Willoughby Seago

Theoretical Physics

Quantum Field Theory

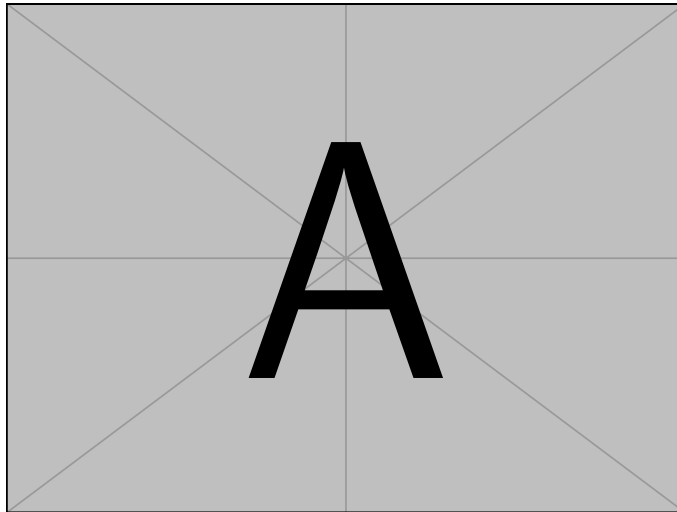# Quantum Field Theory

Willoughby Seago

These are my notes from the course quantum field theory. I took this course as a part of the theoretical physics degree at the University of Edinburgh.

These notes were last updated at 23:05 on September 28, 2022. For notes on other topics see .

Chapters

Page

Contents

Page

List of Figures

One

Introduction

1.1 Course Overview

Quantum field theory (QFT) is combination of relativity and quantum mechanics. It will be our topic of study in this course. We will focus mostly on a single quantum field theory, namely **quantum electrodynamics (QED)**, which is the result of combining electrodynamics, which is inherently relativistic, and quantum mechanics.

There are various courses leading into this course, some of the important ones for which I have notes are

- *Principles of Quantum Mechanics*;

- *Quantum Theory*;

- *Classical Electrodynamics*; and

- *Symmetries of Quantum Mechanics*.

This course has a companion course, *Symmetries of Particles and Fields*, which focuses on the mathematical abstraction of symmetry. Ideas from this companion course, and the related course *Symmetries of Quantum Mechanics* will occur throughout this course. Another related course, focused more on the experimental side of this field, is *Particle Physics*. I have notes for both *Symmetries of Particles and Fields* and *Particle Physics*.

This course, and its companion, lead naturally into the course *Gauge Theories in Particle Physics*, for which I also have notes. Many ideas in this course will be expanded upon in the gauge theories course. The gauge theories course will also treat other quantum field theories, such as **electroweak** interactions (a combination of electromagnetism and the weak interactions), and **quantum chromodynamics (QCD)** (the quantum field theory of the strong force). The course also briefly touches on lattice gauge theory, a particular way of doing calculations in nonperturbative QCD.

This course is roughly divided into four sections:

- Canonical QFT: Here we deal with operators and quantisation, familiar from traditional quantum mechanics, such as in *Principles of Quantum Mechanics*.

- QED: We will use this as an example, to allow us to work in a concrete setting rather than in the abstract. This is in many ways the simplest of quantum

1

field theories, we study mostly interactions of photons and electrons, and since photons aren't charged the interactions are about as simple as they get. Mathematically this "simplicity" is due to the underlying U(1) symmetry, which is an Abelian gauge group, and is significantly easier to work with than, say, the SU(3) symmetry of QCD.

- Path integral formalism: This is a more abstract, but often more powerful, yet equivalent formulation of quantum field theory. This formalism leads itself to gauge theories, and so will be used heavily in *Gauge Theories in Particle Physics*. Nonrelativistic path integrals were treated in *Quantum Theory*.

- Renormalisation: When we do QFT calculations we often get infinite results. Renormalisation is the process of removing these infinities and extracting meaningful results. The renormalisation of QED will be treated in more detail in *Gauge Theories*.

There is some concurrent teaching of these topics, but in these notes I separate them out, so if something's not making sense maybe check to see if its been explained in more detail in a different section.

1.2 Conventions

Quantum field theory is full of conventions. One convention that almost everyone follows is that we work in natural units, where $c = \hbar = 1$. This makes the formulas look much simpler and less cluttered, and we can put c and $\hbar$ back in by dimensional analysis.

The second convention is unfortunately much more varied, its the choice of metric. We will use the $(+---)$ metric,

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{1.2.1}$$

The most common alternative to this is $(-+++)$, where the diagonal has a single negative 1 and the rest is made of positive 1s. There are some even rarer choices to use an imaginary metric, $(i+++)$, which still gives the same relative minus sign when we square each element in the inner product.

Throughout the course, unless specified otherwise, we will use the Einstein summation convention. Specifically, if an index appears exactly twice in a term, once in a raised position and once in a lowered position, then we sum over all values of that index. We also follow the convention where Greek indices, such as μ and ν, run from 0 to the number of dimensions minus one, most commonly meaning $\mu = 0, 1, 2, 3$. On the other hand, Latin indices, such as i and j, run from 1 to the number of dimensions minus one, most commonly, $i = 1, 2, 3$. This means that Greek indices are summed over all components, whereas Latin indices are summed only over spatial components.

We follow the convention that the electromagnetic field strength tensor is

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{1.2.2}$$

An alternative convention is

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

(1.2.3)

which differs from our choice by an overall minus sign. For example, this results in Maxwell's equations being written as $\partial_\mu F^{\mu\nu} = -j^\nu$, instead of $\partial_\mu F^{\mu\nu} = j^\nu$, which is what we'll be using.

Part I

Canonical Quantisation

Two

Classical Fields

2.1 Relativity Basics

In order to understand quantum field theory we will need to quantise fields. In order to quantise fields we should be sure that we understand classical fields, so this is where we start the course. We will briefly recap ideas mostly from the *Classical Electrodynamics* and *Quantum Theory* courses, so check the notes for this course for more details.

2.1.1 Four-Vectors

We can consider a generic four-vector,

$$a^\mu = (a^0, \boldsymbol{a}) = (a^0, a^1, a^2, a^3). \tag{2.1.1}$$

Using the metric, $\eta_{\mu\nu} := \mathrm{diag}(1, -1, -1, -1)$, we can lower the indices to get a covariant four-vector:

$$a_\nu := \eta_{\mu\nu} a^\mu = (a^0, -\boldsymbol{a}) = (a^0, -a^1, -a^2, -a^3) = (a_0, a_1, a_2, a_3). \tag{2.1.2}$$

Given two four vectors, $a^\mu = (a^0, \boldsymbol{a})$ and $b = (b^0, \boldsymbol{b})$, we can take the inner product,

$$a \cdot b := a^\mu b_\mu = a_\mu b^\mu = \eta_{\mu\nu} a^\mu b^\nu = a^0 b^0 - \boldsymbol{a} \cdot \boldsymbol{b}. \tag{2.1.3}$$

Here $\boldsymbol{a} \cdot \boldsymbol{b} = a^i b^i = a^1 b^1 + a^2 b^2 + a^3 b^3$ is the standard dot product for three-vectors.

For the specific case of the inner product of a four-vector with itself we use the shorthand

$$a^2 = a \cdot a = a^0 a^0 - \boldsymbol{a} \cdot \boldsymbol{a} = (a^0)^2 - \boldsymbol{a}^2. \tag{2.1.4}$$

Four vectors transform under Lorentz transformations. The broadest class of such transformations is the **Lorentz group**, $\mathrm{O}(1, 3)$. This includes all boosts and rotations, including those that invert the directions of space or time. In this course we only consider **proper orthochronous Lorentz transformations**, that is Lorentz transformations preserving the orientation of space (proper) and time (orthochronous). These form the **proper orthochronous Lorentz group** $\mathrm{SO}^+(1, 3)$. From now on if we say Lorentz transformation we mean *proper orthochronous* Lorentz transformation.

5

2.1.2 Specific-Four Vectors

The position is a four-vector, $x^\mu = (t, \boldsymbol{x})$, where t is the time coordinate and $\boldsymbol{x}$ is the spatial position. Note that in SI units this would be $x^\mu = (ct, \boldsymbol{x})$, since the components of a four-vector must have the same dimensions. We can also construct a covariant position four-vector, $x_\mu = (t, -\boldsymbol{x})$.

The momentum is a four-vector, $p^\mu = (E, \boldsymbol{p})$, where E is the energy, and $\boldsymbol{p}$ is the relativistic three-momentum. Note that in SI units this would be $p^\mu = (E/c, \boldsymbol{p})$. We can also construct a covariant momentum four-vector, $p_\mu = (E, -\boldsymbol{p})$.

For a (real[1]) particle of mass m the four-momentum squares to the mass squared, that is

$$m^2 = p^2 = E^2 - \boldsymbol{p}^2. \tag{2.1.5}$$

Rearranging this gives us

$$E^2 = \boldsymbol{p}^2 + m^2. \tag{2.1.6}$$

This is the relativistic **energy-momentum relation**, it's perhaps more familiar if we reinstate the factors of c:

$$E^2 = \boldsymbol{p}^2 c^2 + m^2 c^4, \tag{2.1.7}$$

and is most famous in the case where $\boldsymbol{p} = \boldsymbol{0}$:

$$E = mc^2. \tag{2.1.8}$$

We can construct a four-vector derivative by defining the derivative with respect to x^μ:

$$\partial_\mu := \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial t}, \nabla \right). \tag{2.1.9}$$

Note that this is a covariant vector, since the contravariant quantity, x^μ, appears in the denominator, so the derivative transforms in the opposite way to how it would if x^μ were in the numerator. We can similarly define a contravariant operator,

$$\partial^\mu = \frac{\partial}{\partial x_\mu} = \left(\frac{\partial}{\partial t}, -\nabla \right). \tag{2.1.10}$$

The square of these operators is common enough to be given it's own name, it's called the **d'Alembert operator**,

$$\partial^2 = \partial^\mu \partial_\mu = \frac{\partial^2}{\partial t^2} - \nabla^2. \tag{2.1.11}$$

This same quantity is also denoted $\Box^2$, as a sort of four-dimensional (note four sides) Lorentzian-manifold analogue of the Laplacian, ∇^2, and also confusingly sometimes denoted $\Box$, without the superscript 2. We also call this the wave operator, since if f is a field satisfying $\partial^2 f = 0$ then expanding and rearranging this we have

$$\frac{\partial^2 f}{\partial t^2} = \frac{\partial^2 f}{\partial x^2}, \tag{2.1.12}$$

which is the equation of a wave travelling at the speed of light.

[1] as in, not a virtual particle
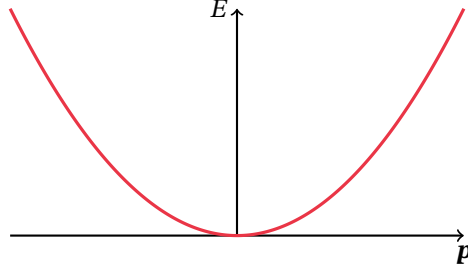
Figure 2.1: The nonrelativistic energy-momentum relation assigns a single energy to each three-momentum. Note that this is really just a slice through a four-dimensional plot, but all that really matters is the magnitude of the three-momentum.

2.2 Nonrelativistic Particle and Wave

A nonrelativistic particle of mass m has the energy-momentum relation

$$E = \frac{\boldsymbol{p}^2}{2m} = \frac{1}{2}m\boldsymbol{v}^2. \tag{2.2.1}$$

To get an equation for this particle we substitute in the usual operators,

$$E \to i\frac{\partial}{\partial t}, \quad \text{and} \quad \boldsymbol{p} \to -i\nabla. \tag{2.2.2}$$

Acting on an arbitrary state, ψ, we then get

$$i\frac{\partial \psi}{\partial t} = \frac{(-i\nabla)^2}{2m}\psi = -\frac{1}{2m}\nabla^2\psi. \tag{2.2.3}$$

This is exactly the **Schrödinger equation** for a free particle.

One solution to this is a plane wave,

$$\psi(x) = e^{-iEt+i\boldsymbol{p}+\boldsymbol{x}} = e^{-ip\cdot x}. \tag{2.2.4}$$

This is only a solution if $E^2 = \boldsymbol{p}^2/2m$. Notice that while we can write the exponent with the relativistic four-vectors this is still nonrelativistic since the energy-momentum is not relativistic. In particular, we have a single energy value for each possible momentum, as seen in Figure 2.1.

We know that solutions to the Schrödinger equation can be interpreted as amplitudes, which we then take the modulus square of, $|\psi|^2$, to get a probability density function. We call a function with this property a **wave function**.

2.3 Relativistic Particle and Wave

A relativistic particle of mass m has the energy-momentum relation

$$E^2 = \boldsymbol{p}^2 + m^2. \tag{2.3.1}$$

Making the operator substitutions and acting on an arbitrary state, φ, we get

$$\left(i\frac{\partial}{\partial t}\right)^2 \varphi = (-i\nabla)^2\varphi + m^2\varphi. \tag{2.3.2}$$
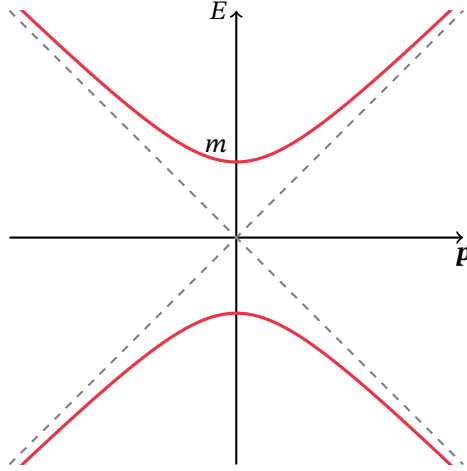
Figure 2.2: The relativistic energy momentum assigns two energies to a each three-momentum. The curve is a hyperbola. Notice that at zero three-momentum there we don't have zero energy, instead we have energy $\pm m$, since we are counting the mass towards the energy in the relativistic formulation. The hyperbola here are the "shell" referred to in "on-shell".

Rearranging this we get

$$(\partial^2 + m^2)\varphi = 0. \tag{2.3.3}$$

This is the **Klein–Gordon equation**.

One solution to this is a plane wave,

$$\varphi(x) = \mathrm{e}^{-ip\cdot x}. \tag{2.3.4}$$

This is a solution provided that $p^2 = m^2$. If $p^2 = m^2$ then we say that p^μ is **on-shell**.

If $e^{-ip\cdot x}$ is a solution then so is

$$\varphi^* = \mathrm{e}^{ip\cdot x}. \tag{2.3.5}$$

We see that we get pairs of solutions. These correspond to the pairs of possible energy values, $E = \pm\sqrt{p^2 + m^2}$, as shown in Figure 2.2.

The negative energy solutions to the Klein–Gordon equation pose a problem, in particular, it is possible for $\varphi^*\varphi$ to be negative, which means there is no probabilistic interpretation of φ, and so φ is *not* a wave function. Instead, we just call φ a field.

States of a system obeying the Klein–Gordon equation correspond to relativistic free particles (plural).

2.4 Maxwell Fields

The most familiar fields are electromagnetic fields. In this section we will briefly recap a relativistic treatment of these fields. For more details see the *Classical Electrodynamics* course.

The electromagnetic "potential" is

$$A^\mu(x) = (\varphi(x), \boldsymbol{A}(x)), \tag{2.4.1}$$

where φ is the electric potential and $\boldsymbol{A}$ is the electromagnetic vector potential. The electric "field" is then

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{2.4.2}$$

> ! An alternative convention,
>
> $$F^{\mu\nu} = \partial^\nu A^\mu - \partial^\mu A^\nu, \tag{2.4.3}$$
>
> differs by a sign.

The electromagnetic current density is

$$j^\mu = (\rho, \boldsymbol{j}). \tag{2.4.4}$$

This current is conserved, mathematically this is expressed as its four-divergence vanishing,

$$\partial_\mu j^\mu = \frac{\partial \rho}{\partial t} + \nabla \cdot \boldsymbol{j} = 0. \tag{2.4.5}$$

This is a continuity equation.

Maxwell's equation is

$$\partial_\mu F^{\mu\nu} = j^\nu. \tag{2.4.6}$$

Writing $F^{\mu\nu}$ in terms of A^μ we get

$$\partial_\mu F^{\mu\nu} = \partial_\mu \partial^\mu A^\nu - \partial_\mu \partial^\nu A^\mu = \partial^2 A^\nu - \partial^\nu \partial_\mu A^\mu. \tag{2.4.7}$$

The **Lorenz gauge** condition is to choose A^μ such that its four-divergence vanishes, $\partial_\mu A^\mu$. Then we have $\partial_\mu F^{\mu\nu} = \partial^2 A^\nu$, so we can write Maxwell's equation as

$$\partial^2 A^\nu = j^\nu. \tag{2.4.8}$$

In QFT we call A^μ the **electromagnetic field** and we call $F^{\mu\nu}$ the **electromagnetic field strength**. This is partly an annoying historical artefact and partly because A^μ is the field appearing in our equations in an analogous way to φ in the Klein–Gordon equation.

In free space, that is when $j^\mu = 0$, Maxwell's equation is

$$\partial^2 A^\mu = 0. \tag{2.4.9}$$

Notice that this is similar to the Klein–Gordon equation with massless particles, $m = 0$. This is good, because we would like electromagnetism to work with photons, which are massless. The one difference from the Klein–Gordon equation is that A^μ is a four-vector, whereas φ is a scalar. The only change in the solutions is that we, in theory, get a different solution for each component of A^μ. The way

that this manifests in the solutions is we get an extra vector out front called the **polarisation vector**, ε^μ:

$$A^\mu = \varepsilon^\mu \mathrm{e}^{-ip\cdot x}. \tag{2.4.10}$$

Assuming the momentum is on-shell, that is $p^2 = 0$ in this massless case. This means the momentum is light-like, more good news for electromagnetism explaining photons.

It can be shown that[2] $\varepsilon_\mu p^\mu = 0$. We are also free to choose ε^μ such that $\varepsilon^0 = 0$, then the condition $\varepsilon_\mu p^\mu = 0$ becomes $\boldsymbol{\varepsilon} \cdot \boldsymbol{p} = 0$. This means that electromagnetic waves are **transverse**, since their polarisation vector, $\boldsymbol{\varepsilon}$, is perpendicular to their direction of travel, $\boldsymbol{p}$.

Since the mathematics of the vector field A^μ is so similar to the mathematics of the scalar field φ we will mostly deal with scalar fields in this course to develop our theory, then work with other fields once we have a good understanding off the maths. Each type of field corresponds to a different spin. Scalar fields describe spin zero particles, like the Higgs boson, and vector fields describe spin one particles, like the photon. Spin 1/2 particles are described by spinors.

[2] See the *Classical Electrodynamics* course, note that there we work with $k^\mu = \hbar p^\mu$.

Three

Lagrangians and Hamiltonians

3.1 Lagrangian Dynamics of a Particle

> *That's all there is to a Lagrangian dynamics course, the rest is just examples.*
>
> Richard Ball

(R) See the *Lagrangian Dynamics* course for more details.

Consider a particle of mass m moving in one dimension in a conservative force field. We can define its position with a **generalised coordinate**, q, which gives us a **generalised velocity**, $\dot{q} := \mathrm{d}q/\mathrm{d}t$. We can then define the **Lagrangian**:

$$L(q, \dot{q}) = T(\dot{q}) - V(q). \tag{3.1.1}$$

Here $T(\dot{q})$ is the kinetic energy of the particle, which we assume depends only on the velocity for simplicity, and $V(q)$ is the potential energy of the particle, which must depend only on the position. The Lagrangian is a function of the generalised position and velocity.

The **action** of the particle between the times t_1 and t_2 is defined to be

$$S[q(t)] := \int_{t_1}^{t_2} \mathrm{d}t \, L(q, \dot{q}). \tag{3.1.2}$$

Hamilton's principle states that the path the particle takes, $q(t)$, is such that the action is extremised, that is the variation, δS, given by varying $q \to q + \delta q$ vanishes. We can compute the variation in a general action, S, by computing the variation in the Lagrangian:

$$\delta L = \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q}. \tag{3.1.3}$$

For a smooth variation, δq, we have

$$\delta \dot{q} = \frac{\partial}{\partial t}(\delta q), \tag{3.1.4}$$

and so

$$\delta L = \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \frac{\mathrm{d}}{\mathrm{d}t}(\delta q) \tag{3.1.5}$$

Integrating this the variation in the action is

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \left[\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \frac{\mathrm{d}}{\mathrm{d}t}(\delta q) \right]. \tag{3.1.6}$$

The product rule tells us that

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}} \delta q \right) = \frac{\partial L}{\partial \dot{q}} \frac{\mathrm{d}}{\mathrm{d}t}(\delta \dot{q}) + \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}} \right) \delta q. \tag{3.1.7}$$

Rearranging this we can rewrite the second term in δS as a total derivative minus a term:

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \left[\frac{\partial L}{\partial q} \delta q + \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}} \delta q \right) - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}} \right) \delta q \right] \tag{3.1.8}$$

$$= \int_{t_1}^{t_2} \mathrm{d}t \left[\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}} \right) \right] \delta q + \int_{t_1}^{t_2} \mathrm{d}t \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}} \delta q \right) \tag{3.1.9}$$

$$= \int_{t_1}^{t_2} \mathrm{d}t \left[\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}} \right) \right] \delta q + \left[\frac{\partial L}{\partial \dot{q}} \delta q \right]_{t_1}^{t_2}. \tag{3.1.10}$$

Now suppose that δq vanishes at t_1 and t_2, so $\delta q(t_1) = \delta q(t_2) = 0$, then the last term above vanishes and we have

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \left[\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}} \right) \right] \delta q. \tag{3.1.11}$$

If δS is to vanish for all variations δq then we must have

$$\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}} \right) = 0. \tag{3.1.12}$$

This is the **Euler–Lagrange equation** for a single particle in one dimension.

This generalises to n-dimensional space by replacing q with n coordinates q_i and $\dot{q}$ with $\dot{q}_i$. The Lagrangian is then

$$L(q_i, \dot{q}_i) = T(\dot{q}_i) - V(q_i), \tag{3.1.13}$$

and the action is

$$S[q_i(t)] = \int_{t_1}^{t_2} \mathrm{d}t \, L(q_i, \dot{q}_i). \tag{3.1.14}$$

We then get n Euler–Lagrange equations:

$$\frac{\partial L}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_i} \right) = 0. \tag{3.1.15}$$

> **Example 3.1.16** The Lagrangian for a simple harmonic oscillator is
>
> $$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2 x^2. \tag{3.1.17}$$
>
> We have
>
> $$\frac{\partial L}{\partial x} = -m\omega^2 x, \quad \text{and} \quad \frac{\partial L}{\partial \dot{x}} = m\dot{x} \implies \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}}\right) = m\ddot{x}. \tag{3.1.18}$$
>
> Hence,
>
> $$m\ddot{x} = -m\omega^2 x \implies \ddot{x} = -\omega^2 x \tag{3.1.19}$$
>
> is the equation of motion for the simple harmonic oscillator, which is exactly what we would expect.

3.2 Hamiltonian Dynamics of a Particle

Let's go back to the one-dimensional case. We can define the **canonical momentum** of the particle as

$$p := \frac{\partial L}{\partial \dot{q}}. \tag{3.2.1}$$

Note that, in general, this is *not* the normal momentum.

We then define the **Hamiltonian**, H, as a function of the generalised position and canonical momentum:

$$H(q, p) := p\dot{q} - L(q, \dot{q}). \tag{3.2.2}$$

Note that we must eliminate any $\dot{q}$ remaining in the expression for the Hamiltonian, this can be done by solving the defining relation $p = \mathrm{d}L/\mathrm{d}\dot{q}$ for $\dot{q}$ and then substituting in the result.

Now consider what happens when we vary H. Varying the left hand side we get

$$\mathrm{d}H = \frac{\partial H}{\partial q}\,\mathrm{d}q + \frac{\partial H}{\partial p}\,\mathrm{d}p. \tag{3.2.3}$$

Varying the right hand side we have

$$\mathrm{d}H = p\,\mathrm{d}q + q\,\mathrm{d}p - \frac{\partial L}{\partial q}\,\mathrm{d}q - \frac{\partial L}{\partial \dot{q}}\,\mathrm{d}\dot{q}. \tag{3.2.4}$$

We can rewrite the last term in terms of the canonical momentum:

$$\mathrm{d}H = p\,\mathrm{d}q + q\,\mathrm{d}p - \frac{\partial L}{\partial q}\,\mathrm{d}q - p\,\mathrm{d}\dot{q} = q\,\mathrm{d}p - \frac{\partial L}{\partial q}\,\mathrm{d}q. \tag{3.2.5}$$

Using the Euler–Lagrange equations the second term can be rewritten to give

$$\mathrm{d}H = q\,\mathrm{d}p - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right)\,\mathrm{d}q = q\,\mathrm{d}p - \dot{p}\,\mathrm{d}q. \tag{3.2.6}$$

Comparing this with the result for the left hand side variation we can read off

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \text{and} \quad \dot{p} = -\frac{\partial H}{\partial q}. \tag{3.2.7}$$

These are **Hamilton's equations**.

This all generalises to n dimensions. First replace q with q_i, and p with $p_i :=$ $\partial L/\partial q_i$. Then the Hamiltonian is

$$H(q_i, p_i) = \sum_i p_i \dot{q}_i - L(q_i, \dot{q}_i). \tag{3.2.8}$$

Hamilton's equations are

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}. \tag{3.2.9}$$

It can be useful when doing Hamiltonian mechanics to define the **Poisson bracket** of two functions, A and B, depending on position, q_i, and momentum, p_i:

$$\{A, B, :=\} \sum_i \left(\frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} - \frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} \right). \tag{3.2.10}$$

This has the advantage of allowing us to express Hamilton's equations in the more symmetric form

$$\dot{q}_i = \{H, q_i\}, \quad \text{and} \quad \dot{p}_i = \{H, p_i\}. \tag{3.2.11}$$

Note the similarity to Heisenberg's operator equation of motion for a time independent operator, A:

$$\dot{A}(t) = i[H, A]. \tag{3.2.12}$$

Replacing Poisson brackets with i times the commutator can get us quite a long way in quantum mechanics.

The Hamiltonian is conserved. To show this consider the time derivative:

$$\frac{dH}{dt} = \frac{\partial H}{\partial p}\dot{p} + \frac{\partial H}{\partial q}\dot{q} = \dot{q}\dot{p} - \dot{p}\dot{q} = 0. \tag{3.2.13}$$

We can often identify the Hamiltonian with the total energy when this is also conserved.

