Relativistic Quantum Fields 1

Mark Hindmarsh
University of Sussex
m.b.hindmarsh@sussex.ac.uk

Autumn Term 1997

Contents

0 Course information

This course forms an introduction to the theory of quantum fields, at a level appropriate for 4th Year MPhys students of physics. With the addition of parallel reading, it is also the first part of the MSc course in quantum fields, which continues next term.

Description

The core MPhys course will cover:

- a recap of Special Relativity and Quantum Mechanics; Introduction to Natural Units; Introduction to Lagrangian and Hamiltonian mechanics;
- the Klein-Gordon and Maxwell equations as relativistic wave equations; their derivation in the Lagrangian approach to field theory;
- Quantisation of Free Fields.
- the Casimir Effect.

MSc and DPhil students will also read in parallel, on:

- the Lorentz Group, its Algebra and representations; spin and spinors; Dirac matrix algebra;
- the Dirac equation; negative energy states;
- quantisation of the free spinor field.

Books

There are many books on field theory at all levels these days, which is good because you should **READ AROUND**. For MSc and DPhil students I recommend

> *An introduction to quantum field theory*, Michael E. Peskin, Daniel V. Schroeder (Addison-Wesley, Reading, Mass; 1995).

There are quite a few errors that will be corrected in the next printing of the book. Meanwhile, a list can be found at
`http://www.slac.stanford.edu/~mpeskin/QFT.html`

Unfortunately, there are not many books at a level suitable for the MPhys course. The two best are

> *Field Quantization*, W. Greiner and J. Reinhardt (Springer, Berlin, 1996),

> *Quantum Field Theory*, F. Mandl and G. Shaw (Wiley, Chichester, 1984).

Of the two, Greiner and Reinhardt is closer to material of this course. However, neither book uses the same conventions as I do, which is a potential source of confusion. Other books which are somewhat complementary and worth having a look at are

> *Quantum Field Theory*, L. Ryder (C.U.P., Cambridge, 1984).

> *Quantum field theory: a modern introduction*, Michio Kaku (O.U.P., Oxford, 1993)

> *Quantum Field Theory*, C. Itzykson and J.-B. Zuber (McGraw-Hill, New York, 1980).

The last is a real encyclopaedia of field theory, but regrettably not very pedagogical.

Prerequisites

These are the things I'm expecting you to have covered:

- Quantum Mechanics (Schrödinger equations, free particle solution, harmonic oscillator).
- Special Relativity (Lorentz transformations, space-time interval, Lorentz invariance).
- Electromagnetism (Maxwell's equations in free space).

Background reading

It will help to have come across the following subjects, as not much time will be spent on them in the course.

- Lagrangian formulation of classical mechanics (e.g. *Classical mechanics*, T.W. Kibble 3rd ed, London: Longman, 1985).
- 4-vector notation in Special Relativity (e.g. *Special relativity*, J. G. Taylor, Oxford: Clarendon Press, 1975).
- Covariant formulation of Maxwell's equations (e.g. Taylor again).
- Natural Units (Mandl and Shaw has a section on natural units).

Assessment

MPhys There will be four problem sheets given out during the term, and the marks will count 40% towards the total for the course. The rest from an exam at the end of the year, which will be a choice of two questions from three.

MSc MSc students will also do the problem sheets, which again count for 40% of the total for the whole RQF course (Parts 1 and 2). There will be an extra set of problems to work through covering the parallel reading component of the course, counting for another 10% of the marks. The remaining 50% will come from the problem sheets set for Part 2 next term.

Other information

I will keep a set of notes available on the web, at

`http://www.pcss.maps.susx.ac.uk/users/markh/RQF1`.

I am updating it from last year's course as I go along, so it is not yet complete. A PostScript version is also available on the web.

Last year's course, which is full of typos, and covers slightly different material, can be found at

`http://www.pcss.maps.susx.ac.uk/users/markh/RQF1_96`.

1 Preliminaries

Why do we study quantum field theory? The reason is that in most areas of current active physics research it is necessary to have some command of at least the basics of the subject. These include:

- the theory of elementary particles and their interactions;
- laser physics and quantum optics;
- condensed matter physics.

What is quantum field theory? It is of course the quantum theory of fields, but its most important features are that it the most convenient expression of the quantum mechanics of many particles in interaction, and it is the only way to make quantum mechanics consistent with the Principle of Special Relativity.

First, a warning should be issued. Quantum field theory is hard. One needs many different concepts in physics and mathematics to be able to understand and use it. However, once these have been taken on board, it gets easier, and one of the main types of calculation that field theory gets used for, the computation of quantum mechanical scattering amplitudes, can be expressed in a beautifully intuitive way through the use of Feynman diagrams. For example, the main contribution to the amplitude for an electron and a positron to collide, annihilate and emerge as a muon and an antimuon, can be graphically denoted by the diagram in Fig. 1.1. The wavy line represents a virtual photon, a photon whose energy is not related to its momentum in the proper way $E = pc$. The idea of virtual particles is one of the cornerstones of field theory, and one which does not appear naturally in ordinary relativistic quantum mechanics.

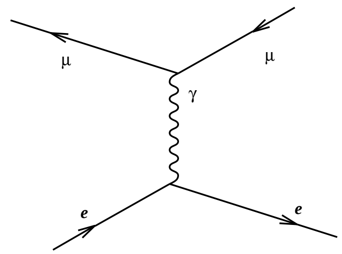

Figure 1.1: The Feynman diagram representing the largest contribution to the scattering amplitude for the process $e^+e^- \to \mu^+\mu^-$. The process proceeds via a virtual photon.

One of the most surprising things about the existence of virtual particles in quantum field theory is that empty space in not quite empty: it can be thought of as filled with pairs of virtual particles which emerge from nothing and return before they can be observed. To talk of the vacuum in this way is more than just a fanciful picture: one can actually detect the effect of these virtual particles in experiments. We shall be studying the Casimir effect: there is an attractive force between conducting plates in a vacuum, inversely proportional to the cube of the separation. This is due to the plates excluding some of the virtual photons in the vacuum from between them. Although this effect was predicted in 1948, it was not measured until 1996. Quantum field theory, although it is more than fifty years old, is still developing.

1.1 Special Relativity

There are many ways of formulating the Principle of Special Relativity: here is one of them.

> Fundamental physical laws are the same for all observers moving with constant velocity relative to one another.

Special Relativity is a rather peculiar theory as it appears to be a theory about other theories. However, when applied to dynamics it does have testable consequences, which have of course corroborated the theory to great accuracy. It is in fact true in only a restricted domain, where gravitational fields are weak and can be neglected. Although it may not seem like it, the gravitational field at the surface of the earth it sufficiently weak to be neglected for most purposes in relativity.

1.1.1 Lorentz transformations and Space-time interval

Suppose a frame of reference F' is moving at velocity v in the x direction relative to another F. Special Relativity tells us that coordinates measured in F' are related to coordinates measured in F by a *Lorentz transformation*

$$\begin{aligned} t \to t' &= \gamma\left(t - \tfrac{v}{c^2}x\right), \\ x \to x' &= \gamma\left(x - vt\right), \\ y \to y' &= y, \\ z \to z' &= z, \end{aligned} \tag{1.1}$$

where

$$\gamma = 1/\sqrt{1 - v^2/c^2}. \tag{1.2}$$

Although neither the distance nor the time interval between two events is observer-independent in special relativity, there is a concept of *space-time interval*, which contains a little of both, and is something that all observers can agree upon.

In its infinitesimal form, the interval ds between two events at (t, x, y, z) and $(t + dt, x + dx, y + dy, z + dz)$ is given by

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2. \tag{1.3}$$

One can easily check that the value of ds doesn not change under this transformation. We say that ds is *Lorentz invariant.*

1.1.2 Relativistic notation

If a theory – which is usually formulated in terms of a set of equations – is consistent with the Principle of Special Relativity we often say that it is Lorentz *co*variant. There is a notation which, if followed correctly, automatically ensures that equations are Lorentz covariant. Quantities are assembled into *4-vectors* (and later on we will encounter *4-tensors*) which transform in a simple linear way when one compares their values between observers in different states of uniform motion. Just as one can form a 3-component vector x^i from three spatial coordinates

$$x^i = (x^1, x^2, x^3) \equiv (x, y, z), \tag{1.4}$$

one can define the 4-vector space-time coordinate for an event by

$$x^\mu = (x^0, x^1, x^2, x^3) \equiv (ct, x, y, z). \tag{1.5}$$

We shall adopt the following conventions when labelling vectors:

- Greek letters from mid-alphabet for space-time indices $(\mu, \nu, \rho, \sigma, \ldots)$;
- Roman letters from mid-alphabet for spatial indices $(i, j, k, \ldots)$;
- bold face will also be used for 3-vectors, e.g. $\mathbf{x}$.

