acts on by $-1/2$, it should also create left-handed antiparticles. For the $u_-(x)$ field, we would find that it annihilates left-handed particles and creates right-handed particles.

In fact, while the Weyl Lagrangian does not describe the familiar massive fermions like electrons or nucleons, there is a particle in nature which exists only as left-handed particles and right-handed antiparticles: the neutrino. As far as anyone

has been able to measure, neutrinos are massless fermions. Right-handed neutrinos and left-handed antineutrinos have never been observed, and so neutrinos are described by the u_- field.

Clearly the Weyl Lagrangian, in distinguishing right and left-handed particles, violates parity. Under a parity transformation the three momentum of a particle flips sign, while its spin is unchanged. Thus parity interchanges left and right-handed particles. Similarly, the Weyl Lagrangian violates charge conjugation invariance, since charge conjugation will turn a left-handed neutrino into a left-handed antineutrino, which does not exist. However, the combined operation of CP will turn a left-handed neutrino into a right-handed antineutrino, both of which exist in the theory. Thus, although we haven't quantized the theory to show this explicitly, we expect that the Weyl Lagrangian violates C and P separately, but conserves the product CP.

7.3 The Dirac Equation

This is all very well, but it's not what we set out to find. We were really looking for a theory of electrons, which are certainly not massless. Furthermore, the strong and electromagnetic interactions of electrons are observed to conserve parity (the weak interactions, which we shall study later, violate parity). Therefore we would like to write down a free field theory of massive spin 1/2 particles which has a parity symmetry.

We have already argued that parity interchanges left and right-handed fields. Thus we can define the action of parity on the $u_\pm$ fields to be

$$P : u_\pm(\vec{x},t) \to u_\mp(-\vec{x},t). \tag{7.60}$$

A parity invariant theory must therefore have both types of spinors. The simplest Lagrangian is just

$$\mathcal{L}_0 = iu_+^\dagger(\partial_0 + \vec{\sigma}\cdot\vec{\nabla})u_+ + iu_-^\dagger(\partial_0 - \vec{\sigma}\cdot\vec{\nabla})u_- \tag{7.61}$$

but this is nothing more than two decoupled massless spinors. However, it is easy to check explicitly that $u_+^\dagger u_-$ and $u_-^\dagger u_+$ transform as scalars under Lorentz transformations. Therefore we can include the parity conserving term

$$\begin{aligned}\mathcal{L} &= \mathcal{L}_0 - m\left(u_+^\dagger u_- + u_-^\dagger u_+\right)\\ &= iu_+^\dagger(\partial_0 + \vec{\sigma}\cdot\vec{\nabla})u_+ + iu_-^\dagger(\partial_0 - \vec{\sigma}\cdot\vec{\nabla})u_- - m\left(u_+^\dagger u_- + u_-^\dagger u_+\right)\end{aligned} \tag{7.62}$$

The coupling multiplying the $u_+^\dagger u_-$ term has dimensions of mass, so we have suggestibly called it m. This is the *Dirac* Lagrangian, and as we shall see it describes massive spin 1/2 fields. In its current form it doesn't look the way Dirac wrote it down, but we will be introducing some slick new notation shortly to put it in a more elegant form.

We can again vary the fields and derive the equations of motion. We find the coupled equations

$$\begin{aligned} i(\partial_0 + \vec{\sigma}\cdot\vec{\nabla})u_+ &= mu_- \\ i(\partial_0 - \vec{\sigma}\cdot\vec{\nabla})u_- &= mu_+. \end{aligned} \tag{7.63}$$

Multiplying the first equation by $\partial_0 - \vec{\sigma}\cdot\vec{\nabla}$, we find

$$(\partial_0^2 - \vec{\nabla}^2)u_+ = -im(\partial_0 - \vec{\sigma}\cdot\vec{\nabla})u_- = -m^2 u_+ \tag{7.64}$$

and so each of the components of u_+ and u_- obeys the massive Klein-Gordon equation

$$(\partial_\mu\partial^\mu + m^2)u_\pm(x) = 0. \tag{7.65}$$

At this point we will introduce some notation to make life easier. We can group the two fields u_+ and u_- into a single four component "bispinor" field ψ:

$$\psi \equiv \begin{pmatrix} u_+ \\ u_- \end{pmatrix}. \tag{7.66}$$

In terms of ψ, the Dirac Lagrangian is

$$\mathcal{L} = i\psi^\dagger\partial_0\psi + i\psi^\dagger\vec{\alpha}\cdot\vec{\nabla}\psi - m\psi^\dagger\beta\psi \tag{7.67}$$

where

$$\vec{\alpha} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & -\vec{\sigma} \end{pmatrix}, \quad \beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{7.68}$$

and each entry represents a two-by-two matrix. The equation of motion is

$$i(\partial_0 + \vec{\alpha}\cdot\vec{\nabla})\psi = \beta m\psi. \tag{7.69}$$

This is the *Dirac equation*. Note that we can get the Dirac equation directly from Eq. (7.67) from the Euler-Lagrange equations for ψ:

$$\begin{aligned} \Pi^\mu_{\psi^\dagger} &= \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi^\dagger)} = 0 \\ \frac{\partial\mathcal{L}}{\partial\psi^\dagger} &= 0 \Rightarrow i\partial_0\psi + \vec{\alpha}\cdot\vec{\nabla}\psi - m\beta\psi = 0. \end{aligned} \tag{7.70}$$

You just have to remember that ψ is now a four component column vector, and so these are now matrix equations. If you prefer, you can leave the spinor indices explicit in this derivation.

In terms of ψ, a parity transformation is now

$$P : \psi(\vec{x},t) \to \beta\psi(-\vec{x},t) \tag{7.71}$$

and under a Lorentz boost

$$\psi \to e^{\vec{\alpha}\cdot\hat{e}\phi/2}\psi. \tag{7.72}$$

Since u_+ and u_- transform the same way under rotations, ψ transforms under rotations as

$$R : \psi \to e^{\vec{L}\cdot\hat{e}\theta}\psi \tag{7.73}$$

where $\vec{L} = \frac{1}{2}\begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}$.

The α's and β obey the relations

$$\{\beta,\alpha_i\} = 0, \ \ \{\alpha_i,\alpha_j\} = 0 \ (i \neq j), \ \ \beta^2 = \alpha_1^2 = \alpha_2^2 = \alpha_3^2 = 1 \tag{7.74}$$

where $\{A,B\}$ is the *anticommutator* of A and B:

$$\{A,B\} = AB + BA. \tag{7.75}$$

Finally, we note that the components of $(\psi^\dagger\psi, \psi^\dagger\vec{\alpha}\psi)$ form a 4-vector.

This representation is not unique. We could, for example, have defined ψ by

$$\psi = \frac{1}{\sqrt{2}}\begin{pmatrix} u_+ + u_- \\ u_+ - u_- \end{pmatrix} \tag{7.76}$$

and all of these results would still hold, except the α's and β would be different. However, they would still obey the anticommutations relations Eq. (7.74). In this basis (the "Dirac" basis), we find

$$\vec{\alpha} = \begin{pmatrix} 0 & \sigma \\ \sigma & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{7.77}$$

We could define other bases as well. The Weyl basis turns out to be convenient for highly relativistic particles $m \ll E$ while the Dirac basis is convenient in the nonrelativistic limit $m \gg E$. However, as we shall see, in most situations we will never have to specify the basis. The anticommutation relations Eq. (7.74), which hold in any basis, will be sufficient.

Very shortly we will introduce some even more slick notation which will allow us to write all of our results in a Lorentz covariant form. However, before proceeding to that let us finish our discussion of the plane wave solutions to the Dirac equation. We will need these solutions to canonically quantize the theory, since the plane wave solutions multiply the creation and annihilation operators in the quantum theory.

7.3.1 Plane Wave Solutions to the Dirac Equation

As in the Weyl equation, take the energy p_0 to be positive, $p_0 = \sqrt{|\vec{p}|^2 + m^2}$. Then we have both positive and negative frequency solutions

$$\psi(x) = u_{\vec{p}} e^{-ip\cdot x}, \;\; \psi(x) = v_{\vec{p}} e^{ip\cdot x} \tag{7.78}$$

where $u_{\vec{p}}$ and $v_{\vec{p}}$ are constant four component bispinors. Substituting the first solution into the Dirac equation, we find

$$(p_0 - \vec{\alpha}\cdot\vec{p})u_{\vec{p}} = \beta m u_{\vec{p}}. \tag{7.79}$$

For definiteness, we will work in the Dirac basis, so $\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. In the rest frame, $\vec{p}$=0 and $p_0 = m$, so we find

$$u_{\vec{p}} = \beta u_{\vec{p}} \Rightarrow u_{\vec{p}} = \begin{pmatrix} a \\ b \\ 0 \\ 0 \end{pmatrix} \tag{7.80}$$

and so two linearly independent solutions are

$$u_{\vec{0}}^{(1)} = \sqrt{2m}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \;\; u_{\vec{0}}^{(2)} = \sqrt{2m}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}. \tag{7.81}$$

(The factor of $\sqrt{2m}$ in the normalization is a convention. Note that Mandl & Shaw do not include this factor in their definition of the plane wave states.) Now, in both the Dirac and Weyl bases the spin operator is

$$S_z = \frac{1}{2}\begin{pmatrix} \sigma_z & 0 \\ 0 & \sigma_z \end{pmatrix} \tag{7.82}$$

so $S_z u_{\vec{p}}^{(1)} = +\frac{1}{2}u_{\vec{p}}^{(1)}$ and $S_z u_{\vec{p}}^{(2)} = -\frac{1}{2}u_{\vec{p}}^{(2)}$. The two solutions correspond to the two spin states. As we expected, both solutions are present for a massive field.

Now we use our knowledge of the transformation properties of ψ to find the the plane wave solutions when $\vec{p} \neq 0$. Instead of solving the Dirac equation for $\vec{p} \neq 0$, we can just boost the coordinate system in the opposite direction

$$u_{\vec{p}}^{(r)} = e^{\vec{\alpha}\cdot\hat{e}\phi/2} u_{\vec{0}}^{(r)} \tag{7.83}$$

where $\hat{e} = \vec{p}/|\vec{p}|$, $\cosh\phi = \gamma = E/m$ and $\sinh\phi = |\vec{p}|/m$. Using $\alpha_i^2 = 1$ and $\alpha_i\alpha_j = -\alpha_j\alpha_i$, we get

$$u_{\vec{p}}^{(r)} = \left[\cosh\frac{\phi}{2} + \vec{\alpha}\cdot\hat{e}\sinh\frac{\phi}{2}\right] u_{\vec{0}}^{(r)}. \tag{7.84}$$

Now, $\cosh\phi/2 = \sqrt{(1+\cosh\phi)/2}$ and $\sinh\phi/2 = \sqrt{(1+\sinh\phi)/2}$, so

$$u_{\vec{p}}^{(r)} = \left[\sqrt{\frac{E+m}{2m}} + \sqrt{\frac{E-m}{2m}}\vec{\alpha}\cdot\hat{e}\right] u_{\vec{0}}^{(r)} \tag{7.85}$$

so in the Dirac basis we find, for $\vec{p}$ in the $\hat{z}$ direction,

$$u_{\vec{p}}^{(1)} = \begin{pmatrix} \sqrt{E+m} \\ 0 \\ \sqrt{E-m} \\ 0 \end{pmatrix}, \;\; u_{\vec{p}}^{(2)} = \begin{pmatrix} 0 \\ \sqrt{E+m} \\ 0 \\ -\sqrt{E-m} \end{pmatrix}. \tag{7.86}$$

The case where $\vec{p}$ is not parallel to $\hat{z}$ is straightforward to compute from Eq. (7.85).

Similar arguments also allow us to find the solutions for the v's, which in the quantum theory we expect to multiply creation operators for antiparticles. We find

$$v_{\vec{0}}^{(1)} = \sqrt{2m}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \;\; v_{\vec{0}}^{(2)} = \sqrt{2m}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$v_{\vec{p}}^{(1)} = \begin{pmatrix} \sqrt{E-m} \\ 0 \\ \sqrt{E+m} \\ 0 \end{pmatrix}, \;\; v_{\vec{p}}^{(2)} = \begin{pmatrix} 0 \\ -\sqrt{E-m} \\ 0 \\ \sqrt{E+m} \end{pmatrix}. \tag{7.87}$$

Notice that we have chosen our solutions to be orthonormal:

$$u_{\vec{0}}^{(r)\dagger} u_{\vec{0}}^{(s)} = 2m\delta^{rs}, \;\; v_{\vec{0}}^{(r)\dagger} v_{\vec{0}}^{(s)} = 2m\delta^{rs} \tag{7.88}$$

or

$$u_{\vec{0}}^{(r)\dagger} \beta u_{\vec{0}}^{(s)} = 2m\delta^{rs}, \;\; v_{\vec{0}}^{(r)\dagger} \beta v_{\vec{0}}^{(s)} = -2m\delta^{rs} \tag{7.89}$$

This second form is useful because we've already noted that $\psi^\dagger\beta\psi$ is a Lorentz scalar. Therefore we can immediately write

$$\begin{aligned} u_{\vec{0}}^{(r)\dagger} \beta u_{\vec{0}}^{(s)} = u_{\vec{p}}^{(r)\dagger} \beta u_{\vec{p}}^{(s)} = 2m\delta^{rs} \\ v_{\vec{0}}^{(r)\dagger} \beta v_{\vec{0}}^{(s)} = v_{\vec{p}}^{(r)\dagger} \beta v_{\vec{p}}^{(s)} = -2m\delta^{rs} \end{aligned} \tag{7.90}$$

since the scalar is unaffected by Lorentz boosts.

7.4 γ Matrices

With all of these α's and β's, the theory doesn't look Lorentz covariant. Time and space appear to be on a different footing, although we know they're not because $\mathcal{L}$ is a scalar. We can clean things up a bit by introducing even more notation which makes everything manifestly Lorentz covariant. It will also allow us to write down combinations of bispinors which transform in simple ways under Lorentz transformations.

We've already seen that for two bispinors ψ_1 and ψ_2, $\psi_1^\dagger\beta\psi_2$ is a Lorentz scalar. It's convenient to make use of this fact and define the *Dirac Adjoint* of a bispinor ψ

$$\overline{\psi} \equiv \psi^\dagger\beta. \tag{7.91}$$

Therefore $\overline{\psi}_1\psi_2$ is a scalar: under a Lorentz transformation

$$\overline{\psi}_1\psi_2 \to \overline{\psi}_1\psi_2 \tag{7.92}$$

(note that since $\beta^2 = 1$, $\psi^\dagger = \overline{\psi}\beta$). Furthermore, we know that the components of $(\psi_1^\dagger\psi_2, \psi_1^\dagger\vec{\alpha}\psi_2) = (\overline{\psi}_1\beta\psi_2, \overline{\psi}_1\beta\vec{\alpha}\psi_2)$ transform like the components of a four-vector. It's convenient then to define the four matrices γ^μ, $\mu = 1..4$, by

$$\gamma^0 \equiv \beta, \ \ \gamma^i \equiv \beta\alpha_i. \tag{7.93}$$

(Note that the label i on the α's is not a Lorentz index and so there is no distinction between upper and lower indices on α. The index i on γ is a Lorentz index, and so this equation defines γ^μ with raised indices.) The components of the four vector are now simply written as $\overline{\psi}_1\gamma^\mu\psi_2$. The γ^μ's are called the Dirac γ matrices. You will learn to know and love them.

Under a Lorentz transformation $\psi \to D(\Lambda)\psi$, and so $\overline{\psi} = \psi^\dagger\beta \to \psi^\dagger D^\dagger(\Lambda)\beta = \psi^\dagger\beta\beta D^\dagger(\Lambda)\beta \equiv \overline{\psi}\overline{D}(\Lambda)$, where we have defined the Dirac adjoint of the operator $D(\Lambda)$

$$\overline{D}(\Lambda) = \gamma^0 D^\dagger(\Lambda)\gamma^0. \tag{7.94}$$

Since under a Lorentz transformation,

$$\overline{\psi}\gamma^\mu\psi \to \overline{\psi}\overline{D}(\Lambda)\gamma^\mu D(\Lambda)\psi = \Lambda^\mu{}_\nu\overline{\psi}\gamma^\nu\psi \tag{7.95}$$

we find that the γ matrices satisfy

$$\overline{D}(\Lambda)\gamma^\mu D(\Lambda) = \Lambda^\mu{}_\nu\gamma^\nu. \tag{7.96}$$

We can now use this technology to construct objects from the bispinors which transform in more complicated ways under the Lorentz group. For example, $\overline{\psi}\gamma^\mu\gamma^\nu\psi$

transforms like a two index tensor:

$$\begin{aligned}\overline{\psi}\gamma^\mu\gamma^\nu\psi &\to \overline{\psi}\overline{D}(\Lambda)\gamma^\mu D(\Lambda)\overline{D}(\Lambda)\gamma^\nu D(\Lambda)\psi \\ &= \Lambda^\mu{}_\alpha\Lambda^\nu{}_\beta\overline{\psi}\gamma^\alpha\gamma^\beta\psi\end{aligned} \tag{7.97}$$

(where we have used $\overline{D}(\Lambda)D(\Lambda) = 1$, which follows from the fact that $\overline{\psi}\psi$ is a scalar).

The commutation relations for the α's and β may now be written in terms of the γ's as

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \tag{7.98}$$

Thus, the γ matrices all anticommute with one another, and $(\gamma^0)^2 = -(\gamma^1)^2 = -(\gamma^2)^2 = -(\gamma^3)^2 = 1$. For any four-vector a_μ, we define $\not{a}$ ("a-slash") by

$$\not{a} = a_\mu\gamma^\mu. \tag{7.99}$$

From the γ algebra it follows that

$$\not{a}\not{b} + \not{b}\not{a} = 2a \cdot b \tag{7.100}$$

and $\not{a}\not{a} = a^2$. The Dirac Lagrangian may be written in a manifestly Lorentz invariant form

$$\overline{\psi}(i\not{\partial} - m)\psi \tag{7.101}$$

and the Dirac equation is

$$(i\not{\partial} - m)\psi = 0. \tag{7.102}$$

Note that these are all four by four matrix equations, where we have suppressed matrix indices. Also, everything is still classical, and the γ matrix algebra is simply a statement about matrix multiplication, not about quantum operators anticommuting.

Another property of the γ matrices is that they are not all Hermitian,

$$\gamma^{\mu\dagger} = \gamma_\mu = g_{\mu\nu}\gamma^\nu = \gamma^0\gamma^\mu\gamma^0 \tag{7.103}$$

but they are self-Dirac adjoint ("self-bar")

$$\overline{\gamma}^\mu = \gamma^\mu. \tag{7.104}$$

The orthonormality conditions on the plane wave solutions are now

$$\overline{u}_{\vec{p}}^{(r)}u_{\vec{p}}^{(s)} = 2m\delta^{rs} = -\overline{v}_{\vec{p}}^{(r)}v_{\vec{p}}^{(s)}, \ \ \overline{u}_{\vec{p}}^{(r)}v_{\vec{p}}^{(s)} = 0 \tag{7.105}$$

and since $i\partial\!\!\!/ u_{\vec{p}}(x) = i(-ip\!\!\!/)u_{\vec{p}}(x) = p\!\!\!/ u_{\vec{p}}(x)$ and $i\partial\!\!\!/ v_{\vec{p}}(x) = i(ip\!\!\!/)v_{\vec{p}}(x) = -p\!\!\!/ v_{\vec{p}}(x)$, the Dirac equation implies that the plane wave solutions satisfy

$$(p\!\!\!/ - m)u_{\vec{p}}^{(r)} = 0 = (p\!\!\!/ + m)v_{\vec{p}}^{(r)}. \tag{7.106}$$

Taking the Dirac adjoint of Eq. (7.106) gives

$$\overline{u}_{\vec{p}}^{(r)}(p\!\!\!/ - m) = 0 = \overline{v}_{\vec{p}}^{(r)}(p\!\!\!/ + m). \tag{7.107}$$

The plane wave bispinors also obey the completeness relations

$$\sum_{r=1}^{2} u_{\vec{p}}^{(r)}\overline{u}_{\vec{p}}^{(r)} = p\!\!\!/ + m, \;\; \sum_{r=1}^{2} v_{\vec{p}}^{(r)}\overline{v}_{\vec{p}}^{(r)} = p\!\!\!/ - m. \tag{7.108}$$

These will be very useful later on when we calculate cross sections.