> **Example 3.2.14** The canonical momentum for a simple harmonic oscillator is
>
> $$p = \frac{\partial L}{\partial \dot{x}} = m\dot{x}, \tag{3.2.15}$$
>
> which is just the normal momentum in this case. The Hamiltonian is then
>
> $$H = p\dot{q} - L = m\dot{x}^2 - \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega^2 x^2 = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega^2 x^2, \tag{3.2.16}$$

> which is exactly the energy of a simple harmonic oscillator.

3.3 Lagrangian Dynamics of a Field

Lagrangian dynamics doesn't change that much when we work with fields. We start by replacing the generalised coordinate $q(t)$ with the field, $\varphi(x)$, which now depends on a position in spacetime, rather than just a time. Instead of a Lagrangian we work with a **Lagrangian density**, $\mathcal{L}(\varphi, \partial_\mu \varphi)$, replacing the time derivative of the coordinate with a four-vector derivative. We then have to integrate over a region of spacetime to get the action:

$$S[\varphi(x)] := \int \mathrm{d}^4 x \, \mathcal{L}(\varphi, \partial_\mu \varphi). \tag{3.3.1}$$

The integral should be taken over some arbitrary region of spacetime. Note that we can write this as

$$S[\varphi(x)] = \int \mathrm{d}t \int \mathrm{d}^3 x \, \mathcal{L}(\varphi, \partial_\mu \varphi) = \int \mathrm{d}t \, L(\varphi, \partial_\mu \varphi) \tag{3.3.2}$$

where

$$L(\varphi, \partial_\mu \varphi) := \int \mathrm{d}^3 x \, \mathcal{L}(\varphi, \partial_\mu \varphi). \tag{3.3.3}$$

So, integrating $\mathcal{L}$ over space gives L, which we call the Lagrangian, which explains the interpretation of $\mathcal{L}$ as a Lagrangian *density*. While this distinction is important the Lagrangian doesn't actually appear that much in QFT, so people often just call $\mathcal{L}$ the Lagrangian, leaving the density part implicit.

The Euler–Lagrange equations for the Lagrangian density don't change much, and can be derived in a very similar way. We start by varying the field, $\varphi \to \varphi + \delta\varphi$, and looking at the resulting variation in the Lagrangian density:

$$\delta\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \varphi} \delta\varphi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \delta(\partial_\mu \varphi). \tag{3.3.4}$$

For a smooth variation we have

$$\delta(\partial_\mu \varphi) = \partial_\mu(\delta\varphi) \tag{3.3.5}$$

and so

$$\delta\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \varphi} \delta\varphi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \partial_\mu(\delta\varphi). \tag{3.3.6}$$

We can use the product rule,

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \delta\varphi \right) = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \right) \delta\varphi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \partial_\mu(\delta\varphi), \tag{3.3.7}$$

to write the last term as a total derivative minus a term:

$$\delta\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \varphi} \delta\varphi + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \delta\varphi \right) - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \right) \delta\varphi. \tag{3.3.8}$$

Integrating over some spacetime region, Ω, to get the variation in the action we have

$$\delta S = \int_\Omega \mathrm{d}^4 x \left[\frac{\partial \mathcal{L}}{\partial \varphi} \delta \varphi + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \delta \varphi \right) - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \right) \delta \varphi \right] \tag{3.3.9}$$

$$= \int_\Omega \mathrm{d}^4 x \left[\frac{\partial \mathcal{L}}{\partial \varphi} \delta \varphi - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \right) \delta \varphi \right] + \int_\Omega \mathrm{d}^4 x \, \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \delta \varphi \right). \tag{3.3.10}$$

Applying the divergence theorem to the second integral we can replace it with an integral of $(\partial \mathcal{L}/\partial(\partial_\mu \varphi))\delta \varphi$ over the boundary, $\partial \Omega$. Choosing a variation, $\delta \varphi$, which vanishes on the boundary this term will vanish, and we will be left with

$$\delta S = \int_\Omega \mathrm{d}^4 x \left[\frac{\partial \mathcal{L}}{\partial \varphi} \delta \varphi - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \right) \delta \varphi \right]. \tag{3.3.11}$$

For this to vanish for all variations of the field we require that

$$\frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \right) = 0. \tag{3.3.12}$$

These are the **Euler–Lagrange equations** for a scalar field.

3.4 Hamiltonian Dynamics of a Field

Following the same logic as for a single particle we can define the **canonical momentum**,

$$\pi(x) \coloneqq \frac{\partial \mathcal{L}}{\partial \dot{\varphi}}. \tag{3.4.1}$$

This is *not* the same as the actual momentum. Notice that this definition treats time differently to position, this will result in the Hamiltonian formulation not being Lorentz invariant, although the results we get still are, the steps between are just frame dependent. This corresponds to the non-Lorentz invariance of the energy, the energy of a particle of mass m is γm.

We then define the Hamiltonian, H, to be

$$H \coloneqq \int \pi \dot{\varphi} \, \mathrm{d}^4 x - L = \int \mathrm{d}^4 x \, (\pi \dot{\varphi} - \mathcal{L}) = \int \mathrm{d}^3 \boldsymbol{x} \, \mathcal{H} \tag{3.4.2}$$

where

$$\mathcal{H} \coloneqq \pi \dot{\varphi} - \mathcal{L} \tag{3.4.3}$$

is the **Hamiltonian density**. Both the Hamiltonian and Hamiltonian density must be functions of the field, φ, and the canonical momentum, π, which is just another field.

The same derivation as before works without modification, so we won't repeat it here, the result is **Hamilton's equations** for a field:

$$\dot{\varphi} = \frac{\partial \mathcal{H}}{\partial \pi}, \quad \text{and} \quad \dot{\pi} = -\frac{\partial \mathcal{H}}{\partial \varphi}. \tag{3.4.4}$$

As before, the Hamiltonian density is conserved:

$$\frac{d\mathcal{H}}{dt} = \int d^3x \left(\frac{\partial \mathcal{H}}{\partial \varphi} \dot{\varphi} + \frac{\partial \mathcal{H}}{\partial \pi} \dot{\pi} \right) \tag{3.4.5}$$

$$= \int d^3x (-\dot{\pi}\dot{\varphi} + \dot{\varphi}\dot{\pi}) \tag{3.4.6}$$

$$= 0. \tag{3.4.7}$$

Again, when energy is conserved we can often identify it with the Hamiltonian.

3.5 Klein–Gordon Equation

> *We have to learn to run before we walk.*
>
> Richard Ball

The Lagrangian for a scalar field, φ, is

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \varphi)(\partial^\mu \varphi) - \frac{1}{2}m^2\varphi^2. \tag{3.5.1}$$

Clearly we have

$$\frac{\partial \mathcal{L}}{\partial \varphi} = -m^2\varphi. \tag{3.5.2}$$

We can also compute the derivative with respect to $\partial_\mu \varphi$, note that $(\partial_\mu \varphi)(\partial^\mu \varphi)$ is just $(\partial \varphi)^2$, so we can treat this just like the derivative of a quantity squared, giving

$$\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} = \partial^\mu \varphi. \tag{3.5.3}$$

We end with an upper index, since the left hand side has a lower index in the denominator, which is an upper index in the numerator. We then have

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \varphi} \right) = \partial_\mu \partial^\mu \varphi = \partial^2 \varphi. \tag{3.5.4}$$

Using the Euler–Lagrange equations we then have

$$\partial^2 \varphi + m^2\varphi = 0, \tag{3.5.5}$$

which is just the Klein–Gordon equation.

The canonical momentum associated with this Lagrangian is

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = \frac{\partial \mathcal{L}}{\partial (\partial_0 \varphi)} = \partial^0 \varphi = \dot{\varphi}. \tag{3.5.6}$$

The Hamiltonian is then

$$\mathcal{H} = \pi\dot{\varphi} - \mathcal{L} = \frac{1}{2}(\pi^2 + (\nabla\varphi)^2 + m^2\varphi^2). \tag{3.5.7}$$

The $\nabla\varphi$ term comes from

$$\pi^2 - \frac{1}{2}(\partial_\mu\varphi)(\partial^\mu\varphi) = \pi^2 - \frac{1}{2}[(\partial_0\varphi)(\partial^0\varphi) - (\partial_i\varphi)(\partial^i\varphi)] \tag{3.5.8}$$

$$= \pi^2 - \frac{1}{2}[\dot{\varphi}^2 - (\nabla\varphi)^2] = \frac{1}{2}[\pi^2 - (\nabla\varphi)^2] \tag{3.5.9}$$

where we've used $\dot{\varphi} = \pi$.

Compare the Hamiltonian density for the Klein–Gordon equation with the Hamiltonian for the harmonic oscillator:

$$H = \frac{1}{2m}(p^2 + m^2\omega^2 x^2). \tag{3.5.10}$$

Up to a factor of ω^2 and $1/m$ this is pretty much the same as the Klein–Gordon Hamiltonian density, minus the $(\nabla\varphi)^2$ term. Similarly the Klein–Gordon Lagrangian density is very similar to the harmonic oscillator Lagrangian. This suggests that we should study the harmonic oscillator, and indeed we shall soon.

Example 3.5.11 — Electromagnetism The Lagrangian density for the electromagnetic field, A^μ, is

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{3.5.12}$$

We can insert the definition of $F^{\mu\nu}$ and expand this out to get

$$\mathcal{L} = -\frac{1}{4}(\partial^\mu A^\nu - \partial^\nu A^\mu)(\partial_\mu A_\nu - \partial_\nu A_\mu) \tag{3.5.13}$$

$$= -\frac{1}{4}[(\partial^\mu A^\nu)(\partial_\mu A_\nu) - (\partial^\nu A^\mu)(\partial_\mu A_\nu) \tag{3.5.14}$$

$$- (\partial^\mu A^\nu)(\partial_\nu A_\mu) + (\partial^\nu A^\mu)(\partial_\nu A_\mu)] \tag{3.5.15}$$

$$= -\frac{1}{2}(\partial^\mu A^\nu)(\partial_\mu A_\nu) + \frac{1}{2}(\partial^\mu A^\nu)(\partial_\nu A_\mu). \tag{3.5.16}$$

We treat each component of A_ν as an independent field. We then have

$$\frac{\partial\mathcal{L}}{\partial A_\nu} = 0, \quad \text{and} \quad \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)} = -\partial^\mu A^\nu + \partial^\nu A^\mu = -F^{\mu\nu}. \tag{3.5.17}$$

Hence, the Euler–Lagrange equations give

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)}\right) = -\partial_\mu\partial^\mu A^\nu + \partial^\nu\partial_\mu A^\mu = 0. \tag{3.5.18}$$

That is,

$$\partial^2 A^\nu - \partial^\nu\partial_\mu A^\mu = 0 \iff -\partial_\mu F^{\mu\nu} = 0 \tag{3.5.19}$$

This is Maxwell's equation in a vacuum in an arbitrary gauge.

The canonical momentum is

$$\pi^\mu(x) = \frac{\partial\mathcal{L}}{\partial\dot{A}_\mu} = \frac{\partial\mathcal{L}}{\partial(\partial_0 A_\mu)} = -\partial^\mu A^0 + \partial^0 A^\mu = -F^{0\mu}. \tag{3.5.20}$$

Antisymmetry of $F^{\mu\nu}$ implies that $\pi^\mu(x) = 0$, and so $\pi^\mu(x) = (0, -\boldsymbol{E}(x))$, where $\boldsymbol{E}(x)$ is the electric field.

The fact that $\pi^0(x) = 0$ will cause problems later when we try to quantise the electromagnetic field.

In the presence of sources the Lagrangian density is instead

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - J^\mu A_\mu. \tag{3.5.21}$$

This doesn't change the term given by differentiating with respect to $\partial_\mu A_\nu$, all that changes is we now have

$$\frac{\partial\mathcal{L}}{\partial A_\mu} = -J^\mu, \tag{3.5.22}$$

and so the result of applying the Euler–Lagrange equations is now

$$\partial^2 A^\nu - \partial^\nu \partial_\mu A^\mu = J^\nu, \tag{3.5.23}$$

which is Maxwell's equation in the presence of a source.

3.6 Symmetries

3.6.1 Discrete Symmetries

There are three discrete symmetries of particular interest in quantum field theory, they are

- **parity**: P : $\boldsymbol{x} \mapsto -\boldsymbol{x}$, P : $x^\mu = (x^0, x^i) \mapsto (x^0, -x^i) = (x_0, x_i) = x_\mu$, so parity acts to swap raised and lowered indices of positions;

- **time reversal**: T : $t \mapsto t$, T : $x^\mu = (x^0, x^i) \mapsto (-x^0, x^i) = -(x^0, -x^i) = -(x_0, x_i) = -x_\mu$, so time reversal acts to swap raised and lowered indices and negate positions;

- **charge conjugation**: C : $e \mapsto -e$, where e is the charge of the particle.

Consider, for example, the Lagrangian of the Klein–Gordon equation:

$$\mathcal{L} = \frac{1}{2}(\partial^\mu\varphi)(\partial_\mu\varphi) - \frac{1}{2}m^2\varphi^2. \tag{3.6.1}$$

Under parity the scalar field is unchanged, and $\partial^\mu \leftrightarrow \partial_\mu$, so the Lagrangian is unchanged under parity. Under time reversal the scalar field is unchanged, and $\partial^\mu \leftrightarrow -\partial_\mu$, so the Lagrangian is unchanged under time reversal. Since this Lagrangian has no charges involved it is trivially invariant under charge conjugation.

Now consider the Electromagnetic Lagrangian,

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - J^\mu A_\mu. \tag{3.6.2}$$

Under parity transformations $F^{\mu\nu} \leftrightarrow F_{\mu\nu}$, $J^\mu \to J_\mu$, and $A_\mu \to A^\mu$, so

$$\text{P}\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - J_\mu A^\mu = \mathcal{L}, \tag{3.6.3}$$

so the Lagrangian is unchanged by parity. Under time reversal transformations $F^{\mu\nu} \leftrightarrow -F_{\mu\nu}$ since the derivatives in the definition of $F^{\mu\nu}$ transform as $\partial^{\mu} \to -\partial_{\mu}$ under time reversal. On the other hand, $J^{\mu} \to J^{\mu}$, and $A_{\mu} \to A_{\mu}$, since time reversal doesn't effect the current or potential. Hence,

$$\mathrm{T}\mathcal{L} = -\frac{1}{4}(-F_{\mu\nu})(-F^{\mu\nu}) - J^{\mu}A_{\mu} = \mathcal{L}, \tag{3.6.4}$$

so the Lagrangian is unchanged under time reversal. Under charge conjugation $F^{\mu\nu} \leftrightarrow -F_{\mu\nu}$, $J^{\mu} \to -J_{\mu}$, and $A^{\mu} \to -J_{\mu}$, so

$$\mathrm{C}\mathcal{L} = -\frac{1}{4}(-F_{\mu\nu})(-F^{\mu\nu}) - (-J_{\mu})(-A^{\mu}) = \mathcal{L}, \tag{3.6.5}$$

so the Lagrangian is invariant under charge conjugation.

Invariance under each of these symmetries means that both of these Lagrangians are invariant under the combined CPT symmetry, given by $x^{\mu} \mapsto -x^{\mu}$ and $e \mapsto -e$.

3.6.2 Continuous Symmetries

(R)   For more details see the *Symmetries of Particles and Fields* course.

If the action is invariant under a group of continuous symmetries then we have a conserved quantity. This is the essence of **Noether's theorem**. The most common examples being

- invariance under translations in time leading to energy conservation;

- invariance under spatial translations leading to momentum conservation;

- invariance under rotation leading to angular momentum conservation.

The simplest case is, perhaps, when the Lagrangian is invariant under translations, then the action is also necessarily translation invariant. Suppose we have a translation

$$\varphi(x) \to \varphi'(x) = \varphi(x) + \delta\varphi(x). \tag{3.6.6}$$

Note that the translation is allowed to depend on x. The variation in $\mathcal{L}$ is

$$\delta\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \varphi}\delta\varphi + \frac{\partial \mathcal{L}}{\partial(\partial_{\mu})}\delta(\partial_{\mu}\varphi) \tag{3.6.7}$$

$$= \partial_{\mu}\left(\frac{\partial \mathcal{L}}{\partial(\partial_{\mu}\varphi)}\right)\delta\varphi + \frac{\partial \mathcal{L}}{\partial(\partial_{\mu}\varphi)}\delta(\partial_{\mu}\varphi) \tag{3.6.8}$$

$$= \partial_{\mu}\left(\frac{\partial \mathcal{L}}{\partial(\partial_{\mu}\varphi)}\delta\varphi\right). \tag{3.6.9}$$

The second equality is simply an application of the Euler–Lagrange equations to rewrite the first term. The last equality is just recognising the product rule.

Imposing that $\mathcal{L}$ is invariant under translations like this we have $\delta\mathcal{L} = 0$, which gives us the continuity equation

$$\partial_{\mu}J^{\mu}(x) = 0, \tag{3.6.10}$$

where

$$J^\mu(x) := \frac{\partial \mathcal{L}}{\partial(\partial_\mu)} \delta\varphi \tag{3.6.11}$$

is the conserved current, a generalisation of the normal current in electrodynamics. There is also a corresponding conserved "charge":

$$Q = \int d^3\boldsymbol{x}\, J^0(x). \tag{3.6.12}$$

To see that this quantity is conserved consider

$$\dot{Q} = \int d^3\boldsymbol{x}\, \partial_0 J^0(x) = - \int d^3\boldsymbol{x}\, \partial_i J^i(x) = 0 \tag{3.6.13}$$

where in the last step we've used the divergence theorem to rewrite the last term as a surface integral, then take this surface to be at infinity, and made the usual assumption that fields vanish sufficiently quickly at infinity.

> **Example 3.6.14 — Probability Current** Consider the Lagrangian
>
> $$\mathcal{L} = i\psi^* \partial_t \psi - \frac{1}{2m}(\nabla\psi^*) \cdot (\nabla\psi) - V(\boldsymbol{x}, t)\psi^*\psi. \tag{3.6.15}$$
>
> If we treat ψ and ψ^* as separate fields then varying one of them gives the Schrödinger equation. This Lagrangian is invariant under the transformation
>
> $$\psi \to e^{i\alpha}\psi, \quad \text{and} \quad \psi^* \to e^{-i\alpha}\psi^* \tag{3.6.16}$$
>
> where α is a constant. Linearising we get the transformation
>
> $$\psi \to \psi(1 + i\alpha), \quad \text{and} \quad \psi^* \to \psi^*(1 - i\alpha). \tag{3.6.17}$$
>
> Since we have two fields, ψ and ψ^*, the conserved current is a sum over field:
>
> $$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \delta\psi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi^*)} \delta\psi^*. \tag{3.6.18}$$
>
> Computing the derivatives we have
>
> $$\frac{\partial \mathcal{L}}{\partial(\partial_t \psi)} = i\psi^*, \qquad \frac{\partial \mathcal{L}}{\partial(\partial_i \psi)} = -\frac{1}{2m}\partial^i\psi^*, \tag{3.6.19}$$
>
> $$\frac{\partial \mathcal{L}}{\partial(\partial_t \psi^*)} = 0, \qquad \frac{\partial \mathcal{L}}{\partial(\partial_i \psi^*)} = -\frac{1}{2m}\partial^i\psi. \tag{3.6.20}$$
>
> Hence, using the variations
>
> $$\delta\psi = i\alpha\psi, \quad \text{and} \quad \delta\psi^* = -i\alpha\psi^* \tag{3.6.21}$$
>
> we have the time component of the conserved current:
>
> $$J^0 = \rho = -\alpha\psi^*\psi, \tag{3.6.22}$$

and the position components:

$$J^i = \frac{i\alpha}{2m}(\psi^*\partial^i\psi - \psi\partial^i\psi^*) \qquad (3.6.23)$$

$$\boldsymbol{J} = \frac{i\alpha}{2m}(\psi\nabla\psi^* - \psi^*\nabla\psi). \qquad (3.6.24)$$

We can then identify J^μ as, up to a constant, the conserved probability current of Section A.3.2.

It can be shown that the Euler–Lagrange equations are invariant under the addition of a four-divergence,

$$\mathcal{L} \to \mathcal{L} + \partial_\mu\Lambda^\mu \qquad (3.6.25)$$

for some four-vector field Λ^μ with a sufficiently smooth derivative. The same derivation as before, but setting $\delta\mathcal{L} = \partial_\mu\Lambda^\mu$ instead of zero, then gives a conserved current

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta\varphi - \Lambda^\mu. \qquad (3.6.26)$$

As well as changing the Lagrangian directly we can instead act on spacetime, for example, by translation:

$$x^\mu \to x'^\mu = x^\mu + a^\mu \qquad (3.6.27)$$

for some infinitesimal four-vector a^μ. This induces a change in the scalar field, given by Taylor expanding:

$$\varphi(x) \to \varphi(x+a) = \varphi(x) + a^\mu\partial_\mu\varphi(x). \qquad (3.6.28)$$

Since the Lagrangian density is also a scalar it must transform the same way:

$$\mathcal{L} \to \mathcal{L} + a^\mu\partial_\mu\mathcal{L} = \mathcal{L} + a^\nu\partial_\mu(\delta^\mu_{\ \nu}\mathcal{L}). \qquad (3.6.29)$$

Comparing this to the transformation $\mathcal{L} \to \mathcal{L} + \partial_\mu\Lambda^\mu$ we can identify

$$\partial_\mu\Lambda^\mu = a^\nu\partial_\mu(\delta^\mu_\nu\mathcal{L}). \qquad (3.6.30)$$

We then have four conserved currents, one for each value of ν, such as

$$J^\mu(x) = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}a^0\partial_0\varphi - a^0\delta^\mu_{\ 0}\mathcal{L}, \qquad (3.6.31)$$

which comes from setting $\nu = 0$. We can combine these four currents into a single rank 2 tensor, scaling out the a^ν parameter, to get

$$T^\mu_{\ \nu} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\mathcal{L})}\partial_\nu\varphi - \mathcal{L}\delta^\mu_{\ \nu}. \qquad (3.6.32)$$

Raising the index we get

$$T^{\mu\nu} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\partial^\nu\varphi - \mathcal{L}\eta^{\mu\nu}. \qquad (3.6.33)$$

This is the **energy-momentum tensor**, also called the **stress-energy tensor**[1].

The conserved "charge" associated with the $\nu = 0$ component is

$$\int \mathrm{d}^3\boldsymbol{x}\, T^{00} = \int \left(\frac{\partial \mathcal{L}}{\partial(\partial_0\varphi)}\partial^0\varphi - \mathcal{L} \right) = \int \mathrm{d}^3\boldsymbol{x}\, \mathcal{H} = H, \tag{3.6.34}$$

so this is an expression of energy conservation. This shouldn't be surprising as translations by a^0 correspond to time translations.

Similarly the conserved charges associated with the $\nu = i$ components are

$$P^i = \int \mathrm{d}^3\boldsymbol{x}\, T^{0i} = \int \mathrm{d}^3\boldsymbol{x}\, \pi\partial^i\varphi. \tag{3.6.35}$$

Since we expect these three components to form a four-vector with time component given by H we interpret them as the actual momentum of the particle, as opposed to the conjugate momentum, π. We can then interpret $\pi\partial_i\varphi$ as the momentum density.

[1] Strictly this is the energy-momentum density tensor, but very rarely does anyone bother to make that distinction.

Four

Quantised Fields

Quantum field theory is a quantum theory of fields. We've seen classical fields, now we need some quantum mechanics before we can start quantising fields. So, in this section we'll give a quick recap of quantum mechanics. For more details see the *Principles of Quantum Mechanics* and *Quantum Theory* courses.

4.1 Quantum Mechanics

It's slightly disturbing that people believe it.

Richard Ball, on quantum mechanics

4.1.1 The Basics

In quantum mechanics we start with coordinates q_i, and a Lagrangian, $L(q_i, \dot{q}_i)$. We then define canonical momenta,

$$p_i := \frac{\partial L}{\partial q_i}. \tag{4.1.1}$$

From this we can find a Hamiltonian,

$$H = \sum_i p + i\dot{q}_i - L. \tag{4.1.2}$$

[1] An operator, A, is Hermitian if $A^\dagger = A$, where $A^\dagger$ is the **Hermitian conjugate**, defined such that $\langle\psi|A|\varphi\rangle = \langle\varphi|A^\dagger|\psi\rangle^*$.

So far this could all be classical. The quantum mechanics begins when we interpret q_i, p_i, and H as **Hermitian operators**[1] acting on a Hilbert space of states. The Hermitian requirement is so that these operators have real eigenvalues, which we can then interpret as the results of measurements. For example, the average energy is given by

$$E = \langle\psi, t|H|\psi, t\rangle, \tag{4.1.3}$$

where $|\psi, t\rangle$ is the state of the particle, that is a vector in the Hilbert space, and $\langle\psi, t| = |\psi, t\rangle^\dagger$ is the corresponding vector in the dual space.

The next step is to realise that operators don't necessarily commute, and so we impose the **canonical commutation relation (CCR)**:

$$[q_i, p_i] := i\delta_{ij}, \tag{4.1.4}$$

where $[A, B] := AB - BA$ is the **commutator**. As well as this we impose that the positions commute, and the momenta commute, so

$$[q_i, q_j] = [p_i, p_j] = 0. \tag{4.1.5}$$

4.1.2 Time Dependence

The time evolution of a state, $|\psi, t\rangle$, under a system with Hamiltonian H is given by the **Schrödinger equation**

$$i\frac{\mathrm{d}}{\mathrm{d}t}|\psi, t\rangle = H|\psi, t\rangle. \tag{4.1.6}$$

This is a fairly simple differential equation, if we ignore the fact that H is an operator and $|\psi, t\rangle$ is a vector, the solution is just

$$|\psi, t\rangle = \mathrm{e}^{-iHt}|\psi\rangle \tag{4.1.7}$$

where $|\psi\rangle = |\psi, 0\rangle$ is the initial state of the system. Fortunately this is perfectly valid, even though H *is* an operator and $|\psi, t\rangle$ *is* vector, we just have to interpret the exponential through its power series:

$$\mathrm{e}^{-iHt} := \sum_{n=0}^{\infty} \frac{1}{n!}(-iHt)^n. \tag{4.1.8}$$

This scheme with which we have worked so far, where operators are time independent and states are time dependent, is called the **Schrödinger picture**. It turns out that its actually easier to do QFT by interpreting operators to be time dependent and states to be time independent, called the **Heisenberg picture**. We distinguish between these two pictures by either including t or not, as appropriate in our operators and states. A state, $|\psi\rangle$, in the Heisenberg picture is related to the state $|\psi, t\rangle$ in the Schrödinger picture by

$$|\psi\rangle = \mathrm{e}^{iHt}|\psi, t\rangle, \tag{4.1.9}$$

which is just what we get rearranging Equation (4.1.7). An operator, $A(t)$, in the Heisenberg picture is related to the operator A in the Schrödinger picture by

$$A(t) = \mathrm{e}^{iHt}A\mathrm{e}^{-iHt}. \tag{4.1.10}$$

Importantly, expectation values are the same in both pictures, and so both pictures describe the same physics:

$$\langle A \rangle_{\mathrm{S}} = \langle \psi, t | A | \psi, t \rangle \tag{4.1.11}$$
$$= \langle \psi | \mathrm{e}^{iHt} A \mathrm{e}^{-iHt} \tag{4.1.12}$$
$$= \langle \psi | A(t) | \psi \rangle \tag{4.1.13}$$
$$= \langle A \rangle_{\mathrm{H}}, \tag{4.1.14}$$

where the subscripts refer to the picture in which we interpret the expectation value.