1.1.3 Space-time metric

It is convenient in special relativity (and fundamental in general relativity!) to define a second rank 4-tensor $g_{\mu\nu}$, which is used in the expression for the space-time interval:

$$ds^2 = \sum_{\mu,\nu=0}^{3} g_{\mu\nu} dx^\mu dx^\nu. \tag{1.6}$$

We call $g_{\mu\nu}$ the *metric*. In Special Relativity, the metric always takes a particular constant value,

$$g_{\mu\nu} = \eta_{\mu\nu} \equiv \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & -1 & \cdot & \cdot \\ \cdot & \cdot & -1 & \cdot \\ \cdot & \cdot & \cdot & -1 \end{pmatrix}. \tag{1.7}$$

We often call $\eta_{\mu\nu}$ the *Minkowski metric.*

In order to save writing a large number of summation signs, we often use the convention that repeated indices in an expression are summed over automatically without the need for a summation sign, so that (1.6) can be written as

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu. \tag{1.8}$$

Note that the position of the indices is important. For example, we may use the metric tensor to lower the index on dx^μ:

$$dx_\mu = \eta_{\mu\nu} dx^\nu = (cdt, -dx, -dy, -dz). \tag{1.9}$$

Thus the spatial components of a 4-vector with its index lowered (or a *covariant* vector) have the opposite sign to its counterpart with a raised index (a *contravariant* vector). To emphasise this trap once more:

$$x_i = -x^i. \tag{1.10}$$

One can also define the metric tensor with raised indices as the matrix inverse of the covariant metric tensor

$$\eta_{\mu\nu}\eta^{\mu\rho} = \delta_\mu^{\ \rho}, \tag{1.11}$$

where δ_μ^ρ is the *Kronecker delta*, defined by

$$\delta_\mu^{\ \nu} = \begin{cases} 1, & \mu = \nu, \\ 0, & \mu \neq \nu. \end{cases} \tag{1.12}$$

1.1.4 Matrix representation of Lorentz transformations

Recall the form of a Lorentz transformation between two frames of reference moving with velocity v in the x direction relative to one another (1.1). This linear transformation can be represented by matrices Λ^μ_ν acting linearly on the coordinates:

$$x^\mu \to x'^\mu = \Lambda^\mu_{\ \nu} x^\nu, \tag{1.13}$$

or in differential form

$$dx^\mu \to dx'^\mu = \Lambda^\mu_{\ \nu} dx^\nu, \tag{1.14}$$

(Remember the implied summation over the index ν). The matrix Λ has entries

$$\Lambda^\mu_{\ \nu} = \begin{pmatrix} \gamma & -\gamma v/c & \cdot & \cdot \\ -\gamma v/c & \gamma & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix}, \tag{1.15}$$

where μ labels the rows and ν the columns. This and indeed any Lorentz transformation leaves the interval ds invariant, so let us see what we can infer about the matrices Λ from this condition. The interval transforms as

$$ds^2 \to ds'^2 = \eta_{\mu\nu} dx'^\mu dx'^\nu = \eta_{\mu\nu}\Lambda^\mu_{\ \rho}\Lambda^\nu_{\ \sigma} dx^\rho dx^\sigma = ds^2. \tag{1.16}$$

Thus

$$\eta_{\mu\nu}\Lambda^\mu_{\ \rho}\Lambda^\nu_{\ \sigma} = \eta_{\rho\sigma}. \tag{1.17}$$

In fact, any matrix Λ which leaves the metric $\eta_{\mu\nu}$ invariant under the transformation (1.17) represents a Lorentz transformation. These matrices form a group of transformations known as the *Lorentz group.* A more detailed exploration of the properties of the Lorentz group, which is important for understanding the concepts of mass and spin in quantum field theory, can be found in Peskin and Schroeder, Chapter 3.

1.1.5 4-vectors

A 4-vector can be defined as any object which transforms the same way as dx^μ under a Lorentz transformation. For example, energy E and momentum $\mathbf{p}$ of a particle belong together in a single momentum 4-vector p^μ, defined by

$$p^\mu = (E/c, \mathbf{p}). \tag{1.18}$$

Using the metric, we can form a Lorentz invariant quantity with this 4-vector also:

$$p \cdot p \equiv p^2 = p^\mu p^\nu \eta_{\mu\nu} = E^2/c^2 - \mathbf{p}^2. \tag{1.19}$$

We call this the *scalar product.* Recalling the relativistic relation between energy, momentum and mass, which is $E^2 = \mathbf{p}^2c^2 + m^2c^4$, where m is the rest mass of the particle, we find that

$$p^2 = m^2c^2. \tag{1.20}$$

In general, the scalar product $a \cdot b$ is invariant for any 4-vectors a, b, e.g.

$$p \cdot x = Et - \mathbf{p}\cdot\mathbf{x}. \tag{1.21}$$

Another important object is the velocity 4-vector. In Newtonian mechanics, the velocity is dx^i/dt, so it might be thought that $d^x\mu/dt$ is the analogous quantity. However, dt is not Lorentz invariant, so the object as a whole does not transform like dx^μ as a 4-vector should. Instead, we should differentiate with respect to a Lorentz invariant quantity. There is one ready-made for us in the form of the space-time interval ds, from which we can define the *proper time* $d\tau = ds/c$. The 4-velocity of a particle can then be written as

$$v^\mu = \frac{dx^\mu}{d\tau}. \tag{1.22}$$

1.1.6 Derivatives

The standard partial derivatives with respect to spatial coordinates are collected together into a 3-vector:

$$\partial_i = \frac{\partial}{\partial x^i} = \left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right). \tag{1.23}$$

We often write $\boldsymbol{\nabla}$ for ∂_i. The relativistic generalisation is a covariant 4-vector ∂_μ, which is defined as

$$\partial_\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x^i}\right). \tag{1.24}$$

There is a second order differential operator which students of electromagnetism will already have come across, called the wave operator or the *d'Alembertian*. It is sometimes given its own symbol $\Box$:

$$\Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2. \tag{1.25}$$

This is already Lorentz invariant, for it may also be written as $\partial_\mu \partial_\nu \eta^{\mu\nu}$. You may recall that it was partly consideration of the theory of electromagnetic waves that led Einstein to formulate the special theory of relativity: electromagnetism (in free space) is automatically a theory which is consistent with the special theory of relativity, a fact which becomes obvious when it is written down in terms of 4-vectors and tensors. This is not at all obvious if it is written down as Maxwell originally did, component by component.

1.1.7 Lorentz covariance

The power of 4-vector formalism is that when one uses it consistently to write down equations, the equation will keep the same form under a Lorentz transformation: that is, it will be *Lorentz covariant.* For example, we might want a relativistic generalisation of Newton's Second Law, $\mathbf{F} = \mathbf{dp/dt}$. We have already seen that differentiations with respect to time should be changed to differentiations with respect to proper time, so the analogous equation in relativistic dynamics should be

$$F^\mu = \frac{dp^\mu}{d\tau}. \tag{1.26}$$

The zeroth component of the force 4-vector, F^0, is the (proper) rate of change of the energy of the particle.

A potential source of confusion is whether physical quantities are represented by covariant or contravariant vectors. As a rule, coordinates and momenta naturally have their indices up and derivatives naturally have their indices down.

1.2 Natural units

(A brief introduction to natural units can also be found in Mandl and Shaw p96.)

Natural units are very convenient to use in particle physics, where having to deal with quantities which are tiny when expressed in SI units quickly becomes tiresome. Recall that SI units are based on standard units of mass M, length L, and time T, which are the kilogram (kg), the metre (m) and the second (s).

Natural units, on the other hand, are based on units of mass, velocity, and *action*, which has the same dimensions as energy E multiplied by time. The natural units of velocity and action (which have no names) are chosen so that the speed of light c (which has dimensions of velocity) and the reduced Planck constant $\hbar = h/2\pi$ (which has dimensions of action) both have value unity. The natural unit of mass is eV$/c^2$, where eV is the symbol for the electron-Volt. We can translate back to the SI unit of energy, the Joule (J) by noting that the electron-Volt is defined as the work done moving a particle with charge e through a potential difference of 1 Volt. Using the value of the electronic charge, approximately 1.602×10^{-19} Coulomb, we find

$$1\,\mathrm{eV} = 1.602 \times 10^{-19}\,\mathrm{J}. \tag{1.27}$$

In particle physics, it is much more usual to use GeV, or 10^9 eV as the unit.

In order to recover the natural unit of mass, we use the famous relation between mass and energy in special relativity to note that a mass of 1 eV$/c^2$ has mass-energy 1 eV, or 1.602×10^{-19} J. We nearly always drop the c^2, as it is equal to 1. Hence the mass of the proton in natural units is equal to 0.9383 GeV, and that of the neutron 0.9396 GeV.

Conversion between natural units and SI units can be a little tricky, especially if one is not just dealing with simple units. The fundamental formulas to use are that the standard natural units of velocity (V) and action (A) are related to SI units by

$$V = L/T, \tag{1.28}$$
$$A = ET = MV^2T. \tag{1.29}$$

The inverse relations can be quickly shown to be

$$T = A/MV^2, \tag{1.30}$$
$$L = A/MV. \tag{1.31}$$

For example, consider the formula $E^2 = \mathbf{p}^2c^2 + m^2c^4$, which we have already met when we discussed 4-vectors. In natural units, $c = 1$, so $E^2 = \mathbf{p}^2 + m^2$: hence momentum and energy have the same units as mass (eV). Moreover, $\hbar = 1$, so we can use the quantum mechanical relation $\mathbf{p} = \hbar\mathbf{k}$ to infer that wavenumber also has units of mass. We write this as $[k] = $ GeV, where the square brackets are short for "the dimensions of". It is also clear from (1.31) that in natural units

$$[L] = [T] = \mathrm{GeV}^{-1}. \tag{1.32}$$

The following relations are very useful when converting units:

$$c = 3.00 \times 10^8\,\mathrm{ms}^{-1}, \quad \hbar = 6.58 \times 10^{-25}\,\mathrm{GeV\ s}, \quad \hbar c = 1.97 \times 10^{-16}\,\mathrm{GeV\ m}. \tag{1.33}$$

From this one can infer that

$$1\,\mathrm{m} \simeq 5 \times 10^{15},\ \mathrm{GeV}^{-1} \quad 1\,\mathrm{s} \simeq 10^{24}\,\mathrm{GeV}^{-1}. \tag{1.34}$$

More accurate values of $\hbar$ and c, and other fundamental constants, are in Appendix A.