7.4.1 Bilinear Forms

We have already seen that $\overline{\psi}\psi$ transforms like a scalar and that the components of $\overline{\psi}\gamma^\mu\psi$ form a 4-vector. We could go on indefinitely and construct n component tensors $\overline{\psi}\gamma^\mu\gamma^\nu ... \gamma^\alpha\psi$, but since any collection of γ matrices is simply a four by four matrix there are only be sixteen independent bilinears which can be constructed out of $\overline{\psi}$ and ψ. We already have five - the one component scalar and the four-vector. We can choose the remaining eleven to transform simply under Lorentz transformations.

Consider first $\overline{\psi}\gamma^\mu\gamma^\nu\psi$. This is a sixteen component object. However, we may split it up into symmetric and antisymmetric pieces: $\overline{\psi}\{\gamma^\mu, \gamma^\nu\}\psi$ and $\overline{\psi}[\gamma^\mu, \gamma^\nu]\psi$. Since $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$, the symmetric combination is simply $2g^{\mu\nu}\overline{\psi}\psi$ and so is not an independent bilinear form. The antisymmetric combination is new. We define

$$\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu] \tag{7.109}$$

and then the six independent components of $\overline{\psi}\sigma^{\mu\nu}\psi$ transform like a two index antisymmetric tensor (note that some books define $\sigma^{\mu\nu}$ with an opposite sign to this). This bring the number of bilinears to eleven, so we need to find five more.

Skipping to four component objects, we next consider $\overline{\psi}\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta\psi$. But if any two indices are the same, this doesn't give us anything new. For example,

$$\gamma^0\gamma^1\gamma^0\gamma^2 = -\gamma^0\gamma^0\gamma^1\gamma^2 = -\gamma^1\gamma^2 = i\sigma^{12}. \tag{7.110}$$

So only the matrix $\gamma^0\gamma^1\gamma^2\gamma^3$ and its various permutations are new. Thus we define a new matrix γ_5:

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \frac{i}{4!}\epsilon_{\mu\nu\alpha\beta}\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta \equiv \gamma^5. \tag{7.111}$$

Here, $\epsilon_{\mu\nu\alpha\beta}$ is a totally antisymmetric four index tensor, and

$$\epsilon_{0123} = 1 = -\epsilon_{1023} = \epsilon_{1032} = ... \tag{7.112}$$

γ_5 is in many ways the "fifth γ matrix." It obeys

$$(\gamma_5)^2 = 1, \;\; \gamma_5 = \gamma_5^\dagger = -\overline{\gamma}_5, \;\; \{\gamma_5, \gamma^\mu\} = 0. \tag{7.113}$$

Since $\epsilon_{\mu\nu\alpha\beta}\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta$ has no free indices, it transforms like a scalar under boosts and rotations, $\overline{\psi}\gamma_5\psi \to \overline{\psi}\gamma_5\psi$. However, its transformation differs from that of $\overline{\psi}\psi$ when we consider parity transformations. Under parity,

$$\begin{aligned}\psi(\vec{x},t) &\to \beta\psi(-\vec{x},t) = \gamma^0\psi(-\vec{x},t)\\ \overline{\psi}(\vec{x},t) &\to \overline{\psi}(-\vec{x},t)\beta = \overline{\psi}(-\vec{x},t)\gamma^0\end{aligned} \tag{7.114}$$

and so under a parity transformation

$$\overline{\psi}\psi(\vec{x},t) \to \overline{\psi}\psi(-\vec{x},t) \tag{7.115}$$

exactly as a scalar should transform. However,

$$\overline{\psi}\gamma_5\psi \to \overline{\psi}\gamma^0\gamma_5\gamma^0\psi(-\vec{x},t) = -\overline{\psi}\gamma_5\gamma^0\gamma^0\psi(-\vec{x},t) = -\overline{\psi}\gamma_5\psi(-\vec{x},t). \tag{7.116}$$

Thus $\overline{\psi}\gamma_5\psi$ changes sign under a parity transformation, and so transforms like a pseudoscalar.

The final four independent bilinear forms are the components of

$$\overline{\psi}\gamma_\mu\gamma_5\psi \tag{7.117}$$

which make up an axial vector. Again, we see that under a parity transformation $\overline{\psi}\gamma^\mu\psi(\vec{x},t) \to \overline{\psi}\gamma^0\gamma^\mu\gamma^0\psi(-\vec{x},t)$, and so

$$\begin{aligned}\overline{\psi}\gamma^0\psi(\vec{x},t) &\to \overline{\psi}\gamma^0\psi(-\vec{x},t)\\ \overline{\psi}\gamma^i\psi(\vec{x},t) &\to -\overline{\psi}\gamma^i\psi(-\vec{x},t), \; i = 1..3.\end{aligned} \tag{7.118}$$

The spatial components of $\overline{\psi}\gamma^\mu\psi$ flip sign under a reflection whereas the time component is unchanged, which is how a vector transforms under parity. On the other hand, the addition of the γ_5 means that the axial vector transforms like

$$\begin{aligned}\overline{\psi}\gamma^0\gamma_5\psi(\vec{x},t) &\to -\overline{\psi}\gamma^0\gamma_5\psi(-\vec{x},t)\\ \overline{\psi}\gamma^i\psi(\vec{x},t) &\to \overline{\psi}\gamma^i\gamma_5\psi(-\vec{x},t), \; i = 1..3\end{aligned} \tag{7.119}$$

which is the correct transformation law for an axial vector.

Thus, we have chosen the sixteen bilinears which can be formed from a Dirac field and its adjoint to transform simply under Lorentz transformations. To summarize, we have

$$\begin{aligned} S &= \overline{\psi}\psi \quad \text{(scalar)} \\ V^\mu &= \overline{\psi}\gamma^\mu\psi \quad \text{(vector)} \\ T^{\mu\nu} &= \overline{\psi}\sigma^{\mu\nu}\psi \quad \text{(tensor)} \\ P &= \overline{\psi}\gamma_5\psi \quad \text{(pseudoscalar)} \\ A^\mu &= \overline{\psi}\gamma^\mu\gamma_5\psi \quad \text{(axial vector)}. \end{aligned} \tag{7.120}$$

Given these transformation laws it will be easy to construct Lorentz invariant interaction terms in the Lagrangian. For example, if we have a vector field A^μ (such as a photon), a Lorentz invariant interaction is $A^\mu\overline{\psi}\gamma_\mu\psi$. An axial vector field B^μ could couple in a parity conserving manner as $B^\mu\overline{\psi}\gamma_\mu\gamma_5\psi$. A scalar field ϕ (such as a meson) could couple like $\phi\overline{\psi}\psi$, whereas the coupling $\phi\overline{\psi}\gamma_5\psi$ conserves parity if ϕ transforms like a pseudoscalar. Finally, in a parity violating theory (such as the weak interactions) a vector field W^μ could couple to some linear combination of vector and axial vector currents: $W^\mu\overline{\psi}\gamma_\mu(a + b\gamma_5)\psi$. This interaction is parity violating because there is no way to define the transformation of W^μ under parity such that this term is parity invariant.

7.4.2 Chirality and γ_5

In the Weyl basis, $\gamma_5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. We can define the projection operators (in any basis)

$$P_R = \tfrac{1}{2}(1+\gamma_5), \;\; P_L = \tfrac{1}{2}(1-\gamma_5). \tag{7.121}$$

These satisfy the requirements for projections operators: $P_R^2 = P_R$, $P_L^2 = P_L$, $P_R P_L = 0$, $P_R + P_L = 1$, and they project out the Weyl spinors u_+ and u_- from the Dirac bispinor:

$$\begin{aligned} \begin{pmatrix} u_+ \\ 0 \end{pmatrix} &= \tfrac{1}{2}(1+\gamma_5)\psi = P_R\psi \equiv \psi_R \\ \begin{pmatrix} 0 \\ u_- \end{pmatrix} &= \tfrac{1}{2}(1-\gamma_5)\psi = P_L\psi \equiv \psi_L. \end{aligned} \tag{7.122}$$

We also find that $\overline{\psi_R} \equiv \overline{\psi P}_R = \overline{\psi}\gamma^0 P_R \gamma^0 = \overline{\psi}P_L$ and $\overline{\psi_L} = \overline{\psi}P_R$. ψ_L and ψ_R are just the left and right-handed pieces of the Dirac bispinor in four component, rather than two component (u_- and u_+), notation. The Weyl Lagrangian for right-handed particles may therefore be written

$$\mathcal{L} = \overline{\psi} i\partial\!\!\!/ P_R\psi = \overline{\psi} i\partial\!\!\!/ P_R^2\psi = \overline{\psi}P_L i\partial\!\!\!/ P_R\psi = \overline{\psi_R} i\partial\!\!\!/ \psi_R. \tag{7.123}$$

Similarly, for left-handed particles we have $\mathcal{L} = \overline{\psi_L} i\partial\!\!\!/ \psi_L$. The Dirac Lagrangian is

$$\mathcal{L} = \overline{\psi_L} i\partial\!\!\!/ \psi_L + \overline{\psi_R} i\partial\!\!\!/ \psi_R - m(\overline{\psi_L}\psi_R + \overline{\psi_R}\psi_L). \tag{7.124}$$

As we noticed before, we see that without the mass term the Dirac Lagrangian just describes two independent helicity eigenstates. The mass term couples the right and left-handed fields, so the helicity of the massive field is no longer a good quantum number. As we argued from physical grounds earlier, this is exactly what must happen for a massive particle, since its helicity is no longer a Lorentz invariant quantity.

We also note that we may write the parity violating Weyl Lagrangian describing left-handed neutrinos in the four-component form

$$\mathcal{L} = \overline{\psi}_L i\partial\!\!\!/ \psi_L \tag{7.125}$$

.

When $m \neq 0$, the Dirac Lagrangian is invariant under the $U(1)$ symmetry $\psi_{L,R} \to e^{-i\lambda}\psi_{L,R}$. Because of the mass term, the left and right handed fields must transform the same way under the internal symmetry. However, when $m = 0$ this is no longer required, and the theory has two independent $U(1)$ symmetries,

$$\psi_L \to e^{-i\lambda}\psi_L, \;\; \psi_R \to \psi_R \tag{7.126}$$

and

$$\psi_R \to e^{-i\lambda}\psi_R, \;\; \psi_L \to \psi_L. \tag{7.127}$$

The independent symmetries are called *chiral* symmetries, where the term chiral denotes the fact that the symmetries has a "handedness", that is, it distinguishes left and right handed particles. Chiral symmetries play an important role in the study of both the strong and weak interactions. For example, the weak interactions involve the coupling of vector fields (the $W^\pm$ and Z bosons) to only the *left-handed* components of spin 1/2 fields. The Z boson, for example, is the quantum of the Z^μ vector field, which has a coupling of the form

$$Z^\mu \overline{\psi}_L \gamma_\mu \psi_L = Z^\mu \overline{\psi} \tfrac{1}{2}\gamma_\mu (1-\gamma_5)\psi. \tag{7.128}$$

Such a theory clearly violates parity.

7.5 Summary of Results for the Dirac Equation

These pages summarize the results we have derived for the Dirac equation, without proofs. You will find many of these results in Appendix A of Mandl & Shaw; however, they use a different normalization for the plane wave states.

7.5.1 Dirac Lagrangian, Dirac Equation, Dirac Matrices

The theory is defined by the Lagrange Density

$$\mathcal{L} = \psi^\dagger \left[i\partial_0 + i\vec{\alpha}\cdot\vec{\nabla} - \beta m \right] \psi. \tag{7.129}$$

where ψ is a set of four complex fields, arranged in a column vector (a *Dirac bispinor*) and the α's and β are a set of 4×4 Hermitian matrices (the *Dirac Matrices*). The corresponding equation of motion is

$$(i\partial_0 + i\vec{\alpha}\cdot\vec{\nabla} - \beta m)\psi = 0. \tag{7.130}$$

The Dirac matrices obey the following algebra,

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij},\ \{\alpha_i, \beta\} = 0,\ \beta^2 = 1. \tag{7.131}$$

Two representations of the Dirac algebra that will be useful to us are the *Weyl representation*

$$\vec{\alpha} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & -\vec{\sigma} \end{pmatrix},\ \ \beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{7.132}$$

and the *standard* (or *Dirac*) representation

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix},\ \ \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{7.133}$$

(where each component represents a 2×2 matrix).

7.5.2 Space-Time Symmetries

The Dirac equation is invariant under both Lorentz transformations and parity. Under a Lorentz transformation characterized by a 4×4 Lorentz matrix Λ,

$$\Lambda : \psi(x) \to D(\Lambda)\psi(\Lambda^{-1}x). \tag{7.134}$$

For a boost characterized by rapidity ϕ in the $\hat{e}$ direction,

$$D\left(A(\hat{e}\phi)\right) = e^{\vec{\alpha}\cdot\hat{e}\phi/2} \tag{7.135}$$

while for a rotation of angle θ about the $\hat{e}$ axis,

$$D\left(R(\hat{e}\theta)\right) = e^{-i\vec{L}\cdot\hat{e}\theta} \tag{7.136}$$

where

$$\vec{L} = \frac{1}{2}\begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \tag{7.137}$$

in both the Weyl and standard representations.

Under parity,

$$P : \psi(\vec{x}, t) \to \beta\psi(-\vec{x}, t). \tag{7.138}$$

7.5.3 Dirac Adjoint, γ Matrices

The *Dirac adjoint* of a Dirac bispinor is defined by

$$\overline{\psi} = \psi^\dagger\beta \tag{7.139}$$

and the Dirac adjoint of a 4×4 matrix is

$$\overline{A} = \beta A^\dagger \beta. \tag{7.140}$$

These obey the usual rules for adjoints, *e.g.*

$$\left(\overline{\psi}A\phi\right)^* = \overline{\phi}\overline{A}\psi. \tag{7.141}$$

The γ *matrices* are defined by

$$\gamma^0 = \beta,\ \ \gamma^i = \beta\alpha_i. \tag{7.142}$$

From these we can define the γ matrices with lowered indices by

$$\gamma_\mu \equiv g_{\mu\nu}\gamma^\nu. \tag{7.143}$$

The γ matrices are not all Hermitian,

$$\gamma^{\mu\dagger} = \gamma_\mu = g_{\mu\nu}\gamma^\nu = \gamma^0\gamma^\mu\gamma^0 \tag{7.144}$$

but they are self-Dirac adjoint ("self-bar")

$$\overline{\gamma}^\mu = \gamma^\mu. \tag{7.145}$$

They obey the γ algebra

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \tag{7.146}$$

and also obey

$$\overline{D}(\Lambda)\gamma^\mu D(\Lambda) = \Lambda^\mu{}_\nu \gamma^\nu. \tag{7.147}$$

For any 4-vector a, we define

$$\not{a} = a_\mu \gamma^\mu \tag{7.148}$$

and from the γ algebra it follows that

$$\not{a}\not{b} + \not{b}\not{a} = 2a \cdot b. \tag{7.149}$$

In this notation, the Dirac Lagrange density is

$$\overline{\psi}(i\not{\partial} - m)\psi \tag{7.150}$$

and the Dirac equation is

$$(i\not{\partial} - m)\psi = 0. \tag{7.151}$$

7.5.4 Bilinear Forms

There are sixteen linearly independent bilinear forms we can make from a Dirac bispinor and its adjoint. We can choose these sixteen to form the components of objects that transform in simple ways under the Lorentz group and parity. These are

$$\begin{aligned} S &= \overline{\psi}\psi \quad \text{(scalar)} \\ V^\mu &= \overline{\psi}\gamma^\mu\psi \quad \text{(vector)} \\ T^{\mu\nu} &= \overline{\psi}\sigma^{\mu\nu}\psi \quad \text{(tensor)} \\ P &= \overline{\psi}\gamma_5\psi \quad \text{(pseudoscalar)} \\ A^\mu &= \overline{\psi}\gamma^\mu\gamma_5\psi \quad \text{(axial vector)} \end{aligned} \tag{7.152}$$

where we have defined

$$\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu] \tag{7.153}$$

and

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \frac{i}{4!}\epsilon_{\mu\nu\alpha\beta}\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta \equiv \gamma^5. \tag{7.154}$$

Here, $\epsilon_{\mu\nu\alpha\beta}$ is a totally antisymmetric four index tensor, and

$$\epsilon_{0123} = 1. \tag{7.155}$$

γ_5 is in many ways the "fifth γ matrix." It obeys

$$(\gamma_5)^2 = 1, \;\; \gamma_5 = \gamma_5^\dagger = -\overline{\gamma}_5, \;\; \{\gamma_5, \gamma^\mu\} = 0. \tag{7.156}$$

7.5.5 Plane Wave Solutions

The positive-frequency solutions of the Dirac equation are of the form

$$\psi = u e^{-ip\cdot x} \tag{7.157}$$

where $p^2 = m^2$ and $p^0 = \sqrt{\vec{p}^2 + m^2}$. The negative-frequency solutions are of the form

$$\psi = v e^{ip\cdot x}. \tag{7.158}$$

There are two positive-frequency and two negative-frequency solutions for each p. The Dirac equation implies that

$$(\not{p} - m)u = 0 = (\not{p} + m)v. \tag{7.159}$$

For a particle at rest, $p = (m, \vec{0})$, we can choose the independent u's and v's in the standard representation to be

$$u^{(1)}_{\vec{0}} = \begin{pmatrix} \sqrt{2m} \\ 0 \\ 0 \\ 0 \end{pmatrix}, u^{(2)}_{\vec{0}} = \begin{pmatrix} 0 \\ \sqrt{2m} \\ 0 \\ 0 \end{pmatrix}, v^{(1)}_{\vec{0}} = \begin{pmatrix} 0 \\ 0 \\ \sqrt{2m} \\ 0 \end{pmatrix}, v^{(2)}_{\vec{0}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \sqrt{2m} \end{pmatrix}. \tag{7.160}$$

(Note that these are normalized differently than in Mandl & Shaw. They omit the $\sqrt{2m}$ from the normalization and instead include it in the definition of D, the invariant phase space factor.) We can construct the solutions for a moving particle, $u^{(r)}_{\vec{p}}$ and $v^{(r)}_{\vec{p}}$, by applying a Lorentz boost.

These solutions are normalized such that

$$\overline{u}^{(r)}_{\vec{p}} u^{(s)}_{\vec{p}} = 2m\delta^{rs} = -\overline{v}^{(r)}_{\vec{p}} v^{(s)}_{\vec{p}}, \;\; \overline{u}^{(r)}_{\vec{p}} v^{(s)}_{\vec{p}} = 0. \tag{7.161}$$

They obey the completeness relations

$$\sum_{r=1}^{2} u^{(r)}_{\vec{p}} \overline{u}^{(r)}_{\vec{p}} = \not{p} + m, \;\; \sum_{r=1}^{2} v^{(r)}_{\vec{p}} \overline{v}^{(r)}_{\vec{p}} = \not{p} - m. \tag{7.162}$$

Another way of expressing the normalization condition is

$$\overline{u}^{(r)}_{\vec{p}} \gamma^\mu u^{(s)}_{\vec{p}} = 2\delta^{rs} p^\mu = \overline{v}^{(r)}_{\vec{p}} \gamma^\mu v^{(s)}_{\vec{p}}. \tag{7.163}$$

This form has a smooth limit as $m \to 0$.