This picture changing is the same for all operators, so in particular, the position and momentum in the Heisenberg picture are given by

$$q_i(t) = \mathrm{e}^{iHt}q_i\mathrm{e}^{-iHt}, \qquad \text{and} \qquad p_i(t) = \mathrm{e}^{iHt}p_i\mathrm{e}^{-iHt}. \tag{4.1.15}$$

We also have

$$H(t) = \mathrm{e}^{iHt} H \mathrm{e}^{-iHt} = \mathrm{e}^{iHt} \mathrm{e}^{-iHt} H = H, \tag{4.1.16}$$

which works since H commutes with itself, and the exponentials are just power series in H, so also commute with H. This shows that the Hamiltonian is time independent in both pictures, which is a statement of energy conservation.

Consider two operators, $A(t)$ and $B(t)$, in the Heisenberg picture. If we know their commutator, $[A, B]$, in the Schrödinger picture then we can also compute it in the Heisenberg picture:

$$[A(t), B(t)] = \langle \mathrm{e}^{iHt} A \mathrm{e}^{-iHt} | \mathrm{e}^{iHt} B \mathrm{e}^{-Ht} \rangle \tag{4.1.17}$$

$$= \mathrm{e}^{iHt} A \mathrm{e}^{-iHt} \mathrm{e}^{iHt} B \mathrm{e}^{-iHt} - \mathrm{e}^{iHt} B \mathrm{e}^{-iHt} \mathrm{e}^{iHt} A \mathrm{e}^{iHt} \tag{4.1.18}$$

$$= \mathrm{e}^{iHt} A B \mathrm{e}^{-iHt} - \mathrm{e}^{iHt} B A \mathrm{e}^{-iHt} \tag{4.1.19}$$

$$= \mathrm{e}^{iHt} (AB - BA) \mathrm{e}^{-iHt} \tag{4.1.20}$$

$$= \mathrm{e}^{iHt} [A, B] \mathrm{e}^{iHt}. \tag{4.1.21}$$

In particular,

$$[q_i(t), p_j(t)] = \mathrm{e}^{iHt} [q_i, p_j] \mathrm{e}^{-iHt} = \mathrm{e}^{iHt} i\delta_{ij} \mathrm{e}^{-iHt} = i\delta_{ij} \mathrm{e}^{iHt} \mathrm{e}^{-iHt} = i\delta_{ij}, \tag{4.1.22}$$

so the canonical commutation relations hold if both operators are evaluated at the same time. Similarly,

$$[q_i(t), q_j(t)] = [p_i(t), p_j(t)] = 0. \tag{4.1.23}$$

In the Heisenberg picture instead of the Schrödinger equation we have the **Heisenberg equation**, which we can derive by considering the time derivative of a generic operator, $A(t)$:

$$\frac{\mathrm{d}}{\mathrm{d}t} A(t) = \frac{\mathrm{d}}{\mathrm{d}t} (\mathrm{e}^{iHt} A \mathrm{e}^{-iHt}) \tag{4.1.24}$$

$$= iH \mathrm{e}^{iHt} A \mathrm{e}^{-iHt} + \mathrm{e}^{iHt} A (-iH) \mathrm{e}^{-iHt} \tag{4.1.25}$$

$$= i \mathrm{e}^{iHt} H A \mathrm{e}^{-iHt} - i \mathrm{e}^{iHt} A H \mathrm{e}^{-iHt} \tag{4.1.26}$$

$$= i \mathrm{e}^{iHt} [H, A] \mathrm{e}^{-iHt} \tag{4.1.27}$$

$$= i [H, A(t)]. \tag{4.1.28}$$

That is,

$$\frac{\mathrm{d}}{\mathrm{d}t} A(t) = i [H, A(t)]. \tag{4.1.29}$$

The Heisenberg equation applied to the position and momentum gives

$$\frac{\mathrm{d}}{\mathrm{d}t} q(t) = i [H, q(t)], \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}t} p(t) = i [H, p(t)]. \tag{4.1.30}$$

Note the similarity to Equation (3.2.11). In general when we work in the Heisenberg picture things can look pretty similar to classical mechanics, but with i times the commutator in place of Poisson brackets.

Example 4.1.31 Consider the Lagrangian for a particle of mass m in a position dependent potential, V, in one dimension:

$$L = \frac{1}{2}m\dot{q}^2 - V(q). \tag{4.1.32}$$

The Hamiltonian for this system is

$$H = \frac{p^2}{2m} + V(q). \tag{4.1.33}$$

We then have

$$\dot{q} = i[H, q(t)] = \frac{i}{2m}[p^2, q]. \tag{4.1.34}$$

Now, consider three operators, A, B, and C. We have

$$[AB, C] = ABC - CBA \tag{4.1.35}$$
$$= ABC - ACB + ACB - CBA \tag{4.1.36}$$
$$= A[B, C] + [A, C]B. \tag{4.1.37}$$

Using this with $A = B = p$ and $C = q$ gives

$$[p^2, q] = p[p, q] + [p, q]p = -2ip, \tag{4.1.38}$$

where the negative comes from the antisymmetry of the commutator, $[p, q] = -[q, p] = -i$. Hence,

$$\dot{q} = \frac{p}{2m}. \tag{4.1.39}$$

Classically we would have $p = mv = m\dot{q}$, and this is just the quantum analogue of the classical nonrelativistic momentum.
We also have

$$\dot{p} = i[H, p(t)] = i[V(q), p]. \tag{4.1.40}$$

We can expand $V(q)$ as a power series in q, we also have the identity

$$[q^n, p] = inq^{n-1}. \tag{4.1.41}$$

This can be proven with induction on n. First, take $n = 0$, then

$$[q^0, p] = [1, p] = 0. \tag{4.1.42}$$

Now suppose that

$$[q^k, p] = ikq^{k-1} \tag{4.1.43}$$

for some nonnegative integer k. Then

$$[q^{k+1}, p] = [qq^k, p] = q[q^k, p] + [q, p]q^k$$
$$= qikq^{k-1} + iq^k = i(k+1)q^k, \tag{4.1.44}$$

and so by induction the identity holds for all natural numbers, n.
Expanding V as a Taylor series we have

$$V(q) = \sum_{n=0}^{\infty} \frac{q^n}{n!} \frac{\partial^n V}{\partial q^n}\bigg|_{q=0}. \tag{4.1.45}$$

Then we have

$$[V(q), p] = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\partial^n V}{\partial q^n}\bigg|_{q=0} [q^n, p] \tag{4.1.46}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\partial^n V}{\partial q^n}\bigg|_{q=0} inq^{n-1} \tag{4.1.47}$$

$$= \sum_{n=1}^{\infty} \frac{iq^{n-1}}{(n-1)!} \frac{\partial^n V}{\partial q^n}\bigg|_{q=0}. \tag{4.1.48}$$

To first order we then have

$$[V(q), p] = i\frac{\partial^n V}{\partial q^n}\bigg|_{q=0} \tag{4.1.49}$$

and so

$$\dot{p} = i[V(q), p] = -\frac{\partial V}{\partial q}. \tag{4.1.50}$$

Compare this to the classical nonrelativistic equation

$$F = -\frac{\partial V}{\partial x}, \tag{4.1.51}$$

expressing the force due to a potential.
Both of these two analogies, known as **Ehrenfest's theorem**, between classical mechanics and quantum mechanics are really due to the similarity between the classical and quantum equations of motion in the Heisenberg picture.

4.2 Quantum Fields

Fundamentally silly

Richard Ball

The procedure for quantising fields is almost identical to the procedure for quantising position and momentum. Instead of the position and momentum we have the fields φ and π. We interpret these as operators in the Heisenberg picture. Since in field theory we deal with densities instead to be integrated over, such as $\mathcal{L}$ instead of L, we replace δ_{ij} in the canonical commutation relations with $\delta^3(\boldsymbol{x}-\boldsymbol{x}')$, note that this treats time differently to space. We postulate that the at some time,

t, we have

$$[\varphi(t,\boldsymbol{x}),\pi(t,\boldsymbol{x}')] = i\delta^3(\boldsymbol{x}-\boldsymbol{x}') \tag{4.2.1}$$

and the fields commute with themselves at all positions:

$$[\varphi(t,\boldsymbol{x}),\varphi(t,\boldsymbol{x}')] = [\pi(t,\boldsymbol{x}),\pi(t,\boldsymbol{x}')] = 0. \tag{4.2.2}$$

This is called the **equal time commutation relation (ETCR)**.

In quantum mechanics as done in the previous section we treat t as a parameter and $\boldsymbol{x}$ as an operator. This is clearly not Lorentz covariant. There are also other problems, such as having only a single $\boldsymbol{x}$ operator which only allows us to talk of a single particle. In quantum field theory we interpret both t and $\boldsymbol{x}$ as parameters and φ as an operator. We can then make this Lorentz covariant and treat systems with multiple particles.

4.2.1 Equations of Motion

The Heisenberg equation,

$$\dot{A}(t,\boldsymbol{x}) = i[H,A(t,\boldsymbol{x})], \tag{4.2.3}$$

applies to any operator, A, and hence applies to φ. We then have

$$\dot{\varphi}(t,\boldsymbol{x}) = i[H,\varphi(t,\boldsymbol{x})] = i\int \mathrm{d}^3\boldsymbol{x}'\,[\mathcal{H}(t,\boldsymbol{x}'),\varphi(t,\boldsymbol{x})] \tag{4.2.4}$$

We've replaced the Hamiltonian with an integral over the Hamiltonian density. We'll now introduce a shorthand notation where primes denote quantities evaluated at $(t,\boldsymbol{x}')$ and an absence of a prime is a quantity evaluated at $(t,\boldsymbol{x})$, so $\varphi' = \varphi(t,\boldsymbol{x}')$ and $\varphi = \varphi(t,\boldsymbol{x})$. The Hamiltonian density is

$$\mathcal{H}' = \frac{1}{2}(\pi'^2 + (\nabla'\varphi')^2 + m^2\varphi'). \tag{4.2.5}$$

The commutator then becomes

$$[\mathcal{H}',\varphi] = \frac{1}{2}\left([\pi'^2,\varphi] + [(\nabla'\varphi')^2,\varphi] + m^2[\varphi'^2,\varphi]\right). \tag{4.2.6}$$

The last two terms vanish, since the fields commute with themselves at all positions. The first term isn't too bad if we use

$$[AB,C] = A[B,C] + [A,C]B, \tag{4.2.7}$$

we have

$$\frac{1}{2}[\pi'^2,\varphi] = \frac{1}{2}\pi'[\pi',\varphi] + \frac{1}{2}[\pi',\varphi]\pi' \tag{4.2.8}$$

$$= -\frac{1}{2}\pi'[\varphi,\pi'] - \frac{1}{2}[\varphi,\pi']\pi' \tag{4.2.9}$$

$$= -\frac{1}{2}\pi'i\delta^3(\boldsymbol{x}-\boldsymbol{x}') - \frac{1}{2}i\delta^3(\boldsymbol{x}-\boldsymbol{x}')\pi' \tag{4.2.10}$$

$$= -i\pi'\delta^3(\boldsymbol{x}-\boldsymbol{x}'). \tag{4.2.11}$$

Hence, we have the equation of motion

$$\dot{\varphi}(t, \boldsymbol{x}) = i \int d^3\boldsymbol{x}'[\mathcal{H}(t, \boldsymbol{x}'), \varphi(t, \boldsymbol{x})] \tag{4.2.12}$$

$$= i \int d^3\boldsymbol{x}'(-i\pi(t, \boldsymbol{x}')\delta^3(\boldsymbol{x} - \boldsymbol{x}')) \tag{4.2.13}$$

$$= \pi(t, \boldsymbol{x}). \tag{4.2.14}$$

So, $\dot{\varphi} = \pi$. This should be no surprise since we found this same relationship for φ satisfying the Klein–Gordon equation before we started quantising.

Similarly we can find the equation of motion for π:

$$\dot{\pi}(t, \boldsymbol{x}) = i[H, \pi(t, \boldsymbol{x})] = i \int d^3\boldsymbol{x}'[\mathcal{H}(t, \boldsymbol{x}'), \pi(t, \boldsymbol{x})]. \tag{4.2.15}$$

The commutator in this case is

$$[\mathcal{H}', \pi] = \frac{1}{2}([\pi'^2, \pi] + [(\nabla'\varphi')^2, \pi] + m^2[\varphi'^2, \pi]). \tag{4.2.16}$$

The first term vanishes this time. The last term can be computed easily:

$$\frac{1}{2}m^2[\varphi'^2, \pi] = \frac{1}{2}m^2\varphi'^2[\varphi', \pi] + \frac{1}{2}m^2[\varphi', \pi]\varphi' \tag{4.2.17}$$

$$= \frac{1}{2}m^2 i\delta^3(\boldsymbol{x} - \boldsymbol{x}') + \frac{1}{2}m^2\varphi' i\delta^3(\boldsymbol{x} - \boldsymbol{x}') \tag{4.2.18}$$

$$= im^2\delta^3(\boldsymbol{x} - \boldsymbol{x}'). \tag{4.2.19}$$

The second term isn't too bad, we just have to notice that ∇' acts on $\boldsymbol{x}'$ and not on $\boldsymbol{x}$, allowing us to pull it outside of commutators like $[\nabla'\varphi', \pi] = \nabla'[\varphi', \pi]$. Proceeding as before we then have

$$\frac{1}{2}[(\nabla'\varphi')^2, \pi] = \frac{1}{2}(\nabla'\varphi') \cdot [\nabla'\varphi', \pi] + \frac{1}{2}[\nabla'\varphi', \pi] \cdot (\nabla'\varphi') \tag{4.2.20}$$

$$= \frac{1}{2}(\nabla'\varphi') \cdot (\nabla'[\varphi', \pi]) + \frac{1}{2}(\nabla'[\varphi', \pi]) \cdot (\nabla'\varphi') \tag{4.2.21}$$

$$= \frac{1}{2}(\nabla'\varphi') \cdot (\nabla' i\delta^3(\boldsymbol{x} - \boldsymbol{x}')) + \frac{1}{2}(\nabla' i\delta^3(\boldsymbol{x} - \boldsymbol{x}')) \cdot (\nabla'\varphi')$$

$$= i(\nabla'\varphi') \cdot (\nabla'\delta^3(\boldsymbol{x} - \boldsymbol{x}')). \tag{4.2.22}$$

We can deal with the gradient of the delta distribution by integrating by parts, and assuming that fields vanish sufficiently quickly at infinity:

$$\int d^3\boldsymbol{x}'(\nabla'\varphi) \cdot (\nabla'\delta^3(\boldsymbol{x} - \boldsymbol{x}')) = \int d^3\boldsymbol{x}' \, \nabla' \cdot [\delta^3(\boldsymbol{x} - \boldsymbol{x}')(\nabla'\varphi')] \tag{4.2.23}$$

$$- \int d^3\boldsymbol{x}' \, (\nabla'^2\varphi')\delta^3(\boldsymbol{x} - \boldsymbol{x}') \tag{4.2.24}$$

$$= \int_{\partial V} [\delta^3(\boldsymbol{x} - \boldsymbol{x}')\nabla'\varphi'] \tag{4.2.25}$$

$$- \int d^3\boldsymbol{x}' \, (\nabla'^2\varphi')\delta^3(\boldsymbol{x} - \boldsymbol{x}') \tag{4.2.26}$$

$$= - \int d^3\boldsymbol{x}' \, (\nabla'^2\varphi')\delta^3(\boldsymbol{x} - \boldsymbol{x}') \tag{4.2.27}$$

where as usual the boundary term vanishes. This means that, as distributions,

$$i(\nabla'\varphi') \cdot (\nabla'\delta^3(\boldsymbol{x} - \boldsymbol{x}')) = -i(\nabla'^2\varphi')\delta^3(\boldsymbol{x} - \boldsymbol{x}'). \tag{4.2.28}$$

Hence, we have

$$\dot{\pi}(t, \boldsymbol{x}) = i \int \mathrm{d}^3\boldsymbol{x}' [\mathcal{H}(t, \boldsymbol{x}'), \varphi(t, \boldsymbol{x})] \tag{4.2.29}$$

$$= i \int \mathrm{d}^3\boldsymbol{x}' \left[-i(\nabla'^2\varphi')\delta^3(\boldsymbol{x} - \boldsymbol{x}') + im^2\delta^3(\boldsymbol{x} - \boldsymbol{x}') \right] \tag{4.2.30}$$

$$= \nabla^2\varphi - m^2\varphi. \tag{4.2.31}$$

Now, we have $\pi = \dot{\varphi}$, so $\dot{\pi} = \ddot{\varphi} = \partial_0\partial^0\varphi$. Hence,

$$\partial_0\partial^0\varphi = \nabla^2\varphi - m^2\varphi = \partial_i\partial^i\varphi - m^2\varphi. \tag{4.2.32}$$

Rearranging this we get

$$0 = (\partial_0\partial^0 - \partial_i\partial^i)\varphi + m^2\varphi = (\partial^2 + m^2)\varphi, \tag{4.2.33}$$

so the quantised fields satisfy the Klein–Gordon equation. This means that the Heisenberg equations for field operators look a lost like classical field equations with operators replacing the fields and i times the commutator replacing Poisson brackets.

Even though the Hamiltonian formulation and the equal time commutation relations aren't Lorentz covariant the final result is, and that's all that we care about. The steps in the derivation, however, are frame dependent. This is an inescapable problem with canonical quantisation.

Five

Quantum Harmonic Oscillator

By understanding the harmonic oscillator you understand everything! Well, accept from all the things we can't solve.

Richard ball

Surely the hydrogen atom isn't a harmonic oscillator? Well it is!

Richard Ball

The quantum harmonic oscillator is one of the most important systems in quantum mechanics. It's everywhere, we use it to model atoms and phonons, molecules, and pretty much anything else. Even other systems we study in quantum mechanics, such as the hydrogen atom or pairs of particles, can be interpreted as a harmonic oscillator by a change of coordinates. After all, any potential in physics can always be Taylor expanded to second order at a minimum to get a harmonic potential. Part of the reason why we study the harmonic oscillator is that its very nice. We can solve it exactly. The results are also nice, evenly spaced energy levels, bounded below.

The harmonic oscillator is the system with the Hamiltonian

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2. \tag{5.0.1}$$

We have previously remarked on the similarity of this and the Klein–Gordon Hamiltonian density, and suggested that study of the harmonic oscillator could be enlightening for the Klein–Gordon equation. In this section we'll study the harmonic oscillator, computing the spectrum of the Hamiltonian, and discussing various interpretations. In the next chapter we'll go back through the same steps but for fields obeying the Klein–Gordon and the Hamiltonian density.

5.1 Solution

There are two ways to solve the harmonic oscillator: the analytic way, involving Hermite polynomials, which is frankly, horrible, and the algebraic way, which is

maybe more mysterious but is much more inline with the quantum field theory interpretation, so this is the way we'll proceed.

Start by defining a new operators from the position and momentum operators:

$$a := \sqrt{\frac{m\omega}{2}}q + \frac{i}{\sqrt{2m\omega}}p. \tag{5.1.1}$$

The factors are chosen so that our final statements have a simple form, note that a is dimensionless. The Hermitian conjugate of this is

$$a^\dagger = \sqrt{\frac{m\omega}{2}}q - \frac{i}{\sqrt{2m\omega}}p. \tag{5.1.2}$$

Note that a is not Hermitian, it's not an observable, for now we can just think of it as a mathematical trick. We'll see later that combinations of a and $a^\dagger$ can be treated as observables with meaningful values. The motivation for this definition is that if a and $a^\dagger$ commuted (they don't, as we'll see soon) we could factorise the Hamiltonian as

$$\omega a a^\dagger. \tag{5.1.3}$$

A bit of algebra allows us to rewrite q and p in terms of a and $a^\dagger$:

$$q = \frac{1}{2m\omega}(a + a^\dagger), \quad \text{and} \quad p = -i\sqrt{\frac{m\omega}{2}}(a - a^\dagger). \tag{5.1.4}$$

We can now compute the commutation relations for a and $a^\dagger$ from the canonical commutation relation:

$$i = [q, p] \tag{5.1.5}$$

$$= -\frac{i}{2}[a + a^\dagger, a - a^\dagger] \tag{5.1.6}$$

$$= -\frac{i}{2}\left([a, a] - [a, a^\dagger] + [a^\dagger, a] - [a^\dagger, a^\dagger]\right) \tag{5.1.7}$$

$$= i[a, a^\dagger]. \tag{5.1.8}$$

Hence, we have that

$$[a, a^\dagger] = 1. \tag{5.1.9}$$

An alternative approach to quantum mechanics is to start with a and $a^\dagger$, motivating them starting with the interpretation we'll see later, and then show that the canonical commutation relations are what they are as a consequence of this choice.

We can write the Hamiltonian in terms of a and $a^\dagger$:

$$H = -\frac{\omega}{4}(a - a^\dagger)^2 + \frac{\omega}{4}(a + a^\dagger)^2 \tag{5.1.10}$$

$$= -\frac{\omega}{4}(a^2 - aa^\dagger - a^\dagger a + (a^\dagger)^2) + \frac{\omega}{4}(a^2 + aa^\dagger + a^\dagger a + (a^\dagger)^2) \tag{5.1.11}$$

$$= \frac{\omega}{2}(aa^\dagger + a^\dagger a). \tag{5.1.12}$$

Now use

$$1 = [a, a^\dagger] = aa^\dagger - a^\dagger a \implies aa^\dagger = a^\dagger a + 1 \tag{5.1.13}$$

and we get

$$H = \frac{\omega}{2}(aa^\dagger + a^\dagger a) = \frac{\omega}{2}(a^\dagger a + 1 + a^\dagger a) = \omega\left(a^\dagger a + \frac{1}{2}\right). \tag{5.1.14}$$

To compute the spectrum of the Hamiltonian we work with the operator $N \coloneqq a^\dagger a$. This *is* a Hermitian operator, $N^\dagger = (a^\dagger a) = a^\dagger(a^\dagger)^\dagger = a^\dagger a = N$. We can then write the Hamiltonian as

$$H = \omega\left(N + \frac{1}{2}\right), \tag{5.1.15}$$

if you remember what the energy eigenvalues of the harmonic oscillator are you may see where this is going.

Let $|n\rangle$ be an eigenstate of N with eigenvalue n, so $N|n\rangle = n|n\rangle$. For now all we know is that $n \in \mathbb{R}$, since N has real eigenvalues as a Hermitian operator. We choose to normalise these states so that

$$\langle n|m\rangle = \delta_{nm} \tag{5.1.16}$$

The use of n and m as labels, as well as the Dirac delta hints at something we will show later, that n is an integer, but for now we just treat n as a real number.

The expectation value of N in the state $|n\rangle$ is

$$\langle N\rangle = \langle n|N|n\rangle = n\langle n|n\rangle = n. \tag{5.1.17}$$

Using $N = a^\dagger a$ we can calculate this expectation value a different way:

$$n = \langle n|N|n\rangle = \langle n|a^\dagger a|n\rangle = \|a|n\rangle\|^2, \tag{5.1.18}$$

and since $\|a|n\rangle\|^2 \geq 0$ we have that $n \geq 0$.

Now consider commutators of N with $a^\dagger$ and a:

$$[N, a^\dagger] = [a^\dagger a, a^\dagger] = a^\dagger[a, a^\dagger] + [a^\dagger, a^\dagger]a = a^\dagger. \tag{5.1.19}$$

Similarly,

$$[N, a] = [a^\dagger a, a] = a^\dagger[a, a] + [a^\dagger, a]a = -a. \tag{5.1.20}$$

Suppose we have an eigenstate, $|n\rangle$. What happens when we act on this with N? Well, it turns out that we get another eigenstate:

$$Na^\dagger|n\rangle = ([N, a^\dagger] + a^\dagger N)|n\rangle \tag{5.1.21}$$

$$= (a^\dagger + a^\dagger N)|n\rangle \tag{5.1.22}$$

$$= a^\dagger|n\rangle + a^\dagger N|n\rangle \tag{5.1.23}$$

$$= a^\dagger|n\rangle + a^\dagger n|n\rangle \tag{5.1.24}$$

$$= (n + 1)a^\dagger|n\rangle. \tag{5.1.25}$$

So, $a^\dagger|n\rangle$ is an eigenstate of N with eigenvalue $n + 1$. That is, $a^\dagger|n\rangle \propto |n + 1\rangle$. We can fairly easily figure out the constant of proportionality by considering $\|a^\dagger|n\rangle\|^2$:

$$\|a^\dagger|n\rangle\|^2 = \langle n|aa^\dagger|n\rangle \tag{5.1.26}$$

$$= \langle n|(a^\dagger a + [a, a^\dagger])|n\rangle \tag{5.1.27}$$

$$= \langle n|(a^\dagger a + 1)|n\rangle \tag{5.1.28}$$

$$= \langle n|a^\dagger a|n\rangle + \langle n|n\rangle \tag{5.1.29}$$

$$= \langle n|N|n\rangle + 1 \tag{5.1.30}$$

$$= n + 1. \tag{5.1.31}$$

Since the overall phase of a state has no physical significance we are free to choose the phase of the proportionality constant, so we make the sensible choice that its real and positive. That is, we choose the positive square root, $\|a^\dagger|n\rangle\| = \sqrt{n + 1}$, and we then have

$$a^\dagger|n\rangle = \sqrt{n + 1}|n\rangle. \tag{5.1.32}$$

We can proceed similarly for $a|n\rangle$, we have

$$Na|n\rangle = ([N, a] + aN)|n\rangle \tag{5.1.33}$$

$$= (-a + aN)|n\rangle \tag{5.1.34}$$

$$= (n - 1)a|n\rangle. \tag{5.1.35}$$

So $a|n\rangle$ is an eigenstate of N with eigenvalue $n - 1$, so $a|n\rangle \propto |n - 1\rangle$. We can, again, work out the constant of proportionality:

$$\|a|n\rangle\|^2 = \langle n|a^\dagger a|n\rangle = \langle n|N|n\rangle = n. \tag{5.1.36}$$

So, again choosing a real, positive, solution we have

$$a|n\rangle = \begin{cases} \sqrt{n}|n - 1\rangle & n > 0, \\ 0 & n = 0, \end{cases} \tag{5.1.37}$$

the reason for the careful distinction when $n = 0$ is that the state $|0 - 1\rangle = |-1\rangle$ is not defined, as the eigenvalues of N are nonnegative. The reason we don't have to exclude, say, $n = 0.5$, which would have $|0.5 - 1\rangle = |-0.5\rangle$, is that, as we will now show, n is a nonnegative integer, that is $n \in \mathbb{N} = \{0, 1, 2, \dots\}$.