In electrodynamics, we have also the fundamental constants ϵ_0 and μ_0, the permeability and permittivity of free space, which are related to the speed of light by $c^2 = 1/\mu_0\epsilon_0$. One can extend natural units by setting $\epsilon_0 = 1$, which then of course implies $\mu_0 = 1$ also. Thus Maxwell's equations take a simplified form when expressed in natural units, as we shall see in the next section.

1.3 Lagrangian and Hamiltonian mechanics

The starting point for classical mechanics is Newton's Second Law of Motion, which states that the rate of change of momentum of a particle is proportional to the applied force. For a system of P particles, we may write

$$\dot{\mathbf{p}}_A = \mathbf{F}_A, \tag{1.1}$$

where the index A runs from 1 to P. The momentum $\mathbf{p}_A = m_A\mathbf{v}_A$, where $\mathbf{v}_A = \dot{\mathbf{x}}_A$. If the mass is constant in time, we obtain the formulation $m_A\ddot{\mathbf{x}}_A = \mathbf{F}_A$.

If the force is conservative, which means that the work done by the force around a closed path is zero, then the force can be written as the gradient of a potential energy V:

$$\mathbf{F}_A = -\boldsymbol{\nabla}_A V(\mathbf{x}_1, \ldots, \mathbf{x}_P). \tag{1.2}$$

There may also be *constraints* on the motion: for example, the particles may make up a rigid body, in which case their relative positions are fixed. This reduces the number of *degrees of freedom* of the system. A very common form of constraint is one which relates some or all of the coordinates through a function (which may also depend explicitly on time):

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_P, t) = 0. \tag{1.3}$$

These are known as *holonomic* constraints. If there are k such constraints, then the effective number of degrees of freedom is reduced by k. For example, P particles moving in 3 dimensions under k constraints have $3P - k$ degrees of freedom. Thus not all of the coordinates are needed to describe the motion of the system, and it is convenient to introduce the notion of *generalised coordinates* q_a which result from the solving of the constraint equations. In the system of particles there will be $N = 3P - k$ of them, and one can express the original coordinates as functions

$$\begin{array}{rcl} \mathbf{x}_1 & = & \mathbf{x}_1(q_1, \ldots, q_N, t) \\ \vdots & & \vdots \\ \mathbf{x}_P & = & \mathbf{x}_P(q_1, \ldots, q_N, t) \end{array} \tag{1.4}$$

Generalised coordinates are also useful in systems without constraints: for example, orbits around central potentials, where it is convenient to use spherical polar coordinates.

1.3.1 Lagrangian Mechanics

There is a later and more sophisticated formulation of classical mechanics due to Lagrange. We define the kinetic and potential energies T and U in terms of the generalised coordinates of a system as

$$T = \frac{1}{2}\sum_a m_a \dot{q}_a^2, \qquad U = U(q_a, \dot{q}_a) \tag{1.5}$$

(we have allowed for the possibility of a velocity-dependent potential). We then introduce a new function called the *Lagrangian*, defined as

$$L = T - U. \tag{1.6}$$

The equations of motion of the system are then

$$\frac{\partial L}{\partial q_a} - \frac{\partial}{\partial t}\left(\frac{\partial L}{\partial \dot{q}_a}\right). \tag{1.7}$$

These equations are known as the *Euler-Lagrange* equations.

For example, suppose there is no velocity dependence in the potential. Then

$$\frac{\partial L}{\partial q_a} = -\frac{\partial U}{\partial q_a}, \quad \frac{\partial L}{\partial \dot{q}_a} = m_a\dot{q}_a, \tag{1.8}$$

so that

$$\frac{\partial}{\partial t}(m_a\dot{q}_a) = -\frac{\partial U}{\partial q_a}. \tag{1.9}$$

Thus we return to Newton's Second Law. The quantities $\partial L/\partial\dot{q}_a$ can be identified as a momentum associated with the coordinate q_a – if the q_a are simply Cartesian coordinates, then they are the ordinary mechanical momenta, but in general we refer to them as *conjugate* or *canonical* momenta, which are defined as the derivative of the Lagrangian with respect to the rate of change of the generalised coordinates:

$$p_a = \frac{\partial L}{\partial \dot{q}_a}. \tag{1.10}$$

The Euler-Lagrange equations can be derived from a variational principle, known as Hamilton's Principle or the Principle of Least Action (which is a misnomer). The action for a system between times t_1 and t_2 is defined as

$$I = \int_{t_1}^{t_2} dt\, L(q_a, \dot{q}_a, t). \tag{1.11}$$

The Principle of Least Action states that the action is an extremum for the path of the motion.

To see this, let us consider a small variation in the path, vanishing at t_1 and t_2, sends $q_a(t) \to q'_a(t) = q_a(t) + \delta q_a(t)$ (see Figure 1.2). The action also changes, $I \to I' = I + \delta I$, where to first order in $\delta q_a(t)$,

$$\delta I = \int_{t_1}^{t_2} dt\, \delta L(q_a, \dot{q}_a, t) = \int_{t_1}^{t_2} dt \left(\frac{\partial L}{\partial q_a} \delta q_a + \frac{\partial L}{\partial \dot{q}_a} \delta \dot{q}_a \right). \quad (1.12)$$

Integrating by parts we find

$$\delta I = \int_{t_1}^{t_2} dt \left(\frac{\partial L}{\partial q_a} - \frac{\partial}{\partial t} \left(\frac{\partial L}{\partial \dot{q}_a} \right) \right) \delta q_a + \left[\frac{\partial L}{\partial \dot{q}_a} \delta q_a \right]_{t_1}^{t_2}. \quad (1.13)$$

The second term vanishes, as we are keeping the end points of the path fixed, and so if the variation of the action δI vanishes we find that

$$\frac{\partial L}{\partial q_a} - \frac{\partial}{\partial t} \left(\frac{\partial L}{\partial \dot{q}_a} \right) = 0. \quad (1.14)$$

Thus we return to the Euler-Lagrange equations (1.7).

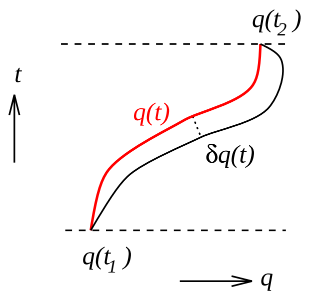

Figure 1.2: A path $q(t)$ for a system with one coordinate, and a small variation $\delta q(t)$ to that path. Classical paths are ones for which the action (see Equation 1.11) is stationary under such small variations.

1.3.2 Hamiltonian Mechanics

There is yet another formulation of classical mechanics due to Hamilton. It may seem excessive to have so many ways of doing the same thing, but each formulation has its own advantages: for example, the Lagrangian and Hamiltonian

formulations are very powerful in bringing out conservation laws in dynamical systems, and relating them to symmetries.

Recall that associated with each coordinate q_a there is also a conjugate momentum p_a, defined by Eq. (1.10). We can therefore express the Euler-Lagrange equations as $\dot{p}_a = \partial L/\partial q_a$. We now define an important function called the *Hamiltonian* from the Lagrangian:

$$H(p_a, q_a, t) = \sum_a p_a \dot{q}_a - L(q_a, \dot{q}_a, t). \quad (1.15)$$

The idea is to replace $\dot{q}_a$ by p_a, having solved equation (1.10). The first of Hamilton's equations follows from the Euler-Lagrange equations (1.7), after noting that $\partial L/\partial q_a = -\partial H/\partial q_a$. It is

$$\dot{p}_a = -\frac{\partial H}{\partial q_a}. \quad (1.16)$$

The second follows straightforwardly from the definition of the Hamiltonian (1.15) upon partial differentiation with respect to p_a:

$$\dot{q}_a = \frac{\partial H}{\partial p_a}. \quad (1.17)$$

If the Lagrangian does not depend explicitly on time, the Hamiltonian is a constant of the motion, as the following calculation shows. Recall that H is a function of q_a, p_a, and possibly t only. Hence

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \sum_a \left(\frac{\partial H}{\partial p_a} p_a + \frac{\partial H}{\partial q_a} q_a \right). \quad (1.18)$$

Using Hamilton's equations (1.16,1.17) we see that

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \quad (1.19)$$

Thus if $\partial L/\partial t = 0$, the Hamiltonian is constant. One can also show that

$$H = T + U, \quad (1.20)$$

where T and U are the kinetic and potential energies introduced in (1.5). Thus the Hamiltonian can be identified with the total energy of the system.

As a trivial example, consider a set of N free particles with mass m. The Lagrangian is

$$L = \frac{1}{2} \sum_a m \dot{q}_a^2, \quad (1.21)$$

from which we can easily derive the Euler-Lagrange equations

$$0 - \frac{\partial}{\partial t}(m \dot{q}_a) = 0, \quad (1.22)$$

or that the acceleration of each particle is zero. The conjugate momentum is $p_a = m\dot{q}_q$, and hence the Hamiltonian becomes

$$H = \frac{1}{2}\sum_a \frac{p_a^2}{m}. \tag{1.23}$$

This is of course the familiar non-relativistic expression for the kinetic energy.