8 Quantizing the Dirac Lagrangian

8.1 Canonical Commutation Relations

or, How Not to Quantize the Dirac Lagrangian

We now wish to construct the quantum theory corresponding to the Dirac Lagrangian, and so we expect to be able to set up canonical commutation relations much in the same way as for the scalar field. The momentum conjugate to ψ is

$$\Pi^0_\psi = \frac{\partial \mathcal{L}}{\partial(\partial_0\psi)} = i\psi^\dagger \tag{8.1}$$

while the momentum conjugate to $\psi^\dagger$ vanishes. Although this seems odd, it is not a problem. The equations of motion from the Dirac equation are first order in time, and so ψ and $\psi^\dagger$ form a complete set of initial value data. That is, if we know ψ and $\psi^\dagger$ at some initial time, we can find the state of the system at any following time (if the equations were second order in time, we would also need the time derivatives of the fields at the initial time). It is only on these fields, which completely define the state of the system, that we need to impose canonical commutation relations. Therefore we take

$$[\psi_a(\vec{x},t), (\Pi^0_\psi)_b(\vec{y},t)] = i\delta_{ab}\delta^{(3)}(\vec{x}-\vec{y}). \tag{8.2}$$

Here we have explicitly displayed the spinor indices a and b. Suppressing the indices, we have

$$[\psi(\vec{x},t), \psi^\dagger(\vec{y},t)] = \delta^{(3)}(\vec{x}-\vec{y}), \;\; [\psi(\vec{x},t), \psi(\vec{y},t)] = [\psi^\dagger(\vec{x},t), \psi^\dagger(\vec{y},t)] = 0. \tag{8.3}$$

Just as in the case of the scalar field, any solution to the free field theory may be written as a sum of plane wave solutions

$$\begin{aligned}
\psi(\vec{x},t) &= \sum_{r=1}^{2}\int \frac{d^3\vec{p}}{(2\pi)^{3/2}\sqrt{2E_p}}\left[b^{(r)}_{\vec{p}}u^{(r)}_{\vec{p}}e^{-ip\cdot x} + c^{(r)\dagger}_{\vec{p}}v^{(r)}_{\vec{p}}e^{ip\cdot x}\right] \\
\psi^\dagger(\vec{x},t) &= \sum_{r=1}^{2}\int \frac{d^3\vec{p}}{(2\pi)^{3/2}\sqrt{2E_p}}\left[b^{(r)\dagger}_{\vec{p}}u^{(r)\dagger}_{\vec{p}}e^{ip\cdot x} + c^{(r)}_{\vec{p}}v^{(r)\dagger}_{\vec{p}}e^{-ip\cdot x}\right].
\end{aligned} \tag{8.4}$$

In the classical theory, the b's and c's are numbers, the Fourier components of the solution, just as in the case of the scalar field. The u's and v's are the four component bispinors we found explicitly in the previous section. Since there are two spin states for the fields, a general solution to the Dirac Equation has components with both spin states, and so the b's and c's carry a spin index.

In the quantum theory, the b's and c's are replaced by operators. We expect that the canonical commutation relations Eq. (8.2) will require that the b's, c's and their conjugates be creation and annihilation operators, so to make things simpler let us make the ansatz

$$\begin{aligned}
[b^{(r)}_{\vec{p}}, b^{(s)\dagger}_{\vec{p}'}] &= B\delta^{rs}\delta^{(3)}(\vec{p}-\vec{p}') \\
[c^{(r)}_{\vec{p}}, c^{(s)\dagger}_{\vec{p}'}] &= C\delta^{rs}\delta^{(3)}(\vec{p}-\vec{p}')
\end{aligned} \tag{8.5}$$

where B and C are constant which we shall solve for. Substituting Eq. (8.5) into the commutation relations gives

$$\begin{aligned}
[\psi(\vec{x},t), \psi^\dagger(\vec{y},t)] &= \sum_{r,s}\int \frac{d^3\vec{p}d^3\vec{p}'}{(2\pi)^3\sqrt{2E_p}\sqrt{2E_{p'}}}\Big\{[b^{(r)}_{\vec{p}}, b^{(s)\dagger}_{\vec{p}'}]u^{(r)}_{\vec{p}}\overline{u}^{(s)}_{\vec{p}'}\gamma^0 e^{ip\cdot x - ip'\cdot y} \\
&\qquad + [c^{(r)\dagger}_{\vec{p}}, c^{(s)}_{\vec{p}'}]v^{(r)}_{\vec{p}}\overline{v}^{(s)}_{\vec{p}'}\gamma^0 e^{-ip\cdot x + ip'\cdot y}\Big\} \\
&= \frac{1}{(2\pi)^3}\int \frac{d^3p}{2E_p}\Big\{B(\not{p}+m)\gamma^0 e^{-i\vec{p}\cdot(\vec{x}-\vec{y})} \\
&\qquad - C(\not{p}-m)\gamma^0 e^{i\vec{p}\cdot(\vec{x}-\vec{y})}\Big\} \\
&= \frac{1}{(2\pi)^3}\int \frac{d^3p}{2E_p}e^{-i\vec{p}\cdot(\vec{x}-\vec{y})}\Big\{B(p_0\gamma^0 + p_i\gamma^i + m)\gamma^0 \\
&\qquad - C(p_0\gamma^0 - p_i\gamma^i - m)\gamma^0\Big\}.
\end{aligned} \tag{8.6}$$

Here we have used the completeness relations $\sum_r u^{(r)}_{\vec{p}}\overline{u}^{(r)}_{\vec{p}} = \not{p}+m$ and $\sum_r v^{(r)}_{\vec{p}}\overline{v}^{(r)}_{\vec{p}} = \not{p}-m$. Clearly if $B=-C$, the $p_i\gamma^i$ and m terms cancel, and the $p_0 = E_p$ in the numerator cancels the denominator. So choosing $B=-C=1$, we obtain

$$[\psi(\vec{x},t), \psi^\dagger(\vec{y},t)] = \frac{1}{(2\pi)^3}\int d^3p\, e^{-i\vec{p}\cdot(\vec{x}-\vec{y})} = \delta^{(3)}(\vec{x}-\vec{y}) \tag{8.7}$$

as required. Note, however, that the sign in the commutator for the c's is opposite to what we might have expected, and suggests that something may not be quite right here.

To see if this is a sensible quantum theory, we should look at the Hamiltonian and see if the energy is bounded below:

$$\mathcal{H} = \Pi^0_\psi\partial_0\psi - \mathcal{L} = i\overline{\psi}\gamma^0\partial_0\psi - i\overline{\psi}\gamma^\mu\partial_\mu\psi + m\overline{\psi}\psi = -i\overline{\psi}\gamma^i\partial_i\psi + m\overline{\psi}\psi. \tag{8.8}$$

Since ψ satisfies the Dirac equation, we can write this as

$$\mathcal{H} = i\overline{\psi}\gamma^0\partial_0\psi = i\psi^\dagger\partial_0\psi. \tag{8.9}$$

In terms of the creation and annihilation operators,

$$i\partial_0\psi = \sum_r \int \frac{d^3p}{(2\pi)^{3/2}}\sqrt{\frac{E_p}{2}}\left[b_{\vec{p}}^{(r)}u_{\vec{p}}^{(r)}e^{-ip\cdot x} - c_{\vec{p}}^{(r)\dagger}v_{\vec{p}}^{(r)}e^{ip\cdot x}\right] \qquad (8.10)$$

and so the Hamiltonian is

$$\begin{aligned} H &= \int d^3x\,\mathcal{H} \\ &= \int d^3x\,\psi^\dagger i\partial_0\psi \\ &= \sum_{r,s}\int d^3x\,\frac{d^3pd^3p'}{(2\pi)^3}\sqrt{\frac{E_p}{E_{p'}}}\left[b_{\vec{p}'}^{(r)\dagger}u_{\vec{p}'}^{(r)\dagger}e^{ip'\cdot x} + c_{\vec{p}'}^{(r)}v_{\vec{p}'}^{(r)\dagger}e^{-ip'\cdot x}\right] \times \\ &\qquad \left[b_{\vec{p}}^{(s)}u_{\vec{p}}^{(s)}e^{-ip\cdot x} - c_{\vec{p}}^{(r)\dagger}v_{\vec{p}}^{(r)}e^{ip'\cdot x}\right]. \qquad (8.11) \end{aligned}$$

As usual, the d^3x integral times the exponential becomes a delta function, and using $u_{\vec{p}}^{(r)\dagger}u_{\vec{p}}^{(s)} = \overline{u}_{\vec{p}}^{(r)}\gamma^0 u_{\vec{p}}^{(s)} = 2\delta^{rs}E_p = v_{\vec{p}}^{(r)\dagger}v_{\vec{p}}^{(s)}$, we arrive at

$$\begin{aligned} H &= \sum_r\int d^3pE_p\left[b_{\vec{p}}^{(r)\dagger}b_{\vec{p}}^{(r)} - c_{\vec{p}}^{(r)}c_{\vec{p}}^{(r)\dagger}\right] \\ &= \sum_r\int d^3pE_p\left[b_{\vec{p}}^{(r)\dagger}b_{\vec{p}}^{(r)} - c_{\vec{p}}^{(r)\dagger}c_{\vec{p}}^{(r)} + \delta^{(3)}(0)\right]. \qquad (8.12) \end{aligned}$$

The $\delta^{(3)}(0)$ will vanish when we normal order, so we just find

$$H = \sum_r \int d^3pE_p\left[N_b(p,r) - N_c(p,r)\right] \qquad (8.13)$$

where $N_b(p,r)$ and $N_c(p,r)$ are the number operators for b and c type particles.

There is indeed something seriously wrong with this theory - the Hamiltonian is unbounded from below! The c-type particles (antiparticles) carry negative energy. The theory therefore has no ground state, since you can always lower the energy of a state by adding antiparticles.

Unlike previous problems with positivity of the energy, we can't fix this problem simply by changing the sign of the Lagrangian. This will simply force the b particles to carry negative energy. There is therefore no way to obtain a sensible quantum theory from the Dirac Lagrangian using canonical commutation relations.

8.2 Canonical Anticommutation Relations

We didn't spend two lectures on spinors and γ matrices just to throw it all in at the first sign of trouble. The theory can be rescued, but the canonical commutation relations must be abandoned and replaced with something else. Recall that

for the scalar field theory we could interpret $a_k^\dagger$ and a_k as creation and annihilation operators because of their commutation relations with the number operator $N = \int d^3k\, a_k^\dagger a_k$ (or equivalently, with the Hamiltonian $H = \int d^3k\,\omega_k a_k^\dagger a_k$.). Let me remind you that this worked because of the following useful identity for commutators:

$$[AB, C] = A[B,C] + [A,C]B. \qquad (8.14)$$

This immediately gives

$$[N, a_k^\dagger] = \int d^3k'\,[a_{\vec{k}'}^\dagger a_{\vec{k}'}, a_k^\dagger] = a_k^\dagger \qquad (8.15)$$

and also

$$[N, a_k] = -a_k. \qquad (8.16)$$

Therefore $a_k^\dagger$ acting on a state raises the eigenvalue of N by one and the energy by ω_k, while a_k acting on the states lowers both eigenvalues, exactly as expected for creation and annihilation operators. However, there is another useful identity for commutators

$$[AB, C] = A\{B,C\} - \{A,C\}B \qquad (8.17)$$

where $\{A,B\} \equiv AB + BA$ is the *anticommutator* of A and B. This is extremely useful, because it means that *if we were to impose anticommutation relations on the* $a_k^\dagger$*'s and* a_k*'s, they could still be interpreted as creation and annihilation operators.* That is, let us impose the relations

$$\begin{aligned} \{a_k, a_{k'}^\dagger\} &= \delta^{(3)}(\vec{k} - \vec{k}') \\ \{a_k, a_{k'}\} &= \{a_k^\dagger, a_{k'}^\dagger\} = 0. \qquad (8.18) \end{aligned}$$

We then find, using Eq. (8.17)

$$\begin{aligned} [N, a_k^\dagger] &= \int d^3k'\,[a_{\vec{k}'}^\dagger a_{\vec{k}'}, a_k^\dagger] = \int d^3k'\,a_{k'}^\dagger\{a_{k'}, a_k^\dagger\} = a_k^\dagger \\ [N, a_k] &= \int d^3k'\,[a_{\vec{k}'}^\dagger a_{\vec{k}'}, a_k] = -\int d^3k'\,\{a_{k'}^\dagger, a_k\}a_{k'} = -a_k \qquad (8.19) \end{aligned}$$

exactly as required. This suggests we try the following anticommutation relations for the b's and c's:

$$\begin{aligned} \{b_{\vec{p}}^{(r)}, b_{\vec{p}'}^{(s)\dagger}\} &= \delta^{rs}\delta^{(3)}(\vec{p} - \vec{p}') \\ \{c_{\vec{p}}^{(r)}, c_{\vec{p}'}^{(s)\dagger}\} &= \delta^{rs}\delta^{(3)}(\vec{p} - \vec{p}') \\ \{b_{\vec{p}}^{(r)}, b_{\vec{p}'}^{(s)}\} &= \{b_{\vec{p}}^{(r)\dagger}, b_{\vec{p}'}^{(s)\dagger}\} = 0 \\ \{c_{\vec{p}}^{(r)}, c_{\vec{p}'}^{(s)}\} &= \{c_{\vec{p}}^{(r)}, c_{\vec{p}'}^{(s)}\} = \{b_{\vec{p}}^{(r)}, c_{\vec{p}'}^{(s)\dagger}\} = ... = 0. \qquad (8.20) \end{aligned}$$

Not surprisingly, substituting these anticommutation relations into the field expansions we find that the equal-time commutation relations are replaced by now obey equal time anticommutation relations

$$\begin{aligned}\{\psi(\vec{x},t),\psi^\dagger(\vec{y},t)\} &= \delta^{(3)}(\vec{x}-\vec{y}) \\ \{\psi(\vec{x},t),\psi(\vec{y},t)\} &= \{\psi^\dagger(\vec{x},t),\psi^\dagger(\vec{y},t)\} = 0. \end{aligned} \tag{8.21}$$

Note that you have to be careful when dealing with anticommuting fields, since the order is always important. For example, $\psi(\vec{x},t)\psi(\vec{y},t) = -\psi(\vec{y},t)\psi(\vec{x},t)$.

The crucial step is now to see if this modification gives us an energy bounded from below. It is easy to see that it does, since the previous derivation of the Hamiltonian goes through completely unchanged up until the last line:

$$\begin{aligned} H &= \sum_r \int d^3p E_p \left[b_{\vec{p}}^{(r)\dagger} b_{\vec{p}}^{(r)} - c_{\vec{p}}^{(r)} c_{\vec{p}}^{(r)\dagger} \right] \\ &= \sum_r \int d^3p E_p \left[b_{\vec{p}}^{(r)\dagger} b_{\vec{p}}^{(r)} + c_{\vec{p}}^{(r)\dagger} c_{\vec{p}}^{(r)} + \delta^{(3)}(0) \right]. \end{aligned} \tag{8.22}$$

The anticommutation relations have given us a crucial sign change in the second term. Throwing away the $\delta^{(3)}(0)$ as usual, we now have

$$H = \sum_r \int d^3p E_p \left[N_b(p,r) + N_c(p,r) \right] \tag{8.23}$$

which is bounded from below. Both b and c particles carry positive energy.

8.3 Fermi-Dirac Statistics

We have saved the theory, but at the price of imposing anticommutation relations on the creation and annihilation operators, and we must now examine the consequences of this. First consider the single particle states in the theory. We label these by the spin r (where $r = 1$ or 2 labels spin up and down, as we did when in the last chapter when writing down the explicit form of the plane wave solutions) as well as the momentum $\vec{p}$. As usual, they are produced by the action of a creation operator on the vacuum (for definiteness, we consider particle states, not antiparticle states, although the arguments will clearly apply in both cases):

$$\begin{aligned} |\,\vec{p},r\rangle &= b_{\vec{p}}^{(r)\dagger}|\,0\rangle \\ \langle \vec{p}', s\,|\vec{p},r\rangle &= \langle 0\,| b_{\vec{p}'}^{(s)} b_{\vec{p}}^{(r)\dagger}|\,0\rangle \\ &= \langle 0\,|\{ b_{\vec{p}'}^{(s)}, b_{\vec{p}}^{(r)\dagger}\}|\,0\rangle \\ &= \delta^{rs}\delta^{(3)}(\vec{p}-\vec{p}') \end{aligned} \tag{8.24}$$

and so the states have the correct normalization, just as they did in the scalar case. However, the multiparticle states are different from the spin 0 case. We find

$$|\,\vec{p}_1, r; \vec{p}_2, s\rangle = b_{\vec{p}_1}^{(r)\dagger} b_{\vec{p}_2}^{(s)\dagger}|\,0\rangle = -b_{\vec{p}_2}^{(s)\dagger} b_{\vec{p}_1}^{(r)\dagger}|\,0\rangle = -|\,\vec{p}_2, s; \vec{p}_1, r\rangle \tag{8.25}$$

and so the states of the theory change sign under the exchange of identical particles. Thus, the particle obey *Fermi-Dirac* statistics, instead of Bose-Einstein statistics. Consistency of the theory has demanded that we quantize the particles as fermions instead of bosons. In particular, the Pauli exclusion principle follows immediately from

$$\left(b_{\vec{p}_1}^{(r)}\right)^2 = -\left(b_{\vec{p}_1}^{(r)}\right)^2 = 0 \tag{8.26}$$

which means that there is no two-particle state made up of two identical particles in the same state

$$|\,\vec{p}_1, r; \vec{p}_1, r\rangle = 0. \tag{8.27}$$

Thus, it is impossible to put two identical fermions in the same state.

It isn't immediately obvious that a theory with Fermi-Dirac statistics will be causal. For bosons, we said that $[\phi(x),\phi(y)] = 0$ for $(x-y)^2 < 0$ guaranteed that spacelike separated observables, which are constructed out of the fields, couldn't interfere with one another:

$$[O(x), O(y)] = 0, \;\; (x-y)^2 < 0. \tag{8.28}$$

However, for fermi fields we now have the relation $\{\psi(x),\psi(y)\} = 0$ as well as $\{\psi(x),\psi^\dagger(y)\} = 0$ for $(x-y)^2 < 0$ (this follows from the analogous calculation to that from which we derived $\Delta_+(x-y) = 0$ for $(x-y)^2 < 0$.) How do we see that this requirement guarantees causality in the quantum theory?

The reason it does it that observables are always *bilinear* in the fields. For example, the energy, momentum and conserved charge are given by

$$\begin{aligned} H &= i\int d^3x\, \psi^\dagger(\vec{x},t)\partial_0\psi(\vec{x},t) \\ P_i &= -i\int d^3x\, \psi^\dagger(\vec{x},t)\partial_i\psi(\vec{x},t) \\ Q &= \int d^3x\, \psi^\dagger(\vec{x},t)\psi(\vec{x},t). \end{aligned} \tag{8.29}$$

This is not surprising. We know that spinors form a double-valued representation of the Lorentz group since they change sign under rotation by 2π. Observables, on the other hand, are unaffected by a rotation by 2π and so must be composed of an even number of spinor fields. Using the anticommutation relations (8.21), we can easily verify that observables bilinear in the fields commute for spacelike separation, as required.

The fact that particles with integer spin must be quantized as bosons while particles with half-integral spin must be quantized as fermions is a general result in field theory, and is known as the spin-statistics theorem. We have, at least for spin 1/2 fields, demonstrated the second part of the theorem. The first part of the theorem, the fact that particles with integral spin must be quantized as bosons, follows from that observation that if we were to attempt to impose canonical anticommutation relations on the creation and annihilation operators for a scalar field we would find that the fields obeyed neither $[\phi(x), \phi(y)]_{(x-y)^2<0} = 0$ nor $\{\phi(x), \phi(y)\}_{(x-y)^2<0} = 0$. The theory would therefore not be causal. This is the gist of the spin-statistics theorem: quantizing integral spin fields as fermions leads to an acausal theory, while quantizing half-integral spin fields as bosons leads to a theory with energy unbounded below (and so with no ground state).