Suppose that $n \notin \mathbb{N}$. Consider some $m \in \mathbb{N} \setminus \{0\}$. Then we can apply a m times to $|n\rangle$:

$$a^m|n\rangle = \sqrt{n}a^{m-1}|n\rangle = \sqrt{n}\sqrt{n - 1}a^{m-2}|n\rangle$$
$$= \dots = \sqrt{n}\sqrt{n - 1}\cdots\sqrt{n - m + 1}|n - m\rangle. \tag{5.1.38}$$

For sufficiently large m we will have $n - m < 0$. However, $n - m$ is an eigenvalue and all eigenvalues of N are nonnegative. This is a contradiction and so we must have $n \in m$. In this case one of the square roots will vanish before we get to negative eigenvalues and we'll avoid the contradiction.

Since $H = \omega(N + 1/2)$ we can see that the eigenvalues of the Hamiltonian are

$$\left(n + \frac{1}{2}\right)\omega \qquad \text{where} \qquad n \in \mathbb{N}. \tag{5.1.39}$$

The eigenvectors of the Hamiltonian can all be written as

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle \tag{5.1.40}$$

where $|0\rangle$ is such that when a acts on it we get zero:

$$a|0\rangle \coloneqq 0. \tag{5.1.41}$$

Note that $|0\rangle$ is an eigenstate with eigenvalue 0, and 0 is the zero vector, these are not the same. The factor of $1/\sqrt{n!}$ is just normalisation, if $|0\rangle$ is normalised then $a^\dagger|0\rangle = |1\rangle = \sqrt{1}|1\rangle$, then $(a^\dagger)^2|0\rangle = \sqrt{1}a^\dagger|1\rangle = \sqrt{1}\sqrt{2}|2\rangle$, and so on, so we can get rid of all these factors with $1/\sqrt{n!}$, and we're left with the normalised states $|n\rangle$.

The most important thing about the harmonic oscillator is that the energy eigenvalues are equally spaced. This makes harmonic oscillators very nice. Another nice property is that the energy eigenvalues are nonnegative. This means that the energy is bounded below, which is good as if it wasn't then systems would quickly decay to having negative infinite energies. The equally spaced eigenvalues is what makes all of quantum field theory possible.

5.2 Interpretations

It's fine, but in a sense, it's just wrong.

Richard Ball

There are two ways to interpret the harmonic oscillator. Historically the first we discuss here was Schrödinger's interpretation, and actually came a year after the second interpretation we'll discuss, which is due to Heisenberg. However, for a long time Schrödinger's interpretation was preferred, mostly because it uses more familiar mathematics, like differential equations, and also because it aligns more similarly with the notion of a classical harmonic oscillator.

5.2.1 Wave Function Interpretation

This is the way you were taught the harmonic oscillator in kindergarten I guess.

Richard Ball

The quantum harmonic oscillator is a quantised classical harmonic oscillator. We take the normal harmonic oscillator Hamiltonian, replace the position and momentum with operators, and solve the equations of motion. The picture that people then have in their head is a harmonic potential with equally spaced energy levels, as seen in Figure 5.1.

The physical variables are q and p, with $q = x$ the position operator, and p chosen to satisfy the canonical commutation relations, which it can be shown is the case if $p = -\mathrm{i}\mathrm{d}/\mathrm{d}x$. We solve the equation $a|0\rangle = 0$ to find the vacuum state $|0\rangle$, and it's wave function $\psi_0(x) = \langle x|0\rangle$:

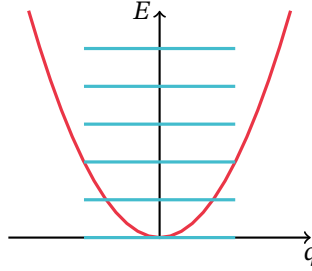$$\psi_0(x) = \sqrt{\frac{m\omega}{2\pi}} \exp\left[-\frac{1}{2}m\omega x^2\right]. \tag{5.2.1}$$

Figure 5.1: The harmonic potential and the evenly spaced energy levels it produces.

We then find the wave equation for the nth excited state, $\sqrt{n!}\psi_n(x) = \langle x|(a^\dagger)^n|0\rangle$:

$$\psi_n(x) = \frac{1}{\sqrt{n!}} \left(0 \frac{i}{\sqrt{m\omega}} \frac{\mathrm{d}}{\mathrm{d}x} + i\sqrt{\frac{m\omega}{2}} x \right)^n \psi_0(x) \tag{5.2.2}$$

where the expression in brackets is just $a^\dagger = i/\sqrt{2m\omega}\,p + \sqrt{m\omega/2}\,q$ with the operators substituted in. After working through lots of derivatives of Gaussians we will find that the solutions are Gaussians times a Hermite polynomial, H_n:

$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}} \left(\frac{m\omega}{\pi} \right)^{1/4} \exp\left[-\frac{1}{2} m\omega x^2 \right] H_n\left(\sqrt{m\omega} x \right). \tag{5.2.3}$$

This interpretation is called **first quantisation**, presumably because it is the first way that most people are taught to interpret the harmonic oscillator.

5.2.2 Quanta

The energy eigenvalues of the harmonic oscillator are evenly spaced. This allows us to interpret the state $|n\rangle$ as consisting of n identical quanta, all of the same energy, ω. The state $|0\rangle$ is then the vacuum state, where there are no quanta, and we have the vacuum energy $\omega/2$.

Rather than try to interpret what a harmonic oscillator is we focus on what it does. Think of it as a black box which can absorb a quanta of energy, so it has $n+1$ quanta, or it can emit a quanta, so it has $n-1$ quanta. Clearly these occurrences align with the action of $|a\rangle^\dagger$ and $|a\rangle$ respectively. This leads to us interpreting $a^\dagger$ as a **creation operator**, producing a quanta of energy in the black box, and a as a **annihilation operator**, destroying a quanta of energy in the black box. Of course this energy must come from somewhere and go somewhere, but just within the black box it may as well be appearing and vanishing.

This interpretation, while a bit odd at first, is actually closer to what we do in experiments. We scatter particles, essentially adding them to the black box, and see what comes out, that is, what leaves the black box. This interpretation also works well when considering atomic spectra, which occur in much the same way.

This interpretation is intrinsically multiparticle, which is good for quantum field theory. This interpretation is called **second quantisation**, presumably because it is the second way that most people are taught to interpret the harmonic oscillator.

5.3   Harmonic Oscillator with Time Dependence

If we include time dependence in the Harmonic oscillator the, working in the Heisenberg picture, this corresponds to replacing q and p with time dependent operators, which in term means introducing a time dependence to a and $a^\dagger$. This time dependence doesn't change the first step of solving the Harmonic oscillator, which is finding q and p in terms of a and $a^\dagger$, which gives us

$$q(t) = \frac{1}{\sqrt{2m\omega}}[a(t) + a^\dagger(t)], \tag{5.3.1}$$

$$p(t) = i\sqrt{\frac{m\omega}{2}}[a(t) - a^\dagger(t)]. \tag{5.3.2}$$

The commutation relations for $a(t)$, $a^\dagger(t)$, and H are $[a(t), a^\dagger(t)] = 1$, $[H, a(t)] = -\omega a(t)$, and $[H, a^\dagger(t)] = \omega a^\dagger(t)$, which are the same as if we replace H with N, since $H = \omega(N + 1/2)$. The time dependence doesn't change the commutators, which we can check quite easily:

$$[a(t), a^\dagger(t)] = [e^{-iHt}ae^{iHt}, e^{-iHt}a^\dagger e^{iHt}] \tag{5.3.3}$$

$$= e^{-iHt}ae^{iHt}e^{-iHt}a^\dagger e^{iHt} - e^{-iHt}a^\dagger e^{iHt}e^{-iHt}ae^{iHt} \tag{5.3.4}$$

$$= e^{-iHt}aa^\dagger e^{iHt} - e^{-iHt}a^\dagger ae^{iHt} \tag{5.3.5}$$

$$= e^{-iHt}[a, a^\dagger]e^{iHt} \tag{5.3.6}$$

$$= e^{-iHt}1e^{iHt} \tag{5.3.7}$$

$$= 1. \tag{5.3.8}$$

Likewise, the commutators with the Hamiltonian are unchanged:

$$[H, a^\dagger(t)] = [H, e^{-iHt}a^\dagger e^{iHt}] \tag{5.3.9}$$

$$= He^{-iHt}a^\dagger e^{iHt} - e^{-iHt}a^\dagger e^{iHt}H \tag{5.3.10}$$

$$= e^{-iHt}Ha^\dagger e^{iHt} - e^{-iHt}a^\dagger He^{iHt} \tag{5.3.11}$$

$$= e^{-iHt}[H, a^\dagger]e^{iHt} \tag{5.3.12}$$

$$= e^{-iHt}\omega a^\dagger e^{iHt} \tag{5.3.13}$$

$$= \omega a^\dagger(t). \tag{5.3.14}$$

An similarly for $[H, a(t)]$.

We can then solve the Heisenberg equation for the creation and annihilation operators:

$$\dot{a}(t) = i[H, a(t)] = -i\omega a(t). \tag{5.3.15}$$

This very simple differential equation, $\dot{a}(t) = -i\omega a(t)$, has a plane wave solution,

$$a(t) = ae^{-i\omega t}, \tag{5.3.16}$$

where $a = a(0)$ is the annihilation operator in the Schrödinger picture. Taking the Hermitian conjugate of this we get

$$\dot{a}^\dagger(t) = a^\dagger e^{i\omega t}. \tag{5.3.17}$$

The position and momentum operators are then

$$q(t) = \frac{1}{\sqrt{2m\omega}}[a\mathrm{e}^{-i\omega t} + a^{\dagger}\mathrm{e}^{i\omega t}], \tag{5.3.18}$$

$$p(t) = i\sqrt{\frac{m\omega}{2}}[a\mathrm{e}^{-i\omega t} - a^{\dagger}\mathrm{e}^{i\omega t}]. \tag{5.3.19}$$

Six

Mode Expansions

6.1 Expanding the Field

The field $\varphi(x)$ satisfies the Klein–Gordon equations. We've seen that this allows plane wave solutions, $e^{\pm ip \cdot x}$. This suggests that we can write a general solution as a superposition of wave solutions, which is just a Fourier transform. Since φ is a function of x this is really an inverse Fourier transform. For comparison consider the inverse transform of a single variable, real, one-dimensional function, f:

$$f(x) = \int \frac{\mathrm{d}p}{2\pi} (\tilde{f}(p)e^{-ipx} + \tilde{f}^*(p)e^{ipx}). \tag{6.1.1}$$

Here $\tilde{f}(p)$ is the Fourier transformed function, which acts as a Fourier coefficient here. We need both the e^{ipx} and conjugate e^{-ipx} terms to get a real function. The factor of 2π is part of the normalisation, but exactly where it appears between the forward and inverse is just convention.

We want our field to have on-shell momentum, that is $p^2 = m^2$. To enforce this we include a factor of $2\pi\delta(p^2 - m^2)$, where the 2π is factor is from the integral representation of the Dirac delta, as the Fourier transform of one:

$$\int \mathrm{d}x \, e^{-ip \cdot x} = 2\pi\delta(p). \tag{6.1.2}$$

We also want to restrict ourselves to positive energy solutions, so $p_0 > 0$, we can enforce this with a factor of $\theta(p_0)$, where θ is the **Heaviside step function** defined by

$$\theta(u) = \begin{cases} 1 & u > 0, \\ 1/2 & u = 0, \\ 0 & u < 0. \end{cases} \tag{6.1.3}$$

The **mode expansion** of φ is then

$$\varphi(x) = \int \frac{\mathrm{d}^4 p}{(2\pi)^4} 2\pi\delta(p^2 - m^2)\theta(p_0)[a(\boldsymbol{p})e^{-ip \cdot x} + a^\dagger(\boldsymbol{p})e^{ip \cdot x}]. \tag{6.1.4}$$

Here $a(\boldsymbol{p})$ and $a^\dagger(\boldsymbol{p})$ are Fourier coefficients, we'll see later that they are the quantum field theory analogues of the creation and annihilation operators for the harmonic oscillator. They depend only on $\boldsymbol{p}$, rather than the four-vector p, since we can take $\delta(p^2 - m^2)$ as fixing the value of p_0.

Define $\omega(\boldsymbol{p}) := +\sqrt{\boldsymbol{p}^2 + m^2}$, which is always nonnegative. This allows us to rewrite the argument of the Dirac delta:

$$(p_0 - \omega(\boldsymbol{p}))(p_0 + \omega(\boldsymbol{p})) = p_0 p_0 - \omega(\boldsymbol{p})^2 = p_0 p_0 - \boldsymbol{p}^2 - m^2 = p^2 - m^2. \quad (6.1.5)$$

So, now we have $\delta(\omega(\boldsymbol{p}))$ in our integral. Recall the distribution identity

$$\delta(f(x)) = \sum_i \frac{\delta(x - a_i)}{|f'(a_i)|} \quad (6.1.6)$$

where a_i are all the zeros of f lying in the range of integration. Since in the integral we are enforcing $p_0 \geq 0$ the only time $p^2 - m^2 = (p_0 - \omega(\boldsymbol{p}))(p_0 + \omega(\boldsymbol{p}))$ vanishes is when $p_0 = \omega(\boldsymbol{p})$, and

$$\left. \frac{\partial}{\partial p}(p^2 - m^2) \right|_{p_0 = \omega(\boldsymbol{p})} = 2\omega(\boldsymbol{p}) \quad (6.1.7)$$

so, as distributions, we have

$$\delta(p^2 - m^2)\theta(p_0) = \frac{1}{2\omega(\boldsymbol{p})}\delta(p^0 - \omega(\boldsymbol{p})). \quad (6.1.8)$$

This lets us write $\varphi(x)$ as

$$\varphi(x) = \int \frac{\mathrm{d}^4 p}{(2\pi)^3} \frac{1}{2\omega(\boldsymbol{p})}\delta(p^0 - \omega(\boldsymbol{p}))[a(\boldsymbol{p})\mathrm{e}^{-ip \cdot x} + a^\dagger(\boldsymbol{p})\mathrm{e}^{ip \cdot x}]. \quad (6.1.9)$$

Performing the p^0 integral using the sifting property of the Dirac delta we have

$$\varphi(x) = \int \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3} \frac{1}{2\omega(\boldsymbol{p})}[a(\boldsymbol{p})\mathrm{e}^{-ip \cdot x} + a^\dagger(\boldsymbol{p})\mathrm{e}^{ip \cdot x}]\bigg|_{p_0 = \omega(\boldsymbol{p})}. \quad (6.1.10)$$

At this point we make some remarks on this expansion:

- The expansion is manifestly Lorentz invariant, while the presence of p^0 may be concerning orthochronous Lorentz transformations (the only type we consider in this course) don't change the sign of p^0, and so $\theta(p^0)$ is Lorentz invariant.

- Since $p^0 = \omega(\boldsymbol{p})$ the energy is always positive, and so is bounded below.

- The integral over $\delta(p^0 - \omega(\boldsymbol{p}))$ fixes $p^0 = \omega(\boldsymbol{p})$, and from now on this will be implicit in the integrals, instead of writing $|_{p_0 = \omega(\boldsymbol{p})}$.

- The operator $\varphi(x)$ is in the Heisenberg picture, but the time dependence comes from the exponential factors, both $a(\boldsymbol{p})$ and $a^\dagger(\boldsymbol{p})$ are time independent operators.

- We need both the a and $a^\dagger$ terms for φ to be Hermitian.

- The measure

$$\frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3} \frac{1}{2\omega(\boldsymbol{p})} = \frac{\mathrm{d}^4 p}{(2\pi)^4} 2\pi \delta(p^2 - m^2)\theta(p^0) \quad (6.1.11)$$

  is Lorentz invariant, but takes a lot of writing, we will use the shorthand notation

$$\bar{\mathrm{d}} p = \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3} \frac{1}{2\omega(\boldsymbol{p})} = \int \frac{\mathrm{d}^4 p}{(2\pi)^4} 2\pi \delta(p^2 - m^2)\theta(p^0). \quad (6.1.12)$$

Now that we've expanded the field φ it is easy to expand the conjugate field, π, using $\pi = \dot{\varphi}$:

$$\pi(x) = \dot{\varphi}(x) = -\frac{i}{2} \int \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3} [a(\boldsymbol{p})\mathrm{e}^{-ip\cdot x} - a^\dagger(\boldsymbol{p})\mathrm{e}^{ip\cdot x}] \tag{6.1.13}$$

$$= -\frac{i}{2} \int đp [a(\boldsymbol{p})\mathrm{e}^{-ip\cdot x} - a^\dagger(\boldsymbol{p})\mathrm{e}^{ip\cdot x}]. \tag{6.1.14}$$

Here we've used

$$\frac{\partial}{\partial x^0} \mathrm{e}^{\pm ip\cdot x} = \frac{\partial}{\partial x^0} \exp[\pm ip^0 x_0 \mp p^i x_i]$$
$$= \pm ip^0 \exp[\pm ip^0 x_0 \mp p^i x_i] = \pm ip^0 \exp[\pm ip \cdot x] \tag{6.1.15}$$

and then set $p^0 = \omega(\boldsymbol{p})$.

6.2 The Fourier Coefficients

As with the Harmonic oscillator we now want to invert these relations to find expressions for $a(\boldsymbol{p})$ and $a^\dagger(\boldsymbol{p})$. Since these are inverse Fourier transformations the inversion happens through Fourier transformation. For a single variable, real, one-dimensional function, f, the Fourier transform is

$$\tilde{f}(p) = \int \mathrm{d}x\, f(x)\mathrm{e}^{ip\cdot x}. \tag{6.2.1}$$

Using this, but in three dimensions, we have

$$\int \mathrm{d}^3\boldsymbol{x}\, \mathrm{e}^{ip'\cdot x}\varphi(x) = \int đp \frac{1}{2\omega(\boldsymbol{p})} \int \mathrm{d}^3\boldsymbol{x} \left[a(\boldsymbol{p})\mathrm{e}^{i(p'-p)\cdot x} + a^\dagger \mathrm{e}^{i(p'+p)\cdot x} \right]$$
$$= \int đp\, (2\pi)^3 \left[a(\boldsymbol{p})\delta^3(\boldsymbol{p}' - \boldsymbol{p})\mathrm{e}^{i(p'^0 - p^0)t} \right.$$
$$\left. + a^\dagger(\boldsymbol{p})\delta^3(\boldsymbol{p}' + \boldsymbol{p})\mathrm{e}^{i(p'^0 + p^0)t} \right]$$
$$= \int \frac{\mathrm{d}^3\boldsymbol{p}}{2\omega(\boldsymbol{p})} \left[a(\boldsymbol{p})\delta^3(\boldsymbol{p}' - \boldsymbol{p})\mathrm{e}^{i(p'^0 - p^0)t} \right.$$
$$\left. + a^\dagger(\boldsymbol{p})\delta^3(\boldsymbol{p}' + \boldsymbol{p})\mathrm{e}^{i(p'^0 + p^0)t} \right].$$

Here we've recognised the integral representation of the Dirac delta:

$$\int \mathrm{d}^3\boldsymbol{x}\, \mathrm{e}^{-ip\cdot x} = (2\pi)^3 \delta^3(\boldsymbol{x})\mathrm{e}^{ip^0 t}, \tag{6.2.2}$$

and made the usual assumption that everything is sufficiently convergent for us to be able to swap the order of integrals. Performing the integrals now using the Dirac deltas, and the fact that $\omega(-\boldsymbol{p}) = \omega(\boldsymbol{p})$, we have

$$\int \mathrm{d}^3\boldsymbol{x}\, \mathrm{e}^{ip'\cdot x}\varphi(x) = \frac{1}{2\omega(\boldsymbol{p}')} a(\boldsymbol{p}') + \frac{1}{2\omega(-\boldsymbol{p}')} a^\dagger(-\boldsymbol{p}')\mathrm{e}^{2ip'^0 t} \tag{6.2.3}$$

$$= \frac{1}{2\omega(\boldsymbol{p}')} [a(\boldsymbol{p}') + a^\dagger(-\boldsymbol{p}')\mathrm{e}^{2i\omega(\boldsymbol{p}')t}]. \tag{6.2.4}$$

Here we've used the Dirac delta's to set $\boldsymbol{p} = \pm\boldsymbol{p}'$ in each term as appropriate. Restricting ourselves to on-shell, positive energy, solutions this then fixes the value of p^0 to $\sqrt{\boldsymbol{p}'^2 + m^2} = \omega(\boldsymbol{p}')$.

We can do exactly the same for π, the only difference is signs and a factor of $1/\omega(\boldsymbol{p})$, so we get

$$\int \mathrm{d}^3\boldsymbol{x}\, \mathrm{e}^{ip'\cdot x}\pi(x) = -\frac{i}{2}[a(\boldsymbol{p}') - a(-\boldsymbol{p}')^\dagger \mathrm{e}^{2i\omega(\boldsymbol{p}')t}]. \tag{6.2.5}$$

Adding i times this to $\omega(\boldsymbol{p})$ times the result for φ we get

$$a(\boldsymbol{p}) = \int \mathrm{d}^3\boldsymbol{x}\, \mathrm{e}^{ip\cdot x}[\omega(\boldsymbol{p})\varphi(t, \boldsymbol{x}) + i\pi(t, \boldsymbol{x})]. \tag{6.2.6}$$

Taking the conjugate we get

$$a^\dagger(\boldsymbol{p}) = \int \mathrm{d}^3\boldsymbol{x}\, \mathrm{e}^{-ip\cdot x}[\omega(\boldsymbol{p})\varphi(t, \boldsymbol{x}) - i\pi(t, \boldsymbol{x})]. \tag{6.2.7}$$

Now we have expressions for $a(\boldsymbol{p})$ and $a^\dagger(\boldsymbol{p})$ we can compute their commutator:

$$[a(\boldsymbol{p}), a^\dagger(\boldsymbol{p}')] = \int \mathrm{d}^3\boldsymbol{x}\int \mathrm{d}^3\boldsymbol{x}'\, \mathrm{e}^{ip\cdot x - ip'\cdot x'} \tag{6.2.8}$$
$$[\omega(\boldsymbol{p})\varphi(t, \boldsymbol{x}) + i\pi(t, \boldsymbol{x}), \omega(\boldsymbol{p}')\varphi(t, \boldsymbol{x}') - i\pi(t, \boldsymbol{x}')].$$

Using the equal time commutation relations only the cross terms don't vanish, and so this becomes

$$[a(\boldsymbol{p}), a^\dagger(\boldsymbol{p}')] = \int \mathrm{d}^3\boldsymbol{x}\int \mathrm{d}^3\boldsymbol{x}'\, \mathrm{e}^{ip\cdot x - ip'\cdot x'}\big(-i\omega(\boldsymbol{p})[\varphi(t, \boldsymbol{x}), \pi(t, \boldsymbol{x}')] \tag{6.2.9}$$
$$+ i\omega(\boldsymbol{p}')[\pi(t, \boldsymbol{x}), \varphi(t, \boldsymbol{x}')]\big).$$

$$= \int \mathrm{d}^3\boldsymbol{x}\int \mathrm{d}^3\boldsymbol{x}'\, \mathrm{e}^{ip\cdot x - ip'\cdot x'}(\omega(\boldsymbol{p})\delta^3(\boldsymbol{x} - \boldsymbol{x}') + \omega(\boldsymbol{p}')\delta^3(\boldsymbol{x} - \boldsymbol{x}')) \tag{6.2.10}$$

$$= \int \mathrm{d}^3\boldsymbol{x}\, \mathrm{e}^{i(p-p')\cdot x}(\omega(\boldsymbol{p}) + \omega(\boldsymbol{p}')) \tag{6.2.11}$$

$$= (2\pi)^3\delta(\boldsymbol{p} - \boldsymbol{p}')2\omega(\boldsymbol{p}) \tag{6.2.12}$$

where again we've recognised the integral representation of the Dirac delta, and made use of the resulting Dirac delta to set $\boldsymbol{p} = \pm\boldsymbol{p}'$, and hence $p^0 = p'^0$ choosing to stay on-shell and positive energy. We introduce the shorthand notation,

$$(2\pi)^3\delta(\boldsymbol{p} - \boldsymbol{p}')2\omega(\boldsymbol{p}) = \delta(\boldsymbol{p} - \boldsymbol{p}'). \tag{6.2.13}$$

This allows us to write

$$[a(\boldsymbol{p}), a^\dagger(\boldsymbol{p}')] = \delta(\boldsymbol{p} - \boldsymbol{p}'). \tag{6.2.14}$$

Compare this to $[a, a^\dagger] = 1$ for the harmonic oscillator. We also have

$$[a(\boldsymbol{p}), a(\boldsymbol{p}')] = [a^\dagger(\boldsymbol{p}'), a^\dagger(\boldsymbol{p}')] = 0. \tag{6.2.15}$$

Note that $\delta(\boldsymbol{p} - \boldsymbol{p}')$ is just a number, not an operator. This notation is nice, since if we combine the invariant measure, $\mathrm{d}p$, and $\delta(\boldsymbol{p} - \boldsymbol{p}')$ then we can just pretend we have $\mathrm{d}p$ and $\delta^3(\boldsymbol{p} - \boldsymbol{p})$:

$$\int \mathrm{d}p\, f(\boldsymbol{p})\delta(\boldsymbol{p} - \boldsymbol{p}') = f(\boldsymbol{p}'). \tag{6.2.16}$$

6.3 The Hamiltonian

To complete the analogy with the harmonic oscillator we need to find the Hamiltonian in terms of $a(\boldsymbol{p})$ and $a^\dagger(\boldsymbol{p})$. Starting with the hamiltonian

$$H = \frac{1}{2} \int \mathrm{d}^3\boldsymbol{x} \left[\pi(x)^2 + (\nabla\varphi(x))^2 + m^2\varphi(x)^2 \right] \tag{6.3.1}$$

we can split this into three components,

$$H_1 = \frac{1}{2} \int \mathrm{d}^3\boldsymbol{x} \, \pi(x)^2, \tag{6.3.2}$$

$$H_2 = \frac{1}{2} \int \mathrm{d}^3\boldsymbol{x} \, (\nabla\varphi(x))^2, \tag{6.3.3}$$

$$H_3 = \frac{1}{2} \int \mathrm{d}^3\boldsymbol{x} \, m^2\varphi(x)^2, \tag{6.3.4}$$

$$\tag{6.3.5}$$

We'll then treat each of these separately.