1.4 Quantum mechanics

In accepting quantum mechanics, we move to a higher level of abstraction about how we think about dynamical systems. In classical mechanics, we specify the state of a system by giving its coordinates and their rate of change, q_a and $\dot{q}_a$. Given a Lagrangian function, the subsequent evolution in then in principle completely determined by the Euler-Lagrange equations. In practice, of course, it may not be easy to solve them, and there may be great complexity in the dynamics.

1.4.1 Basic wave mechanics

In the version of quantum mechanics formulated by Schrödinger, we specify the state of a system at time t by the value of a complex *wave function* $\psi(q_a, t)$. We are restricted to making probabilistic statements about the values of dynamical variables: for example, the probability of the system having coordinates in a small interval dq_a around q_a is $dP(q_a) = |\psi(q_a,t)|^2 d^N q_a$.

The time evolution is determined by the Schrödinger equation,

$$i\hbar\frac{\partial}{\partial t}\psi(q_a,t) = \left[\sum_a \frac{\hbar^2}{2m_a}\frac{\partial^2}{\partial q_a^2} + U(q_a)\right]\psi(q_a,t) \tag{1.1}$$

(I shall leave $\hbar$ in for this section.) Physical quantities are associated with operators $\hat{\Omega}$ acting on the wave function. For example,

$$\begin{aligned} \text{Energy:}\quad \hat{E} &= i\hbar\frac{\partial}{\partial t}, \\ \text{Momentum:}\quad \hat{p}_a &= -i\hbar\frac{\partial}{\partial q_a}, \\ \text{Position:}\quad \hat{q}_a &= q_a. \end{aligned}$$

The position and momentum operators satisfy the *canonical commutation relations*

$$[\hat{q}_a, \hat{p}_b] = i\hbar\delta_{ab}, \tag{1.2}$$

where the *commutator bracket* of any two quantities A and B is defined as

$$[A,B] = AB - BA. \tag{1.3}$$

The Schrödinger equation may be written

$$i\hbar\frac{\partial}{\partial t}\psi(q_a,t) = \hat{H}\psi(q_a,t), \tag{1.4}$$

where the Hamiltonian operator is given by

$$\hat{H} = \left[\sum_a \frac{1}{2m_a}\hat{p}_a^2 + U(\hat{q}_a)\right]. \tag{1.5}$$

A quantum system is essentially defined by a Hamiltonian (which determines the time evolution of the wave function) plus the canonical commutation relations.

If $\hat{\Omega}$ is an operator associated with some physical quantity ω, then the average value of ω in an ensemble of systems prepared with a normalised wave function $\psi(q_a,t)$ is

$$\bar{\omega} \equiv \langle\hat{\Omega}\rangle = \int d^N q_a\, \psi^\dagger\hat{\Omega}\psi. \tag{1.6}$$

Usually there is some spread around $\bar{\omega}$, unless the wave function is an *eigenstate* of $\hat{\Omega}$, that is, $\psi = v_\mu(q_a)$, with

$$\hat{\Omega}v_\mu(q_a) = \omega_\mu v_\mu(q_a). \tag{1.7}$$

One can decompose wave functions into linear superpositions of eigenfunctions (eigenstates) of operators, for example the Hamiltonian, whose eigenfunctions satisfy

$$Hu_\alpha(q_a) = E_\alpha u_\alpha(q_a), \tag{1.8}$$

where k labels the eigenfunctions $u_\alpha(q_a)$ and their associated eigenvalues E_α. These labels may be discrete (represented by a set of integers) or continuous (represented by a set of real numbers). Suppose they are discrete, then we may write

$$\psi(q_a,t) = \sum_\alpha a_\alpha u_\alpha(q_a)e^{-iE_\alpha t/\hbar}, \tag{1.9}$$

where

$$a_\alpha = e^{+iE_\alpha t/\hbar}\int d^N q_a u_\alpha^\dagger(q_a)\psi(q_a,t). \tag{1.10}$$

One could just as well express the wave function as a superposition of eigenfunctions of $v_\mu(q_a)$ of any operator Ω. The crucial property of the eigenfunctions which allows this is that they form a *complete set*. They are also orthogonal, in the sense that

$$\int d^N q_a v_{\mu'}^\dagger(q_a)v_\mu(q_a) = 0, \quad \text{if } \mu' \neq \mu. \tag{1.11}$$

It is usual to normalise eigenfunctions with a discrete label to unity:

$$\int d^N q_a |v_\mu(q_a)|^2 = 1. \tag{1.12}$$

We say that they form a complete *orthonormal* set.

Eigenfunctions with a continuous label need special handling and the introduction of the *Dirac δ-function* (see Appendix B).

As an example, we take a single free particle moving in 3-dimensional space with coordinate $\mathbf{x}$. The momentum operator is $\hat{\mathbf{p}} = -i\hbar\boldsymbol{\nabla}$, which has eigenfunctions

$$v_{\mathbf{p}}(\mathbf{x}) = e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}. \tag{1.13}$$

The normalisation condition that these eigenfunctions satisfy is

$$\int d^3x v^\dagger_{\mathbf{p}'}(\mathbf{x})v_{\mathbf{p}}(\mathbf{x}) = (2\pi)^3\delta(\mathbf{p}' - \mathbf{p}). \tag{1.14}$$

1.4.2 Matrix formulation: Dirac bra and ket notation

There is another formulation of quantum mechanics, which goes back to Heisenberg, which represents the states of a system and the operators in terms of vectors and matrices acting on them. Dirac developed an extremely compact notation which is very useful in quantum field theory, so it is worth going over here.

A state function $\psi_\alpha(q_a, t)$ is represented by a *ket vector* $|\alpha\rangle$, while its adjoint $\psi^\dagger_\alpha$ is represented by a *bra vector* $\langle\alpha|$. The *inner product* puts two state vectors together in a *bracket*:

$$\int d^N q_a \psi_\alpha(q^a)^\dagger \psi_\beta(q^a) = \langle\alpha|\beta\rangle, \tag{1.15}$$

which is a complex number. By analogy with matrices, one can think of the coordinates q_a as a kind of row index for a column vector $\psi(q_a)$, and as a column index for a row vector $\psi^\dagger(q_a)$.

Operators act on ket vectors from the left to give another ket vector, and similarly on bra vectors from the right:

$$\Omega|\beta\rangle = |\beta'\rangle,$$
$$\langle\alpha|\Omega = \langle\alpha'|.$$

The matrix elements of an operator Ω are denoted

$$\int d^N q_a \psi^\dagger_\alpha(q_a)\Omega\psi_\beta(q_a) = \langle\alpha|\Omega|\beta\rangle. \tag{1.16}$$

As a trivial example, the matrix elements of the unit operator 1 are just the inner products of the state vectors concerned:

$$\langle\alpha|1|\beta\rangle = \langle\alpha|\beta\rangle. \tag{1.17}$$

We can use the eigenvalues of operators to label states. For example,

Bra vector	eigenstate of	eigenvalue
$\lvert E\rangle$	energy	E
$\lvert\mathbf{p}\rangle$	momentum	$\mathbf{p}$
$\lvert\mathbf{x}\rangle$	position	$\mathbf{x}$
$\lvert\mu\rangle$	Ω	ω_μ

In the last line, Ω represents some general operator. The normalisation conditions on the position and momentum eigenstates are expressed as

$$\langle\mathbf{x}'|\mathbf{x}\rangle = \delta(\mathbf{x}' - \mathbf{x}), \qquad \langle\mathbf{p}'|\mathbf{p}\rangle = (2\pi\hbar)^3\delta(\mathbf{p}' - \mathbf{p}). \tag{1.18}$$

(Note that the difference in normalisation between position and momentum eigenstates is a matter of convention: it is perfectly possible to normalise them the same way.) To convert between the bra and ket notation and the wave function representation, we can use the following relations:

$$\langle\mathbf{x}|E\rangle = u_E(\mathbf{x}), \qquad \langle\mathbf{x}|\mathbf{p}\rangle = v_{\mathbf{p}}(\mathbf{x}). \tag{1.19}$$

The fact that the eigenfunctions of an operator form a complete set can be expressed as the following identity, known as the *completeness relation*:

$$\sum_\mu |\mu\rangle\langle\mu| = 1. \tag{1.20}$$

For operators with continuous eigenvalues, the sum goes over to an integral. For 3-dimensional position and momentum eigenfunctions, for example, we have

$$\int d^3x|\mathbf{x}\rangle\langle\mathbf{x}| = 1, \tag{1.21}$$

$$\int \frac{d^3p}{(2\pi\hbar)^3}|\mathbf{p}\rangle\langle\mathbf{p}| = 1. \tag{1.22}$$

Again, note the difference in the completeness relations, which stem from the choice of normalisation for momentum eigenfunctions.

1.4.3 The harmonic oscillator

The harmonic oscillator is one of the most important quantum systems, as it is one of the very few that can be solved exactly. Its importance extends to field theory, because we will find that a field can be thought of as a collection of an infinite number of oscillators in interaction.