8.4 Perturbation Theory for Spinors

Now that we understand free field theory for spin 1/2 fields, we can introduce interaction terms into the Lagrangian and build up the Feynman rules for perturbation theory. Let us consider a simple nucleon-meson theory (now that the nucleons are fermions, we no longer need to enclose the word in quotations)

$$\mathcal{L} = \overline{\psi}(i\partial\!\!\!/ - m)\psi + \frac{1}{2}(\partial_\mu \phi)^2 - \frac{\mu^2}{2}\phi^2 - g\overline{\psi}\Gamma\psi\phi \tag{8.30}$$

where we either take $\Gamma = 1$, in which case ϕ is a scalar, or $\Gamma = i\gamma_5$, in which case ϕ is a pseudoscalar (we include the i so that the Lagrangian is Hermitian, $\mathcal{L} = \mathcal{L}^\dagger$.). The theory with $\Gamma = 1$ is known as *Yukawa theory*; it was originally invented by Yukawa to describe the interaction between real pions and nucleons. It turns out that Yukawa theory does not, in fact, provide the correct description of nucleon-meson interactions even at low energies (where the internal structure of the nucleons and pions is irrelevant, so they may be treated as fundamental particles). However, in modern particle theory the Standard Model contains Yukawa interaction terms coupling the scalar Higgs field to the quarks and leptons.

Dyson's formula and Wick's theorem go through for fermi fields in almost the same way as for scalars. However, the anticommutation relations introduce a crucial difference. Recall that when $(x-y)^2 < 0$, time ordering is not a Lorentz invariant concept. In one frame $x_0 > y_0$ while in another $y_0 > x_0$. Nevertheless, the T-product of two scalar fields *is* Lorentz invariant because the fields commute when $(x-y)^2 < 0$, so $\phi(x)\phi(y) = \phi(y)\phi(x)$ and the order is unimportant. However, for fermions this no longer holds. If $(x-y)^2 < 0$, fermi fields anticommute.

So for spacelike separation,

$$T(\psi(x)\psi(y)) = \psi(x)\psi(y) \tag{8.31}$$

in the frame where $x_0 > y_0$, but

$$T(\psi(x)\psi(y)) = \psi(y)\psi(x) = -\psi(x)\psi(y) \tag{8.32}$$

in the frame where $y_0 > x_0$. Therefore we must modify our definition of the T-produce of fermi fields to make it Lorentz invariant. The solution is simple: just define the T-product to include a factor of $(-1)^N$, where N is the number of exchanges of fermi fields required to time order the fields. Thus, for two fields

$$T(\psi(x)\psi(y)) = \begin{cases} \psi(x)\psi(y), & x_0 > y_0; \\ -\psi(y)\psi(x), & y_0 > x_0. \end{cases} \tag{8.33}$$

since for $y_0 > x_0$ we must perform one exchange of fermi fields to time order them. When $(x-y)^2 < 0$ the fields anticommute and the T-product is the same in any frame. Also, from Eq. (8.33) and the anticommutation relations, we have

$$T(\psi_1(x_1)\psi_2(x_2)) = -T(\psi_2(x_2)\psi_1(x_1)). \tag{8.34}$$

Therefore we treat fermi fields as anticommuting inside the time ordering symbol. (Note that in this discussion of T-products I am using ψ to represent any generic fermi field, including $\psi^\dagger$).

The normal-ordered product is defined as before. Writing $\psi = \psi^{(+)} + \psi^{(-)}$, where $\psi^{(+)}$ multiplies an annihilation operator and $\psi^{(-)}$ multiplies a creation operator, the normal-ordered product $:\psi_1\psi_2:$ is

$$\begin{aligned} :\psi_1\psi_2: &= :\left[\psi_1^{(+)}\psi_2^{(+)} + \psi_1^{(+)}\psi_2^{(-)} + \psi_1^{(-)}\psi_2^{(-)} + \psi_1^{(-)}\psi_2^{(+)}\right]: \\ &= \left[\psi_1^{(+)}\psi_2^{(+)} - \psi_2^{(-)}\psi_1^{(+)} + \psi_1^{(-)}\psi_2^{(-)} + \psi_1^{(-)}\psi_2^{(+)}\right] \end{aligned} \tag{8.35}$$

where the second term has picked up a factor of (-1) because of the interchange of two fermi fields. Just as for the T-product, fermi fields can be treated as anticommuting inside a normal ordered product,

$$:\psi_1\psi_2: = -:\psi_2\psi_1:. \tag{8.36}$$

(Recall that bose fields commuted inside T-products and N-products; that is, their order was unimportant).

With this modified definition of the time-ordered product, Dyson's formula and Wick's theorem go through as before. Note, however, that we must be careful with contractions in Wick's theorem, for example for fermion fields $A_1 - A_4$ we have

$$:\overbrace{A_1 A_2 A_3} A_4: = -:\overbrace{A_1 A_3} A_2 A_4: = -\overbrace{A_1 A_3}:A_2 A_4: \tag{8.37}$$

and so pulling this particular contraction out of the normal-ordered product introduces a minus sign. In general, pulling a contraction out of a normal-ordered product introduces a factor of $(-1)^N$, where N is the number of interchanges of fermi fields required.

8.4.1 The Fermion Propagator

The fermion propagator is obtained from the contraction $\overbrace{\psi(x)\overline{\psi}}(y)$ (note that this is a four by four matrix: $S_{ab} = \overbrace{\psi(x)_a\overline{\psi}}(y)_b$. As with scalar fields, this is number (or rather a matrix of numbers) instead of an operator, so

$$\begin{aligned}\overbrace{\psi(x)\overline{\psi}}(y) &= \langle 0\,|\overbrace{\psi(x)\overline{\psi}}(y)|\,0\rangle \\ &= \langle 0\,|T(\psi(x)\overline{\psi}(y)) - :\psi(x)\overline{\psi}(y):|\,0\rangle \\ &= \langle 0\,|T(\psi(x)\overline{\psi}(y))|\,0\rangle. \end{aligned} \tag{8.38}$$

First consider the case $x_0 > y_0$. Then $T(\psi(x)\overline{\psi}(y)) = \psi(x)\overline{\psi}(y)$. Putting in the explicit expressions for the fields and using the completeness relations for the plane wave states $\sum_r u^{(r)}_{\vec{p}}\overline{u}^{(r)}_{\vec{p}} = \not{p} + m$ we find

$$\begin{aligned}\overbrace{\psi(x)\overline{\psi}}(y) &= \int \frac{d^3p}{(2\pi)^3 2E_p} e^{-ip\cdot(x-y)}(\not{p}+m) \\ &= (i\not{\partial}_x + m) i\Delta_+(x-y) \;\; (x_0 > y_0) \end{aligned} \tag{8.39}$$

where we recall that

$$i\Delta_+(x-y) = \int \frac{d^3p}{(2\pi)^3 2E_p} e^{-ip\cdot(x-y)} \tag{8.40}$$

and $\Delta_+(x-y) = 0$ when $x_0 = y_0$. Performing a similar calculation for $x_0 < y_0$, we find[12]

$$\begin{aligned}\overbrace{\psi(x)\overline{\psi}}(y) &= \theta(x_0-y_0)(i\not{\partial}_x + m)i\Delta_+(x-y) + \theta(y_0-x_0)(i\not{\partial}_x + m)i\Delta_+(y-x) \\ &= (i\not{\partial}_x + m)\left(\theta(x_0-y_0)i\Delta_+(x-y) + \theta(y_0-x_0)i\Delta_+(y-x)\right) \\ &= (i\not{\partial}_x + m)\overbrace{\phi(x)\phi}(y) \end{aligned} \tag{8.41}$$

[12]Note that it is legitimate to pull the derivative outside of the θ function because the additional term which arises when a time derivative acts on the θ function vanishes: $\partial\theta(x_0-y_0)/\partial x_0 \Delta_+(x-y) = \delta(x_0-y_0)\Delta_+(x-y) = 0$.

where

$$\overbrace{\phi(x)\phi}(y) = \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)} \frac{i}{p^2 - m^2 + i\epsilon}. \tag{8.42}$$

We have now related the fermion propagator to the scalar propagator. Moving the derivative back inside the integral we have

$$\overbrace{\psi(x)_a\overline{\psi}}(y)_b = \int \frac{d^4p}{(2\pi)^4} \frac{i(\not{p}_{ab} + mI_{ab})}{p^2 - m^2 + i\epsilon} e^{-ip\cdot(x-y)}. \tag{8.43}$$

(where we have explicitly included the matrix indices, and I is the four by four identity matrix). We immediately see that this gives the Feynman rule for the fermion propagator shown in Fig. (8.1). Note that the propagator is odd in p, so it

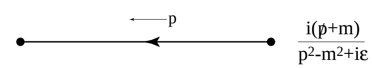

Figure 8.1: The fermion propagator.

matters that p and the conserved charge (the arrow on the propagator) are poiting in the same direction. When they point in opposite directions the sign of p is reversed (Fig. (8.2)). Note that $p^2 - m^2 + i\epsilon = (\not{p} + m - i\epsilon)(\not{p} - m + i\epsilon)$, so the propagator

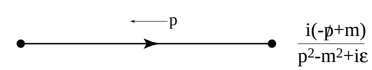

Figure 8.2: The fermion propagator is odd in p.

is often written as

$$\frac{i(\not{p}+m)}{(\not{p}+m-i\epsilon)(\not{p}-m+i\epsilon)} = \frac{i}{\not{p}-m+i\epsilon} \tag{8.44}$$

(the $i\epsilon$ in the $\not{p}+m$ term in the denominator does not affect the location of the pole, so in the limit $\epsilon \to 0$ we may cancel this against the numerator).

Of course, just as in the scalar theory, contractions of fields which don't create and then annihilate the same particle vanish,

$$\begin{aligned}\overbrace{\phi(x)\psi}(y)_a &= 0 \\ \overbrace{\psi(x)_a\psi}(y)_b &= 0 \\ \overbrace{\overline{\psi}(x)_a\overline{\psi}}(y)_b &= 0. \end{aligned} \tag{8.45}$$

8.4.2 Feynman Rules

We can deduce the Feynman rules for this theory by explicitly calculating the amplitudes for several scattering processes. The $\mathcal{O}(g^2)$ term in Dyson's formula is

$$\frac{(-ig)^2}{2!}\int d^4x_1\, d^4x_2\, T\left[\overline{\psi}_a(x_1)\Gamma_{ab}(x_1)\psi_b(x_1)\phi(x_1)\overline{\psi}_c(x_2)\Gamma_{cd}\psi_d(x_2)\phi(x_2)\right] \tag{8.46}$$

where we have included the spinor indices, and we are using the convention that repeated spinor indices are summed over (so this is just matrix multiplication). This term contributes to a variety of processes. First we consider nucleon-meson scattering, $N+\phi\to N+\phi$. There are two contractions which contribute:

$$\begin{aligned}&:\overline{\psi}_a(x_1)\Gamma_{ab}\overbrace{\psi_b(x_1)\phi(x_1)\overline{\psi}_c}(x_2)\Gamma_{cd}\psi_d(x_2)\phi(x_2):\\ &:\overbrace{\overline{\psi}_a(x_1)\Gamma_{ab}\psi_b(x_1)\phi(x_1)\overline{\psi}_c(x_2)\Gamma_{cd}\psi_d}(x_2)\phi(x_2):.\end{aligned} \tag{8.47}$$

Anticommuting the fields inside the normal-ordered product, we can rewrite the second term as

$$:\overline{\psi}_c(x_2)\Gamma_{cd}\overbrace{\psi_d(x_2)\phi(x_2)\overline{\psi}_a}(x_1)\Gamma_{ab}\psi_b(x_1)\phi(x_1):(-1)^4 \tag{8.48}$$

since there are four permutations of fermion fields required (note that fermi fields commute with bose fields). This only differs from the first time by the interchange of x_1 and x_2, and since we are symmetrically integrating over x_1 and x_2, the two terms give identical contributions. Just as before, this cancels the $1/2!$ in Dyson's formula.

We can pull the propagator out of the first term, and since this involves an even number of exchanges of fermi fields (two), we get

$$\begin{aligned}&:\overline{\psi}_a(x_1)\Gamma_{ab}\overbrace{\psi_b(x_1)\phi(x_1)\overline{\psi}_c}(x_2)\Gamma_{cd}\psi_d(x_2)\phi(x_2):=\\ &\qquad\overbrace{\psi_b(x_1)\overline{\psi}_c}(x_2):\overline{\psi}_a(x_1)\Gamma_{ab}\phi(x_1)\Gamma_{cd}\psi_d(x_2)\phi(x_2):.\end{aligned} \tag{8.49}$$

The ψ field inside the normal product must now annihilate the nucleon. For the relativistically normalized nucleon state $|\,N(p,r)\rangle$ (momentum p, spin r) we have

$$\langle 0\,|\psi(x_2)|\,N(p,r)\rangle = e^{-ip\cdot x_2}\times u^{(r)}_{\vec{p}} \tag{8.50}$$

and similarly

$$\langle N(p',r')\,|\overline{\psi}(x_1)|\,0\rangle = e^{ip'\cdot x_1}\times \overline{u}^{(r')}_{\vec{p}'} \tag{8.51}$$

This immediately gives us two additional Feynman rules:

- For each incoming fermion with momentum p and spin r, include a factor of $u^{(r)}_{\vec{p}}$
- For each outgoing fermion with momentum p and spin r, include a factor of $\overline{u}^{(r)}_{\vec{p}}$

(see Fig. (8.3).) Finally, each interaction vertex corresponds to a factor of $-ig\Gamma$

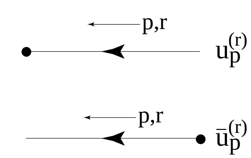

Figure 8.3: Feynman rules for external fermion legs.

(see Fig. (8.4).) The vertices and fermion propagators are four by four matrices

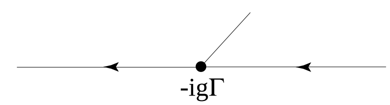

Figure 8.4: Fermion-scalar interaction vertex.

while the bispinors are four component column vectors. The amplitude is given by multiplying all of these factors together. From Eq. (8.49) we see that the matrices are multiplied together in the order $\overline{u}^{(r')}_{\vec{p}'}\Gamma S\Gamma u^{(r)}_{\vec{p}}$, where S_{ab} is the fermion propagator. Diagrammatically, this just corresponds to starting at the head of the arrow and working back to the start, including each matrix as it is encountered along the fermion line.

That last point is important, so I'm going to say it again. *When calculating Feynman diagrams for spinors, the order of matrix multiplication is given by starting at the head of an arrow and working back to the start, including each matrix as it is encountered along the fermion line.*

The ϕ fields act as they always did on the meson states, and as before we get two Feynman diagrams corresponding to the two choices of which ϕ field creates or annihilates which meson, as shown in Fig. (8.5). Applying the Feynman rules, we find the invariant amplitude for this process to be

$$i\mathcal{A} = (-ig)^2\overline{u}^{(r')}_{\vec{p}'}\Gamma\left[\frac{i(\not{p}+\not{q}+m)}{(p+q)^2-m^2+i\epsilon}+\frac{i(\not{p}-\not{q}'+m)}{(p-q')^2-m^2+i\epsilon}\right]\Gamma u^{(r)}_{\vec{p}}. \tag{8.52}$$

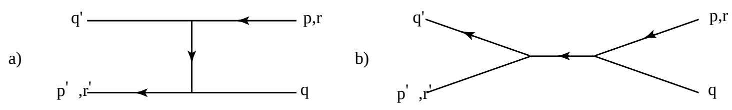

Figure 8.5: Feynman diagrams contributing to nucleon-meson scattering.

We next consider antinucleon-meson scattering, $\overline{N}+\phi \to \overline{N}+\phi$. This is almost identical to the previous process, but now the $\overline{\psi}$ field annihilates the incoming antinucleon and the ψ field creates the outgoing antinucleon. So in the normal-ordered product $:\overline{\psi}(x_1)\Gamma\phi(x_1)\Gamma\psi(x_2)\phi(x_2):$ the $\overline{\psi}$ field has to be moved to the right of the ψ field in order to annihilate the incoming state, introducing a factor of -1 because of the interchange of the two fermion fields. When acting on the external states, the fields now include factors of v and $\overline{v}$:

$$\begin{aligned}\langle 0\,|\overline{\psi}(x_1)|\,\overline{N}(p,r)\rangle &= e^{-ip\cdot x_1}\times\overline{v}^{(r)}_{\vec{p}} \\ \langle\overline{N}(p',r')\,|\psi(x_2)|\,0\rangle &= e^{ip'\cdot x_2}\times v^{(r')}_{\vec{p}'}\end{aligned} \tag{8.53}$$

This leads to three more Feynman rules:

- For each incoming antifermion with momentum p and spin r, include a factor of $\overline{v}^{(r)}_{\vec{p}}$.
- For each outgoing antifermion with momentum p and spin r, include a factor of $v^{(r)}_{\vec{p}}$.

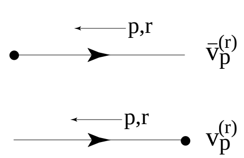

Figure 8.6: Feynman rules for external fermion legs.

- Include the appropriate minus signs from Fermi statistics.

This time the matrices are multiplied by starting at the incoming antinucleon and working back to the outgoing nucleon. Diagrammatically, it's the same as before: start at the end of the arrow (with a factor of $\overline{v}^{(r)}_{\vec{p}}$ this time, instead of $\overline{u}^{(r)}_{\vec{p}}$)

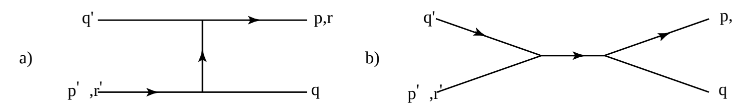

Figure 8.7: Feynman diagrams contributing to antinucleon-meson scattering.

and work along the line to the start of the arrow. The two diagrams contributing to the process are shown in Fig. (8.7). The amplitude is

$$i\mathcal{A} = -(-ig)^2\overline{v}^{(r)}_{\vec{p}}\Gamma\left[\frac{i(-\not{p}-\not{q}+m)}{(p+q)^2-m^2+i\epsilon}+\frac{i(-\not{p}+\not{q}'+m)}{(p-q')^2-m^2+i\epsilon}\right]\Gamma v^{(r')}_{\vec{p}'}. \tag{8.54}$$

The overall -1 is clearly irrelevant, since it vanishes when the amplitude is squared. However, the fermi minus signs will be significant in the next example, because they will differ between the two diagrams.

Finally, we consider nucleon-nucleon scattering, $N+N \to N+N$. In this case the ϕ fields are contracted, leaving us with the matrix element

$$\langle N(p',r');N(q',s')\,|:\overline{\psi}\Gamma\psi(x_1)\overline{\psi}\Gamma\psi(x_2):|\,N(p,r);N(q,s)\rangle. \tag{8.55}$$

Now we have to be careful. The (relativistically normalized) state $|\,N(p,r);N(q,s)\rangle$ can be defined either as

$$(2\pi)^3(2\omega_q)^{1/2}(2\omega_p)^{1/2}b^{(s)\dagger}_{\vec{q}}b^{(r)\dagger}_{\vec{p}}|\,0\rangle \tag{8.56}$$

or

$$(2\pi)^3(2\omega_q)^{1/2}(2\omega_p)^{1/2}b^{(r)\dagger}_{\vec{p}}b^{(s)\dagger}_{\vec{q}}|\,0\rangle. \tag{8.57}$$

Because of Fermi statistics, the two definitions differ by a relative minus sign. So for definiteness, let us choose the first definition. This sets the convention, and so we have now choice but to define the corresponding bra as

$$\langle N(p',r');N(q',s')\,| = (2\pi)^3(2\omega_{q'})^{1/2}(2\omega_{p'})^{1/2}\langle 0\,|b^{(r')}_{\vec{p}'}b^{(s')}_{\vec{q}'} \tag{8.58}$$

and so the matrix element in Eq. (8.55) becomes

$$4(2\pi)^6(\omega_{q'}\omega_{p'}\omega_q\omega_p)^{1/2}\langle 0\,|b^{(r')}_{\vec{p}'}b^{(s')}_{\vec{q}'}:\overline{\psi}\Gamma\psi(x_1)\overline{\psi}\Gamma\psi(x_2):b^{(s)\dagger}_{\vec{q}}b^{(r)\dagger}_{\vec{p}}|\,0\rangle. \tag{8.59}$$

There are now two possibilities: either $\psi(x_1)$ or $\psi(x_2)$ can annihilate the nucleon with momentum q. First we choose the case where $\psi(x_2)$ annihilates the nucleon.