Start with the first term. Substituting in the expansion for $\pi(x)$, being careful to use different integration variables for each factor of π, we get

$$H_1 = -\frac{1}{8} \int \mathrm{d}^3\boldsymbol{x} \int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3} \left[a(\boldsymbol{p})\mathrm{e}^{-ip\cdot x} - a^\dagger(\boldsymbol{p})\mathrm{e}^{ip\cdot x} \right]$$
$$\times \int \frac{\mathrm{d}^3\boldsymbol{p}'}{(2\pi)^3} \left[a(\boldsymbol{p}')\mathrm{e}^{-ip'\cdot x} - a^\dagger(\boldsymbol{p}')\mathrm{e}^{ip'\cdot x} \right]. \tag{6.3.6}$$

Expanding this all the integrand becomes

$$a(\boldsymbol{p})a(\boldsymbol{p}')\mathrm{e}^{-i(p+p')\cdot x} - a(\boldsymbol{p})a^\dagger(\boldsymbol{p}')\mathrm{e}^{-i(p-p')\cdot x}$$
$$- a^\dagger(\boldsymbol{p})a(\boldsymbol{p}')\mathrm{e}^{i(p-p')\cdot x} + a^\dagger(\boldsymbol{p})a^\dagger(\boldsymbol{p}')\mathrm{e}^{i(p+p')\cdot x}. \tag{6.3.7}$$

We can perform the integration over position first, and, recalling that $p^0 = \omega(\boldsymbol{p})$, we get two types of terms, time dependent and time independent, these come from recognising the integral representation of the Dirac delta:

$$\int \mathrm{d}^3\boldsymbol{x} \, \mathrm{e}^{\pm i(p+p')\cdot x} = (2\pi)^3 \delta(\boldsymbol{p} + \boldsymbol{p}')\mathrm{e}^{\pm 2i\omega(\boldsymbol{p})t}, \tag{6.3.8}$$

$$\int \mathrm{d}^3\boldsymbol{x} \, \mathrm{e}^{\pm i(p-p')\cdot x} = (2\pi)^3 \delta(\boldsymbol{p} - \boldsymbol{p}'). \tag{6.3.9}$$

We should expect that the time dependent terms in H will all cancel, since H is time independent.

After performing the integral over positions we are left with the following integrand

$$(2\pi)^3 a(\boldsymbol{p})a(\boldsymbol{p}')\delta(\boldsymbol{p} - \boldsymbol{p}')\mathrm{e}^{-2i\omega(\boldsymbol{p})t} - (2\pi)^3 a(\boldsymbol{p})a^\dagger(\boldsymbol{p}')\delta(\boldsymbol{p} - \boldsymbol{p}')$$
$$- (2\pi)^3 a^\dagger(\boldsymbol{p})a(\boldsymbol{p}')\delta(\boldsymbol{p} - \boldsymbol{p}') + (2\pi)^3 a^\dagger(\boldsymbol{p})a^\dagger(\boldsymbol{p}')\delta(\boldsymbol{p} + \boldsymbol{p}')\mathrm{e}^{2i\omega(\boldsymbol{p})t}.$$

Performing the integral over $\boldsymbol{p}'$, and cancelling one factor of $(2\pi)^3$ we are left with

$$H_1 = -\frac{1}{8} \int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3} \left[a(\boldsymbol{p})a(\boldsymbol{p})\mathrm{e}^{-2i\omega(\boldsymbol{p})t} - a(\boldsymbol{p})a^\dagger(\boldsymbol{p}) \right. \tag{6.3.10}$$
$$\left. - a^\dagger(\boldsymbol{p})a(\boldsymbol{p}) + a^\dagger(\boldsymbol{p})a^\dagger(\boldsymbol{p})\mathrm{e}^{2i\omega(\boldsymbol{p})t} \right].$$

Keeping only the time independent terms, since the time independent terms will cancel with terms in H_2 and H_3, we get the contribution from the first term:

$$H_1 \to \frac{1}{8} \int \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3} \left[a(\boldsymbol{p}) a^\dagger(\boldsymbol{p}) + a^\dagger(\boldsymbol{p}) a(\boldsymbol{p}) \right]. \tag{6.3.11}$$

Doing the same for the other two terms we get their time independent contributions:

$$H_2 \to \frac{1}{8} \int \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3} \frac{\boldsymbol{p}^2}{\omega(\boldsymbol{p})^2} \left[a(\boldsymbol{p}) a(\boldsymbol{p})^\dagger + a(\boldsymbol{p})^\dagger a(\boldsymbol{p}) \right], \tag{6.3.12}$$

$$H_3 \to \frac{1}{8} \int \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3} \frac{m^2}{\omega(\boldsymbol{p})^2} \left[a(\boldsymbol{p}) a^\dagger(\boldsymbol{p}) + a^\dagger(\boldsymbol{p}) a(\boldsymbol{p}) \right]. \tag{6.3.13}$$

Noticing that $\boldsymbol{p}^2 + m^2$ cancels with $1/\omega(\boldsymbol{p})^2$ in the full expression we get

$$H = \frac{1}{4} \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \left[a(\boldsymbol{p}) a^\dagger(\boldsymbol{p}) + a(\boldsymbol{p}) a^\dagger(\boldsymbol{p}) \right]. \tag{6.3.14}$$

Compare this to the Harmonic oscillator result,

$$H = \frac{1}{2} \omega (a a^\dagger + a^\dagger a). \tag{6.3.15}$$

We can now compute the commutators of H with $a(\boldsymbol{p})$ and $a^\dagger(\boldsymbol{p})$:

$$[H, a^\dagger(\boldsymbol{p})] = \frac{1}{4} \int \frac{\mathrm{d}^3 p'}{(2\pi)^3} [a(\boldsymbol{p}') a^\dagger(\boldsymbol{p}') + a^\dagger(\boldsymbol{p}') a(\boldsymbol{p}'), a^\dagger(\boldsymbol{p})] \tag{6.3.16}$$

$$= \frac{1}{4} \int \frac{\mathrm{d}^3 p'}{(2\pi)^3} \left([a(\boldsymbol{p}') a^\dagger(\boldsymbol{p}'), a^\dagger(\boldsymbol{p})] + [a^\dagger(\boldsymbol{p}') a(\boldsymbol{p}'), a^\dagger(\boldsymbol{p})] \right)$$

$$= \frac{1}{4} \int \frac{\mathrm{d}^3 p'}{(2\pi)^3} \left(a(\boldsymbol{p}')[a^\dagger(\boldsymbol{p}'), a^\dagger(\boldsymbol{p})] + [a(\boldsymbol{p}'), a^\dagger(\boldsymbol{p})] a^\dagger(\boldsymbol{p}') \right.$$
$$\left. + a^\dagger(\boldsymbol{p}')[a(\boldsymbol{p}'), a^\dagger(\boldsymbol{p})] + [a^\dagger(\boldsymbol{p}'), a^\dagger(\boldsymbol{p})] \right) \tag{6.3.17}$$

$$= \frac{1}{4} \int \mathrm{d}^3 p' \, 2\omega(\boldsymbol{p}) [\delta(\boldsymbol{p} - \boldsymbol{p}') a^\dagger(\boldsymbol{p}') + a^\dagger(\boldsymbol{p}') \delta(\boldsymbol{p}' - \boldsymbol{p})] \tag{6.3.18}$$

$$= \omega(\boldsymbol{p}) a^\dagger(\boldsymbol{p}). \tag{6.3.19}$$

Similarly, with $a(\boldsymbol{p})$ we get

$$[H, a(\boldsymbol{p})] = -\omega(\boldsymbol{p}) a(\boldsymbol{p}). \tag{6.3.20}$$

Compare these to the harmonic oscillator results, $[H, a^\dagger] = \omega a^\dagger$ and $[H, a] = -\omega a$.

We can view the quantum fields as being a quantum harmonic oscillator of energy $\omega(\boldsymbol{p}) = \sqrt{\boldsymbol{p}^2 + m^2}$ for every Fourier mode, $\boldsymbol{p}$.

6.4 Particle Interpretation

Consider an energy eigenstate, $|E\rangle$, with eigenvalue E, so $H|E\rangle = E|E\rangle$. Then the state $a^\dagger(\boldsymbol{p})|E\rangle$ is also an energy eigenstate, with energy $E + \omega(\boldsymbol{p})$:

$$H a^\dagger(\boldsymbol{p})|E\rangle = \left([H, a^\dagger(\boldsymbol{p})] + a^\dagger(\boldsymbol{p}) H \right) |E\rangle \tag{6.4.1}$$

$$= (\omega(\boldsymbol{p}) a^\dagger(\boldsymbol{p}) + a^\dagger(\boldsymbol{p}) E) \tag{6.4.2}$$

$$= (\omega(\boldsymbol{p}) + E) a^\dagger(\boldsymbol{p})|E\rangle. \tag{6.4.3}$$

Similarly, $a(\boldsymbol{p})|E\rangle$ is an energy eigenstate with energy $E - \omega(\boldsymbol{p})$.

This suggests that, following the interpretation of the harmonic oscillator, $a^\dagger(\boldsymbol{p})$ is a **creation operator**, creating a quantum with energy $\omega(\boldsymbol{p}) = \sqrt{\boldsymbol{p}^2 + m^2}$, and $a(\boldsymbol{p})$ is an **annihilation operator**, destroying a quantum with energy $\omega(\boldsymbol{p})$. Note that all quanta in this interpretation have positive energy. There is another interpretation, which we'll discuss later, where $a^\dagger(\boldsymbol{p})$ creates a quantum with energy $\omega(\boldsymbol{p})$ and $a(\boldsymbol{p})$ creates a quantum with energy $-\omega(\boldsymbol{p})$.

Since the Hamiltonian is nonnegative, its defined as a sum of squares of real quantities, there must be some lowest energy state, $|0\rangle$, which we call the **vacuum state**. This state is such that any annihilation operator acting on it gives the zero vector:

$$a(\boldsymbol{p})|0\rangle = 0 \quad \forall \boldsymbol{p}. \tag{6.4.4}$$

Starting with $|0\rangle$ we can act with creation operators to make new states. Each creation operator, $a^\dagger(\boldsymbol{p})$, makes a new particle with energy $\omega(\boldsymbol{p})$ and momentum $\boldsymbol{p}$. For example,

$$|\psi\rangle = a^\dagger(\boldsymbol{p})a^\dagger(\boldsymbol{q})a^\dagger(\boldsymbol{r})|0\rangle \tag{6.4.5}$$

is a state with three particles with total energy $\omega(\boldsymbol{p}) + \omega(\boldsymbol{q}) + \omega(\boldsymbol{r})$, and total momentum $\boldsymbol{p} + \boldsymbol{q} + \boldsymbol{r}$.

Since all $a^\dagger(\boldsymbol{p})$ commute the state

$$a^\dagger(\boldsymbol{p})a^\dagger(\boldsymbol{q})a^\dagger(\boldsymbol{r})|0\rangle \tag{6.4.6}$$

is the same as $|\psi\rangle$. We interpret this as meaning that the particles are bosons, swapping any two particles leaves the state unchanged. This means the particles follow Bose–Einstein statistics.

The space of all states which can be constructed in this way, called **Fock space**, is the space given by all complex-linear combinations of

$$a^\dagger(\boldsymbol{p_1}) \cdots a^\dagger(\boldsymbol{p_n})|0\rangle \tag{6.4.7}$$

for any values of $n \in \mathbb{N}$.

Formally, the Fock space (for bosons) is defined to be

$$F(\mathbb{H}) = \overline{\bigoplus_{n=0}^{\infty} S\mathbb{H}^{\otimes n}} = \overline{\mathbb{C} \oplus \mathbb{H} \oplus S(\mathbb{H} \otimes \mathbb{H}) \oplus S(\mathbb{H} \otimes \mathbb{H} \otimes \mathbb{H}) \oplus \cdots} \tag{6.4.8}$$

where $\mathbb{H}$ is the single particle Hilbert space,

$$\mathbb{H}^{\otimes n} := \bigotimes_{i=1}^{n} \mathbb{H} = \underbrace{\mathbb{H} \otimes \cdots \otimes \mathbb{H}}_{n \text{ times}}, \tag{6.4.9}$$

the operator S symmetrises all tensors, the direct sum, $\oplus$, runs from $n = 0$, for states with no particles, to infinity, with each term representing an n-particle state space, and the line represents that we complete the space to form a Hilbert space, this entails adding in the values of all absolutely convergent series.

We make the normalisation choice that $\langle 0|0\rangle = 1$, then, if $|\boldsymbol{p}\rangle = a^\dagger(\boldsymbol{p})|\boldsymbol{p}\rangle$ is a single particle state we have

$$\langle \boldsymbol{p}|\boldsymbol{q}\rangle = \langle 0|a(\boldsymbol{p})a^\dagger(\boldsymbol{q})|0\rangle \tag{6.4.10}$$

$$= \langle 0|[a(\boldsymbol{p}), a^\dagger(\boldsymbol{q})]|0\rangle + \langle 0|a^\dagger(\boldsymbol{q})a(\boldsymbol{p})|0\rangle \tag{6.4.11}$$

$$= \delta(\boldsymbol{p} - \boldsymbol{q}). \tag{6.4.12}$$

Note that $a(\boldsymbol{p})|0\rangle$ and $\langle 0|a^\dagger(\boldsymbol{q}) = [a(\boldsymbol{q})|0\rangle]^\dagger$ both vanish, since they are annihilators acting on the vacuum state.

As with the harmonic oscillator we can define a **number density operator**, $N(\boldsymbol{p}) := a^\dagger(\boldsymbol{p})a(\boldsymbol{p})$, the eigenvalues of which integrate to give the number of particles in the state with momentum $\boldsymbol{p}$.

6.5 Ground State Energy

We can rewrite the Hamiltonian in terms of the number density operator:

$$H = \frac{1}{4} \int đp\, \omega(\boldsymbol{p}) \big(a^\dagger(\boldsymbol{p})a(\boldsymbol{p}) + a(\boldsymbol{p})a^\dagger(\boldsymbol{p}) \big) \tag{6.5.1}$$

$$= \frac{1}{2} \int đp\, \omega(\boldsymbol{p}) \left(a^\dagger(\boldsymbol{p})a(\boldsymbol{p}) + \frac{1}{2}[a(\boldsymbol{p}), a^\dagger(\boldsymbol{p})] \right) \tag{6.5.2}$$

$$= \frac{1}{2} \int đp\, \omega(\boldsymbol{p}) \big(N(\boldsymbol{p}) + \omega(\boldsymbol{p})(2\pi)^3 \delta^3(\boldsymbol{p} - \boldsymbol{p}) \big) \tag{6.5.3}$$

$$= \frac{1}{2} \int đp\, \omega(\boldsymbol{p}) \big(N(\boldsymbol{p}) + \omega(\boldsymbol{p})(2\pi)^3 \delta^3(\boldsymbol{0}) \big). \tag{6.5.4}$$

Clearly we have $\langle 0|N(\boldsymbol{p})|0\rangle = \langle 0|a^\dagger(\boldsymbol{p})a(\boldsymbol{p})|0\rangle = 0$, and so

$$\langle 0|H|0\rangle = \frac{1}{2} \int \mathrm{d}^3\boldsymbol{p}\, \omega(\boldsymbol{p})\delta^3(\boldsymbol{0}). \tag{6.5.5}$$

This is infinite. In fact, its *very* infinite. Each $\delta(0)$ is infinite, and we have three of them multiplied together, then we integrate over all momentum space, which is infinite, to make matters worse $\omega(\boldsymbol{p})$ grows arbitrarily large with $|\boldsymbol{p}|$.

The solution, as with so many similar problems in physics, is just not to think about it. All that we can measure is the energy of particles absorbed and emitted by the system. We cannot measure the ground state energy, so it doesn't matter that our theory gives a nonsensical answer when we try to compute the expected ground state energy.

6.5.1 Normal Ordering

One solution to the infinity arising in the ground state energy is to be careful about how we order $a(\boldsymbol{p})$ and $a^\dagger(\boldsymbol{p})$. If we keep all $a^\dagger(\boldsymbol{p})$ on the left of all $a(\boldsymbol{p})$ then, when we take expectation values, one of our $a(\boldsymbol{p})$ will act on $|0\rangle$ giving the zero vector, and one of our $a^\dagger(\boldsymbol{p})$ will act on $\langle 0|$, also giving the zero vector.

Mathematically what happens is that when we change to order of operators to have all $a^\dagger(\boldsymbol{p})$ on the left we'll get commutators of $a(\boldsymbol{p})$ and $a^\dagger(\boldsymbol{p}')$, which give rise to Dirac deltas which cancel with the existing ones, to leave a finite result.

This idea that careful ordering of operators can avoid infinities goes beyond the expectation value of the ground state energy. Consider our field, $\varphi(x)$. We can split this into parts,

$$\varphi(x) = \varphi^+(x) + \varphi^-(x) \tag{6.5.6}$$

where

$$\varphi^+(x) = \int đp\, a(\boldsymbol{p})\mathrm{e}^{-ip\cdot x}, \tag{6.5.7}$$

and

$$\varphi^-(x) = \int \text{d}p \, a^\dagger(\boldsymbol{p}) \text{e}^{ip\cdot x} = (\varphi^+)^\dagger. \tag{6.5.8}$$

If these were single particle states in quantum mechanics then they would correspond to positive and negative energy states respectively, but they aren't, although we'll see an interpretation along these lines later.

Now suppose we want to consider the product $\varphi(x)\varphi(y)$ for two arbitrary spacetime points, x and y. We can expand this as follows:

$$\varphi(x)\varphi(y) = [\varphi^+(x) + \varphi^-(x)][\varphi^+(y) + \varphi^-(y)] \tag{6.5.9}$$
$$= \varphi^+(x)\varphi^+(y) + \varphi^+(x)\varphi^-(y) + \varphi^-(x)\varphi^+(y) + \varphi^-(x)\varphi^-(y). \tag{6.5.10}$$

The **normal ordering** of $\varphi(x)\varphi(y)$, denoted $:\varphi(x)\varphi(y):$, is defined as the result of taking this product, but swapping terms so that all $a^\dagger(\boldsymbol{p})$ appear on the left of all $a(\boldsymbol{p}')$, that is all φ^+ appear on the left of all φ^-, so

$$:\varphi(x)\varphi(y): = \varphi^+(x)\varphi^+(y) + \varphi^+(x)\varphi^-(y) + {\color{red}\varphi^+(y)\varphi^-(x)} + \varphi^-(x)\varphi^-(y). \tag{6.5.11}$$

Then we have

$$\langle 0 | :\varphi(x)\varphi(y): | 0 \rangle = 0. \tag{6.5.12}$$

In general, if X is any nontrivial[1] polynomial in $a^\dagger(\boldsymbol{p})$ and $a(\boldsymbol{p})$ we'll have

$$\langle 0 | :X: | 0 \rangle = 0. \tag{6.5.13}$$

Subtracting the zero point energy of H from H is equivalent to normal ordering

$$:H: = H - \langle 0 | H | 0 \rangle, \tag{6.5.14}$$

since by normal ordering H we swap exactly those terms which give infinity, and all others give zero, so we can think of normal ordering as subtracting them off.

From now on we'll assume that H and $\boldsymbol{p}$ are normal ordered, often without writing it.

6.6 Ground State Momentum

The physical momentum operator is given by

$$\boldsymbol{P} = \int \text{d}^3\boldsymbol{x} \, \pi(x)\nabla\varphi(x). \tag{6.6.1}$$

Recall that this comes from the T^{i0} component of the energy-momentum tensor. Expressing the fields in terms of $a(\boldsymbol{p})$ and $a^\dagger(\boldsymbol{p})$ we get

$$\boldsymbol{P} = \frac{1}{2}\text{d}p \, \boldsymbol{p}(a(\boldsymbol{p})a^\dagger(\boldsymbol{p}) + a^\dagger(\boldsymbol{p})a(\boldsymbol{p})) \tag{6.6.2}$$

where the factor of $\boldsymbol{p}$ comes from computing $\nabla\varphi$. Rewriting this using the commutator of $a(\boldsymbol{p})$ and $a^\dagger(\boldsymbol{p})$ we get

$$\boldsymbol{P} = \int \text{d}p \, \boldsymbol{p}(N(\boldsymbol{p}) + \omega(\boldsymbol{p})(2\pi)^3\delta(\boldsymbol{0})). \tag{6.6.3}$$

---

[1] That is, X is not just a scalar, in which case $\langle 0 | :X: | 0 \rangle = X\langle 0 | 0 \rangle = X$.

Unlike with the ground state energy the $\delta(\mathbf{0})$ factor doesn't cause a problem. Integrating over all modes in the modes with momenta $\mathbf{p}$ and $-\mathbf{p}$ cancel, since $\omega(\mathbf{p}) = \omega(-\mathbf{p})$. So, we have

$$\langle 0|\mathbf{P}|0\rangle = \langle 0|{:}\,\mathbf{P}{:}\,|0\rangle = 0, \tag{6.6.4}$$

that is, the vacuum state has no momentum, which really ought to be the case.

Seven

Covariant Commutators

7.1 What are They

The equal time commutation relations treat time specially. This is generally undesirable in relativity. So, in this section we derive a commutator, $[\varphi(x), \varphi(y)]$ at two arbitrary spacetime points, x and y. We start by writing

$$\varphi(x) = \varphi^+(x) + \varphi^-(x), \tag{7.1.1}$$

where φ^+ and φ^- are as in Equations (6.5.6) and (6.5.8). Since φ^+ only consists of annihilation operators and φ^- only consists of creation operators we have

$$[\varphi^+(x), \varphi^+(y)] = [\varphi^-(x), \varphi^-(y)] = 0. \tag{7.1.2}$$

Hence, expanding out the commutator we have

$$[\varphi(x), \varphi(y)] = [\varphi^-(x), \varphi^+(y)] + [\varphi^-(x), \varphi^+(y)]. \tag{7.1.3}$$

Consider the first of these,

$$[\varphi^+(x), \varphi^-(y)] = \int đp \int đp' \, [a(\boldsymbol{p}), a^\dagger(\boldsymbol{p}')] e^{-ip\cdot x + ip'\cdot y} \tag{7.1.4}$$

$$= \int đp \int đp' \, \delta(\boldsymbol{p} - \boldsymbol{p}') e^{-ip\cdot x + ip'\cdot y} \tag{7.1.5}$$

$$= \int đp \, e^{-ip\cdot(x-y)} \tag{7.1.6}$$

$$=: i\Delta^+(x - y), \tag{7.1.7}$$

where the last line defines Δ^+:

$$\Delta^+(x) := i \int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3} \frac{1}{2\omega(\boldsymbol{p})} e^{ip\cdot x} \tag{7.1.8}$$

The factor of i is a convention to make Δ^+ a real function.

The second commutator can now easily be worked out using antisymmetry:

$$[\varphi^-(x), \varphi^+(y)] = -[\varphi^+(y), \varphi^-(x)] = -i\Delta^+(y - x) =: i\Delta^-(x - y), \tag{7.1.9}$$

where the last equality defines Δ^-.

Combining these results we have

$$[\varphi(x), \varphi(y)] = i\Delta^+(x - y) + i\Delta^-(x - y) =: i\Delta(x - y) \tag{7.1.10}$$

where

$$\Delta(x) := \Delta^+(x) + \Delta^-(x) \tag{7.1.11}$$

$$= i \int đp(e^{ip\cdot x} - e^{-ip\cdot x}) \tag{7.1.12}$$

$$= -2 \int đp \sin(p \cdot x). \tag{7.1.13}$$

Note that Δ is a real $(\Delta(x) = \Delta(x)^*)$, odd $(\Delta(-x) = -\Delta(x))$ function. This is what we would expect since the field is real and the commutator is antisymmetric.

Consider the full form of Δ^+:

$$i\Delta^+(x) = \int \frac{\mathrm{d}^4 p}{(2\pi)^4} 2\pi\delta(p^2 - m^2)\theta(p_0)e^{-ip\cdot x}. \tag{7.1.14}$$

We can also write Δ^- similarly:

$$i\Delta^-(x) = -\int \frac{\mathrm{d}^4 p}{(2\pi)^4} 2\pi\delta(p^2 - m^2)\theta(p_0)e^{ip\cdot x}. \tag{7.1.15}$$

If we make a change of variables, $p \to -p$, then $\mathrm{d}^4 p \to \mathrm{d}^4 p$, and

$$i\Delta^-(x) = -\int \frac{\mathrm{d}^4 p}{(2\pi)^4} 2\pi\delta(p^2 - m^2)\theta(-p_0)e^{-ip\cdot x}. \tag{7.1.16}$$

This allows us to write Δ as

$$i\Delta(x) = \int \frac{\mathrm{d}^4 p}{(2\pi)^4} 2\pi\delta(p^2 - m^2)\varepsilon(p_0)e^{-ip\cdot x} \tag{7.1.17}$$

where

$$\varepsilon(p_0) := \theta(p_0) - \theta(-p_0) = \begin{cases} +1 & p_0 > 0, \\ 0 & p_0 = 0, \\ -1 & p_0 < 0, \end{cases} \tag{7.1.18}$$

is the sign function.

Note that for proper orthochronous Lorentz transformations Δ is Lorentz invariant, since the sign of p_0 doesn't change and all other terms are manifestly Lorentz invariant.