The Hamiltonian for a harmonic oscillator of mass m moving in one dimension with displacement q and restoring force kq, is

$$H = \frac{1}{2}\frac{p^2}{2m} + \frac{1}{2}kq^2. \tag{1.23}$$

If we make the redefinitions

$$p \to p/\sqrt{m}, \quad q \to \sqrt{m}q, \quad \omega^2 = k/m, \tag{1.24}$$

we may write

$$H = \frac{1}{2}(p^2 + \omega^2 q^2). \tag{1.25}$$

In the quantum oscillator, p, q and H become operators, and the system possesses a discrete series of energy eigenfunctions $u_n(q)$, with n a non-negative integer, and energy eigenvalues

$$E_n = (n + \frac{1}{2})\hbar\omega. \tag{1.26}$$

The $u_n(q)$ turn out to be proportioanl to Hermite polynomials multiplied by a Gaussian:

$$u_n(q) \propto H_n(\alpha q)e^{-\frac{1}{2}\alpha^2 q^2}, \tag{1.27}$$

where $\alpha = m\omega/\hbar^2$.

In the bra and ket formalism, the energy eigenstates are denoted $|n\rangle$, and can be normalised to satisfy $\langle m|n\rangle = \delta_{mn}$. We form two important operators known as *ladder operators*, a and $a^\dagger$, by

$$a = (i\hat{p} + \omega\hat{q})/\sqrt{2\omega}, \quad a^\dagger = (-i\hat{p} + \omega\hat{q})/\sqrt{2\omega}. \tag{1.28}$$

One can show (see Problem Sheet 1) that the ladder operators satisfy the commutation relation

$$[a, a^\dagger] = 1. \tag{1.29}$$

One can also show that when the ladder operators act on energy eigenstates they return eigenstates with quantum number n differing by 1:

$$a|n\rangle = \sqrt{n}|n-1\rangle, \qquad a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle. \tag{1.30}$$

This accounts for their collective name: individually, they are termed lowering and raising operators respectively.

2 Relativistic Wave Equations

2.1 The Klein-Gordon equation

The Klein-Gordon equation was originally thought up by Schrödinger, who wanted a relativistic wave equation describing the electron, even before he settled on what we now call the Schrödinger equation,

$$i\hbar\frac{\partial}{\partial t}\psi = -\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2\psi. \tag{2.1}$$

Of course, Klein and Gordon published first, so they got the credit. We will see later why Schrödinger dropped this equation. One of the ways of understanding the Schrödinger equation is to recall that in quantum mechanics, physical quantities are represented by operators:

$$E \leftrightarrow i\hbar\frac{\partial}{\partial t}, \qquad \mathbf{p} \leftrightarrow -i\hbar\boldsymbol{\nabla}, \tag{2.2}$$

(where the double-headed arrow symbol means "is represented by"). Hence the Schrödinger equation (2.1) represents the non-relativistic equation for the kinetic energy $E = \mathbf{p}^2/2m$.

As we know, the relativistic relation between energy and momentum is $E^2 = \mathbf{p}^2 + m^2$, which seems to suggest that a relativistic version of the Schrd̈inger equation ought to be

$$-\frac{\partial^2}{\partial t^2}\phi = -\boldsymbol{\nabla}^2\phi + m^2\phi, \tag{2.3}$$

or, in a manifestly Lorentz covariant form,

$$(\partial^2 + m^2)\phi = 0. \tag{2.4}$$

(It is traditional to use ϕ rather than ψ in this context).

The solutions to this equation are

$$\phi(x) = ae^{-iEt+\mathbf{p}\cdot\mathbf{x}}, \tag{2.5}$$

with $E^2 = \mathbf{p}^2 + m^2$. We can represent the general solution as a Fourier integral:

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E}\left[a(\mathbf{p})e^{-ip\cdot x} + b^*(\mathbf{p})e^{ip\cdot x}\right], \tag{2.6}$$

where $p^0 = E = +(\mathbf{p}^2 + m^2)^{1/2}$. (The factors $(2\pi)^3$ and $(2E)^{-1}$ are included for later convenience.) Here, I have allowed ϕ to be complex: if we take $b(\mathbf{p}) = a(\mathbf{p})$ we can force ϕ to be real – it would however not have the interpretation as a wave function. But in fact ϕ *cannot* be interpreted as a wave function anyway, as Scrödinger discovered. A wave function is a probability amplitude, whose modulus squared is the probability of finding the particle at a particular position (or with a particular momentum). For example, in non-relativistic quantum mechanics, the probability density $\rho = |\psi|^2$, which is associated with a probability *current* $\mathbf{j} = -\frac{i}{2m}(\psi^*\boldsymbol{\nabla}\psi - \boldsymbol{\nabla}\psi^*\psi)$, in the sense that together they make up a probability conservation equation

$$\dot{\rho} + \boldsymbol{\nabla}\cdot\mathbf{j} = 0, \tag{2.7}$$

an equation which one can check by differentiating ρ with respect to time and using the Schrödinger equation.

The KG equation also has a conserved density and a 3-vector current, which are

$$\rho = -\frac{i}{2}(\phi^* \partial_t \phi - \partial_t \phi^* \phi), \qquad \mathbf{j} = -\frac{i}{2}(\phi^* \boldsymbol{\nabla} \phi - \boldsymbol{\nabla} \phi^* \phi). \tag{2.8}$$

However, this density cannot be interpreted as a probability density, as it is not positive definite. This was the reason that Schrödinger chose the non-relativistic form for his equation.

2.2 Maxwell's equations in covariant form

This section recaps some important results, and introduces the formulation of the theory in an explicitly Lorentz covariant manner.

Firstly, we recall Maxwell's equations in free space, writing them down in natural units, in which the permittivity and permeability of free space μ_0 and ϵ_0 are both unity:

$$\begin{array}{cc} \textbf{Homogeneous} & \textbf{Inhomogeneous} \\ \boldsymbol{\nabla}\cdot\mathbf{B} = 0 & \boldsymbol{\nabla}\cdot\mathbf{E} = \rho \\ \frac{\partial}{\partial t}\mathbf{B} + \boldsymbol{\nabla} \wedge \mathbf{E} = 0 & -\frac{\partial}{\partial t}\mathbf{E} + \boldsymbol{\nabla} \wedge \mathbf{B} = \mathbf{j} \end{array} \tag{2.9}$$

The homogeneous and inhomogeneous (i.e. having a source term on the right hand side) equations have differing status. The homogeneous equations imply the existence of potentials ϕ and $\mathbf{A}$ from which the physically measureable quantities $\mathbf{E}$ and $\mathbf{B}$ can be calculated. These potentials are specified only up to a *gauge transformation*, $\phi \to \phi - \dot{\Lambda}$ and $\mathbf{A} \to \mathbf{A} + \boldsymbol{\nabla}\Lambda$, where Λ is an arbitrary function of space and time. The inhomogeneous equations imply that the source terms must obey a current conservation equation, $\dot{\rho} + \boldsymbol{\nabla}\cdot\mathbf{j} = 0$. To summarise:

$$\begin{array}{cc} \textbf{Homogeneous} & \textbf{Inhomogeneous} \\ \text{potentials } \phi, \mathbf{A} & \text{current conservation} \\ \begin{array}{rcl} \mathbf{B} &=& \boldsymbol{\nabla} \wedge \mathbf{A} \\ \mathbf{E} &=& -\dot{\mathbf{A}} - \boldsymbol{\nabla}\phi \\ \phi &\to& \phi - \dot{\Lambda} \\ \mathbf{A} &\to& \mathbf{A} + \boldsymbol{\nabla}\Lambda \end{array} & \dot{\rho} + \boldsymbol{\nabla}\cdot\mathbf{j} = 0 \end{array} \tag{2.10}$$

All these quantities can be assembled into 4-vectors and a 4-tensor. The gauge potentials belong together in a 4-vector potential $A^\mu = (\phi, \mathbf{A})$, while the charge density ρ and the current density $\mathbf{j}$ can be put together into a 4-vector current density $j^\mu = (\rho, \mathbf{j})$. Recall that putting quantities together into 4-vectors is not just a matter of notation: it means that the quantities transform just like the space-time coordinates x^μ under a Lorentz transformation.

The electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$ also belong together in a Lorentz covariant object, as they are mixed up by Lorentz transformations (an observer

moving through a magnetic field also sees an electric field). However, this object cannot be a 4-vector as there are a total of 6 components of the electric and magnetic fields, when they are taken together. In fact, the object is an *antisymmetric tensor*, the field strength tensor $F^{\mu\nu}$, whose entries are as follows:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}. \tag{2.11}$$

Note that E^1 represents the first component of the electric field vector, which in Cartesian coordinates is the x component. The reverse relations may be written

$$E^i = F^{i0}, \qquad B^i = -\frac{1}{2}\epsilon^{ijk} F^{jk}, \tag{2.12}$$

which introduces the *Levi-Civita* tensor ϵ^{ijk}. The Levi-Civita tensor is defined by

$$\epsilon^{ijk} = \begin{cases} +1 & \text{if } i \neq j \neq k \text{ cyclic,} \\ -1 & \text{if } i \neq j \neq k \text{ anticyclic,} \\ 0 & \text{otherwise.} \end{cases} \tag{2.13}$$

The field strength tensor has a neat expression in terms of the gauge potential 4-vector A^μ:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{2.14}$$

Let us check this explicitly for the components corresponding to the electric field:

$$F^{i0} = \partial^i A^0 - \partial^0 A^i = \partial^i A^0 - \dot{A}^i = -\partial_i \phi - \dot{A}^i = E^i, \tag{2.15}$$

where we have used the expression for the electric field in terms of the gauge potentials in equation (2.10).