Using the field expansion for ψ, we then obtain

$$\begin{aligned} & \langle 0\,|b_{\vec{p}'}^{(r')} b_{\vec{q}'}^{(s')} : \overline{\psi}\Gamma\psi(x_1)\overline{\psi}(x_2)\Gamma u_{\vec{q}}^{(s)} b_{\vec{p}}^{(r)\dagger}|\,0\rangle e^{-iq\cdot x_2} \\ = & -\langle 0\,|b_{\vec{p}'}^{(r')} b_{\vec{q}'}^{(s')} : \overline{\psi}(x_1)\Gamma\overline{\psi}(x_2)\Gamma u_{\vec{q}}^{(s)}\psi(x_1) b_{\vec{p}}^{(r)\dagger}|\,0\rangle e^{-iq\cdot x_2} \\ = & -\langle 0\,|b_{\vec{p}'}^{(r')} b_{\vec{q}'}^{(s')} : \overline{\psi}(x_1)\Gamma u_{\vec{p}}^{(r)}\overline{\psi}(x_2)\Gamma u_{\vec{q}}^{(s)}|\,0\rangle e^{-iq\cdot x_2 - ip\cdot x_1} \end{aligned} \tag{8.60}$$

where in the last line we have put the factor of $u_{\vec{p}}^{(r)}$ (which is not an operator, so it commutes with the fields) in the correct position as far as matrix multiplication goes.

Now there are two choices for which $\overline{\psi}$ field creates which nucleon. The crucial observation is that the two choices differ by a relative minus sign. If $\overline{\psi}(x_1)$ creates the nucleon with q', then there is no relative minus sign. However, if $\overline{\psi}(x_2)$ creates the nucleon, the fields must be anticommuted, and there is an additional minus sign. Thus, we find two terms

$$-\overline{u}_{\vec{q}'}^{(s')}\Gamma u_{\vec{q}}^{(s)}\overline{u}_{\vec{p}'}^{(r')}\Gamma u_{\vec{p}}^{(r)} e^{-i((q-q')\cdot x_2+(p-p')\cdot x_1)} \tag{8.61}$$

and

$$\overline{u}_{\vec{q}'}^{(s')}\Gamma u_{\vec{p}}^{(r)}\overline{u}_{\vec{p}'}^{(r')}\Gamma u_{\vec{q}}^{(s)} e^{-i((q-p')\cdot x_2+(p-q')\cdot x_1)}. \tag{8.62}$$

We could now follow through the same line or reasoning, except choosing $\psi(x_1)$ to annihilate the nucleon with momentum q, and we would find the same result with the interchange $x_1 \leftrightarrow x_2$. Since we are integrating over x_1 and x_2 symmetrically, once again this cancels the $1/2!$ in Dyson's formula.

The two terms clearly correspond to the diagrams in Fig. (8.8), and the two graphs have a relative minus sign. Therefore the amplitude for the process is

$$-ig^2\left[\frac{\overline{u}_{\vec{q}'}^{(s')}\Gamma u_{\vec{q}}^{(s)}\overline{u}_{\vec{p}'}^{(r')}\Gamma u_{\vec{p}}^{(r)}}{(q-q')^2-\mu^2+i\epsilon} - \frac{\overline{u}_{\vec{q}'}^{(s')}\Gamma u_{\vec{p}}^{(r)}\overline{u}_{\vec{p}'}^{(r')}\Gamma u_{\vec{q}}^{(s)}}{(q-p')^2-\mu^2+i\epsilon}\right] \tag{8.63}$$

Note that since the overall sign of the graphs is unimportant. It is only the relative minus sign which is significant. Also note that in this case there are two sets of arrowed lines, corresponding to two independent matrix multiplications. Again, the rule is to follow each arrowed line from finish to start, multiplying matrices as you come to them. In these diagram you simply have to do this twice.

The two graphs differ only by the interchange of identical fermions in the final state. As expected, our theory automatically incorporates fermi statistics. The two amplitudes interfere with a relative minus sign.

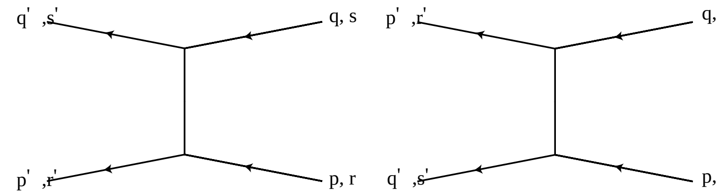

Figure 8.8: Feynman diagrams contributing to nucleon-nucleon scattering.

A similar situation arises in nucleon-antinucleon scattering. It is straightforward to show, using the techniques of this section, that two diagrams which differ by the exchange of a fermion in the final state and an antifermion in the initial state (or vice versa) also interfere with a relative minus sign.

8.5 Spin Sums and Cross Sections

Our Feynman rules allow us to calculate amplitudes in terms of plane-wave solutions u and v. To calculate the rate for a process with given initial and final spins, we could simply use the explicit forms of the u's and v's that we found earlier. However, in many cases this is not necessary. In a large number of experimental situations the spins of the initial particles are unknown, and so a given particle has a 50% chance to be in either spin state. Similarly, the final spins are often not measured, so we are interested in cross sections or decay rates in which we sum over all possible spins of the final particles. In such situations we can use the completeness relations for the spinors to perform the averaging and summing without ever writing down the explicit form of the plane wave states. We will demonstrate this in a worked example, nucleon-meson scattering in the "pseudoscalar" theory, $\Gamma = i\gamma_5$. This example also shows you several other tricks which are useful for evaluating amplitudes with Fermi fields.

From Eq. (8.52) the invariant Feynman amplitude for nucleon-meson scattering is

$$i\mathcal{A} = ig^2\overline{u}_{\vec{p}'}^{(r')}\gamma_5\left[\frac{(\not{p}+\not{q}+m)}{(p+q)^2-m^2+i\epsilon} + \frac{(\not{p}-\not{q}'+m)}{(p-q')^2-m^2+i\epsilon}\right]\gamma_5 u_{\vec{p}}^{(r)}. \tag{8.64}$$

Using $\gamma_5^2 = 1$ and $\{\gamma_5, \gamma_\mu\} = 0$, we can anticommute the second γ_5 through the propagators, where it hits the first γ_5 and the two square to one. Also, by conservation of momentum, $p - q' = p' - q$, so we may rewrite the amplitude as

$$i\mathcal{A} = ig^2 \overline{u}_{\vec{p}'}^{(r')} \left[\frac{(-\not{p} - \not{q} + m)}{(p+q)^2 - m^2 + i\epsilon} + \frac{(-\not{p}' + \not{q} + m)}{(p'-q)^2 - m^2 + i\epsilon} \right] u_{\vec{p}}^{(r)}. \qquad (8.65)$$

Next we use the fact that the spinors obey $(\not{p} - m) u_{\vec{p}}^{(r)} = \overline{u}_{\vec{p}}^{(r)} (\not{p} - m) = 0$ as well as the mass-shell conditions $p^2 = p'^2 = m^2$, $q^2 = q'^2 = \mu^2$ to write this as

$$i\mathcal{A} = ig^2 \overline{u}_{\vec{p}'}^{(r')} \not{q} u_{\vec{p}}^{(r)} \left[\frac{1}{2p \cdot q + \mu^2 + i\epsilon} + \frac{1}{2p' \cdot q - \mu^2 + i\epsilon} \right]. \qquad (8.66)$$

Let us call the expression in the square bracket $F(p, p', q)$. Squaring the amplitude we get

$$\begin{aligned} |\mathcal{A}|^2 &= g^4 F(p,p',q)^2 |\overline{u}_{\vec{p}'}^{(r')} \not{q} u_{\vec{p}}^{(r)}|^2 \\ &= g^4 F(p,p',q)^2 q_\mu q_\nu \overline{u}_{\vec{p}'}^{(r')} \gamma^\mu u_{\vec{p}}^{(r)} u_{\vec{p}}^{(r)\dagger} \gamma^0 \gamma^0 \gamma^{\nu\dagger} \gamma^0 u_{\vec{p}'}^{(r')} \\ &= g^4 F(p,p',q)^2 q_\mu q_\nu \overline{u}_{\vec{p}'}^{(r')} \gamma^\mu u_{\vec{p}}^{(r)} \overline{u}_{\vec{p}}^{(r)} \gamma^\nu u_{\vec{p}'}^{(r')} \qquad (8.67) \end{aligned}$$

where we have use the relations $\gamma_0^2 = 1$ and $\gamma_0 \gamma^\mu \gamma_0 = \gamma^{\mu\dagger}$. The collection of spinors and gamma matrices is simply a number (a one by one matrix) and so is equal to its trace. The reason for writing it in this way is that a trace of a product of matrices is invariant under cyclic permutations of the factors. Therefore

$$\begin{aligned} |\mathcal{A}^2| &= g^4 F(p,p',q)^2 q_\mu q_\nu \mathrm{Tr} \left[\overline{u}_{\vec{p}'}^{(r')} \gamma^\mu u_{\vec{p}}^{(r)} \overline{u}_{\vec{p}}^{(r)} \gamma^\nu u_{\vec{p}'}^{(r')} \right] \\ &= g^4 F(p,p',q)^2 q_\mu q_\nu \mathrm{Tr} \left[u_{\vec{p}'}^{(r')} \overline{u}_{\vec{p}'}^{(r')} \gamma^\mu u_{\vec{p}}^{(r)} \overline{u}_{\vec{p}}^{(r)} \gamma^\nu \right]. \qquad (8.68) \end{aligned}$$

Now the completeness relations comes in. Averaging over initial spins (corresponding to $\frac{1}{2}\sum_r |\mathcal{A}|^2$) and summing over final spins (corresponding to $\sum_{r'} |\mathcal{A}|^2$) we obtain

$$\begin{aligned} \frac{1}{2} \sum_{r,r'} |\mathcal{A}|^2 &= \frac{1}{2} \sum_{r,r'} g^4 F(p,p',q)^2 q_\mu q_\nu \mathrm{Tr} \left[u_{\vec{p}'}^{(r')} \overline{u}_{\vec{p}'}^{(r')} \gamma^\mu u_{\vec{p}}^{(r)} \overline{u}_{\vec{p}}^{(r)} \gamma^\nu \right] \\ &= \frac{1}{2} g^4 F(p,p',q)^2 q_\mu q_\nu \mathrm{Tr} \left[(\not{p}' + m) \gamma^\mu (\not{p} + m) \gamma^\nu \right]. \qquad (8.69) \end{aligned}$$

The traces of products of γ matrices have simple expressions, which are straightforward to prove (you can find a discussion of traces in Appendix A of Mandl & Shaw). Some useful formulas are:

$$\begin{aligned} &\mathrm{Tr}\left[\gamma^\mu \gamma^\nu\right] = 4 g^{\mu\nu} \\ &\mathrm{Tr}\left[\gamma^\mu \gamma^\nu \gamma^\alpha \gamma^\beta\right] = 4(g^{\mu\nu} g^{\alpha\beta} + g^{\mu\beta} g^{\nu\alpha} - g^{\mu\alpha} g^{\nu\beta}) \\ &\mathrm{Tr}\left[\text{odd number of } \gamma \text{ matrices}\right] = 0 \\ &\mathrm{Tr}\left[\gamma^\mu \gamma_5\right] = \mathrm{Tr}\left[\gamma^\mu \gamma^\nu \gamma_5\right] = \mathrm{Tr}\left[\gamma^\mu \gamma^\nu \gamma^\alpha \gamma_5\right] = 0 \\ &\mathrm{Tr}\left[\gamma^\mu \gamma^\nu \gamma^\alpha \gamma^\beta \gamma_5\right] = 4i\epsilon^{\mu\nu\alpha\beta}. \qquad (8.70) \end{aligned}$$

Applying these trace theorems to our expression gives

$$\begin{aligned} \frac{1}{2} \sum_{r,r'} |\mathcal{A}|^2 &= 2g^4 F(p,p',q)^2 q_\mu q_\nu \left[p'^\mu p^\nu + p'^\nu p^\mu - g^{\mu\nu} p \cdot p' + m^2 g^{\mu\nu} \right] \\ &= 2g^4 F(p,p',q)^2 \left[2(p' \cdot q)(p \cdot q) - p \cdot p' \mu^2 + m^2 \mu^2 \right]. \qquad (8.71) \end{aligned}$$

At this point it is straightforward, if somewhat tedious, to go to the centre of mass frame, substitute explicit expressions for the external momenta and perform the phase space integrals to obtain the total cross section for meson nucleon scattering.

9 Vector Fields and Quantum Electrodynamics

Quantizing a scalar field theory led to a theory of mesons, while the quantized spinor field allowed us to describe the interactions of spin 1/2 fermions. In this section we will see that a classical free field theory of a massless vector field is simply Maxwell's equations in free space. Quantizing the theory will give us the theory of the quantized electromagnetic field, *Quantum Electrodynamics.* The particles associated with the quantized vector field will be photons. However, quantizing a massless vector field is a delicate procedure, due to complications arising from gauge invariance of the classical theory. In this section we will finesse these problems by quantizing the theory of a massive vector field and then taking the massless limit and tackling any problems that arise at that stage.

9.1 The Classical Theory

A vector field is a four component field whose components transform in the familiar way under Lorentz transformations,

$$A'^{\mu}(x) = \Lambda^{\mu}{}_{\nu} A^{\nu}(\Lambda^{-1}x). \tag{9.1}$$

Since we already know how products of four-vectors transform, we can go straight to writing down Lagrangians. As before, we want to construct the simplest $\mathcal{L}$ which is quadratic in the fields (so that the resulting equations of motion are linear), has no more than two derivatives (a simplifying assumption) and is Lorentz invariant. This gives the following terms:

- 0 derivatives: there is only one possibility,

$$A^{\mu}A_{\mu}.$$

- 1 derivative: there are no possible Lorentz invariant terms in four dimensions.

- 2 derivatives: there are two independent terms,

$$\partial_{\mu}A^{\nu}\partial^{\mu}A_{\nu}, \;\; \partial_{\mu}A^{\mu}\partial^{\nu}A_{\nu}.$$

  Any other term may be written as a sum of these terms and a total derivative, and so gives the same contribution to the action. For example, up to total derivatives, $\partial_{\mu}A^{\nu}\partial_{\nu}A^{\mu} \sim A^{\nu}\partial_{\mu}\partial_{\nu}A^{\mu} \sim \partial_{\nu}A^{\nu}\partial_{\mu}A^{\mu}$.

The most general Lagrangian satisfying these requirements is then

$$\mathcal{L} = \pm\frac{1}{2}\left[\partial_{\mu}A^{\nu}\partial_{\nu}A^{\nu} + a\partial_{\mu}A^{\mu}\partial_{\nu}A^{\nu} + bA_{\mu}A^{\mu}\right] \tag{9.2}$$

for some constants a and b. This leads to the equations of motion

$$-\Box A_{\nu} - a\partial_{\nu}\partial_{\mu}A^{\mu} + bA_{\nu} = 0. \tag{9.3}$$

As before, we look for plane wave solutions of the form

$$A_{\nu}(x) = \varepsilon_{\nu}e^{-ik\cdot x} \tag{9.4}$$

for some constant 4-vector ϵ_{ν}. This leads to

$$k^2\varepsilon_{\nu} + ak_{\nu}k\cdot\varepsilon + b\varepsilon_{\nu} = 0. \tag{9.5}$$

The solutions to Eq. (9.5) may be classified in a Lorentz invariant manner into two classes,

1. $\varepsilon \propto k$ (4-D longitudinal)
2. $\varepsilon \cdot k = 0$ (4-D transverse).

In the rest frame of the field, these two types of solution correspond to $\varepsilon = (\varepsilon_0, \vec{0})$ and $\varepsilon = (0, \vec{\varepsilon})$, respectively. The lead to the equations of motion

1. (4-D longitudinal)

$$\begin{aligned} k^2k_{\nu} + ak^2k_{\nu} + bk_{\nu} = 0 \quad &\Rightarrow \quad (k^2(1+a)+b)k_{\nu} = 0 \\ &\Rightarrow \quad k^2 = \frac{-b}{1+a} \equiv \mu_L^2 \end{aligned} \tag{9.6}$$

   This solution has the right dispersion relation for a particle of mass μ_L^2.

2. (4-D transverse)

$$k^2\varepsilon_{\nu} + b\varepsilon_{\nu} = 0 \Rightarrow k^2 = -b \equiv \mu_T^2. \tag{9.7}$$

   This solution describes a field of mass μ_T^2.

The 4-D transverse solution appears to be what we are looking for, since the ε's clearly correspond to the three polarization state of a massive spin one particle. The 4-D longitudinal solution, however, isn't very interesting. This type of solution looks exactly like a scalar field. Since we already know how to quantize scalar field theory, this doesn't lead to anything new. It would be nice to get rid of this solution altogether. This is simple enough to do: if $b \neq 0$ (that is, if the 4-D transverse field is massive), setting $a = -1$ takes μ_L to ∞, removing it from the spectrum. Or, if you prefer, when $a = -1$ and $b \neq 0$, the equation of motion Eq. (9.6) has no solutions[13]. Therefore the longitudinal solutions are absent from the Lagrangian

$$\mathcal{L} = \pm\frac{1}{2}\left[(\partial_\mu A_\nu)^2 - (\partial_\mu A^\mu)^2 - \mu^2 A^2\right] \tag{9.8}$$

where $\mu^2 \equiv \mu_T^2$. This leads to the equations of motion

$$\Box A_\nu - \partial_\nu \partial_\mu A^\mu + \mu^2 A_\nu = 0. \tag{9.9}$$

This can be written in a more compact form by introducing some more notation. Define the *field strength* tensor

$$F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{9.10}$$

In terms of $F^{\mu\nu}$, the Lagrangian is

$$\mathcal{L} = \pm\left[\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}\mu^2 A_\mu A^\mu\right] \tag{9.11}$$

and the equations of motion are

$$\partial_\mu F^{\mu\nu} + \mu^2 A^\nu = 0. \tag{9.12}$$

Equation (9.12) is known as the *Proca Equation*. Using the fact that $F_{\mu\nu}$ is anti-symmetric, $F_{\mu\nu} = -F_{\nu\mu}$, we derive the requirement that the field is transverse

$$\partial_\mu \partial_\nu F^{\mu\nu} = 0 \Rightarrow \partial_\mu A^\mu = 0. \tag{9.13}$$

Substituting this condition into the Proca equation, each component of A^μ is found to satisfy the massive Klein-Gordon equation,

$$(\Box + \mu^2)A_\nu = 0. \tag{9.14}$$

Equations (9.13) and (9.14) are equivalent to the Proca equation, although in this form it is not clear how to derive them from a Lagrangian. They look promising, however. At the level of these two equations the $\mu^2 \to 0$ limit is smooth,

$$\Box A^\mu = 0, \quad \partial_\mu A^\mu = 0. \tag{9.15}$$

[13]When $a = -1$ and $b = 0$, *any* k is a solution to Eq. (9.6). It is this arbitrariness in the solution to the classical theory which makes the massless theory diffi cult to quantize.