7.2 Microcausality

Suppose x is space-like, that is $x^2 < 0$. Then there exists some proper orthochronous Lorentz transformation, Λ, taking x to $-x$. Space-like vectors lie outside the light cone, and we can view this transformation as rotating around the light cone. Since $\Delta^\pm(x)$ are individually Lorentz invariant we can apply a Lorentz transformation to one of them and not the other when computing Δ and get the same result, so

$$\Delta(x) = \Delta^+(x) + \Delta^-(x) \tag{7.2.1}$$

$$= \Delta^+(x) + \Delta^-(\Lambda x) \tag{7.2.2}$$

$$= \Delta^+(x) + \Delta^-(-x) \tag{7.2.3}$$

$$= \Delta^+(x) - \Delta^+(x) \tag{7.2.4}$$

$$= 0, \tag{7.2.5}$$

so Δ vanishes for all space-like points. Hence, $[\varphi(x), \varphi(y)]$ vanishes for all space-like separations, $x - y$.

On the other hand no such transformation exists for time-like or light-like points, since it would have to map from one half of the light cone to the other, which is a discontinuous transformation.

We can interpret this as a requirement for causality, specifically for a nonzero commutator between fields at space-like separated points which happen to have the same time coordinate we'd need some form of communication at this time between these points, and that can't happen between space-like points at the same time without violating causality.

This analysis tells us that Δ is not analytic, it has a sudden jump from zero to nonzero at light-like points.

7.3   Equal Time Commutation Relations

The equal time commutation relations can now be viewed as a special case of the covariant commutation relations. Since we have $\pi = \dot{\varphi}$ we want to compute $[\varphi(t, \boldsymbol{x}), \pi(t, \boldsymbol{x}')] = [\varphi(t, \boldsymbol{x}), \dot{\varphi}(t, \boldsymbol{x}')]$. To do this take the time derivative of Δ:

$$\frac{\partial}{\partial x^0} \Delta(x - y) = i \int \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3} \frac{1}{2\omega(\boldsymbol{p})} \frac{\partial}{\partial x^0} \left[\mathrm{e}^{ip\cdot(x-y)} - \mathrm{e}^{-ip\cdot(x-y)} \right] \tag{7.3.1}$$

$$= i \int \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3} \frac{1}{2\omega(\boldsymbol{p})} i\omega(\boldsymbol{p}) \left[\mathrm{e}^{ip\cdot(x-y)} + \mathrm{e}^{-ip\cdot(x-y)} \right] \tag{7.3.2}$$

$$= -\frac{1}{2} \int \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3} \left[\mathrm{e}^{ip\cdot(x-y)} + \mathrm{e}^{-ip\cdot(x-y)} \right]. \tag{7.3.3}$$

For the equal time commutation relations we have $x^0 = y^0$ and so

$$p \cdot (x - y) = \boldsymbol{p} \cdot (\boldsymbol{x} - \boldsymbol{y}). \tag{7.3.4}$$

Hence,

$$\frac{\partial}{\partial x^0} \Delta(x - y) = -\frac{1}{2} \int \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3} \left[\mathrm{e}^{i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{y})} + \mathrm{e}^{-i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{y})} \right] \tag{7.3.5}$$

$$= -\delta^3(\boldsymbol{x} - \boldsymbol{y}), \tag{7.3.6}$$

where we've recognised the integral representation of the Dirac delta:

$$\int \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3} \mathrm{e}^{\pm i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{y})} = \delta^3(\boldsymbol{x} - \boldsymbol{y}). \tag{7.3.7}$$

We then have

$$\frac{\partial}{\partial x^0} [i\Delta(x - y)] \Big|_{x^0=y^0=t} = \frac{\partial}{\partial x^0} [\varphi(t, \boldsymbol{x}), \varphi(t, \boldsymbol{y})] \tag{7.3.8}$$

$$= [\dot{\varphi}(t, \boldsymbol{x}), \varphi(t, \boldsymbol{y})] \tag{7.3.9}$$

$$= [\pi(t, \boldsymbol{x}), \varphi(t, \boldsymbol{y})] \tag{7.3.10}$$

$$= -i\delta^3(\boldsymbol{x} - \boldsymbol{y}) \tag{7.3.11}$$

Hence,

$$[\varphi(t, \boldsymbol{x}), \pi(t, \boldsymbol{y})] = i\delta^3(\boldsymbol{x} - \boldsymbol{y}). \tag{7.3.12}$$
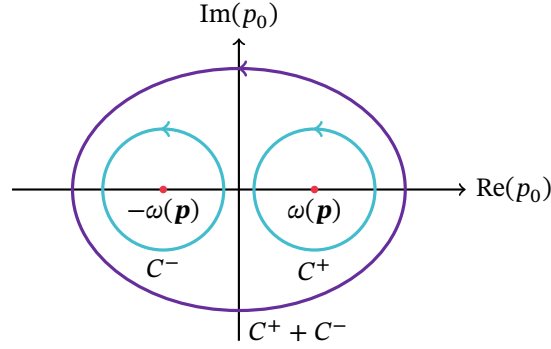
Figure 7.1: The contours used in the contour representation of $\Delta^{\pm}$ and Δ.

So, we can view the equal time commutation relations as a specific case of the covariant commutation relations. This is why the equal time commutation relations can still produce Lorentz invariant physics.

7.4 Contour Representation

(R) This section makes use of some complex analysis, in particular contour integrals and the residue theorem. See either *Methods of Theoretical Physics* or *Methods of Mathematical Physics* for details.

The contour representation of $\Delta^{\mp}$ is

$$\Delta^{\pm}(x) = -\int_{C^{\pm}} \frac{\mathrm{d}^4 p}{(2\pi)^4} \frac{\mathrm{e}^{-ip\cdot x}}{p^2 - m^2}. \tag{7.4.1}$$

The contours $C^{\pm}$ are anticlockwise circles containing the poles at $\pm\omega(\boldsymbol{p})$. This is shown in Figure 7.1.

To see why this works start with

$$p^2 - m^2 = (p_0 + \omega(\boldsymbol{p}))(p_0 - \omega(\boldsymbol{p})), \tag{7.4.2}$$

which shows there are poles at $p_0 = \pm\omega(\boldsymbol{p})$. We can then apply the residue theorem, which says that

$$\oint_{\gamma} f(z)\,\mathrm{d}z = 2\pi i \sum_i \mathrm{Res}(f, a_i) \tag{7.4.3}$$

where γ is a closed contour, f is analytic on and inside γ, a_i are the poles of f inside γ, and $\mathrm{Res}(f, a_i)$ is the residue of f at a_i.

We have a particularly easy case of two simple poles. In the case of a simple pole the residue at a_i is given by

$$\mathrm{Res}(f, a_i) = (z - a_i)f(z)|_{z=a_i}, \tag{7.4.4}$$

so

$$\mathrm{Res}\left(\frac{\mathrm{e}^{-ip\cdot x}}{(p_0 + \omega(\boldsymbol{p}))(p_0 - \omega(\boldsymbol{p}))}, \pm\omega(\boldsymbol{p})\right) = \left.\frac{\mathrm{e}^{-ip\cdot x}}{p_0 \mp \omega(\boldsymbol{p})}\right|_{p_0=\pm\omega(\boldsymbol{p})} = \mp\frac{\mathrm{e}^{-ip\cdot x}}{2\omega(\boldsymbol{p})}. \tag{7.4.5}$$

Hence,

$$-\int_{C^+}\frac{\mathrm{d}p_0}{2\pi}\frac{\mathrm{e}^{-ip\cdot x}}{p^2-m^2}=-i\frac{\mathrm{e}^{-ip\cdot x}}{2\omega(\boldsymbol{p})}\bigg|_{p_0=\omega(\boldsymbol{p})},\tag{7.4.6}$$

and

$$-\int_{C^+}\frac{\mathrm{d}p_0}{2\pi}\frac{\mathrm{e}^{-ip\cdot x}}{p^2-m^2}=i\frac{\mathrm{e}^{ip\cdot x}}{2\omega(\boldsymbol{p})}\bigg|_{p_0=-\omega(\boldsymbol{p})}.\tag{7.4.7}$$

That is,

$$-\int_{C^\pm}\frac{\mathrm{d}p_0}{2\pi}\frac{\mathrm{e}^{-ip\cdot x}}{p^2-m^2}=\mp i\frac{\mathrm{e}^{\mp ip\cdot x}}{2\omega(\boldsymbol{p})}\bigg|_{p_0=\omega(\boldsymbol{p})}.\tag{7.4.8}$$

Having performed the integral over p_0 we are left with an integral over $\boldsymbol{p}$ so

$$-\int_{C^\pm}\frac{\mathrm{d}^4p}{(2\pi)^4}\frac{\mathrm{e}^{-ip\cdot x}}{p^2-m^2}=\mp i\int\frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3}\frac{\mathrm{e}^{mpip\cdot x}}{2\omega(\boldsymbol{p})}=\Delta^\pm(x).\tag{7.4.9}$$

The contour representation of Δ is then simply given by

$$\Delta(x)=-\int_{C}\frac{\mathrm{d}^4p}{(2\pi)^4}\frac{\mathrm{e}^{-ip\cdot x}}{p^2-m^2}\tag{7.4.10}$$

where $C=C^++C^-$ is the combined contour, with the middle parts cancelling out, to give a contour containing both poles as shown in Figure 7.1.

7.5 Δ as Green's Functions

To interpret Δ and $\Delta^\pm$ consider the Klein–Gordon equation,

$$0=(\partial^2+m^2)\varphi.\tag{7.5.1}$$

Taking the Fourier transform and using the rule that $\partial_x\to-ip$ we have

$$0=((-ip)^2+m^2)\tilde{\varphi}\implies(p^2-m^2)\tilde{\varphi}=0.\tag{7.5.2}$$

Transforming back then gives us the contour representation of Δ or $\Delta^\pm$, depending on the boundary conditions.

This means we can interpret Δ and $\Delta^\pm$ as classical Green's functions for the Klein–Gordon equation, so

$$(\partial^2+m^2)\Delta(x)=\delta^4(x).\tag{7.5.3}$$

We call Δ and $\Delta^\pm$ propagators, since they tell us how a quantum propagates between states.

Part II

Interactions

Eight

Scattering

8.1 Interactions

So far, we have only considered Lagrangians which are quadratic in the field, we call these **free field Lagrangians**, since quadratic Lagrangians give linear equations of motion with plane wave solutions, which we interpret as many independent harmonic oscillators, each corresponding to a noninteracting particle. In order to have our particles do anything, such as be measured, we need interactions. This means introducing a nonquadratic term to our Lagrangian.

8.1.1 φ^3 Theory

The simplest nonquadratic term that we can add to our Lagrangian is a cubic term, giving the Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \varphi)(\partial^\mu \varphi) - \frac{1}{2}m^2\varphi^2 - \frac{1}{6}g\varphi^3. \tag{8.1.1}$$

The first two terms are just the usual Klein–Gordon Lagrangian, call it $\mathcal{L}$. We can treat this as the Lagrangian for a free particle. The extra term, $-g\varphi^3/6$, is our interaction term, call it $\mathcal{L}_\mathrm{I}$. We call a theory with a cubic term like this a $\boldsymbol{\varphi^3}$ **theory** The factor of 1/6 in this factor is just conventional, its there to make equations simpler later. A φ^n theory would equivalently have a factor of $1/n!$, each time we differentiate this factor becomes closer to 1. The factor of g is called the **coupling constant**. It measures how strong the interaction is.

We can easily apply the Euler–Lagrange equations to this modified Lagrangian, without much changing. We have

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} = \partial^2 \varphi, \qquad \text{and} \qquad \frac{\partial \mathcal{L}}{\partial \varphi} = -m^2\varphi - \frac{1}{3}g\varphi^2. \tag{8.1.2}$$

Hence, we have the equations of motion

$$(\partial^2 + m^2)\varphi = -\frac{1}{2}g\varphi^2. \tag{8.1.3}$$

This is just the Klein–Gordon equation with some sort of source term.

This is now a nonlinear differential equation. These are, in general, hard to solve. There is no general theory and we are usually restricted to specific examples with solutions. Because of this we assume that the interaction isn't very strong, that is that $g \ll 1$, and then we solve by expanding in g and truncating at some appropriate point in our calculations.

8.2 The S Matrix

Most processes in physics are, when it comes down to it, scattering. We see light after it scatters off an object, two objects collide in classical mechanics, two protons collide in the LHC, a whole bunch of particles collide all over the place and a statistical study of this gives statistical mechanics, and so on. Scattering is a very simple process, there are three steps:

- We start in some initial state with free particles of definite momentum, spin, etc.

- The particles undergo some scattering process, which we treat as a black box.

- We finish in some final state with free particles of definite momentum, spin, etc.

In order to allow us to consider free particles we assume that the initial state happens a long long time before the scattering, taking $t \to -\infty$. Similarly we assume the final state is a long long time after the scattering, taking $t \to +\infty$. Call the initial state $|\Psi, -\infty\rangle$ and the final state $|\Psi, +\infty\rangle$. Then an in between state at time t is $|\Psi, t\rangle$.

The initial and final state must somehow be related if we are to pass from one to the other. We take this relation to be of the form

$$|\Psi, +\infty\rangle = S|\Psi, -\infty\rangle \tag{8.2.1}$$

where S, known as the **S matrix**, is the operator describing the scattering process.

Now suppose that we start with some fixed, known initial state, $|i\rangle$, so $|\Psi, -\infty\rangle = |i\rangle$. We can choose some complete orthonormal basis of possible final states, $\{|f\rangle\}$, which allows us to expand the final state as

$$|\Psi, +\infty\rangle = \sum_f |f\rangle\langle f|\Psi, +\infty\rangle. \tag{8.2.2}$$

We interpret

$$|\langle f|\Psi, +\infty\rangle|^2 = \mathbb{P}(i \to f) \tag{8.2.3}$$

as the probability that when we start in state $|i\rangle$ our scattering process finishes in state $|f\rangle$.

Using the definition of S we can write the amplitude as

$$\langle f|\Psi, +\infty\rangle = \langle f|S|\Psi, -\infty\rangle = \langle f|S|i\rangle =: S_{fi}, \tag{8.2.4}$$

where $S_{fi} := \langle f|S|i\rangle$ is a matrix element of the scattering operator. So the matrix elements of the scattering operator tell us the amplitude for a particular scattering process.

We can expand the final state as

$$|\Psi, +\infty\rangle = \sum_f |f\rangle S_{fi}, \tag{8.2.5}$$

and take the conjugate to get

$$\langle \Psi, +\infty| = \sum_f S_{if}^* \langle f|. \tag{8.2.6}$$

Combining this, and working with normalised states, we get

$$1 = \langle \Psi, +\infty|\Psi, +\infty\rangle = \sum_{f,f'} S_{fi} S_{if'}^* \langle f|f'\rangle = \sum_{f,f'} S_{fi} S_{if'}^* \delta_{ff'} = \sum_f S_{if}^* S_{fi} \tag{8.2.7}$$

or in terms of operators,

$$1 = S^\dagger S, \tag{8.2.8}$$

so the scattering operator is unitary.

We can interpret this as conservation of probability, since $\{|f\rangle\}$ is a complete set of states we must have

$$\sum_f \mathbb{P}(i \to f) = 1, \tag{8.2.9}$$

and hence if we have a definite start state that is also a complete set, so the evolution must be unitary.

Note that the notion of the S matrix as discussed here is more general than in relativistic quantum mechanics since in QFT we can create and destroy particles, so our scattering can include processes like annihilation or pair production.

8.3   Interaction Picture

We now introduce the third, and final, picture of quantum mechanics, called the **interaction picture**, or **Dirac picture**. Suppose we have a Hamiltonian

$$H = H_0 + H_I \tag{8.3.1}$$

where H_0 describes a free field and H_I an interaction.

In the Dirac picture both operators and states have time dependence. The states evolve as if they were in the Schrödinger picture, but under only the free Hamiltonian, so

$$|\psi, t\rangle_D := e^{iH_0 t}|\psi, t\rangle_S = e^{iH_0 t}e^{-iHt}|\psi\rangle_H \tag{8.3.2}$$

where subscripts D, S, and H denote the Dirac, Schrödinger, and interaction pictures respectively. The first equality is the definition of a state in the Dirac picture, and the second comes from inverting the definition of a state in the Heisenberg picture:

$$|\psi\rangle_H := e^{iHt}|\psi, t\rangle_S. \tag{8.3.3}$$

Operators in the Dirac picture transform as if they were in the Heisenberg picture, but only under the free Hamiltonian, so

$$A_D(t) = e^{iH_0 t} A_S e^{-iH_0 t}. \tag{8.3.4}$$

Take the derivative of the equation defining a state in the Dirac picture, this gives

$$i\frac{\partial}{\partial t}|\psi, t\rangle_D = -H_0 e^{iH_0 t}|\psi, t\rangle_S + e^{iH_0 t} i\frac{\partial}{\partial t}|\psi, t\rangle_S. \tag{8.3.5}$$

Now recognise the final term as $e^{iH_0 t}$ times one side of the Schrödinger equation:

$$i\frac{\partial}{\partial t}|\psi, t\rangle_S = H|\psi, t\rangle_S = H_0|\psi, t\rangle_S + H_I|\psi, t\rangle_S. \tag{8.3.6}$$

Substituting this into the previous equation the first term cancels leaving us with

$$i\frac{\partial}{\partial t}|\psi, t\rangle_D = e^{iH_0 t} H_I|\psi, t\rangle_S = e^{iH_0 t} H_I e^{-iH_0 t}|\psi, t\rangle_D = H_I^D|\psi, t\rangle_D. \tag{8.3.7}$$

This shows that states evolve in time according to the free Hamiltonian, H_0. The time evolution due to the interaction is exhibited by the operators.

We derived the Heisenberg equation by differentiating an arbitrary operator in both pictures. Doing the same here, between the Dirac and Schrödinger picture gives rise to the equation of motion

$$\frac{\partial}{\partial t}A_D(t) = i[H_0, A_D(t)]. \tag{8.3.8}$$

Note that $H_0^S = H_0^D = H_0^H$, but in general $H_I^D \neq H_I^S$, which comes from the general fact that $[H_0, H_I] \neq 0$.

From now on we will work entirely in the Dirac picture. Accordingly we will drop the labels D, S, and H.

8.4 Dyson Series

The **Dyson series** is a formal solution to the evolution equation

$$i\frac{\partial}{\partial t}|\Psi, t\rangle = H_I(t)|\Psi, t\rangle. \tag{8.4.1}$$

Now we make the time dependence of the interaction Hamiltonian explicit. The solution to this is

$$|\Psi, t\rangle = |\Psi, -\infty\rangle - i\int_{-\infty}^{t} dt_1\, H_I(t_1)|\Psi, t_1\rangle. \tag{8.4.2}$$

To see this just take the derivative:

$$i\frac{\partial}{\partial t}|\Psi, t\rangle = i\frac{\partial}{\partial t}|\Psi, -\infty\rangle + \frac{\partial}{\partial t}\int_{-\infty}^{t} dt_1\, H_I(t_1)|\Psi, t_1\rangle = H_I(t)|\Psi, t\rangle. \tag{8.4.3}$$

Here the first term vanishes as it is constant and the second term is easily computed by an application of the fundamental theorem of calculus.

The problem is that this solution for $|\Psi, t\rangle$ has $|\Psi, t_1\rangle$ in it, in particular we integrate up to $t_1 = t$. This means that we need to know the state at time t in order to evaluate the state at time t. This seems like a problem. The solution is that we are assuming weak interactions. This means that the operator $H_I(t)$ is, in some sense, "small", and the operator $H_I(t)H_I(t_1)$ is "smaller". Of course, operators don't have sizes in this way, but we can think about specific matrix elements and it all works out. This suggests that we can solve this problem recursively with each level of recursion being a smaller and smaller correction. The first level of recursion is to substitute $|\Psi, t_1\rangle$ with the entire right hand side of Equation (8.4.2), being careful to avoid reusing integration variables, giving

$$|\Psi, t\rangle = |\Psi, -\infty\rangle - i \int_{-\infty}^{t} dt_1\, H_I(t_1) \left[|\Psi, -\infty\rangle - i \int_{-\infty}^{t_1} dt_2\, H_I(t_2)|\Psi, t_2\rangle \right]. \quad (8.4.4)$$

Note that the second integral is quadratic in the Hamiltonian, so is "smaller" than the rest of the terms. We can repeat this process, writing $|i\rangle = |\Psi, -\infty\rangle$ for compactness, and we get

$$|\Psi, t\rangle = |i\rangle - i \int_{-\infty}^{t} dt_1\, H_I(t_1) \left[|i\rangle - i \int_{-\infty}^{t_1} dt_2\, H_I(t_2) \left\{ |i\rangle - i \int_{-\infty}^{t_2} dt_3\, H_I(t)|\Psi, t_3\rangle \right\} \right].$$

Continuing on like this we have a series

$$|\Psi, t\rangle = |i\rangle + (-i) \int_{-\infty}^{t} dt_1\, H_I(t_1)|i\rangle + (-i)^2 \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2 H_I(t_1)H_I(t_2)|i\rangle$$
$$+ (-i)^3 \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2 \int_{-\infty}^{t_2} dt_3\, H_I(t_1)H_I(t_2)H_I(t_3)|i\rangle + \cdots. \quad (8.4.5)$$

Each term in this series is "smaller" than the previous term and so we can truncate at some point once we've achieved the required level of accuracy.

Taking t to infinity we get

$$|\Psi, +\infty\rangle = |i\rangle + (-i) \int_{-\infty}^{\infty} dt_1\, H_I(t_1)|i\rangle + (-i)^2 \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2 H_I(t_1)H_I(t_2)|i\rangle$$
$$+ (-i)^3 \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2 \int_{-\infty}^{t_2} dt_3\, H_I(t_1)H_I(t_2)H_I(t_3)|i\rangle + \cdots. \quad (8.4.6)$$

Recognising that this has the form of an operator, albeit an infinite series of integrals, acting on the initial state to give the final state we can identify this operator as the S matrix:

$$S = 1 + (-i) \int_{-\infty}^{\infty} dt_1\, H_I(t_1) + (-i)^2 \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2 H_I(t_1)H_I(t_2)$$
$$+ (-i)^3 \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2 \int_{-\infty}^{t_2} dt_3\, H_I(t_1)H_I(t_2)H_I(t_3) + \cdots. \quad (8.4.7)$$

Or, more compactly,

$$S = \sum_{n=0}^{\infty} (-i)^n \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{t_1} dt_2 \cdots \int_{-\infty}^{t_{n-1}} dt_n\, H_I(t_1)H_I(t_2) \cdots H_I(t_n). \quad (8.4.8)$$

Notice that the times, t_1, t_2, and so on appear in increasing order, that is in the integral $t_1 < t_2 < \cdots < t_n$, we can neglect the endpoints of the integration ranges where we may have equality since individual points don't contribute to an integral.

To proceed we make use of the following identity:

$$\int_{t_a}^{t_b} dt_1 \int_{t_a}^{t_1} dt_2 \int_{t_a}^{t_2} \cdots \int_{t_a}^{t_{n-1}} dt_n A(t_1)A(t_2)\cdots A(t_n)$$

$$= \frac{1}{n!} \int_{t_a}^{t_b} dt_1 \int_{t_a}^{t_b} dt_2 \cdots \int_{t_a}^{t_b} dt_n\, \mathrm{T}[A(t_1)A(t_2)\cdots A(t_n)]. \quad (8.4.9)$$

Here t_a and t_b are arbitrary initial and final times with $t_a < t_b$, and A is an arbitrary operator. The operator T is called the **time ordering**. We take the arguments, which are operators, and order according to evaluation time in order of increasing time from left to right, so that the first operator to act on some state is the operator evaluated at the earliest time. That is, for two operators

$$\mathrm{T}[A(t_1), B(t_2)] := \begin{cases} A(t_1)B(t_2) & t_1 > t_2 \\ B(t_2)A(t_1) & t_2 > t_1 \end{cases} \quad (8.4.10)$$

$$= \theta(t_1 - t_2)A(t_1)B(t_2) + \theta(t_2 - t_1)B(t_2)A(t_1). \quad (8.4.11)$$

This can then be extended to n operators in the obvious way, although it becomes harder to write in terms of Heaviside step functions.

We'll show how this works for the $n = 2$ case, and then the general case follows by induction. Start with the right hand side of the identity. For $n = 2$ we have

$$I = \int_{t_a}^{t_b} dt_1 \int_{t_a}^{t_b} dt_2\, \mathrm{T}[A(t_1)A(t_2)]. \quad (8.4.12)$$

We can split the integral over t_2 into two parts, one running from t_a to t_1, and the other from t_1 to t_b:

$$I = \int_{t_a}^{t_b} dt_1 \left[\int_{t_a}^{t_1} dt_2\, \mathrm{T}[A(t_1)A(t_2)] + \int_{t_2}^{t_b} dt_2\, \mathrm{T}[A(t_1)A(t_2)] \right]. \quad (8.4.13)$$

Consider the second term,

$$\int_{t_a}^{t_b} dt_1 \int_{t_2}^{t_b} dt_2\, \mathrm{T}[A(t_1)A(t_2)]. \quad (8.4.14)$$

This is a double integral. This corresponds to integrating over the triangle in the plane bounded by $t_1 = t_2$, $t_1 = t_a$, and $t_2 = t_b$. We can integrate over the same region in a different way. Instead integrate along t_2 from t_a to t_1 and then along t_1 from t_a to t_b. This corresponds to the integral

$$\int_{t_a}^{t_b} dt_2 \int_{t_a}^{t_2} dt_1\, \mathrm{T}[A(t_1)A(t_2)]. \quad (8.4.15)$$

Now, t_1 and t_2 are just integration variables, and the integrand is symmetric in t_1 and t_2, as $\mathrm{T}[A(t_1)A(t_2)] = \mathrm{T}[A(t_2)A(t_1)]$, the ordering of operators at different times within a time ordering is not important, since we're going to time order
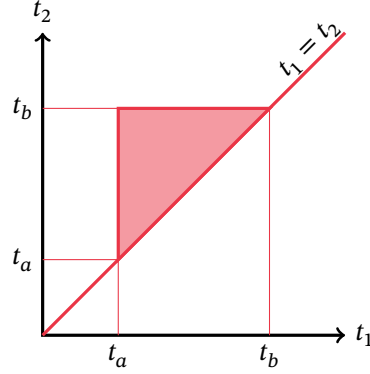
Figure 8.1: The region of integration for the $n = 2$ case of the identity.

them. We are then free to rename t_1 and t_2, and in particular swap them without changing anything else,

$$\int_{t_a}^{t_b} dt_1 \int_{t_a}^{t_1} dt_2 \ T[A(t_1)A(t_2)]. \tag{8.4.16}$$

Combing this back with the first term we get

$$I = 2 \int_{t_a}^{t_b} dt_1 \int_{t_a}^{t_1} dt_2 \ T[A(t_1)A(t_2)]. \tag{8.4.17}$$

Noticing with these integration ranges the operators are already time ordered, so we can drop the explicit time ordering, and that $2 = 2!$ and dividing through by 2 we have proven the identity for $n = 2$.