When expressed in terms of 4-vectors and tensors, electromagnetism looks very simple and beautiful. For example, the covariant expression of the current conservation equation is simply

$$\partial_\mu j^\mu = 0. \tag{2.16}$$

Maxwell's equations become:

$$\begin{array}{cc} \textbf{Homogeneous} & \textbf{Inhomogeneous} \\ \partial^\lambda F^{\mu\nu} + \partial^\mu F^{\nu\lambda} + \partial^\nu F^{\lambda\mu} = 0, & \partial_\mu F^{\mu\nu} = j^\nu. \end{array} \tag{2.17}$$

The first of these expressions looks at first sight as if it contains many more equations than the 4 of the original homogeneous Maxwell equations. However, the antisymmetry of the field strength tensor ($F^{\mu\nu} = -F^{\nu\mu}$) means that the expression is trivial if any of the two indices are equal. Thus all the indices λ,

μ and ν must take different values, and the number of ways of choosing three different numbers from a set of four is itself four, precisely the number of equations we started with.

We can use a four-dimensional version of the Levi-Civita tensor to re-express the homogeneous equations more compactly. This tensor has four indices, and is defined by

$$\epsilon^{\mu\nu\rho\sigma} = \begin{cases} +1 & \text{if } \mu \neq \nu \neq \rho \neq \sigma, \text{ symmetric,} \\ -1 & \text{if } \mu \neq \nu \neq \rho \neq \sigma, \text{ antisymmetric,} \\ 0 & \text{otherwise.} \end{cases} \tag{2.18}$$

A *symmetric* permutation of the indices in one in which an even number of indices are exchanged, while an *antisymmetric* permutation is one for which an odd number of indices are exchanged. Thus, for example, $\epsilon^{0123} = \epsilon^{1032} = +1$, $\epsilon^{1023} = \epsilon^{0132} = -1$, but $\epsilon^{0012} = 0$.

There is also a version with the indices lowered:

$$\epsilon_{\mu\nu\rho\sigma} = \eta_{\mu\alpha}\eta_{\nu\beta}\eta_{\rho\gamma}\eta_{\sigma\delta}\epsilon^{\alpha\beta\gamma\delta}, \tag{2.19}$$

whose symmetric permutations take the value -1 and antisymmetric ones $+1$. Using this totally antisymmetric (i.e. antisymmetric on the exchange of any two indices) tensor, the homogeneous Maxwell equations become

$$\epsilon_{\mu\nu\rho\sigma}\partial^\rho F^{\mu\nu} = 0. \tag{2.20}$$

One of the problems uses this expression to give you practice with the summation convention as applied to these tensors.

2.2.1 Gauges

Recall that the gauge potential A^μ is specified only up to an arbitrary function of space and time Λ (see equation 2.10). The 4-vector version of a gauge transformation is written

$$A^\mu \to A^\mu - \partial^\mu\Lambda, \tag{2.21}$$

under which one can check that $F^{\mu\nu}$ is unchanged. Indeed, using (2.14) one finds

$$F^{\mu\nu} \to \partial^\mu(A^\nu - \partial^\nu\Lambda) - \partial^\nu(A^\mu - \partial^\mu\Lambda) = F^{\mu\nu} - (\partial^\mu\partial^\nu - \partial^\nu\partial^\mu)\Lambda = F^{\mu\nu}. \tag{2.22}$$

The last step follows because we can take the partial derivatives in any order.

In order to use the gauge potential to solve the field equations, one should eliminate this freedom to make gauge transformations, otherwise one can get misled into thinking that the solutions one obtains are all physically distinct. There are many ways to do this, but some of the most common are as follows.

Coulomb gauge This is also called radiation gauge, and is defined by

$$\nabla\cdot\mathbf{A} = 0, \tag{2.23}$$

which implies that the first of the inhomogeneous equations in (2.9) becomes

$$\nabla^2\phi = -\rho. \tag{2.24}$$

This gauge is therefore very useful in solving electrostatics problems, and we shall use it later on when studying the Casimir effect. However, it is not so often used in relativistic applications, as the gauge condition does not respect Lorentz invariance.

The second of the inhomogeneous Maxwell equations becomes, on using the gauge condition,

$$-\frac{\partial}{\partial t}(-\dot{\mathbf{A}} - \nabla\phi) - \nabla^2\mathbf{A} = \mathbf{j}, \tag{2.25}$$

which we may write as

$$(\frac{\partial^2}{\partial t^2} - \nabla^2)\mathbf{A} = \mathbf{j}_T, \tag{2.26}$$

where $\mathbf{j}_T = \mathbf{j} + \nabla\dot{\phi}$. One can show, using the equation of current conservation and Eq. (2.24), that

$$\nabla\cdot\mathbf{j}_T = 0. \tag{2.27}$$

This of course is necessary for the consistency of Eq. (2.26), because if one takes the divergence of the left hand side one gets zero as well.

In free space, that is, in the absence of charges and currents, we may take $\phi = 0$, and the equation for the vector potential becomes

$$(\frac{\partial^2}{\partial t^2} - \nabla^2)\mathbf{A} = 0. \tag{2.28}$$

One of the important discoveries of the last century was that this equation has plane wave solutions, which carry energy and momentum: electromagnetic waves. These solutions may be written

$$\mathbf{A}(t,\mathbf{x}) = \mathbf{a}e^{-i\omega t + i\mathbf{k}\cdot\mathbf{x}}, \tag{2.29}$$

where $\mathbf{a}$ is a constant 4-vector, and $\omega = \pm|\mathbf{k}|$.

The gauge condition (2.23) imposes a condition on the vector $\mathbf{a}$, for

$$\nabla\cdot\mathbf{A} = i\mathbf{k}\cdot\mathbf{a}e^{-i\omega t + i\mathbf{k}\cdot\mathbf{x}} = 0, \tag{2.30}$$

which implies that

$$\mathbf{k}\cdot\mathbf{a} = 0. \tag{2.31}$$

Hence the gauge potential $\mathbf{A}$ is orthogonal to the wave vector $\mathbf{k}$. This means that both the electric and magnetic fields are also orthogonal: we say that the waves are *transverse*.

Lorentz gauge This is very often the gauge to choose if one is interested in wave propagation in electromagnetism. It is defined by

$$\partial_\mu A^\mu = 0, \tag{2.32}$$

and is manifestly Lorentz invariant. In this gauge the equation of motion for the gauge field is simplified, for

$$\partial_\mu F^{\mu\nu} = \partial_\mu \partial^\mu A^\nu - \partial_\mu \partial^\nu A^\mu = \partial^2 A^\mu = j^\mu. \tag{2.33}$$

However, an irritating feature of this gauge is that it does not quite specify A^μ fully. One can still make a gauge transformation $A^\mu \to A^\mu - \partial^\mu \Lambda$ which satisfies the Lorentz gauge condition (2.32), as long as the function Λ satisfies $\partial^2 \Lambda = 0$. Such functions are called *harmonic*. In classical field theory this is not too much of a problem, but in setting up the quantum theory of gauge fields care must be taken.

In free space, the field equation is (2.33)

$$\Box A^\mu \equiv (\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2) A^\mu = 0. \tag{2.34}$$

Its solutions may be written

$$A^\mu(t, \mathbf{x}) = a^\mu e^{-i\omega t + i\mathbf{k}\cdot\mathbf{x}} = a^\mu e^{-ik\cdot x}, \tag{2.35}$$

where a^μ is a 4-vector independent of x^μ, and $\omega = \pm|\mathbf{k}|$.

The condition on a^μ imposed by the Lorentz gauge (2.32) is

$$\partial_\mu A^\mu = -ik_\mu a^\mu e^{-ik_\mu x^\mu} = 0, \tag{2.36}$$

which implies that

$$k \cdot a = 0. \tag{2.37}$$

Hence the field is orthogonal, in the 4-vector sense, to the wave vector k^μ.

2.3 Classical Lagrangian field theory

A field is an object which takes a value (e.g. a real number) at every point in space-time. We are probably most familiar with the electric and magnetic fields $\mathbf{E}(t, \mathbf{x})$ and $\mathbf{B}(t, \mathbf{x})$, which we have seen can be unified into the electromagnetic field strength tensor $F^{\mu\nu}(x)$. The Klein-Gordon field $\phi(t, \mathbf{x})$ is another example.

A very powerful way of classifying fields is to organise them according to their properties under Lorentz transformations. The electromagnetic gauge potential $A_\mu(x)$ is an example of a *vector field*, which is a field which transforms like a covariant 4-vector. (As a matter of notation, I will often write (x) as shorthand for $(t, \mathbf{x})$ in the arguments of functions.) Its transformation law is

$$A_\mu(x) \to A'^\mu(x') = \Lambda_\mu^{\ \nu} A_\nu(x), \tag{2.38}$$

where $x'^\mu = \Lambda^\mu_{\ \nu} x^\nu$. The Klein-Gordon field is an example of a *scalar field*, which transforms as a scalar (a pure number):

$$\phi(x) \to \phi'(x') = \phi(x). \tag{2.39}$$

In this part of the course we will unfortunately not have time to study the other important field in particle physics, which is the *spinor field*, again defined by its Lorentz transformation property.

In field theory, the dynamical coordinates, the analogues of the q_a of classical Lagrangian mechanics, are the values of the fields at every point. A field therefore has an uncountably infinite number of degrees of freedom, which is the source of many of the difficulties in field theory. Sums over the label a are replaced by integrals over space: for example, the Lagrangian can be expressed as the integral over space of an object called the *Lagrangian density* $\mathcal{L}$. For a scalar field, the Lagrangian density can be expressed as a function of the field, its time derivative, and also its spatial derivatives:

$$L = \int d^3x \ \mathcal{L}(\phi, \dot{\phi}, \boldsymbol{\nabla}\phi, t). \tag{2.40}$$

The appearance of spatial derivatives may at first sight be puzzling, but arises naturally from terms coupling neighbouring space points, in the limit that the separation goes to zero. The puzzle is perhaps more that higher derivatives don't appear: in practice, they seem not be relevant for most applications in particle physics.