These are just Maxwell's equations in free space. Recall that in classical electro-magnetism the scalar and vector potentials ϕ and $\vec{A}$ make up the components of the four-vector $A^\mu = (\phi, \vec{A})$. In the gauge where $\partial_\mu A^\mu = 0$, each component of A^μ satisfies the massless wave equation. The vector field A^μ is thus the familiar vector potential of classical electrodynamics, while the components of the field strength tensor are the electric and magnetic fields

$$\begin{aligned} \vec{E} &= -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t} \\ \vec{B} &= \vec{\nabla} \times \vec{A}. \end{aligned} \tag{9.16}$$

By direct substitution, we easily find

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{9.17}$$

We may also verify directly that the massless Proca equation, $\partial_\mu F^{\mu\nu} = 0$, corre-sponds to the free-space Maxwell Equations

$$\vec{\nabla} \times \vec{B} = \frac{\partial \vec{E}}{\partial t}, \quad \vec{\nabla} \cdot \vec{E} = 0 \tag{9.18}$$

while the remaining two equations,

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \quad \vec{\nabla} \cdot \vec{B} = 0 \tag{9.19}$$

immediately follow from the definitions Eq. (9.16). However, things aren't quite so simple. The condition $\partial_\mu A^\mu = 0$ could only be derived when $\mu^2 \neq 0$. Therefore we will stick with finite μ^2 for a while longer.

Returning to the plane wave solutions to the Proca equation, $A_\mu = \varepsilon_\mu e^{-ik\cdot x}$, there are three linearly independent polarization vectors $\varepsilon_\mu^{(r)}$, $r = 1..3$. In the rest frame, we could choose the basis

$$\varepsilon^{(1)} = (0,1,0,0),\ \varepsilon^{(2)} = (0,0,1,0),\ \varepsilon^{(3)} = (0,0,0,1) \tag{9.20}$$

but in fact it is usually more convenient to choose the basis vectors to be eigenstates of J_z:

$$\varepsilon^{(1)} = \frac{1}{\sqrt{2}}(0,1,i,0),\ \varepsilon^{(2)} = \frac{1}{\sqrt{2}}(0,1,-i,0),\ \varepsilon^{(3)} = (0,0,0,1) \tag{9.21}$$

which have $J_z = +1, -1$ and 0 respectively. In any basis, the basis states are chosen to obey orthonormality

$$\varepsilon_\mu^{(r)} \varepsilon^{\mu(s)*} = -\delta^{rs} \tag{9.22}$$

and completeness

$$\sum_{r=1}^{3} \varepsilon_\mu^{(r)} \varepsilon_\nu^{(r)*} = -g_{\mu\nu} + \frac{k_\mu k_\nu}{\mu^2} \tag{9.23}$$

relations. The minus sign in Eq. (9.22) arises because the polarization vectors are spacelike. The orthonormality and completeness relations are Lorentz covariant, so are true in any frame, not just the rest frame.

The sign of the Lagrangian may be fixed by demanding that the energy be bounded below, as usual. Denoting spatial indices by Roman characters, the Lagrangian may be written as

$$\mathcal{L} = \pm \left[\frac{1}{2} F_{0i} F^{0i} + \frac{1}{4} F_{ij} F^{ij} - \frac{\mu^2}{2} A_i A^i - \frac{\mu^2}{2} A_0 A^0 \right] \tag{9.24}$$

and so the time components of the canonical momenta are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial(\partial_0 A_i)} &= \pm F^{0i} \\ \frac{\partial \mathcal{L}}{\partial(\partial_0 A_0)} &= 0. \end{aligned} \tag{9.25}$$

The fact that the momentum conjugate to A_0 vanishes does not constitute a problem. Because $\partial_\mu A^\mu = 0$, there are fewer degrees of freedom than one would naïvely expect, and the spatial A_i's and their canonical momenta are sufficient to define the state of the system.

The Hamiltonian is

$$\begin{aligned} \mathcal{H} &= \pm F_{0i} \partial^0 A^i - \mathcal{L} \\ &= \pm F_{0i} F^{0i} \pm F_{0i} \partial^i A^0 - \mathcal{L} \\ &= \pm F_{0i} F^{0i} \mp \partial^i F_{0i} A^0 - \mathcal{L} \\ &= \pm F_{0i} F^{0i} \mp \mu^2 A_0 A^0 - \mathcal{L} \\ &= \pm \left[\frac{1}{2} F_{0i} F^{0i} - \frac{1}{4} F_{ij} F^{ij} + \frac{\mu^2}{2} A_i A^i - \frac{\mu^2}{2} A_0 A^0 \right] \end{aligned} \tag{9.26}$$

where we have integrated by parts and used the equation of motion $\partial_\mu F^{\mu 0} = \partial_i F^{i0} = -\mu^2 A^0$. The metric tensor obscures it, but each term in the square brackets is a sum of squares with a negative coefficient and so is negative (for example, $-A_i A^i = A^i A^i > 0$), and so the Lagrangian has an overall minus sign,

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{\mu^2}{2} A_\mu A^\mu. \tag{9.27}$$

9.2 The Quantum Theory

Canonically quantizing the theory is a straightforward generalization of the scalar field theory case, so we will skip some of the steps. Since the spatial components A_i and their conjugate momenta form a complete set of initial conditions, it is only on these fields that we impose the canonical commutation relations

$$\begin{aligned} [A_i(\vec{x}, t), F^{j0}(\vec{y}, t)] &= i\delta_i^j \delta^{(3)}(\vec{x} - \vec{y}) \\ [A_i(\vec{x}, t), A_j(\vec{y}, t)] &= [F^{i0}(\vec{x}, t), F^{j0}(\vec{y}, t)] = 0. \end{aligned} \tag{9.28}$$

Expanding the field in terms of plane wave solutions times operator-valued coefficients ${a_k}^{(r)}$ and $a_k^{\dagger\,(r)}$

$$A_\mu(x) = \sum_{r=1}^{3} \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}} \left[{a_k}^{(r)} \varepsilon_\mu^{(r)}(k) e^{-ik\cdot x} + a_k^{\dagger\,(r)} \varepsilon_\mu^{(r)*}(k) e^{ik\cdot x} \right] \tag{9.29}$$

and substituting this into the canonical commutation relations gives, not unexpectedly, the commutation relations

$$\begin{aligned} [{a_k}^{(r)}, a_{k'}^{\dagger\,(s)}] &= \delta^{rs} \delta^{(3)}(\vec{k} - \vec{k}') \\ [{a_k}^{(r)}, a_{k'}^{(s)}] &= [a_k^{\dagger\,(r)}, a_{k'}^{\dagger\,(s)}] = 0. \end{aligned} \tag{9.30}$$

The Hamiltonian also has the expected form

$$: H := \sum_r \int d^3k \, \omega_k a_k^{\dagger\,(r)} {a_k}^{(r)} \tag{9.31}$$

and so we can interpret $a_k^{\dagger\,(r)}$ and ${a_k}^{(r)}$ as creation and annihilation operators for spin one particles with polarization r.

The propagator $\overbrace{A_\mu(x) A_\nu}(y)$ may be calculated in a similar manner as the spinor case. Proceeding as before, we write

$$\overbrace{A_\mu(x) A_\nu}(y) = \langle 0 \,| T(A_\mu(x) A_\nu(y)) |\, 0 \rangle \tag{9.32}$$

If $x_0 > y_0$,

$$\begin{aligned} \overbrace{A_\mu(x) A}_\nu(y) &= \langle 0 \,| A_\mu^{(+)}(x) A_\nu^{(-)}(y) |\, 0 \rangle \\ &= \langle 0 \,| [A_\mu^{(+)}(x), A_\nu^{(-)}(y)] |\, 0 \rangle \\ &= [A_\mu^{(+)}(x), A_\nu^{(-)}(y)] \qquad (9.33) \end{aligned}$$

where we have split A_μ into the piece containing the creation operator, $A_\mu^{(-)}$ and a piece $A_\mu^{(+)}$ containing the annihilation operator. From the expansion of $A_\mu(x)$, it is straightforward to show that

$$\begin{aligned} [A_\mu^{(+)}(x), A_\nu^{(-)}(y)] &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{-ik\cdot(x-y)} \sum_r \varepsilon_\mu^{(r)}(k) \varepsilon_\nu^{(r)*}(k) \\ &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{-ik\cdot(x-y)} \left(-g_{\mu\nu} + \frac{k_\mu k_\nu}{\mu^2} \right) \\ &= \left(-g_{\mu\nu} - \frac{\partial_\mu^y \partial_\nu^y}{\mu^2} \right) i\Delta_+(x-y) \qquad (9.34) \end{aligned}$$

where

$$i\Delta_+(x-y) = \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{-ik\cdot(x-y)} \qquad (9.35)$$

and $\partial_\mu^y \equiv \partial/\partial y^\mu$. After including the $y_0 > x_0$ term we obtain

$$\begin{aligned} \overbrace{A_\mu(x) A}_\nu(y) &= \theta(x_0 - y_0) \left(-g_{\mu\nu} - \frac{\partial_\mu^y \partial_\nu^y}{\mu^2} \right) i\Delta_+(x-y) \\ &\quad + \theta(y_0 - x_0) \left(-g_{\mu\nu} - \frac{\partial_\mu^y \partial_\nu^y}{\mu^2} \right) i\Delta_+(y-x). \qquad (9.36) \end{aligned}$$

Now, the scalar propagator is

$$\begin{aligned} \overbrace{\phi(x)\phi}(y) &= \theta(x_0 - y_0) i\Delta_+(x-y) + \theta(y_0 - x_0) i\Delta_+(y-x) \\ &= \int \frac{d^4k}{(2\pi)^4} e^{-ik\cdot(x-y)} \frac{i}{k^2 - \mu^2 + i\epsilon} \qquad (9.37) \end{aligned}$$

and so we would like to commute the θ functions and derivatives in Eq. (9.36) to obtain

$$\begin{aligned} \overbrace{A_\mu(x) A}_\nu(y) &= \left(-g_{\mu\nu} - \frac{\partial_\mu^y \partial_\nu^y}{\mu^2} \right) \left(\theta(x_0 - y_0) i\Delta_+(x-y) + \theta(y_0 - x_0) i\Delta_+(y-x) \right) \\ &= \int \frac{d^4k}{(2\pi)^4} \left(-g_{\mu\nu} + \frac{k_\mu k_\nu}{\mu^2} \right) e^{-ik\cdot(x-y)} \frac{i}{k^2 - \mu^2 + i\epsilon}. \qquad (9.38) \end{aligned}$$

This leads to the propagator for a massive vector field, which is represented by a wavy line:

$$\frac{-i\left(g^{\mu\nu} - \frac{k^\mu k^\nu}{\mu^2} \right)}{k^2 - \mu^2 + i\epsilon}. \qquad (9.39)$$

Note that the vector propagator carries Lorentz indices: one end of the line corre-

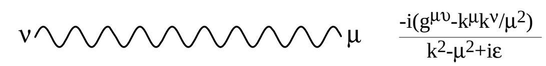

Figure 9.1: The propagator for a massive vector field.

sponds to a field created by A_μ while the other corresponds to the field created by A_ν.

While this is correct, the derivation was not quite right when $\mu = \nu = 0$. In this case, the time derivatives don't commute with the θ functions and there is the additional term when one of the derivatives acts on the θ function, giving a factor of $\delta(x_0 - y_0)$, and the other acts on the Δ_+ function. This wasn't a problem in the spinor case because there was only a single time derivative, and the term vanished because $\Delta(x-y) = 0$ when $x_0 = y_0$. In this case, however, the time derivative of $\Delta(x-y)$ does *not* vanish when $x_0 = y_0$ and so there is an additional term. The fact that this term does not contribute is not obvious in the canonical quantization procedure. The path integral formulation of quantum field theory, which we will not discuss in this course, puts this derivation on sounder footing. If you like, you can use the derivation above for $(\mu,\nu) \neq (0,0)$, and then argue that by Lorentz invariance the result must have this form for $(\mu,\nu) = (0,0)$ as well.[14]

Now consider adding a fermion such as an electron to the theory. A simple interaction term between the fermi field ψ and A_μ is

$$\mathcal{L}_I = -g\overline{\psi}\gamma^\mu \Gamma \psi A_\mu = -g\overline{\psi} \not{A} \Gamma \psi \qquad (9.40)$$

where Γ has the general form $\Gamma = a + b\gamma_5$ by Lorentz Invariance. As we discussed before, when both a and b are nonzero this theory violates parity, since there is no choice of transformation for A^μ under which the interaction term Eq. (9.40) is invariant. A parity conserving theory may have either $\Gamma = 1$ (vector coupling) or $\Gamma = \gamma_5$ (axial vector coupling), in which case the components of A^μ transform under parity as a vector or an axial vector, respectively.

[14]This sort of diffi culty arises in the canonical quantization procedure because it breaks manifest Lorentz invariance, by treating temporal indices different from spatial indices. The path integral formulation treats space and time in a symmetric fashion.

From our previous experience with interacting theories, the interaction term Eq. (9.40) leads to the interaction vertex shown in Fig. (9.2). We note at this stage

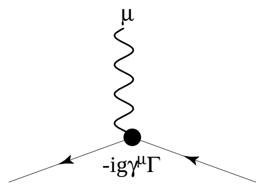

Figure 9.2: Fermion-vector interaction vertex

that there is a simple rule for writing down the Feynman rule associated with an interaction term in $\mathcal{L}$. When all the fields in the interaction term are different, the resulting Feynman rule is just $-i$ times the interaction Hamiltonian, or i times the interaction Lagrangian. A term with n identical fields has a combinatoric factor of $n!$ in the Feynman rule, corresponding to the $n!$ different way of choosing which field corresponds to which line in the vertex.

Finally, evaluating Dyson's formula requires matrix elements of the A_μ field between incoming and outgoing vector meson states and the vacuum. From the field expansion,

$$\begin{aligned}\langle 0\,|A_\mu(x)|\,V(k,r)\rangle &= \sum_{r'=1}^{3}\int \frac{d^3k'}{(2\pi)^{3/2}\sqrt{2\omega_{k'}}}\varepsilon_\mu^{(r')}(k')e^{-ik'\cdot x}\langle 0\,|a_{k'}^{(r')}\sqrt{2\omega_k}(2\pi)^{3/2}a_k^{\dagger\,(r)}|\,0\rangle \\ &= \sum_{r'=1}^{3}\int d^3k'\,\sqrt{\frac{\omega_k}{\omega_{k'}}}\,\varepsilon_\mu^{(r')}(k')e^{-ik'\cdot x}\langle 0\,|a_{k'}^{(r')}a_k^{\dagger\,(r)}|\,0\rangle \\ &= \sum_{r'=1}^{3}\int d^3k'\,\sqrt{\frac{\omega_k}{\omega_{k'}}}\,\varepsilon_\mu^{(r')}(k')e^{-ik'\cdot x}\langle 0\,|[a_{k'}^{(r)},a_k^{\dagger\,(r)}]|\,0\rangle \\ &= \varepsilon_\mu^{(r)}(k)e^{-ik\cdot x} \end{aligned} \tag{9.41}$$

where $|\,V(k,r)\rangle$ is a relativistically normalized single particle state containing a vector meson with momentum k and polarization r. Therefore, each incoming vector meson contributes a factor of $\varepsilon_\mu^{(r)}$ to the amplitude in addition to the usual exponential factor. Equation (9.41) and its complex conjugate lead to the Feynman rule

- For every $\left\{\begin{matrix}\text{incoming}\\ \text{outgoing}\end{matrix}\right\}$ vector meson with momentum k and polarization r, include a factor of $\left\{\begin{matrix}\varepsilon_\mu^{(r)}(k)\\ \varepsilon_\mu^{(r)*}(k)\end{matrix}\right\}$.

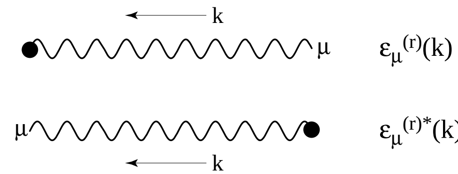

Figure 9.3: Feynman rules for external vector particles.

9.3 The Massless Theory

To obtain a quantum theory of electromagnetism, the limit $\mu \to 0$ must be taken of the results in the previous section. This limit looks bad for several reasons. In the quantum theory, there is a factor of $k^\mu k^\nu/\mu^2$ in the vector propagator. This will turn out to be closely related to a problem which arises at the classical level. Consider $\mathcal{L}$ coupled to a source $\mathcal{J}^\mu(x)$

$$\mathcal{L} = \mathcal{L}_0 - A_\mu \mathcal{J}^\mu(x). \tag{9.42}$$

The equations of motion in this theory are

$$\partial_\mu F^{\mu\nu} + \mu^2 A^\nu = \mathcal{J}^\nu \tag{9.43}$$

which leads to

$$\partial_\mu A^\mu = \frac{1}{\mu^2}\partial_\mu \mathcal{J}^\mu. \tag{9.44}$$

Again, this looks bad in the limit $\mu \to 0$. However, it gives a clue to how to obtain a theory with a sensible $\mu \to 0$ limit: the limit exists only if $\mathcal{A}_\mu$ couples to a *conserved current*. In this case, $\partial_\mu \mathcal{J}^\mu = 0$ and the $\mu \to 0$ limit of Eq. (9.44) is well defined.

Fortunately, we're old hands at finding conserved currents. Recall that Noether's theorem ensures that any internal symmetry has an associated conserved current and charge, the simplest example being a $U(1)$ symmetry associated with the transformation

$$\phi_a(x) \to e^{-i\lambda q_a}\phi_a(x) \tag{9.45}$$

for some set of fields (not necessarily scalar fields) $\{\phi_a\}$. There is no implied sum over a in Eq. (9.45), and the q_a's are numbers (the charge of each field), not operators. Note that the q_a's are arbitrary up to a multiplicative constant; that is, if $\phi_a \to \exp(-iq_a\lambda)\phi_a$ is a symmetry, so is (for example) $\phi_a \to \exp(-2iq_a\lambda)\phi_a$. There is no physics in this ambiguity - if j^μ is a conserved current, so is any multiple of j^μ.

If Eq. (9.45) is a symmetry, $DL = 0$ and the current

$$j^\mu = \sum_a \Pi^\mu_a D\phi_a = -i\sum_a \Pi^\mu_a q_a\phi_a \tag{9.46}$$

is conserved. For example, the Dirac Lagrangian is invariant under the transformation

$$\psi \to e^{-i\lambda}\psi,\ \ \overline{\psi} \to e^{i\lambda}\overline{\psi}. \tag{9.47}$$

Therefore the corresponding q_a's are

$$q_\psi = 1,\ \ q_{\overline{\psi}} = -1 \tag{9.48}$$

and $D\psi = -i\psi$, $D\overline{\psi} = i\overline{\psi}$. The conjugate momenta are

$$\Pi^\mu_\psi = i\overline{\psi}\gamma^\mu,\ \ \Pi^\mu_{\overline{\psi}} = 0 \tag{9.49}$$

and so the conserved current is

$$j^\mu = \overline{\psi}\gamma^\mu\psi. \tag{9.50}$$

For a charged *scalar* field[15] φ we have

$$\Pi^\mu_\varphi = \partial^\mu\varphi^*,\ \ \Pi^\mu_{\varphi^*} = \partial^\mu\varphi \tag{9.51}$$

and so the conserved current is

$$j^\mu = -i(\partial^\mu\varphi^*)\varphi + i(\partial^\mu\varphi)\varphi^*. \tag{9.52}$$

Therefore, we might hope that if we couple a vector field A^μ only to these conserved currents we will obtain a theory with a well-defined $\mu \to 0$ limit. So we might try the following interaction terms:

[15]To avoid confusion with fermi fields ψ, we switch our notation for charged scalars at this point from ψ to φ.

- Fermions:

$$\mathcal{L}_I = -g\overline{\psi}\gamma^\mu A_\mu\psi = -g\overline{\psi}\not{A}\psi \tag{9.53}$$

This is the interaction we had written down earlier, but with $\Gamma = 1$. For massive fermions, only the vector current $\overline{\psi}\gamma^\mu\psi$ is conserved; the axial vector current $\overline{\psi}\gamma^\mu\gamma_5\psi$ isn't associated with an internal symmetry and is not conserved.[16] Therefore we expect that only the theory where the vector field couples to the vector current will have a smooth $\mu \to 0$ limit.