Now, suppose that the identity holds for k integrals, and we have $k+1$ integrals. We can use the induction hypothesis to swap k integrals, and then we pairwise swap integrals as above before the unswapped integral has also been swapped.

We can use this identity to write

$$S = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \cdots \int_{-\infty}^{\infty} dt_n \ T[H_I(t_1)H_I(t_2) \cdots H_I(t_n)]. \tag{8.4.18}$$

We can now replace H_I with an integral over the Hamiltonian density giving

$$S = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int dx_1 \int dx_2 \cdots \int dx_n \ T[\mathcal{H}_I(x_1)\mathcal{H}_I(x_2) \cdots \mathcal{H}_I(x_n)]. \tag{8.4.19}$$

This is called the **Dyson series**.

Note that the interaction Hamiltonian density is given by $\mathcal{H}_I = -\mathcal{L}_I$. This follows directly from the definition

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_I = \pi\dot{\varphi} - \mathcal{L} = \pi\dot{\varphi} - \mathcal{L}_0 - \mathcal{L}_I \tag{8.4.20}$$

where the $\pi\dot{\varphi} - \mathcal{L}_0$ term is the unperturbed Hamiltonian density. Using this we can write

$$S = \sum_{n=0}^{\infty} \frac{i^n}{n!} \int dx_1 \int dx_2 \cdots \int dx_n \; \mathrm{T}[\mathcal{L}\mathrm{I}(x_1)\mathcal{L}\mathrm{I}(x_2)\cdots\mathcal{L}\mathrm{I}(x_n)] \qquad (8.4.21)$$

$$= \exp\left[i \int d^4x \mathcal{L}_{\mathrm{I}} \right]. \qquad (8.4.22)$$

So the S matrix can be viewed as the phase factor given by the interaction action. This last step is simply a formal rewriting recognising the exponential series. When expanding make sure to have each integral in each power be over a different variable, and to time order the integrand. From this form we can see that, assuming the action is Lorentz invariant, then so is the S matrix. This works because time ordering is Lorentz invariant, because it can be done using only terms like $\theta(t - t')$, and each such term is invariant under proper orthochronous Lorentz transformations.

8.5 Interpretation

For a given initial state, $|i\rangle$, and final state, $|f\rangle$, we need to find terms in the interaction Lagrangian, $\mathcal{L}_{\mathrm{I}}$ which give a nonzero transition amplitude for $|i\rangle$ to $|f\rangle$. The entire effect of the interaction is in $\mathcal{L}_{\mathrm{I}}$, which is an operator given by a polynomial in the fields. This means that really it's just a collection of creation and annihilation operators. Therefore the entire effect of the interaction is to create and destroy particles. We then look for terms with

- annihilation operators destroying the initial state,

- creation operators creating the final state,

- paired creation/annihilation operators creating and then destroying intermediate states.

In terms of particles, we look for terms which destroy the incoming particles, potentially create and then destroy one or more new particles, then create some particles which form the final state. The initial and final particles must be on-shell for us to observe them. The intermediate particles however needn't be on-shell, since we cannot observe them. These are called **virtual particles**. Don't put too much stock in the existence and off-shell nature of these virtual particles, they're really just a mathematical construct with different combinations of intermediate particles corresponding to different terms in the Dyson series.

In the interaction picture the fields satisfy the free field equation, and therefore have the usual creation and annihilation operators. We can therefore use normal ordering, placing all creation operators on the left and all annihilation operators on the right, in order to ensure that incoming particles are destroyed and then new particles are created, rather than the other way round.
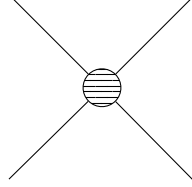
8.6 φ^3 Interactions

In φ^3 theory the interaction Hamiltonian is

$$\mathcal{H}_{\mathrm{I}} = -\mathcal{L}_{\mathrm{I}} = \frac{1}{6}g{:}\varphi^3{:}, \qquad (8.6.1)$$

where we now normal order the φ^3 factor in order to destroy existing particles and then create new particles.

We will consider 2–2 scattering, that is two particles come in, interact in some way, and two particles leave. This can be represented by a diagram like

$$(8.6.2)$$

Here the blob represents some unknown black box type interaction. Time flows, in our convention, from left to right, so the left two legs represent the incoming particles, and the right two the outgoing particles. In order to destroy the incoming particles we start with a $\varphi^+\varphi^+$ term. Then to create two new particles we need a $\varphi^-\varphi^-$ term. Recall that we order operators from right to left, as we want the $\varphi^+\varphi^+$ term to act on the initial before the $\varphi^-\varphi^-$ term. We could therefore look for a

$$\varphi^-\varphi^-\varphi^+\varphi^+ \tag{8.6.3}$$

term in the series. However, we won't find such a term since in φ^3 theory we can only have terms with a multiple of 3 factors of $\varphi^\pm$. This means that there are no direct 2–2 scattering events in φ^3 theory, such events must always go via at least one intermediate virtual particle state. So, next we look at the term in the series which is quadratic in the Hamiltonian, and hence has six factors of $\varphi^\pm$. There are two possible orderings of terms with the required starting point of $\varphi^+\varphi^+$ to annihilate the initial state and end point of $\varphi^-\varphi^-$ to create the final state. They are

$$\varphi^-\varphi^-\varphi^-\varphi^+\varphi^+\varphi^+, \qquad \text{and} \qquad \varphi^-\varphi^-\varphi^+\varphi^-\varphi^+\varphi^+. \tag{8.6.4}$$

When acting on the initial state the the first vanishes, since after annihilating the initial state we are left with the vacuum state, which we then act on with the annihilator φ^+. So, the only nonzero second term quadratic in the Hamiltonian is

$$\varphi^-\varphi^-\varphi^+\varphi^-\varphi^+\varphi^+. \tag{8.6.5}$$

We interpret this as annihilating the initial state, creating a new particle, annihilating it, and then destroying it again. The amplitude for this particular term is proportional to g^2, since each Hamiltonian introduces a factor of g. Further, by assuming that energy is conserved and everything is local when we annihilate the two initial particles we must immediately create the virtual particle, and all of this must happen at the same position, so this is a single spacetime event, x_2. Then when the virtual particle is annihilated we must immediately, and at the same location, create the new particles, so this is a second spacetime event, x_1. These correspond to the two points at which the interaction Hamiltonians are evaluated so we'll have a nonzero term

$$-\frac{g^2}{36} \int_{-\infty}^{\infty} d^4x_1 \int_{-\infty}^{\infty} d^4x_2 \, \varphi^-(x_2)\varphi^-(x_2)\varphi^+(x_1)\varphi^-(x_2)\varphi^+(x_2)\varphi^+(x_2). \tag{8.6.6}$$

Pictorially we can represent this as

$$(8.6.7)$$

Here the left and right legs represent the incoming and outgoing particles, and the single line in the middle is the virtual particle.

Next, we would consider terms cubic in the Hamiltonian. However, things quickly get out of hand if we proceed as we did here. The cubic term will have nine factors of φ. So, we want a way to do this all more systematically. This is the theory we will develop in the next chapter.

Nine

Wick's Theorem

Appendices

A

Pre-Course Revision

The content of this appendix is based on the notes released before the course began to recap key ideas from previous courses. In particular, it covers material from Quantum Theory and Classical Electrodynamics.

A.1 Special Relativity

A.1.1 Four Vectors

Consider some event occurring at the spatial position $\boldsymbol{x}$ and at time t. The coordinates of this event, in four-dimensional Minkowski space, form a **contravariant** four-vector, which we write with upper indices:

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, \boldsymbol{x}). \tag{A.1.1}$$

The **covariant** four-vector has lower indices:

$$x_\mu = (x_0, x_1, x_2, x_3) = (ct, -\boldsymbol{x}). \tag{A.1.2}$$

Similarly a general four-vector has a contravariant and covariant form:

$$a^\mu = (a^0, a^1, a^2, a^3) = (a^0, \boldsymbol{a}), \quad \text{and} \quad a_\mu = (a_0, a_1, a_2, a_3) = (a^0, -\boldsymbol{a}). \tag{A.1.3}$$

In particular, $a^0 = a_0$ and $a^i = -a_i$, where we follow the convention that Greek indices run from 0 to 3 and Latin indices run from 1 to 3.

We move between upper and lower indices using the **metric tensor**, η, also denoted g:

$$a^\mu = \eta^{\mu\nu} a_\nu, \quad \text{and} \quad a_\mu = \eta_{\mu\nu} a^\nu, \tag{A.1.4}$$

where we apply the summation convention, summing over repeated indices with one up and one down. The covariant and contravariant components of the Minkowski metric are given by

$$\eta^{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{A.1.5}$$

and the mixed components are given by

$$\eta_\mu{}^\nu = \eta^\nu{}_\mu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \delta_\mu{}^\nu = \delta^\nu{}_\mu. \qquad (A.1.6)$$

Here $\delta^\mu{}_\nu$ is the Kronecker delta, defined to be 1 if $\mu = \nu$ and 0 otherwise. Note that we have made a choice of signs here.

> ! Sometimes the Minkowski metric is instead chosen to be
>
> $$\eta^{\mu\nu} \stackrel{!}{=} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \qquad (A.1.7)$$
>
> These two conventions are known as $(+---)$ and $(-+++)$ for short, where the signs refer to the signs along the diagonal.

The **scalar product** of two four-vectors, a and b, in Minkowski space is defined to be

$$a \cdot b := a^\mu b_\mu = a_\mu b^\mu = a_\mu b_\nu \eta^{\mu\nu} = a^\mu b^\nu \eta_{\mu\nu} = a^0 b^0 - \boldsymbol{a} \cdot \boldsymbol{b} \qquad (A.1.8)$$

where $\boldsymbol{a} \cdot \boldsymbol{b} = a^1 b^1 + a^2 b^2 + a^3 b^3$ is the normal three-dimensional scalar product.

A.1.2 Lorentz Transformations

Lorentz transformations are linear transformations on four-vectors which leave the scalar product invariant. If Λ is a Lorentz transformation then the four-vector a transforms as

$$a^\mu \to a'^\mu = \Lambda^\mu{}_\nu a^\nu. \qquad (A.1.9)$$

These are the **homogeneous Lorentz transformations**, and they form a group called the **homogeneous Lorentz group**, or just the Lorentz group, denoted $O(1, 3)$, where $(1, 3)$ is the signature of the metric, one time component and three spatial components. Including translations as well we get the Poincaré group, $\mathbb{R}^{1,3} \rtimes O(1, 3)$, where $\mathbb{R}^{1,3}$ denotes translations in Minkowski space[1].

The "standard" Lorentz transformation is a **boost** along the x-axis, where the boosted frame moves at a constant speed, v, along the x-axis:

[1] for more details on groups see Symmetries of Quantum Mechanics or Symmetries of Particles and Fields

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \cosh\omega & -\sinh\omega & 0 & 0 \\ -\sinh\omega & \cosh\omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad (A.1.10)$$

where

$$\tanh\omega = \beta := \frac{v}{c}, \qquad (A.1.11)$$

$$\cosh\omega = \gamma := \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-v^2/c^2}}, \qquad (A.1.12)$$

$$\sinh\omega = \gamma\beta. \qquad (A.1.13)$$

In terms of coordinates we have

$$ct' = \gamma\left(ct - \frac{v}{c}x\right), \tag{A.1.14}$$

$$x' = \gamma(x - vt), \tag{A.1.15}$$

$$y' = y, \tag{A.1.16}$$

$$z' = z. \tag{A.1.17}$$

Note that the Lorentz group contains ordinary three-dimensional rotations as a subgroup, $O(3)$, for example, the following Lorentz transformation is a clockwise rotation by φ about the z-axis, leaving time unaffected:

$$\Lambda^{\mu}{}_{\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi & 0 \\ 0 & -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{A.1.18}$$

By definition the scalar product is invariant under the action of the Lorentz group. This means that for some Lorentz transformation Λ and four-vectors a and b we have

$$b'^{\mu}b'^{\nu}\eta_{\mu\nu} = \Lambda^{\mu}{}_{\rho}a^{\rho}\Lambda^{\nu}{}_{\sigma}b^{\sigma}\eta_{\mu\nu} = a^{\rho}b^{\sigma}\eta_{\rho\sigma}. \tag{A.1.19}$$

The last equality is given by computing the scalar product after the transformation. Since a and b are arbitrary this means we must have

$$\Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}\eta_{\mu\nu} = \eta_{\rho\sigma}. \tag{A.1.20}$$

We can rewrite this as

$$(\Lambda^{\mathsf{T}})_{\rho}{}^{\mu}\eta_{\mu\nu}\Lambda^{\nu}{}_{\sigma} = \eta_{\rho\sigma}. \tag{A.1.21}$$

In matrix notation we then have $\Lambda^{\mathsf{T}}\eta\Lambda = \eta$. Taking the determinant we have

$$\det(\Lambda^{\mathsf{T}}\eta\Lambda) = \det(\Lambda^{\mathsf{T}})\det(\eta)\det(\Lambda) = \det(\eta) \implies (\det\Lambda)^2 = 1, \tag{A.1.22}$$

where we've used $\det(AB) = \det(A)\det(B)$, and $\det(A^{\mathsf{T}}) = \det(A)$. Hence, we have

$$\det\Lambda = \pm 1 \tag{A.1.23}$$

for all $\Lambda \in O(1,3)$. Transformations with $\det\Lambda = +1$ are called **proper Lorentz transformations**, and form a subgroup, $SO(1,3)$. Lorentz transformations with $\det\Lambda = -1$ are called **improper Lorentz transformations**.

Now set the free indices, ρ and σ, to 0 in and we get

$$\Lambda^{\mu}{}_{0}\Lambda^{\nu}{}_{0}\eta_{\mu\nu} = \eta_{00} = 1. \tag{A.1.24}$$

Writing out the components this becomes

$$\Lambda^{0}{}_{0}\Lambda^{0}{}_{0} - \Lambda^{1}{}_{0}\Lambda^{1}{}_{0} - \Lambda^{2}{}_{0}\Lambda^{2}{}_{0} - \Lambda^{3}{}_{0}\Lambda^{3}{}_{0} = 1 \tag{A.1.25}$$

which gives

$$(\Lambda^{0}{}_{0})^2 = 1 + \sum_{i=1}^{3}(\Lambda^{i}{}_{0})^2 \geq 1, \tag{A.1.26}$$

where the inequality follows from the fact that $\Lambda^i{}_0$ is real, so its square is nonnegative. We can conclude from this that either

$$\Lambda^0{}_0 \geq 1 \quad \text{or} \quad \Lambda^0{}_0 \leq -1. \tag{A.1.27}$$

We call transformations with $\Lambda^0{}_0 \geq 1$ **orthochronous Lorentz transformations**, these form a subgroup, $O^+(1,3)$. There is also a subgroup of proper orthochronous Lorentz transformations, $SO^+(1,3)$.

We can divide homogeneous Lorentz transformations into four classes, or in more technical language, the Lie group $O(1,3)$ has four connected components:

$$\begin{array}{ccc} & \det \Lambda = +1 & \det \Lambda = -1 \\ \Lambda^0{}_0 \geq +1 & \text{I} & \text{II} \\ \Lambda^0{}_0 \leq -1 & \text{III} & \text{IV} \end{array} \tag{A.1.28}$$

It is component I, the proper orthochronous Lorentz transformations, which contains the identity and so is the most important. In particular, it is the Lie group generated by the Lie algebra $\mathfrak{so}(1,3)$.

Any transformation of type II, III, or IV can be generated by combing a proper orthochronous transformation of type I with one of the following:

- **Spatial inversion**: The parity operation $x^0 \to x^0$ and $\boldsymbol{x} \to -\boldsymbol{x}$:

$$\Lambda^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{A.1.29}$$

- **Time reversal**: the operation $x^0 \to -x^0$ and $\boldsymbol{x} \to \boldsymbol{x}$:

$$\Lambda^\mu{}_\nu = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{A.1.30}$$

- **Spacetime inversion**: The operation $x^\mu \to -x^\mu$:

$$\Lambda^\mu{}_\nu = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{A.1.31}$$

If x and y are four vectors then the **spacetime interval** $s^2 := (x-y)^\mu(x-y)_\mu$ is invariant under Lorentz transformations, since its just a scalar product of some four-vector $x - y$.

A.1.3 Classification of Four-Vectors

We classify four-vectors, a, by whether their scalar product with themselves is positive, negative, or zero:

- if $a^2 > 0$ we call a **time-like**,

- if $a^2 = 0$ we call a **light-like**,

- if $a^2 < 0$ we call a **space-like**.

A.1.4 Differential Operators

The gradient operator in normal $\mathbb{R}^n$ consists of spatial derivatives at each component. The generalisation to Minkowski space includes a time derivative, and a factor of c for dimensional consistency:

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \nabla\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right). \tag{A.1.32}$$

Note that the derivative with respect to a contravariant vector gives a covariant operator. As with normal four-vectors we can raise the index to get the operator

$$\partial^\mu = \frac{\partial}{\partial x_\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}\right). \tag{A.1.33}$$

The Laplacian, $\nabla^2 = \nabla \cdot \nabla$, can also be generalised to Minkowski space as the scalar product of ∂ with itself, we call the result the **d'Alembertian**:

$$\partial^2 := \partial_\mu \partial^\mu = \partial^\mu \partial_\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}. \tag{A.1.34}$$

This is also denoted $\Box^2$ or $\Box$. Note that the wave equation for some field, φ, can then be written as $\partial^2 \varphi = 0$.

A.1.5 Momentum and Energy

The four-momentum, p, is a four-vector given by

$$p^\mu = \left(\frac{E}{c}, \boldsymbol{p}\right) \tag{A.1.35}$$

where E is the energy of the particle and $\boldsymbol{p}$ is the relativistic three-momentum of the particle.

For a free particle of mass m we have that

$$p^2 = \frac{E^2}{c^2} - \boldsymbol{p} \cdot \boldsymbol{p} = m^2 c^2. \tag{A.1.36}$$

This can be rewritten as

$$E^2 = |\boldsymbol{p}|^2 c^2 + m^2 c^4, \tag{A.1.37}$$

which is known as the **energy-momentum relation**, or the **mass-shell condition**.

Particles for which $p^\mu p_\mu = m^2 c^2$ are said to be on their **mass-shell**.

A.2 Electromagnetic Theory

A.2.1 Maxwell's Equations

We use **Heaviside–Lorentz units**, where $\varepsilon_0 = 1$. In these units **Maxwell's equations** are

$$\nabla \cdot \boldsymbol{B} = 0, \tag{ME1}$$

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}, \tag{ME2}$$

$$\nabla \times \boldsymbol{B} = \boldsymbol{j} + \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t}, \tag{ME3}$$

$$\nabla \cdot \boldsymbol{E} = \rho. \tag{ME4}$$

Here $\boldsymbol{E}$ is the electric field, $\boldsymbol{B}$ is the magnetic field, $\boldsymbol{j}$ is the current density, and ρ is the charge density.

Any vector field with vanishing divergence, including $\boldsymbol{B}$, can be written as the curl of another field, say $\boldsymbol{A}$, then $\nabla \cdot \boldsymbol{B} = 0$ follows from the identity that $\nabla \cdot (\nabla \times \boldsymbol{A}) = 0$ for any vector field $\boldsymbol{A}$. Hence, we can rewrite the first equation as

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}. \tag{ME1'}$$

Similarly we can write $\boldsymbol{E}$ as

$$\boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} - \nabla \Phi \tag{ME2'}$$

since $\nabla \times (\nabla \Phi) = \boldsymbol{0}$ for all scalar fields Φ and the curl of the first term just gets us back to $\boldsymbol{B}$, so this equation can replace the second of Maxwell's equations.

Combing the third and fourth equation it is possible to derive the continuity equation

$$\nabla \cdot \boldsymbol{j} + \frac{1}{c}\frac{\partial \rho}{\partial t} = 0. \tag{A.2.1}$$

Defining the electromagnetic **four-current**, $j_{\text{em}}^{\mu} := (\rho, \boldsymbol{j})$ then we can write this as

$$\partial_{\mu} j_{\text{em}}^{\mu} = 0. \tag{A.2.2}$$

This is a statement of charge conservation.

A.2.2 Four-Vector Potential

We can define the **four-potential** as

$$A^{\mu} := (\Phi, \boldsymbol{A}). \tag{A.2.3}$$

Recall the identity

$$\nabla \times (\nabla \times \boldsymbol{A}) = \nabla(\nabla \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A}. \tag{A.2.4}$$

Substitute the modified Maxwell equations into the third Maxwell equation, giving

$$\nabla \times (\nabla \times \boldsymbol{A}) = \boldsymbol{j} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} - \frac{1}{c}\nabla\frac{\partial \Phi}{\partial t}. \tag{A.2.5}$$

On the left we can use the identity to get

$$\nabla(\nabla \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A} = \boldsymbol{j} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} - \frac{1}{c}\nabla\frac{\partial \Phi}{\partial t}. \tag{A.2.6}$$

Rearranging this gives

$$\frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} - \nabla^2 \boldsymbol{A} + \nabla\left[\nabla \cdot \boldsymbol{A} + \frac{1}{c}\frac{\partial \Phi}{\partial t}\right] = \partial^2 \boldsymbol{A} + \nabla\left[\nabla \cdot \boldsymbol{A} + \frac{1}{c}\frac{\partial \Phi}{\partial t}\right] = \boldsymbol{j}. \tag{A.2.7}$$

Doing something similar with the fourth Maxwell equation we get

$$-\frac{1}{c}\frac{\partial}{\partial t}(\nabla \cdot \boldsymbol{A}) - \nabla^2 \Phi = \rho. \tag{A.2.8}$$

We can combine these two equations into a covariant form:

$$\partial^2 A^{\mu} - \partial^{\mu}\partial_{\nu}A^{\nu} = j_{\text{em}}^{\mu}. \tag{A.2.9}$$

A.2.3   Gauge Transformations

We can define two new potentials, $\tilde{\boldsymbol{A}}$ and $\tilde{\Phi}$, via the **gauge transformation**

$$\tilde{\boldsymbol{A}} = \boldsymbol{A} - \nabla\chi \quad \text{and} \quad \tilde{\Phi} = \Phi + \frac{1}{c}\frac{\partial\chi}{\partial t}, \tag{A.2.10}$$

or in covariant form,

$$\tilde{A}^\mu = A^\mu + \partial^\mu\chi. \tag{A.2.11}$$

Here χ is any scalar field (with second derivatives). Importantly the fields $\boldsymbol{E}$ and $\boldsymbol{B}$ don't change if we make this transformation. We can use this to choose potentials satisfying the **Lorenz gauge condition**

$$\nabla\cdot\boldsymbol{A} + \frac{1}{c}\frac{\partial\Phi}{\partial t} = 0, \tag{A.2.12}$$

or in covariant form,

$$\partial_\nu A^\nu = 0. \tag{A.2.13}$$

In this case Maxwell's equations reduce to two decoupled equations,

$$\partial^2\boldsymbol{A} = \boldsymbol{j}, \quad \text{and} \quad \partial^2\Phi = \rho, \tag{A.2.14}$$

or in covariant form,

$$\partial^2 A^\mu = j^\mu_{\text{em}}. \tag{A.2.15}$$

Note that the Lorenz gauge doesn't fix the potentials uniquely, since

$$\nabla\cdot\tilde{\boldsymbol{A}} + \frac{1}{c}\frac{\partial\tilde{\Phi}}{\partial t} = \nabla\cdot\boldsymbol{A} + \frac{1}{c}\frac{\partial\Phi}{\partial t} - \nabla^2\chi + \frac{1}{c^2}\frac{\partial^2\chi}{\partial t^2}. \tag{A.2.16}$$

So if the original potential, A^μ, satisfies $\partial_\mu A^\mu = 0$ then we can choose to have $\partial_\mu\tilde{A}^\mu = 0$ so long as χ is such that $\partial^2\chi = 0$.

A.2.4   Covariant Form of Maxwell's Equations

The **Maxwell tensor**, or **electromagnetic field strength tensor** is defined to be

$$F^{\mu\nu} := \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{A.2.17}$$

> (!) Note that there is a different sign convention used, for example, by Mandl and Shaw, where
>
> $$F^{\mu\nu} \stackrel{!}{=} \partial^\nu A^\mu - \partial^\mu A^\nu. \tag{A.2.18}$$

We can think of the electromagnetic field strength tensor as the four-dimensional "curl" of the four-potential[2]

The electromagnetic field strength tensor is manifestly gauge invariant, since any gauge transformation cancels between the two terms.

[2]this becomes precise in the language of geometric algebra, where we replace curl with the wedge product, and the wedge product of two vectors gives a bivector, which four our purposes is a two index tensor.

Considering the space-time and space-space components of F we can fairly simply show that they correspond to the normal electric and magnetic fields:

$$F^{i0} = \partial^i A^0 - \partial^0 A^i = \frac{\partial \Phi}{\partial x^i} - \frac{1}{c}\frac{\partial A^i}{\partial t} = E^i, \tag{A.2.19}$$

$$F^{ij} = \partial^i A^j - \partial^j A^i = -\frac{\partial A^i}{\partial x^i} + \frac{\partial A^i}{\partial x^j} = -\varepsilon^{ijk} B^k. \tag{A.2.20}$$

Here ε^{ijk} is the Levi-Civita symbol, which is defined to be 1 if ijk is an even permutation of 123, -1 if it's an odd permutation, and 0 otherwise.

We can write F as a 4×4 matrix:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_z \\ E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{A.2.21}$$

We can then write Equation (A.2.9) as

$$\partial_\mu F^{\mu\nu} = j^\nu_{\text{em}}. \tag{A.2.22}$$

Conservation of the electromagnetic four-current is then very obvious, since

$$\partial_\nu j^\nu_{\text{em}} = \partial_\nu \partial_\mu F^{\mu\nu} = 0, \tag{A.2.23}$$

since $F^{\mu\nu}$ is antisymmetric in μ and ν and $\partial_\nu \partial_\mu$ is symmetric in μ and ν.