From a Lagrangian we can derive Euler-Lagrange equations, by the application of Hamilton's Principle to the action for a scalar field,

$$I = \int_{t_1}^{t_2} L dt = \int_{t_1}^{t_2} dt d^3x \ \mathcal{L}(\phi, \dot{\phi}, \boldsymbol{\nabla}\phi, t). \tag{2.41}$$

Consider an arbitrary variation in the field, $\phi(x) \to \phi(x) + \delta\phi(x)$, where $\delta\phi(x)$ vanishes at t_1 and t_2, and outside a spatial region R with boundary B. This variation causes a small change in the action:

$$I \to I' = I + \delta I. \tag{2.42}$$

Taylor expanding the Lagrangian density, we find

$$I' \simeq \int_{t_1}^{t_2} dt d^3x \ [\mathcal{L} + \frac{\partial \mathcal{L}}{\partial \phi}\delta\phi + \frac{\partial \mathcal{L}}{\partial \boldsymbol{\nabla}\phi}\cdot\boldsymbol{\nabla}\delta\phi + \frac{\partial \mathcal{L}}{\partial \dot{\phi}}\delta\dot{\phi}], \tag{2.43}$$

where

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\nabla}\phi} = \left(\frac{\partial \mathcal{L}}{\partial(\partial_x \phi)}, \frac{\partial \mathcal{L}}{\partial(\partial_y \phi)}, \frac{\partial \mathcal{L}}{\partial(\partial_z \phi)}\right). \tag{2.44}$$

We can separate out the variation in the action δI, and integrate the last term by parts, to obtain

$$\delta I = \int_{t_1}^{t_2} dt d^3x \, [\frac{\partial \mathcal{L}}{\partial \phi}\delta\phi + \frac{\partial \mathcal{L}}{\partial \nabla\phi}\cdot\nabla\phi\delta\phi - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\delta\phi] + \left[\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\delta\phi\right]_{t_1}^{t_2}. \tag{2.45}$$

In this last expression the last term vanishes because of the conditions placed on $\delta\phi$ at t_1 and t_2. Hence, upon integrating by parts again, we get

$$\delta I = \int_{t_1}^{t_2} dt d^3x \, [\frac{\partial \mathcal{L}}{\partial \phi}\delta\phi - \nabla\cdot\frac{\partial \mathcal{L}}{\partial \nabla\phi}\delta\phi - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\delta\phi] + \int_{t_1}^{t_2} dt \int_B \mathbf{dS}\cdot\left[\frac{\partial \mathcal{L}}{\partial \nabla\phi}\delta\phi\right]. \tag{2.46}$$

Again, the last term vanishes, as we supposed that the variation $\delta\phi$ vanished on the boundary B. Hence we are left with

$$\delta I = \int_{t_1}^{t_2} dt d^3x \, [\frac{\partial \mathcal{L}}{\partial \phi} - \nabla\cdot\frac{\partial \mathcal{L}}{\partial \nabla\phi} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{\phi}}]\delta\phi. \tag{2.47}$$

The argument proceeds by noting that not only is $\delta\phi$ arbitrary, but so is the region R. Hence the condition that the action be stationary reduces to

$$\frac{\partial \mathcal{L}}{\partial \phi} - \nabla\cdot\frac{\partial \mathcal{L}}{\partial \nabla\phi} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{\phi}} = 0. \tag{2.48}$$

This is the Euler-Lagrange equation for a scalar field.

We can put it into a more obviously relativistic form by noting that $\partial_\mu\phi = (\partial_t\phi, \partial_i\phi)$, so that the Euler-Lagrange equation becomes

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_i\frac{\partial \mathcal{L}}{\partial(\partial_i\phi)} - \partial_0\frac{\partial \mathcal{L}}{\partial(\partial_0\phi)} = 0, \tag{2.49}$$

or

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)} = 0. \tag{2.50}$$

If there are many fields ϕ_A we can derive an Euler-Lagrange equation for each of them,

$$\frac{\partial \mathcal{L}}{\partial \phi_A} - \partial_\mu\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi_A)} = 0. \tag{2.51}$$

We are now in a position to ask the question: from what Lagrangian density does the Klein-Gordon equation follow? The Klein-Gordon field may be split into its real and imaginary parts, $\phi(x) = \phi_R(x) + i\phi_I(x)$, each of which obey the same equation

$$(\partial^\mu\partial_\mu + m^2)\phi_R(x) = 0, \quad (\partial^\mu\partial_\mu + m^2)\phi_I(x) = 0. \tag{2.52}$$

Consider first the following Lagrangian density for the real part,

$$\mathcal{L}_R = \frac{1}{2}\partial^\mu\phi_R\partial_\mu\phi_R - \frac{1}{2}m^2\phi_R^2. \tag{2.53}$$

We can immediately derive

$$\frac{\partial \mathcal{L}}{\partial \phi_R} = -m^2\phi_R, \quad \frac{\partial \mathcal{L}}{\partial(\partial_i\phi_R)} = \partial^i\phi_R, \quad \frac{\partial \mathcal{L}}{\partial(\partial_0\phi_R)} = \partial^0\phi. \tag{2.54}$$

Hence,

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi_R)} = (\partial^0\phi_R, \partial^i\phi_R), \tag{2.55}$$

and so

$$\frac{\partial \mathcal{L}}{\partial \phi_R} - \partial_\mu\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi_R)} = -m^2\phi_R - \partial_\mu\partial^\mu\phi = 0. \tag{2.56}$$

Thus the Lagrangian (2.53) implies the correct equation for the real part of the Klein-Gordon field: exactly the same argument may be made for the imaginary part of the field, so we can write for the total lagrangian

$$\begin{aligned} \mathcal{L} = \mathcal{L}_R + \mathcal{L}_L &= \frac{1}{2}\partial^\mu\phi_R\partial_\mu\phi_R - \frac{1}{2}m^2\phi_R^2 + \frac{1}{2}\partial^\mu\phi_I\partial_\mu\phi_I - \frac{1}{2}m^2\phi_I^2 \\ &= \frac{1}{2}pa^\mu\phi^*\partial_\mu\phi - \frac{1}{2}m^2\phi^*\phi. \end{aligned} \tag{2.57}$$

Lastly in this section, we will write down the Hamiltonian for a scalar field (sticking with a single real field for the time being). Drawing on our experience with finite numbers of degrees of freedom, it seems natural to define a conjugate momentum

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(x)}. \tag{2.58}$$

Thus for our example (2.53), $\pi(x) = \dot{\phi}(x)$. The Hamiltonian is then defined

$$H = \int d^3x \, \pi(x)\dot{\phi}(x) - \int d^3x \mathcal{L}(\phi, \partial_i\phi, \dot{\phi}). \tag{2.59}$$

The Hamiltonian for a field obeying the Klein-Gordon equation is then

$$H = \int d^3x \left[\frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2\right]. \tag{2.60}$$

We talk of the three terms in the Hamiltonian as *kinetic*, *gradient* and *potential* terms respectively. One can allow more general functions of ϕ in the potential term than just a quadratic:

$$H = \int d^3x \left[\frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + V(\phi)\right]. \tag{2.61}$$

In classical field theory, V can be any function of ϕ, as long as it is bounded below. Quantum field theory is another story however.

2.3.1 Classical Lagrangian for electromagnetism

In this section we will just write down a Lagrangian density, and show that the resulting Euler-Lagrange equations are indeed just Maxwell's equations. The Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - j^\mu A_\mu, \tag{2.62}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Let us expand, to get

$$\mathcal{L} = -\frac{1}{4}(F_{0i}F^{0i} + F^{i0}F^{i0} + F_{ij}F^{ij}) - j^\mu A_\mu, \tag{2.63}$$

The antisymmetry of $F^{\mu\nu}$ immediately gives us

$$\mathcal{L} = -\frac{1}{2}(F_{i0}F^{i0} + \frac{1}{2}F_{ij}F^{ij}) - j^\mu A_\mu. \tag{2.64}$$

Lowering the indices

$$F^{i0} = F^i_0 = -F_{i0}, \tag{2.65}$$
$$F^{ij} = -F^i_j = F_{ij}, \tag{2.66}$$

we arrive at

$$\mathcal{L} = -\frac{1}{2}(F_{i0}F_{i0} + \frac{1}{2}F_{ij}F_{ij}) - j^\mu A_\mu. \tag{2.67}$$

The generalised coordinates of this system are the gauge potentials A_0 and A_i. We can therefore derive the Maxwell's equations by looking at their Euler-Lagrange equations:

$$\frac{\partial\mathcal{L}}{\partial A_0} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_0)} = 0, \quad \frac{\partial\mathcal{L}}{\partial A_i} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_i)} = 0. \tag{2.68}$$

Examining first the equations for A_0, we find

$$\frac{\partial\mathcal{L}}{\partial A_0} = -j^0, \quad \frac{\partial\mathcal{L}}{\partial(\partial_0 A_0)} = 0. \tag{2.69}$$

The next piece of the Euler-Lagrange equation deserves a closer look:

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial(\partial_j A_0)} &= \frac{\partial F_{i0}}{\partial(\partial_j A_0)}F_{i0}, \\ &= \delta ji F_{i0} = F_{j0}.\end{aligned} \tag{2.70}$$

Hence we have

$$-j^0 - \partial_j F_{j0} = 0, \tag{2.71}$$

which can be recognised as the equation

$$\partial_\mu F^{\mu 0} = j^0, \tag{2.72}$$

which is one of the inhomogeneous Maxwell equations.