- Charged scalars:

$$\mathcal{L}_I = -ig\left[(\partial^\mu\varphi^*)\varphi - (\partial^\mu\varphi)\varphi^*\right]A_\mu \tag{9.54}$$

The situation here is not as nice as it was for fermions. The interaction term contains derivatives of the fields, so it changes the canonical momenta of the theory, thereby changing the expression for j^μ. So although this theory still has a $U(1)$ symmetry and a conserved current, the conserved current is no longer given by Eq. (9.51), and therefore this theory is not expected to have a smooth $\mu \to 0$ limit.

We see from the scalar case that it's not always so easy to ensure that A_μ always couples to a conserved current, because the coupling itself will in general change the expression for the current. Fortunately, there is a magic prescription which guarantees that A_μ always couples to a conserved current. It is called *minimal coupling*.

9.3.1 Minimal Coupling

The minimal coupling prescription is very simple. Given a Lagrangian as a function of the fields and their derivatives, $\mathcal{L}_M(\phi_a, \partial_\mu\phi_a)$, which is invariant under the $U(1)$ transformation $\phi_a \to e^{-i\lambda q_a}\phi_a$, replace it by $\mathcal{L}_M(\phi_a, D_\mu\phi_a)$, where

$$D^\mu\phi_a \equiv \partial^\mu\phi_a + ieA^\mu q_a\phi_a \tag{9.55}$$

(no sum on a). D^μ is called the *gauge covariant derivative*. (Note that again there is an ambiguity in the q_a's; this just corresponds to the freedom to choose the overall coupling constant for the interaction term. For quantum electrodynamics,

[16]For massless fermions, we saw in the chapter on the Dirac Lagrangian that the theory has *two* $U(1)$ symmetries and therefore two conserved currents, $j^\mu_{L,R} = \overline{\psi}_{L,R}\gamma^\mu\psi_{L,R} = \frac{1}{2}\overline{\psi}\gamma^\mu(1 \mp \gamma_5)\psi$. Since any linear combinations of these currents are conserved, both $\overline{\psi}\gamma^\mu\psi$ and $\overline{\psi}\gamma^\mu\gamma_5\psi$ are conserved. Thus, it is possible to couple a massless vector field to the axial vector current in the special case of massless fermions. The mass term breaks the axial $U(1)$ symmetry associated with $\overline{\psi}\gamma^\mu\gamma_5\psi$ but not the vector $U(1)$.

if we choose the dimensionless coupling constant e to be the fundamental electric charge, then q will be the electric charge of the field measured in units of e.) The resulting Lagrangian has the following two properties:

1. $\mathcal{L}_M$ is still invariant under the $U(1)$ transformation, and

2. A^μ is coupled to a conserved current. That is

$$\frac{\partial \mathcal{L}_I}{\partial A_\mu} = -ej^\mu \tag{9.56}$$

and

$$\partial_\mu j^\mu = 0. \tag{9.57}$$

This is straightforward to show. Under a $U(1)$ transformation,

$$\begin{aligned} D_\mu \phi_a &\rightarrow D_\mu \left(e^{-i\lambda q_a}\phi_a\right) \\ &= \partial_\mu \left(e^{-i\lambda q_a}\phi_a\right) + ieA_\mu q_a \left(e^{-i\lambda q_a}\phi_a\right) \\ &= e^{-i\lambda q_a} D_\mu \phi_a \end{aligned} \tag{9.58}$$

and so $D_\mu\phi_a$ transforms in the same way as $\partial_\mu\phi_a$. Therefore if $\mathcal{L}(\phi_a, \partial_\mu\phi_a)$ is invariant under the $U(1)$ symmetry, so is $\mathcal{L}(\phi_a, D_\mu\phi_a)$. This proves the first assertion.

In terms of the canonical momenta, the conserved current is

$$j^\mu = \sum_a \Pi_a^\mu D\phi_a = \sum_a \Pi_a^\mu(-iq_a\phi_a). \tag{9.59}$$

From the definition of the gauge covariant derivative, we also have

$$\frac{\partial(D_\nu\phi_a)}{\partial A_\mu} = ieq_a\phi_a\delta_\nu^\mu \tag{9.60}$$

and so we find

$$\begin{aligned} \frac{\partial \mathcal{L}_I}{\partial A_\mu} = \frac{\partial \mathcal{L}_M}{\partial A_\mu} &= \sum_a \frac{\partial \mathcal{L}_M}{\partial(D_\nu\phi_a)}\frac{\partial(D_\nu\phi_a)}{\partial A_\mu} \\ &= \sum_a \frac{\partial \mathcal{L}_M}{\partial(\partial_\nu\phi_a)} ieq_a\phi_a \\ &= \sum_a \Pi_a^\mu ieq_a\phi_a \\ &= -ej^\mu \end{aligned} \tag{9.61}$$

as required, proving the second assertion.

Going back to our examples, the minimally coupled Dirac Lagrangian for a fermion with charge q (in units of the elementary charge e) is

$$\mathcal{L} = \overline{\psi}(i\not{D} - m)\psi = \overline{\psi}(i\not{\partial} - eq\not{A} - m)\psi \tag{9.62}$$

which is just what we had before. However, the minimally coupled scalar Lagrangian for a scalar with charge q is

$$\begin{aligned} \mathcal{L} &= D_\mu\varphi^* D^\mu\varphi - m^2 A^\mu A_\mu \\ &= (\partial_\mu - ieqA_\mu)\varphi^*(\partial^\mu + ieqA^\mu)\varphi - m^2 A^\mu A_\mu \\ &= \partial_\mu\varphi^*\partial^\mu\varphi - ieqA_\mu(\varphi^*\partial^\mu\varphi - \varphi\partial^\mu\varphi^*) + e^2q^2A_\mu A^\mu\varphi^*\varphi. \end{aligned} \tag{9.63}$$

The term linear in A_μ is what we had before, but there is a new term quadratic in A_μ. This will lead to a new kind of vertex, with the Feynman rule in Fig. (9.4), the so-called "seagull graph" (to avoid confusion with fermion lines, we will denote charged scalars by dashed lines in this chapter). Only the theory defined by

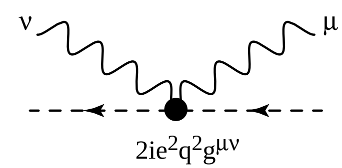

Figure 9.4: The "seagull graph" for charged scalar-photon interactions.

Lagrangian Eq. (9.63) with all the interaction terms given by minimal coupling has a well-defined limit as $\mu \rightarrow 0$.

The Feynman rule for the term in Eq. (9.63) linear in A_μ is slightly subtle, but it turns out that the naïve approach gives the correct answer. Naïvely, we notice that a derivative ∂_μ acting on the piece of the field which annihilates an incoming state (and so has a factor of $\exp(-ip\cdot x)$) brings down a factor of $-ip_\mu$. Similarly, when acting on the piece of the field which creates an outgoing state, it brings down a factor of ip_μ. Therefore, we expect the Feynman rule shown in Fig. (9.5). There are two problems with this derivation. First of all, the derivative interaction changes the canonical momenta in the theory, and so changes the canonical commutation relations. Second, in Dyson's formula the derivative cannot be pulled out of the time ordered product. However, it turns out (we won't prove this here) that these two problems cancel one another, and that the naïve Feynman rule is actually correct.

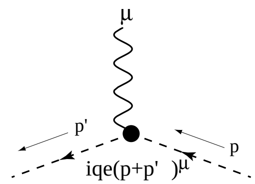

Figure 9.5: Feynman rule for the derivately coupled charged scalar.

We now justify the assertion that the coupling constant e arising in the covariant derivative is just the elementary charge. We will show this for the Dirac field. Recall that in the scalar case the $U(1)$ charge was just proportional to the number of particles minus the number of antiparticles, indicating that particle and antiparticle carried opposite charges. The same is true for Dirac fields: the conserved charge is

$$\begin{aligned} Q &= \int d^3x\, \overline{\psi}\gamma^0\psi \\ &= \int d^3x\, \psi^\dagger\psi. \end{aligned} \tag{9.64}$$

Substituting the field expansion in terms of creation and annihilation operators, it is straightforward to show that this is

$$\begin{aligned} Q &= \sum_r \int d^3k\, \left(b^{(r)\dagger}_{\vec{k}} b^{(r)}_{\vec{k}} - c^{(r)\dagger}_{\vec{k}} c^{(r)}_{\vec{k}}\right) \\ &= N_b - N_c. \end{aligned} \tag{9.65}$$

If the particles annihilated by ψ have electric charge q in units of the elementary charge e, the total electric charge of the system is therefore $Q_{\text{e.m.}} = qeQ$, and the electromagnetic four-current is therefore $j^\mu_{\text{e.m.}} = qe\overline{\psi}\gamma^\mu\psi$. Thus, for electrons ($q = -1$), the electric charge and electromagnetic current are

$$\begin{aligned} Q_{\text{e.m.}} &= -e\int d^3x\, \psi^\dagger(x)\psi(x) \\ j^\mu_{\text{e.m.}} &= -e\overline{\psi}\gamma^\mu\psi. \end{aligned} \tag{9.66}$$

The Euler-Lagrange equation for a massless vector field A^μ is therefore

$$\begin{aligned} \partial_\mu F^{\mu\nu} &= -e\overline{\psi}\gamma^\nu\psi \\ &= j^\nu_{e.m} \end{aligned} \tag{9.67}$$

which is just Maxwell's equations in the presence of an electromagnetic current $j^\nu_{e.m.}$.

The elementary charge is often expressed in terms of the "fine-structure constant" α, where

$$\alpha = \frac{e^2}{4\pi\hbar c} = \frac{1}{137.035} \tag{9.68}$$

and so in natural units

$$e = \sqrt{4\pi\alpha}. \tag{9.69}$$

9.3.2 Gauge Transformations

The minimally coupled Lagrangian $\mathcal{L}_M(\phi_a, D_\mu\phi_a)$ is invariant under a much larger group of symmetries than $\mathcal{L}_M(\phi_a, \partial_\mu\phi_a)$. It is invariant under the strange-looking transformation

$$\begin{aligned} \lambda(x) : \phi_a(x) &\to e^{-iq_a\lambda(x)}\phi_a(x) \\ A_\mu(x) &\to A_\mu(x) + \frac{1}{e}\partial_\mu\lambda(x) \end{aligned} \tag{9.70}$$

for any space-time dependent function $\lambda(x)$. Note that when $\lambda(x)$ is a constant and not a function of space-time this is just the usual $U(1)$ transformation on the ϕ_a's (note that the A^μ field is invariant if λ is constant). This kind of symmetry is called a *global* $U(1)$ symmetry, since λ is the same at all points.

The transformation Eq. (9.70) is called a *local* or *gauge* transformation, and $\mathcal{L}_M$ is said to have a $U(1)$ gauge symmetry. Since $\lambda(x)$ is now a function of space-time, the theory is invariant under *different* $U(1)$ transformations at each point in space-time. The odd transformation law of the A^μ fields is crucial here: the Dirac Lagrangian is not invariant under gauge transformations, since $\overline{\psi}\partial\!\!\!/\psi$ picks up a term proportional to $(\partial_\mu\lambda)\,\overline{\psi}\gamma^\mu\psi$. The transformation property of A^μ is chosen precisely to cancel this term:

$$\begin{aligned} D_\mu\phi_a &= (\partial_\mu + ieA_\mu q_a)\phi_a \\ &\to (\partial_\mu + ieA_\mu q_a + iq_a\partial_\mu\lambda(x))\left(e^{-iq_a\lambda(x)}\phi_a\right) \\ &= e^{-iq_a\lambda(x))}(\partial_\mu - iq_a\partial_\mu\lambda(x) + ieA_\mu q_a + iq_a\partial_\mu\lambda(x))\phi_a \\ &= e^{-iq_a\lambda(x)}D_\mu\phi_a. \end{aligned} \tag{9.71}$$

Therefore unlike the usual derivative $\partial_\mu\phi_a$, the gauge covariant derivative $D_\mu\phi_a$ transforms in the same way under a gauge transformation as it does under a global transformation. Thus, if $\mathcal{L}_M(\phi_a, \partial_\mu\phi_a)$ is invariant under a global $U(1)$ transformation, $\mathcal{L}_M(\phi_a, D_\mu\phi_a)$ is invariant under a gauged $U(1)$ transformation. Therefore, every time we use the minimal coupling prescription we end up with a theory in which $\mathcal{L}_M$ is invariant under a gauge symmetry.

So far we have just looked at $\mathcal{L}_M$, the "matter" (fermions and scalars) Lagrangian, and ignored the free part of the vector Lagrangian, $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}+\frac{\mu^2}{2}A_\mu A^\mu$. Since $F_{\mu\nu}$ is antisymmetric in its indices, it is also gauge invariant,

$$\lambda(x) : F_{\mu\nu} \to F_{\mu\nu} + \frac{1}{e}\left(\partial_\mu\partial_\nu\lambda(x) - \partial_\nu\partial_\mu\lambda(x)\right) = F_{\mu\nu}. \tag{9.72}$$

However, the vector meson mass term $\frac{\mu^2}{2}A_\mu A^\mu$ is *not* gauge invariant:

$$\lambda(x) : A_\mu A^\mu \to A_\mu A^\mu + \frac{2}{e}\partial_\mu\lambda(x)A^\mu + \frac{1}{e}\partial_\mu\lambda(x)\partial^\mu\lambda(x). \tag{9.73}$$

The complete Lagrangian is only gauge invariant when $\mu = 0$.

In quantum electrodynamics, which is the vector theory we are really interested in, the photon is massless and so the theory has exact gauge invariance. Rather than being a help in solving the theory, this gauge invariance complicates things tremendously, making it difficult to quantize the massless theory directly. The problem arises at the classical level: if $\{A_\mu(x), \phi_a(x)\}$ is a set of fields which form a solution to the equations of motion then so is the set

$$\left\{A_\mu(x) + \frac{1}{e}\partial_\mu\lambda(x), e^{-i\lambda(x)q_a}\phi_a(x)\right\} \tag{9.74}$$

for some arbitrary function $\lambda(x)$. Therefore the problem of finding the time-evolution of the fields from some initial values is ill-defined. No matter how much initial value data I have at $t = 0$ (the fields, their first, second, third ... derivatives), I can never uniquely predict the field configuration at some later time, since their exist an infinite number of gauge transformed solutions of the equations of motion which also have the same initial value data. These field configurations just differ by a gauge transformation which vanishes at $t = 0$.

Furthermore, in the massless theory the condition $\partial_\mu A^\mu = 0$ implied by the Proca equation is no longer implied by the equations of motion. If $A_\mu(x)$ is a solution to the equations of motion satisfying $\partial_\mu A^\mu = 0$, then another solution to the equations of motion is $A'_\mu(x) = A_\mu(x) + \partial_\mu\lambda(x)/e$, which satifies

$$\partial_\mu A^{\mu\prime}(x) = \frac{1}{e}\Box\lambda(x) \neq 0. \tag{9.75}$$

The four dimensionally longitudinal mode which we had banished has come back to haunt us. $\partial_\mu A^\mu$ is no longer zero, but arbitrary.

Things are not actually so badly defined. $F^{\mu\nu}$ is gauge invariant, and therefore so are the electric and magnetic fields $\vec{E}$ and $\vec{B}$. In fact, any observable is gauge invariant. Two systems different by a gauge transformation contain identical physics; they just differ in the choice of description. So we can fix the description by fixing the gauge once and for all. Some popular gauge choices are

$\vec{\nabla} \cdot \vec{A} = 0$ (Coulomb gauge)

$\partial_\mu A^\mu = 0$ (Lorentz gauge)

$A_0 = 0$ (temporal gauge)

$A_3 = 0$ (axial gauge)

The trick is then to canonically quantize the theory in the given gauge, that is, subject to the corresponding constraint[17]. In perturbation theory, different choices of gauge result in extra terms in the photon propagator proportional to $k^\mu k^\nu$. However, as we will see in the next section, these terms in the propagator do not contribute in a minimally coupled theory and so, as expected, physical amplitudes are independent of gauge.

9.4 The Limit $\mu \to 0$

Since we are avoiding quantizing the gauge invariant massless theory directly, we will instead derive the Feynman rules for Quantum Electrodynamics by examining at the $\mu \to 0$ of the theory of a minimally coupled massive vector field. With any luck the minimal coupling prescription will have solved the problems we previously noted in taking this limit. Indeed, in this section we will see by direct calculation that the factors of $1/\mu^2$ in the quantum theory do not contribute to amplitudes when the theory is minimally coupled, and that the massless theory does in fact have sensible Feynman rules.

First consider the process $e^+e^- \to \mu^+\mu^-$, where e and μ are two different fermion fields (electrons and muons), minimally coupled to a massive gauge boson. In the limit $\mu \to 0$ this is just the pair production process $e^+e^- \to \mu^+\mu^-$ in QED. There is only one graph at $\mathcal{O}(g^2)$ which contributes to this process, shown in Fig. (9.6). The $1/\mu^2$ term in the amplitude is

$$\begin{aligned} & i\frac{e^2}{\mu^2}\overline{v}^{(r)}_{\vec{p}_+}\gamma^\mu u^{(s)}_{\vec{p}_-}\frac{k_\mu k_\nu}{k^2-\mu^2+i\epsilon}\overline{u}^{(r')}_{\vec{p}'_-}\gamma^\nu v^{(s')}_{\vec{p}'_+} \\ = \; & i\frac{e^2}{\mu^2(k^2-\mu^2+i\epsilon)}\overline{v}^{(r)}_{\vec{p}_+}\not{k}u^{(s)}_{\vec{p}_-}\overline{u}^{(r')}_{\vec{p}'_-}\not{k}v^{(s')}_{\vec{p}'_+}. \end{aligned} \tag{9.76}$$

[17] See Chapter 5 of Mandl & Shaw for a discussion of the Gupta-Bleurer method of canonically quantizing the massless theory.

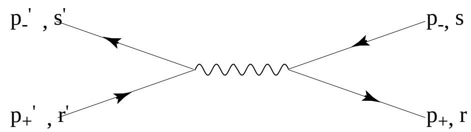

Figure 9.6: Feynman diagram contributing to $e^+e^- \to \mu^+\mu^-$.

But by the Dirac equation,

$$\overline{v}^{(r)}_{\vec{p}_+} \not{k} u^{(s)}_{\vec{p}_-} = \overline{v}^{(r)}_{\vec{p}_+} (\not{p}_- + \not{p}_+) u^{(s)}_{\vec{p}_-} = \overline{v}^{(r)}_{\vec{p}_+} (m_e - m_e) u^{(s)}_{\vec{p}_-} = 0. \tag{9.77}$$

So this term vanishes before taking μ to zero. Of course, this is no accident (there are no accidents in field theory). It just follows from current conservation:

$$\begin{aligned} \partial_\mu j^\mu = 0 \Rightarrow \langle 0 \,|\partial_\mu j^\mu|\, e^+e^- \rangle &= \langle 0 \,|\partial_\mu \overline{e}\gamma^\mu e|\, e^+e^- \rangle \\ &= \partial_\mu (\overline{v}^{(r)}_{\vec{p}_+} \gamma^\mu u^{(s)}_{\vec{p}_-} e^{-i(p_+ + p_-)\cdot x}) \\ &= -i(p_+ + p_-) \overline{v}^{(r)}_{\vec{p}_+} \gamma^\mu u^{(s)}_{\vec{p}_-} e^{-i(p_+ + p_-)\cdot x} \\ &= -i \overline{v}^{(r)}_{\vec{p}_+} \not{k} u^{(s)}_{\vec{p}_-} e^{-i(p_+ + p_-)\cdot x} \end{aligned} \tag{9.78}$$

and so $\overline{v}\not{k}u = 0$. Although we have just demonstrated it in one process, this is a very general feature of minimal coupling, and it means that in such theories we can completely ignore the piece of the propagator proportional to $k^\mu k^\nu$. Therefore, in the $\mu \to 0$ limit the vector boson is the photon of quantum electrodynamics, with the propagator shown in Fig. (9.7).