A.2.5 Interaction of Matter with EM Fields

The classical Hamiltonian for a free, non-relativistic particle is

$$H = \frac{|\boldsymbol{p}|^2}{2m}. \tag{A.2.24}$$

For a particle with charge q in an electromagnetic field we modify the Hamiltonian through the **minimal coupling prescription**,

$$H \to H - q\Phi, \quad \text{and} \quad \boldsymbol{p} \to \boldsymbol{p} - \frac{q}{c}\boldsymbol{A}, \tag{A.2.25}$$

giving the classical Hamiltonian for a non-relativistic particle in an electromagnetic field:

$$H = \frac{|\boldsymbol{p} - q\boldsymbol{A}/c|^2}{2m} + q\Phi. \tag{A.2.26}$$

Note that this equation is in Heaviside–Lorentz units, in SI units there is no factor of c.

This Hamiltonian is chosen such that the force on a particle due to the electromagnetic field is

$$\boldsymbol{F} = q\left(\boldsymbol{E} + \frac{1}{c}\boldsymbol{v} \times \boldsymbol{B}\right), \tag{A.2.27}$$

which is the Lorentz force in Heaviside–Lorentz units, again the factor of c is absent in SI units.

A.3   Relativistic Quantum Mechanics

A.3.1   Klein–Gordon Equation

The **Schrödinger equation** for a free particle of mass m is

$$-\frac{\hbar^2}{2m}\nabla^2\Psi(\boldsymbol{r},t) = i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t). \tag{A.3.1}$$

One way to obtain this is from the non-relativistic energy-momentum relation,

$$E = \frac{|\boldsymbol{p}|^2}{2m} = H, \tag{A.3.2}$$

by substituting in the relevant operators:

$$E \to i\hbar\frac{\partial}{\partial t}, \quad \text{and} \quad \boldsymbol{p} \to -i\hbar\nabla. \tag{A.3.3}$$

To obtain a relativistic equation then we should use the relativistic energy-momentum relation,

$$E^2 = |\boldsymbol{p}|^2 c^2 + m^2 c^4. \tag{A.3.4}$$

Then after the operator substitution we have

$$-\hbar^2\frac{\partial^2}{\partial t^2}\varphi(\boldsymbol{r},t) = -\hbar^2 c^2\nabla^2\varphi(\boldsymbol{r},t) + m^2 c^2\varphi(\boldsymbol{r},t). \tag{A.3.5}$$

This is the **Klein–Gordon equation**.

We can write the Klein–Gordon equation in a manifestly covariant form as

$$\left(\partial^2 + \frac{m^2 c^2}{\hbar^2}\right)\varphi(x) = 0 \qquad \text{or} \qquad (\partial^2 + \mu^2)\varphi(x) = 0 \tag{A.3.6}$$

where $\mu = mc/\hbar$. We can jump straight to this covariant form by instead using the operator prescription

$$p^\mu \to i\hbar\frac{\partial}{\partial x^\mu} = i\hbar\left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right). \tag{A.3.7}$$

Note that if $m = 0$ then the Klein–Gordon equation reduces to a wave equation, $\partial^2\varphi(x) = 0$.

The Klein–Gordon equation is valid for any spin 0 particle.

A.3.1.1   Free Particle Solutions

The Klein–Gordon equation has plane-wave solutions,

$$\varphi(\boldsymbol{r},t) = \exp[i\boldsymbol{k}\cdot\boldsymbol{r} - i\omega t]. \tag{A.3.8}$$

Substituting this into the Klein–Gordon equation we see it is a solution as long as ω, $\boldsymbol{k}$, and m are related by

$$\hbar^2\omega^2 = \hbar^2 c^2|\boldsymbol{k}|^2 + m^2 c^4. \tag{A.3.9}$$

Taking the square root we have

$$\hbar\omega = \pm\sqrt{\hbar^2 c^2 |\boldsymbol{k}|^2 + m^2 c^4}. \tag{A.3.10}$$

These solutions are eigenfunctions of the momentum operator and the energy operator, with eigenvalues $\boldsymbol{p} = \hbar\boldsymbol{k}$ and $E = \hbar\omega$ respectively. Interpreting $\hbar\omega$ as the allowed energy of the free particle there is ambiguity in the sign of the total energy. We have both positive and negative energy solutions. Initially this was viewed as a problem, but later people realised that both sorts of solutions were important, with negative energy solutions corresponding to antiparticles.

Define the four-vector $k^\mu = (\omega/c, \boldsymbol{k})$. Then the solutions to the Klein–Gordon equation can be written in a covariant form as

$$\varphi(x) = \exp[-ik \cdot x] = \exp[-ik^\mu x_\mu] = \exp[-ip^\mu x_\mu/\hbar]. \tag{A.3.11}$$

This allows us to interpret the four-momentum as $p^\mu = \hbar k^\mu$.

If φ is a solution of the Klein–Gordon equation then so is its complex conjugate, φ^*. The general solution is a superposition of plane-wave solutions:

$$\varphi(x) = \int \frac{\mathrm{d}^3 x}{2E_p (2\pi)^3} [f(p) \exp[-ip \cdot x/\hbar] + g(p) \exp[ip \cdot x/\hbar]]. \tag{A.3.12}$$

Here the factors in the denominator of the measure are included for future convenience, and

$$E_p := +\sqrt{|\boldsymbol{p}|^2 c^2 + m^2 c^4} \tag{A.3.13}$$

is defined to be positive.

A.3.2 Continuity Equation and Probability Interpretation

Denote the Schrödinger equation by (SE) and its conjugate by (SE)*, that is

$$(\text{SE}) = -\frac{\hbar^2}{2m} \nabla^2 \Psi(\boldsymbol{r}, t) - i\hbar \frac{\partial}{\partial t} \Psi(\boldsymbol{r}, t) = 0. \tag{A.3.14}$$

Then by considering

$$\Psi^*(\text{SE}) - \Psi(\text{SE}) = 0 \tag{A.3.15}$$

we can derive a continuity equation

$$\nabla \cdot \boldsymbol{j} + \frac{\partial \rho}{\partial t} = 0, \tag{A.3.16}$$

where

$$\rho := \Psi^* \Psi, \quad \text{and} \quad \boldsymbol{j} := -\frac{i\hbar}{2m} (\Psi^* \nabla \Psi - \Psi \nabla \Psi^*). \tag{A.3.17}$$

These are the **probability density** and **probability current** respectively.

We can do the same for the Klein–Gordon equation and obtain the same continuity equation, but now with

$$\rho := \frac{i\hbar}{2mc^2} \left(\varphi^* \frac{\partial \varphi}{\partial t} - \varphi \frac{\partial \varphi^*}{\partial t} \right), \quad \text{and} \quad \boldsymbol{j} := -\frac{i\hbar}{2m} (\varphi^* \nabla \varphi - \varphi \nabla \varphi^*). \tag{A.3.18}$$

Here we chose to make $\boldsymbol{j}$ identical to the non-relativistic Schrödinger current. We can show that in the non-relativistic limit ρ reduces to $\varphi^* \varphi$. There is a problem however, the candidate for the probability density, $\rho(x)$, is no longer positive definite, the negative energy solutions will have $\rho(x) < 0$. This means that there is no simple probability density interpretation for ρ from the Klein–Gordon equation.

A.3.3   Dirac Equation

Dirac, aiming to avoid the problems with negative energy/probability solutions, looked for a relativistic equation linear in $\partial/\partial t$. He then proceeded to argue that the equation must be linear in spatial derivatives, since in relativity we have to treat space and time on an equal footing. We therefore start with an equation of the form

$$i\hbar\frac{\partial}{\partial t}\psi(\boldsymbol{r},t) = \hat{H}\psi(\boldsymbol{r},t) \tag{A.3.19}$$

where $\hat{H}$ is linear in space and time derivatives. These requirements mean that we can write out a candidate Hamiltonian with some unknown coefficients:

$$i\hbar\frac{\partial}{\partial t}\psi(\boldsymbol{r},t) = -i\hbar c\left(\alpha^1\frac{\partial}{\partial x^1} + \alpha^2\frac{\partial}{\partial x^2} + \alpha^3\frac{\partial}{\partial x^3}\right)\psi(\boldsymbol{r},t) + \beta mc^2\psi(\boldsymbol{r},t). \tag{A.3.20}$$

We choose to have α^i and β be dimensionless by including the required factors of c, $\hbar$, and m, we also include a conventional factor of $-i$. Using the operator prescription $\hat{\boldsymbol{p}} = -i\hbar\nabla$ we can write this more compactly as

$$i\hbar\frac{\partial}{\partial t}\psi(\boldsymbol{r},t) = (c\boldsymbol{\alpha}\cdot\hat{\boldsymbol{p}} + \beta mc^2)\psi(\boldsymbol{r},t). \tag{A.3.21}$$

This is the **Dirac equation** for a free particle. We assume that the Hamiltonian is independent of position and time, which is what we would want for free particle solutions. Hence α^i and β must be independent of position and time, so commute with the derivatives, such as $\hat{\boldsymbol{p}}$, but not necessarily each other.

Next, we impose the relativistic energy-momentum equation:

$$\hat{H}^2\psi(\boldsymbol{r},t) = (c^2|\hat{\boldsymbol{p}}|^2 + m^2c^4)\psi(\boldsymbol{r},t). \tag{A.3.22}$$

We know that

$$\hat{H}^2\psi(\boldsymbol{r},t) = (c\boldsymbol{\alpha}\cdot\hat{\boldsymbol{p}} + \beta mc^2)(c\boldsymbol{a}\cdot\hat{\boldsymbol{p}} + \beta mc^2)\psi(\boldsymbol{r},t). \tag{A.3.23}$$

Expanding the right hand side, while being careful about the ordering of terms, we get

$$\hat{H}^2\psi(\boldsymbol{r},t) = \left\{c^2[(\alpha^1)^2(\hat{p}^1)^2 + (\alpha^2)^2(\hat{p}^2)^2 + (\alpha^3)^2(\hat{p}^3)^2] + m^2c^4\beta^2\right\}\psi(\boldsymbol{r},t) \tag{A.3.24}$$

$$+ \left\{(\alpha^1\alpha^2 + \alpha^2\alpha^1)\hat{p}^1\hat{p}^2 + (\alpha^2\alpha^3 + \alpha^3\alpha^2)\hat{p}^2\hat{p}^3\right\}\psi(\boldsymbol{r},t) \tag{A.3.25}$$

$$+ mc^3\left\{(\alpha^1\beta + \beta\alpha^1)\hat{p}^1 + (\alpha^2\beta + \beta\alpha^2)\hat{p}^2 + (\alpha^3\beta + \beta\alpha^3)\hat{p}^3\right\}\psi(\boldsymbol{r},t). \tag{A.3.26}$$

The relativistic energy-momentum relation is satisfied if

$$(\alpha^1)^2 = (\alpha^2)^2 = (\alpha^3)^2 = \beta^2 = 1, \tag{A.3.27}$$

$$\alpha^i\alpha^j + \alpha^j\alpha^i = 0, \tag{A.3.28}$$

$$\alpha^i\beta + \beta\alpha^i = 0, \tag{A.3.29}$$

which can be written more compactly sa

$$\{\alpha^i, \alpha^j\} = 2\delta^{ij}, \quad \{\alpha^i, \beta\} = 0, \quad \text{and} \quad \beta^2 = 1, \tag{A.3.30}$$

where $\{A, B\} = AB + BA$ is the **anticommutator**.

Clearly α^i and β cannot just be numbers for these requirements to hold. Instead, we assume they are matrices. Since $\hat{H}$ is Hermitian they must be Hermitian, and hence are square matrices. This means they must have real eigenvalues. Since these matrices square to the identity their squares have eigenvalues whose product is 1, since the determinant is the product of the eigenvalues and the determinant of the unit matrix is 1. Since the eigenvalues are real the eigenvalues of α^i and β must be ± 1.

> **Lemma A.3.31** We have that $\operatorname{tr} \alpha^i = \operatorname{tr} \beta = 0$.
>
> *Proof.* Consider $\operatorname{tr} \alpha^i$. Since $\beta^2 = 1$ we can insert β^2 into the argument of the trace and then use the cyclic property to get
>
> $$\operatorname{tr} \alpha^i = \operatorname{tr}(\beta^2 \alpha^i) = \operatorname{tr}(\beta \alpha^i \beta). \tag{A.3.32}$$
>
> Now using the anticommutation relations we can replace $\alpha^i \beta$ with $-\beta \alpha^i$ and then take the negative outside the trace to get
>
> $$\operatorname{tr} \alpha^i = \operatorname{tr}(\beta \alpha^i \beta) = \operatorname{tr}(\beta(-\beta \alpha^i)) = -\operatorname{tr}(\beta^2 \alpha^i) = -\operatorname{tr}(\alpha^i) \tag{A.3.33}$$
>
> since $\beta^2 = 1$, and so $\operatorname{tr} \alpha^i = -\operatorname{tr} \alpha^i$, which must mean $\operatorname{tr} \alpha^i = 0$. Similarly, consider $\operatorname{tr} \beta$, we have
>
> $$\operatorname{tr} \beta = \operatorname{tr}((\alpha^i)^2 \beta) = \operatorname{tr}(\alpha^i \beta \alpha^i) = -\operatorname{tr}((\alpha^i)^2 \beta) = -\operatorname{tr} \beta, \tag{A.3.34}$$
>
> and so $\operatorname{tr} \beta = 0$. $\qquad \square$

So, if α^i and β are $n \times n$ matrices with eigenvalues ± 1 since the trace is the sum of the eigenvalues and the trace vanishes n must be even so that the eigenvalues can cancel pairwise. There is no 2×2 set of Hermitian traceless 2×2 matrices satisfying the anticommutation relations, although the Pauli matrices come close, satisfying

$$\{\sigma^i, \sigma^j\} = 2\delta^{ij}, \tag{A.3.35}$$

but there is no viable candidate for β.

The simplest representation is 4×4. The **standard representation** is

$$\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \alpha^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \tag{A.3.36}$$

$$\alpha^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \qquad \alpha^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}. \tag{A.3.37}$$

These can be written in a 2×2 block form as

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \text{and} \qquad \alpha = \begin{pmatrix} 0 & \sigma \\ \sigma & 0 \end{pmatrix} \tag{A.3.38}$$

where $\boldsymbol{\sigma} = (\sigma^1, \sigma^2, \sigma^3)$ are the Pauli matrices. We can also write these as $\alpha^i = \sigma^1 \otimes \sigma^i$. It's not too difficult to check that these matrices have the desired properties.

Since the Hamiltonian is a 4×4 matrix the wave function, $\psi(\boldsymbol{r}, t)$, must be a four component column vector:

$$\psi(\boldsymbol{r}, t) = \begin{pmatrix} \psi_1(\boldsymbol{r}, t) \\ \psi_2(\boldsymbol{r}, t) \\ \psi_3(\boldsymbol{r}, t) \\ \psi_4(\boldsymbol{r}, t) \end{pmatrix}. \tag{A.3.39}$$

A.3.4 Probability Density

The Dirac equation for a free particle is

$$i\hbar \frac{\partial}{\partial t} \psi(\boldsymbol{r}, t) = (-i\hbar c \boldsymbol{\alpha} \cdot \vec{\nabla} + \beta m c^2) \psi(\boldsymbol{r}, t). \tag{A.3.40}$$

Here $\vec{\nabla}$ is the normal gradient operator acting to the right. From this we can construct a probability density. Take the Hermitian conjugate of this equation:

$$-i\hbar \frac{\partial}{\partial t} \psi^\dagger(\boldsymbol{r}, t) = \psi^\dagger(\boldsymbol{r}, t) \left(i\hbar c \boldsymbol{\alpha} \cdot \overleftarrow{\nabla} + \beta m c^2 \right). \tag{A.3.41}$$

Here $\overleftarrow{\nabla}$ denotes a gradient operator acting to the left. Note that $\psi^\dagger$ is a row vector. Multiplying the first equation by $\psi^\dagger$ on the left, and the second by ψ on the right we obtain

$$i\hbar \frac{\partial}{\partial t} (\psi^\dagger \psi) = -i\hbar c (\psi^\dagger \boldsymbol{\alpha} \cdot \vec{\nabla} \psi + \psi^\dagger \boldsymbol{\alpha} \cdot \overleftarrow{\nabla} \psi) \tag{A.3.42}$$

$$= -i\hbar c (\psi^\dagger \boldsymbol{\alpha} \cdot \nabla \psi + (\nabla \psi^\dagger) \cdot \boldsymbol{\alpha} \psi) \tag{A.3.43}$$

$$= -i\hbar c \nabla \cdot (\psi^\dagger \boldsymbol{\alpha} \psi). \tag{A.3.44}$$

We can write this as a continuity equation,

$$\nabla \cdot \boldsymbol{j} + \frac{1}{c} \frac{\partial \rho}{\partial t} = 0, \tag{A.3.45}$$

where

$$\rho := \psi^\dagger \psi, \qquad \text{and} \qquad \boldsymbol{j} = \psi^\dagger \boldsymbol{\alpha} \psi. \tag{A.3.46}$$

We can write this in a covariant form:

$$\partial_\mu j^\mu = 0, \tag{A.3.47}$$

where $j = (\rho, \boldsymbol{j})$ is the **four-probability current**.

This tells us that $\psi^\dagger \psi$ transforms as the time component of a four-vector, and $\psi^\dagger \boldsymbol{\alpha} \psi$ transforms as the corresponding space part.

Note that $\rho = |\psi|^2$ is positive definite, solving one of our earlier problems with the Klein–Gordon equation.

A.3.5 Free Particle Solutoins

The Dirac equation has plane-wave solutions:

$$\psi(\mathbf{r}, t) = \psi(x) = \exp[-ik_\mu x^\mu]u(\mathbf{p}) = \exp\left[-\frac{i}{\hbar}(cp_0 t - \mathbf{p} \cdot r)\right]u(\mathbf{p}). \quad \text{(A.3.48)}$$

Here $u(\mathbf{p})$ is a four component column vector.

Substituting this into the Dirac equation we get

$$p^0 u = (\boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc)u. \quad \text{(A.3.49)}$$

We interpret this solution as representing a particle of energy cp^0 and momentum $\mathbf{p}$. Putting in the matrices α^i and β this becomes

$$\begin{pmatrix} -p^0 + mc & 0 & p^3 & p^1 - ip^2 \\ 0 & -p^0 + mc & p^1 + ip^2 & -p^3 \\ p^3 & p^1 - ip^2 & -(p^0 + mc) & 0 \\ p^1 + ip^2 & -p^3 & 0 & -(p^0 + mc) \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = 0. \quad \text{(A.3.50)}$$

The condition for a nontrivial solution for u_i to exist is that the determinant of the matrix vanishes. Computing the determinant we find that the determinant vanishes exactly when

$$[m^2 c^2 + |\mathbf{p}|^2 - (p^0)^2]^2 = 0, \quad \text{(A.3.51)}$$

which is exactly the energy-momentum relation. Taking the square root we have

$$p^0 = \pm\sqrt{m^2 c^2 + |\mathbf{p}|^2}, \quad \text{(A.3.52)}$$

so we still have negative energy solutions.

A.3.5.1 Positive Energy Solutions

First suppose $p^0 > 0$, that is

$$p^0 = +\sqrt{m^2 c^2 + |\mathbf{p}|^2} =: p_+^0 = \frac{E}{c}, \quad \text{(A.3.53)}$$

where we choose to take $E > 0$ as the magnitude of the energy. Substituting into the first two rows of Equation (A.3.50) we get

$$0 = (-p_+^0 + mc)u_1 + p_3 u_3 + (p^1 - ip^2)u_4, \quad \text{(A.3.54)}$$
$$0 = (-p_+^0 + mc)u_2 + (p^1 + ip^2)u_3 - p^3 u_4. \quad \text{(A.3.55)}$$

Only two of the four u_i are linearly independent, so we are free to fix two values. The conventional choice is either $u_1 = 1$ and $u_2 = 0$, or $u_1 = 0$ and $u_2 = 1$. The first choice gives

$$u_3 = \frac{p^3}{p_+^0 + mc}, \quad \text{and} \quad u_4 = \frac{p^1 + ip^2}{p_+^0 + mc}, \quad \text{(A.3.56)}$$

whereas the second gives

$$u_3 = \frac{p^1 - ip^2}{p_+^0 + mc}, \quad \text{and} \quad u_4 = \frac{-p^3}{p_+^0 + mc}. \quad \text{(A.3.57)}$$

A.3.5.2 Negative Energy Solutions

Now suppose $p^0 < 0$, that is

$$p^0 = -\sqrt{m^2c^2 + |\boldsymbol{p}|^2} =: p^0_- = -\frac{E}{c}, \tag{A.3.58}$$

where again, we choose to have $E > 0$. We can follow a similar process to the positive energy solutions, but for the lower two rows of Equation (A.3.50). Again, we make an arbitrary choice for two components, the conventions being either $u_3 = 1$ and $u_4 = 0$, or $u_3 = 0$ and $u_4 = 1$. The first choice leads to

$$u_1 = \frac{p^3}{p^0_- - mc}, \quad \text{and} \quad u_2 = \frac{p^1 + ip^2}{p^0_- - mc}, \tag{A.3.59}$$

and the second gives

$$u_1 = \frac{p^1 - ip^2}{p^0_- - mc}, \quad \text{and} \quad u_2 = \frac{-p^3}{p^0_- - mc}. \tag{A.3.60}$$

A.3.5.3 Summary

The positive energy solutions with four-momentum $p^\mu_+ = (p^0_+, \boldsymbol{p})$ are

$$\omega^1(\boldsymbol{p}) = \begin{pmatrix} 1 \\ 0 \\ \frac{p^3}{p^0_+ + mc} \\ \frac{p^1 + ip^2}{p^0_+ + mc} \end{pmatrix}, \quad \text{and} \quad \omega^2(\boldsymbol{p}) = \omega^1(\boldsymbol{p}) = \begin{pmatrix} 0 \\ 1 \\ \frac{p^1 - ip^2}{p^0_+ + mc} \\ \frac{-p^3}{p^0_+ + mc} \end{pmatrix}. \tag{A.3.61}$$

The negative energy solutions with four-momentum $p^\mu_- = (p^0_-, -\boldsymbol{p})$ are

$$\omega^1(\boldsymbol{p}) = \begin{pmatrix} -\frac{p^1 - ip^2}{p^0_+ + mc} \\ \frac{p^3}{p^0_+ + mc} \\ 0 \\ 1 \end{pmatrix}, \quad \text{and} \quad \omega^2(\boldsymbol{p}) = \omega^1(\boldsymbol{p}) = \begin{pmatrix} -\frac{p^3}{p^0_+ + mc} \\ -\frac{p^1 + ip^2}{p^0_+ + mc} \\ 1 \\ 0 \end{pmatrix}. \tag{A.3.62}$$

Note that there are varying conventions on how the ω^μ are labelled.

A.3.5.4 Rest Frame Solutions

When the particle is at rest, that is $\boldsymbol{p} = \boldsymbol{0}$, we have

$$\omega^1(\boldsymbol{0}) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \omega^2(\boldsymbol{0}) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \omega^3(\boldsymbol{0}) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \text{and} \quad \omega^4(\boldsymbol{0}) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}. \tag{A.3.63}$$

These are such that the solutions

$$\psi^{(1)} = \exp\left[-\frac{imc^2t}{\hbar}\right] \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \psi^{(2)} = \exp\left[-\frac{imc^2t}{\hbar}\right] \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{A.3.64}$$

are degenerate in energy. There should therefore be another operator which commutes with the Hamiltonian and whose eigenvalues label the states. One such operator is

$$\Sigma^3 := \begin{pmatrix} \sigma^3 & 0 \\ 0 & \sigma^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{A.3.65}$$

The rest frame four-component spinors, $\omega^\mu(\mathbf{0})$, are eigenvectors of Σ^3 with eigenvalues ± 1. This suggests that we can interpret the Dirac equation as describing a spin 1/2 particle, and the eigenvalues of Σ^3 are interpreted as the 3rd component of the spin in the rest frame.

The notation Σ^3 suggests we should define

$$\boldsymbol{\Sigma} := \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}, \quad \text{or} \quad \Sigma^i := \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}. \tag{A.3.66}$$

We then have

$$\left(\frac{1}{2}\hbar\boldsymbol{\Sigma}\right) \cdot \left(\frac{1}{2}\hbar\boldsymbol{\Sigma}\right) = \frac{3}{4}\hbar^2\hat{\mathbb{1}} \tag{A.3.67}$$

and $\hbar\Sigma^3/2$ has eigenvalues of $\pm\hbar/2$ suggesting that the spin operator is

$$\hat{\boldsymbol{s}} = \frac{1}{2}\hbar\boldsymbol{\Sigma}. \tag{A.3.68}$$

Note that $\boldsymbol{\Sigma}$ does not commute with the Hamiltonian in frames other than the rest frame, when $\boldsymbol{p} = \mathbf{0}$. The operator $\hat{\boldsymbol{L}} = \hat{\boldsymbol{r}} \times \hat{\boldsymbol{p}}$ also doesn't commute with the Hamiltonian. However, the operator

$$\hat{\boldsymbol{J}} = \hat{\boldsymbol{L}} + \hat{\boldsymbol{s}} = \hat{\boldsymbol{r}} \times \hat{\boldsymbol{p}} + \frac{1}{2}\hbar\boldsymbol{\Sigma} \tag{A.3.69}$$

does commute with the Hamiltonian. This suggests we can interpret $\hat{\boldsymbol{J}}$ as the total angular momentum.

When $\boldsymbol{p} \neq \mathbf{0}$ there are two independent states for any fixed four-momentum. Therefore there exists some operator which commutes with $c\boldsymbol{\alpha} \cdot \hat{\boldsymbol{p}} + \beta mc^2$ and labels the states. One option is the **helicity operator**

$$\hat{h}(\boldsymbol{p}) = \begin{pmatrix} \frac{\sigma \cdot \boldsymbol{p}}{|\boldsymbol{p}|} & 0 \\ 0 & \frac{\sigma \cdot \boldsymbol{p}}{|\boldsymbol{p}|} \end{pmatrix}. \tag{A.3.70}$$

This commutes with $c\boldsymbol{\alpha} \cdot \hat{\boldsymbol{p}} + \beta mc^2$ and has eigenvalues ± 1. We can therefore chose general plane-wave states with $\boldsymbol{p} \neq \mathbf{0}$ to be helicity states.

Acronyms

Index