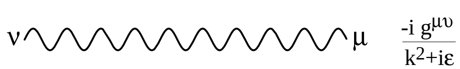

Figure 9.7: The photon propagator.

In a similar vein, you might worry about the factor of $1/\mu^2$ in the polarization sum, Eq. (9.23), but a similar thing happens here. We can demonstrate this by looking at Compton scattering of a massive vector boson off an electron, $Ve^- \to Ve^-$. Two diagrams contribute to this process, giving

$$\begin{aligned} i\mathcal{A} &= -ie^2 \overline{u}^{(s')}_{\vec{p}'} \left[\frac{\gamma^\mu(\not{p}' + \not{k}' + m)\gamma^\nu}{(p'+k')^2 - m^2} + \frac{\gamma^\nu(\not{p}' - \not{k} + m)\gamma^\mu}{(p'-k)^2 - m^2} \right] u^{(s)}_{\vec{p}} \varepsilon^{(r')^*}_\mu (k') \varepsilon^{(r)}_\nu (k) \\ &\equiv \mathcal{M}^{\mu\nu} \varepsilon^{(r')^*}_\mu (k') \varepsilon^{(r)}_\nu (k). \end{aligned} \tag{9.79}$$

Squaring and summing over final spins of the vector particles and averaging over initial spins will give a result proportional to

$$\mathcal{M}^{\mu\nu}\mathcal{M}^{\alpha\beta} \left[-g_{\mu\alpha} + \frac{k'_\mu k'_\alpha}{\mu^2} \right] \left[-g_{\nu\beta} + \frac{k_\nu k_\beta}{\mu^2} \right] \tag{9.80}$$

and so the terms proportional to $k'_\mu M^{\mu\nu}$ and $k_\nu M^{\mu\nu}$ look bad as $\mu \to 0$. However, just as before, the contributions from these terms vanish:

$$\begin{aligned} k'_\mu \mathcal{M}^{\mu\nu} &\propto \overline{u}^{(s')}_{\vec{p}'} \left[\frac{\not{k}'(\not{p}' + \not{k}' + m)\gamma^\nu}{2p'\cdot k'} + \frac{\gamma^\nu(\not{p}' - \not{k} + m)\not{k}'}{-2p\cdot k'} \right] u^{(s)}_{\vec{p}} \\ &= \overline{u}^{(s')}_{\vec{p}'} \left[\frac{(-\not{p}'\not{k}' + m\not{k}' + 2p'\cdot k')\gamma^\nu}{2p'\cdot k'} - \frac{\gamma^\nu(m\not{k}' - \not{k}'\not{p} + 2p\cdot k')}{2p\cdot k'} \right] u^{(s)}_{\vec{p}} \\ &= \overline{u}^{(s')}_{\vec{p}'} \left[\frac{(-m\not{k}' + m\not{k}' + 2p'\cdot k')\gamma^\nu}{2p'\cdot k'} - \frac{\gamma^\nu(m\not{k}' - \not{k}'m + 2p\cdot k')}{2p\cdot k'} \right] u^{(s)}_{\vec{p}} \\ &= \overline{u}^{(s')}_{\vec{p}'} \left[\gamma^\nu - \gamma^\nu \right] u^{(s)}_{\vec{p}} = 0. \end{aligned} \tag{9.81}$$

Similarly, $k_\nu \mathcal{M}^{\mu\nu} = 0$, and so the $k^\mu k^\nu$ term doesn't contribute to the polarization sum.

9.4.1 Decoupling of the Helicity 0 Mode

The result that $k'_\mu \mathcal{M}^{\mu\nu} = 0$ has another consequence in the $\mu \to 0$ limit. A massive spin 1 particle has three spin states, $J_z = \pm 1, 0$, whereas a massless gauge boson like the photon only has two helicity states, ± 1 (once again, this is only possible because the photon is massless. For a massive particle you can always boost to its rest frame and perform a rotation to change a $J_z = 1$ state to a $J_z = 0$ state.) This corresponds to the fact that classical electromagnetic waves are always transverse. The absence of a longitudinal mode corresponds to the absence of a (3-dimensionally) longitudinal photon, $\vec{\varepsilon} \propto \vec{k}$. (We call mode satisfying $\vec{\varepsilon} \propto \vec{k}$ "3D longitudinal" to distinguish it from the 4D longitudinal mode discussed earlier, $\epsilon \propto k$.) How is this apparently discontinuous behaviour possible if the $\mu \to 0$ limit of the theory is smooth?

A massive vector particle travelling in the z direction has four-momentum $k^\mu = (\sqrt{k^2 + \mu^2}, 0, 0, k)$ and three possible polarization states $\varepsilon^{(r)}_\mu$, where

$$\begin{aligned} \varepsilon^{(1)} &= (0, 1, i, 0) \\ \varepsilon^{(2)} &= (0, 1, -i, 0) \\ \varepsilon^{(3)} &= \frac{1}{\mu}(k, 0, 0, \sqrt{k^2 + \mu^2}) \end{aligned} \tag{9.82}$$

$\varepsilon^{(3)}$ is the 3D longitudinal polarization state, satisfying $\varepsilon^{(3)}\cdot\varepsilon^{(3)}=-1$, $\varepsilon^{(3)}\cdot k=0$, $\vec{\varepsilon}^{(3)}\propto\vec{k}$. The amplitude for a longitudinal vector boson to be produced in any process (like the Compton scattering process from the previous section) is proportional to

$$\varepsilon_\mu^{(3)*}\mathcal{M}^\mu \tag{9.83}$$

where the tensor $\mathcal{M}^\mu$ satisfies

$$k^\mu\mathcal{M}_\mu=0\Rightarrow\mathcal{M}_0=-\mathcal{M}_3\frac{k}{\sqrt{k^2+\mu^2}}=-\mathcal{M}_3\left(1+\mathcal{O}\left(\frac{\mu^2}{k^2}\right)\right). \tag{9.84}$$

The amplitude to produce the helicity 0 state is therefore proportional to

$$\begin{aligned}\varepsilon_\mu^{(e)*}\mathcal{M}^\mu &= \frac{1}{\mu}M_3\left[-k\left(1+\mathcal{O}\left(\frac{\mu^2}{k^2}\right)\right)+k\left(1+\mathcal{O}\left(\frac{\mu^2}{k^2}\right)\right)\right]\\ &= -\frac{k}{\mu}\mathcal{O}\left(\frac{\mu^2}{k^2}\right)\to 0 \text{ as } \frac{\mu}{k}\to 0.\end{aligned} \tag{9.85}$$

Therefore, as a result of current conservation, the amplitude to produce a 3D longitudinally polarized vector meson vanishes as $\mu\to 0$. The helicity zero mode smoothly decouples in this limit, and for $\mu=0$ is absent as a physical state. (Nothing special happens to the transverse modes in this limit). Therefore, in the massless theory, there are only two physical polarization states of the vector boson, both of which are 3D transverse. The appropriate form of the polarization sum is

$$\sum_{r=1}^{2}\varepsilon_\mu^{(r)}\varepsilon_\nu^{(r)*}=-g_{\mu\nu}. \tag{9.86}$$

9.5 QED

At this point it is worth summarizing our results. We set out to find a quantum theory of a massless vector field, the photon. We discovered that the massless limit is in general ill-defined, unless the vector field couples to a conserved current. This requirement gave us the interaction between the photon and charged scalars and Dirac fields (up to the overall coupling constant). The resulting theory is *quantum electrodynamics*, the quantum theory of electromagnetism. The QED Lagrangian for a theory with a single charged scalar φ and a single charged fermion ψ with charges q_φ and q_ψ respectively, is

$$\begin{aligned}\mathcal{L} &= \mathrm{D}_\mu\varphi^*\mathrm{D}^\mu\varphi-\mu^2\varphi^*\varphi+\overline{\psi}\left(i\not{\mathrm{D}}-m\right)\psi-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}\\ &= \partial_\mu\varphi^*\partial^\mu\varphi-ieq_\varphi A_\mu(\varphi^*\partial^\mu\varphi-\varphi\partial^\mu\varphi^*)+e^2q_\varphi^2A_\mu A^\mu\varphi^*\varphi\\ &\quad+\overline{\psi}\left(i\not{\partial}-m\right)\psi-eq_\psi\overline{\psi}\not{A}\psi-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}\end{aligned} \tag{9.87}$$

The Feynman rules for the Feynman amplitude $i\mathcal{A}$ in QED are illustrated in Fig. (9.8).

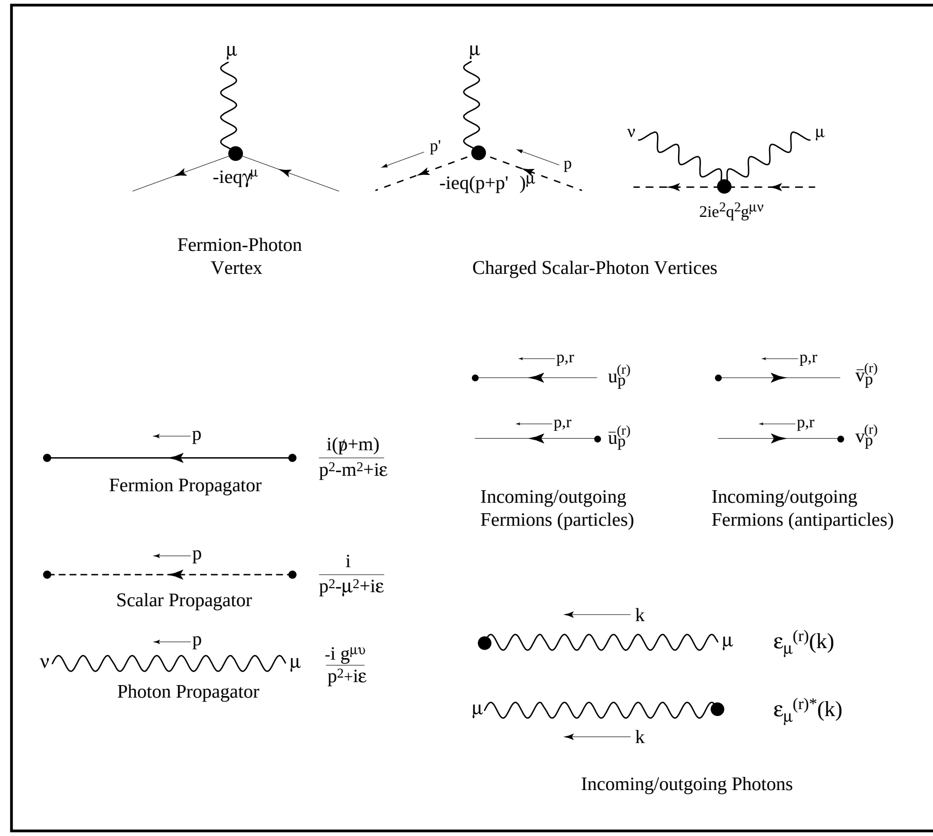

Figure 9.8: Feynman rules for QED

1. For each interaction vertex (fermion-fermion-photon, scalar-scalar-photon, or scalar-scalar-photon-photon) write down the appropriate factor.
2. For each internal line, include a factor of the corresponding propagator.
3. For each external fermion or photon, include the appropriate factor of the four-spinor or polarization vector.

4. The spinor factors (γ matrices, four-spinors) for each fermion line are ordered so that, reading from left to right, they follow the fermion line from the end of an arrow to the start.

5. The four-momenta associated with the lines meeting at each vertex satisfy energy-momentum conservation.

6. Multiply the expression by a phase factor δ which is equal to $+1$ (-1) if an even (odd) number of interchanges of neighbourhing fermion operators is required to write the fermion operators in the correct normal order.

There are additional rules for diagrams with loops, which we have not considered because we are just working at tree-level in this theory.

In some future version of these notes, I will include a few worked examples for QED processes. In the meantime, you should read Chapter 8 of Mandl & Shaw and work through Sections 8.4 (lepton pair production), 8.5 (Bhabha scattering) and 8.6 (Compton scattering). Note than M&S use differently normalized fermion fields than we do; however, the final answers are independent of normalization.

9.6 Renormalizability of Gauge Theories

In this chapter we started with the theory of a massive vector boson and showed that, despite appearances, it was possible to take the $\mu \to 0$ limit, in which case the theory had a larger symmetry, that of gauge invariance. Now we will go one step further and assert that gauge-invariance is *required* in order for a theory of vector bosons to make sense as a fundamental theory (I will explain what I mean by "fundamental" in a moment). To see why this is so, let's go back to the discussion of the previous section.

If a vector boson is coupled to a non-conserved current, the cancellation in Eq. (9.85) does not occur, and instead of being suppressed, the amplitude to produce a helicity 0 mode *grows* like k/μ. Thus, the probability of producing this mode grows with increasing energy without bound. This is in fact a Bad Thing, because at some energy the probability will become greater than 1! (This is known as "unitarity violation"). At this energy the theory has clearly stopped making sense (at least perturbatively).

There is nothing *a priori* wrong with this; we just have to interpret our theory a bit differently - not as a fundamental theory (valid up to arbitrary energy scales), but as an *effective* field theory. This kind of thing is very familiar in physics. If we are interested in fluid dynamics, for example, we don't have to consider the single atoms which make up the fluid. It makes much more sense to consider the fluid as a continuous medium. Similarly, if we are interested in the hydrogen atom we

can treat the proton as a point object, despite the fact that we know it is made up of quarks and gluons. Many aspects of nuclear physics can be described by a field theory of nucleons and pions, despite the fact that we know that these particles aren't "fundamental." It all depends on the scale of physics we're interested in.

In our case, what the theory is telling us is that a theory of massive vector mesons coupled to a nonconserved current is a perfectly fine theory at low energies, but that it *can't* be valid up to arbitrarily high energy scales (that is, down to arbitrarily short distances). At some scale (typically set by the mass of the particle, since that's the only dimensionful parameter in the theory), this theory has to break down in some way - for example, the vector boson could be revealed not to be fundamental, but to be a composite particle, and so its dynamics would change at the scale of order the size of the particle. What is fascinating is that the theory carries within itself the seeds of its own destruction.

This property of a theory, that it be valid up to arbitrarily high energies and so not predict its own demise, is related to a property known as "renormalizability." Roughly speaking, renormalizability is the extension of the above discussion to radiative corrections. You recall that internal loops in a Feynman diagram come with a factor of

$$\int \frac{d^4k}{(2\pi)^4}$$

since the momentum running through the loop is unconstrained. As a result, arbitrarily high momenta run through loop graphs, even if the process being considered is a low-energy process. It is perhaps not surprising, then, that if the theory breaks down at a certain scale, this will affect loop graphs even for low-energy processes. Without getting into the details of radiative corrections, I will just assert that if one attempts to calculate loop graphs in a theory of a massive vector boson coupled to a non-conserved current one is again led to the conclusion that the theory cannot be fundamental.

So what are the requirements for a theory to be renormalizable? It is easy to get an idea, just by unitarity arguments. Imagine a theory with a four-fermion interaction term

$$\mathcal{L}_I = -\frac{g}{M^2}\overline{\psi}\psi\overline{\psi}\psi.$$

Now let's do some dimensional analysis. You showed on an early problem set that in 4 dimensions, the Lagrangian density has dimensions of [mass]^4. From the Dirac Lagrangian, we conclude that the dimensions of ψ are [mass]$^{3/2}$ (so that $m\overline{\psi}\psi$ has the right units). $\overline{\psi}\psi\overline{\psi}\psi$ has dimension [mass]^6, so the coupling must have dimensions [mass]$^{-2}$, which I have made explicit by writing it as g/M^2 (g is dimensionless).

Now, a cross-section has units of area, or [mass]$^{-2}$. Since the amplitude from the four-fermion interaction goes like $1/M^2$, the cross section for fermion-fermion scattering in this theory must be proportional to $1/M^4$. By dimensional analysis, at high energies where we may ignore the fermion masses, we must therefore have

$$\sigma \propto \frac{s}{M^4} \tag{9.88}$$

where $s = (p_1+p_2)^2$ is the squared centre of mass energy of the collision. Since the cross-section grows without bound, once again the probability must grow without bound, and again the theory must break down at some energy scale set by M.

Just by dimensional analysis, you can see that this will happen in ANY theory with coupling constants which are *inverse* power of a mass. Thus, for a theory to be renormalizable, all terms in the Lagrangian must have mass dimension ≤ 4. This answers a question which may have been bothering you all along in this course: why do we always study theories with such simple interaction terms? Why can't we have an interaction term like

$$-g\overline{\psi}\psi \cos\ln\left(1 + \varphi/M\right)?$$

The answer is that this is not a renormalizable interaction. Therefore interaction terms like ϕ^4, $\overline{\psi}\psi\phi$ and ϕ^3 (dimension 4, 4 and 3, respectively) are allowed in a fundamental theory, but interactions like $\overline{\psi}\psi\overline{\psi}\psi$, ϕ^5, $\phi^2\overline{\psi}\psi$ (dimension 6, 5 and 5, respectively) are not. This is why we only considered very simple interaction terms - anything more complicated leads to a non-renormalizable theory.

Of course, there is no reason to only consider theories which are valid up to arbitrary energy scales. After all, we can only do experiments at finite energies. However, since higher-dimension operators come with coupling constants which are proportional to inverse powers of the scale at which the theory breaks down, the effects of these terms are negligible at low energies. Their effects are proportional to the ratio of the momentum of the process to the scale of new physics. Unless they break symmetries which are preserved by the renormalizable terms (such as parity in the case of the weak interactions, or baryon number in GUTs) they can usually be safely ignored.

Finally, it can also be shown that theories with fields of spin > 1 are also non-renormalizable. This is at the root of the difficulty of quantizing gravity: the graviton is a spin-2 field (corresponding to quanitizing the metric tensor $g_{\mu\nu}$), and so the corresponding quantum theory is nonrenormalizable.

It is because of renormalizability that gauge symmetries are so crucial in field theory: the only way to couple a vector field to other fields in a renormalizable way is through a gauge covariant derivative. Furthermore, since a vector meson

mass term breaks the gauge symmetry, only theories with massless vector bosons are renormalizable. The theory of a massive vector boson which we studied in this section, while useful for obtaining the Feynman rules for QED, is not a renormalizable theory.

Now, as you may be aware, there certainly are massive vector bosons coupled to nonconserved currents in the world. The gauge bosons associated with the weak interactions, the $W^{\pm}$ and Z^0, have masses of 80.2 GeV and 91.2 GeV, respectively. Experimentally, they couple to electrons and electron-neutrinos via the following interaction:

$$\begin{aligned}\mathcal{L}_I &= -g_1\left(\overline{\nu}\gamma^\mu(1-\gamma_5)eW^+_\mu + \overline{e}\gamma^\mu(1-\gamma_5)\nu W^-_\mu\right) \\ &\quad -g_2 Z^\mu\left(\overline{e}\gamma_\mu(g_V - g_A\gamma_5)e + \overline{\nu}\gamma_\mu(1-\gamma_5)\nu\right)\end{aligned} \tag{9.89}$$

where g_1, g_2, g_V and g_A are coupling constants which are related to the electric charge and the ratio of the $W^{\pm}$ and Z^0 boson masses. Since the electron is massive, the current $\overline{e}\gamma^\mu\gamma_5\nu$ is not conserved, so we have a theory of gauge bosons coupled to nonconserved currents. So the W and Z clearly can't be fundamental. Furthermore, the theory predicts that unitarity violation due to excessive production of longitudinal W's and Z's will occur at a scale of about 3 TeV=3×10^3 GeV.

The question of what the W's and Z's are made of is the foremost question in particle theory at the moment. The simplest possibility which leads to a renormalizable theory is that of the minimal Weinberg-Salam model, in which the transverse components of the W and Z are fundamental (corresponding to massless vector bosons), while the longitudinal components are made of a scalar particle, known as the "Higgs Boson." In the minimal theory, there are four Higgs bosons, three of which are incorporated into the two W's and the Z, and the fourth of which is just waiting to be experimentally observed. But this is only the simplest possibility - there are many others.