Next we take a look at the equations for A_i:

$$\frac{\partial\mathcal{L}}{\partial A_i} = -j^i, \quad \frac{\partial\mathcal{L}}{\partial(\partial_0 A_i)} = -F_{i0} = F_{0i}. \tag{2.73}$$

Again, we have a piece that is not quite so straightforward to work out:

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial(\partial_j A_i)} &= \frac{1}{2}\frac{\partial F_{kl}}{\partial(\partial_j A_i)}F_{kl}, \\ &= \frac{1}{2}(\delta kj\delta li - \delta ki\delta lj)F_{kl} = F_{ji}.\end{aligned} \tag{2.74}$$

Putting the pieces together, we find

$$-j^i - \partial_0 F_{0i} - \partial_i F_{ji} = 0, \tag{2.75}$$

or upon raising indices and rearranging

$$\partial_\mu F^{\mu i} = j^i. \tag{2.76}$$

Putting (2.72) and (2.76) together, we find that we have all the inhomogeneous Maxwell equations,

$$\partial_\mu F^{\mu\nu} = j^\nu. \tag{2.77}$$

The homogeneous Maxwell equations are in fact an identity, and do not follow from a variational principle. This can be seen from their formulation as

$$\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\partial^\nu F^{\rho\sigma} = 0. \tag{2.78}$$

Substituting the expression for the field strength tensor in terms of the gauge potential, we obtain

$$\epsilon_{\mu\nu\rho\sigma}\partial^\nu\partial^\rho A^\sigma = 0. \tag{2.79}$$

By exchanging the labels ν and ρ (which we can do as they are summed over), we can deduce

$$\epsilon_{\mu\nu\rho\sigma}\partial^\nu\partial^\rho A^\sigma = \epsilon_{\mu\rho\nu\sigma}\partial^\rho\partial^\nu A^\sigma. \tag{2.80}$$

The antisymmetry of the Levi-Civita tensor then shows that

$$\epsilon_{\mu\nu\rho\sigma}\partial^\nu\partial^\rho A^\sigma = -\epsilon_{\mu\nu\rho\sigma}\partial^\nu\partial^\rho A^\sigma. \tag{2.81}$$

The only number that is equal to its negative is zero, hence

$$\epsilon_{\mu\nu\rho\sigma}\partial^\nu\partial^\rho A^\sigma = 0. \tag{2.82}$$

Thus we have demonsrated what we set out to show: that the Principle of Least Action applied to the Lagrangian (2.62) results in Maxwell's equations, half of which are seen to be a identities rather than dynamical equations.

3 Quantum Theory of Free Fields

3.1 The real scalar field

In this section we shall proceed with quantising the real scalar field. Firstly, we will go through the quantisation procedure for a system with a finite number of degrees of freedom, to show that quantising a field is not unlike quantising a more familiar system. Suppose the system has N coordinates q_a with associated velocities $\dot{q}_a$, and Lagrangian

$$L = \frac{1}{2}\sum_{a=1}^{N} m\dot{q}_a^2 - U(q_a) \tag{3.1}$$

(we are supposing for simplicity's sake that all the masses are the same). Each coordinate is also associated with a conjugate momentum p_a, defined as the differential of the Lagrangian with respect to the velocity $\dot{q}_a$, which in this case has the value $m\dot{q}_a$. The Hamiltonian, defined as $\frac{1}{2}\sum_a \dot{q}_a p_a - L$, is then

$$H = \sum_{a=1}^{N} \frac{p_a^2}{2m} + U(q_a). \tag{3.2}$$

The procedure of *canonical quantisation* replaces the classical conjugate variables q_a, p_a by *operators* obeying the following commutation relations

$$[\hat{q}_a, \hat{p}_b] = i\delta_{ab}, \quad [\hat{q}_a, \hat{q}_b] = 0 = [\hat{p}_a, \hat{p}_b]. \tag{3.3}$$

The Hamiltonian is also replaced by an operator $\hat{H}$, which generates the time evolution of the system. We are probably used to thinking about time evolution as the change in time of the state of the system, governed by the *Schrödinger equation*

$$i\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle. \tag{3.4}$$

However, there is another way of thinking about the time evolution of a quantum mechanical system known as the *Heisenberg picture* (to distinguish it from the *Schrödinger picture* above).

In the Heisenberg picture, the states $|\psi\rangle_H$ that are time-independent, and the time evolution of the system is transferred to the operators $O_H(t)$, which satisfy the Heisenberg equations

$$i\frac{d}{dt}O_H(t) = [O_H(t), H]. \tag{3.5}$$

The states and operators in these two pictures can be related by a unitary transformation e^{iHt} (a unitary operator is one whose inverse is its hermitean conjugate). Thus, if we suppose that the two sets of operators and states coincide at

$t = 0$, they are related at all times by

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

This ensures that the matrix elements are the same. To see this, let us consider the matrix elements between any two states $|\psi\rangle_H$ and $|\psi'\rangle_H$ in the Heisenberg picture, and apply these transformations:

$$\begin{aligned} {}_H\langle\psi|O_H(t)|\psi'\rangle_H &= {}_S\langle\psi(t)|e^{-iHt}e^{iHt}O_S e^{-iHt}e^{iHt}|\psi'(t)\rangle_S \\ &= {}_S\langle\psi(t)|O_S|\psi'(t)\rangle_S. \end{aligned} \tag{3.7}$$

This demonstrates the point.

Without any attempt at rigour, we will go through the quantisation procedure for free fields. By "free" we mean non-interacting: once a state has been set up there are no transitions to any other states. This may sound somewhat irrelevant, but essentially all quantum field theories are treated as small perturbations away from free field theories, so it is an appropriate place to start. We will also be able to see the Casimir effect in the free electromagnetic field.

We begin with the Lagrangian density

$$\mathcal{L} = \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - V(\phi), \tag{3.8}$$

from which we can define the conjugate momentum

$$\pi(x) = \frac{\partial\mathcal{L}}{\partial\dot{\phi}(x)}. \tag{3.9}$$

The Hamiltonian density is then found from

$$\mathcal{H} = \pi(x)\dot{\phi}(x) - \mathcal{L} = \frac{1}{2}\pi^2 + \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 + V(\phi). \tag{3.10}$$

In free quantum field theory we usually work in the Heisenberg picture, where the fields carry explicit time dependence. The field theory is quantised by constructing the *equal time canonical commutation relations*

$$[\phi(t,\mathbf{x}), \pi(t,\mathbf{x}')] = i\hbar\delta^3(\mathbf{x}-\mathbf{x}'), \tag{3.11}$$
$$[\phi(t,\mathbf{x}), \phi(t,\mathbf{x}')] = 0, \tag{3.12}$$
$$[\pi(t,\mathbf{x}), \pi(t,\mathbf{x}')] = 0, \tag{3.13}$$

where $\delta^3(\mathbf{x}-\mathbf{x}') = \delta(x^1 - x'^1)\delta(x^2 - x'^2)\delta(x^3 - x'^3)$. One can see the analogy between the field commutation relations and (3.3) if one recalls that the space coordinate $\mathbf{x}$ is like the label a in the system with a finite number of degrees of freedom. The δ-function acts in the same way for a continuous label $\mathbf{x}$ as the Kronecker δ does for a discrete label.

So in quantum field theory, the fields become *operators*, and the field equation becomes an operator equation. For example, we already know the classical equation of motion for a scalar field when $V(\phi) = \frac{1}{2}m^2\phi^2$ – it is just the Klein-Gordon equation. In the quantum field theory we have

$$\partial^2\phi + m^2\phi = 0, \tag{3.14}$$

where the fields are now *operators*. We already know the general solution to this equation, except that we now need to expand in Fourier coefficients which are also operators:

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}\left(a(\mathbf{k})e^{-ik\cdot x} + a^*(\mathbf{k})e^{ik\cdot x}\right). \tag{3.15}$$

Recall that the equations of motion require that $k^0 = \omega_{\mathbf{k}} = |(\mathbf{k}^2 + m^2)^{1/2}|$.

We need to know the properties of $a(\mathbf{k})$ and $a^*(\mathbf{k})$ to proceed further. We start with their commutation relations, which we can deduce from those of ϕ and π. We therefore need expressions for $a(\mathbf{k})$ and $a^*(\mathbf{k})$ in terms of ϕ and π, where

$$\pi(x) = \dot{\phi}(x) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}\left(-i\omega_{\mathbf{k}}a(\mathbf{k})e^{-ik\cdot x} + i\omega_{\mathbf{k}}a^*(\mathbf{k})e^{ik\cdot x}\right). \tag{3.16}$$

Firstly, we take the Fourier transform of the field operator and its conjugate momentum:

$$\phi(\mathbf{k}') = \int d^3x\phi(x)e^{-i\mathbf{k}'\cdot\mathbf{x}}, \quad \pi(\mathbf{k}') = \int d^3x\pi(x)e^{-i\mathbf{k}'\cdot\mathbf{x}}, \tag{3.17}$$

and use the relation

$$\int d^3x e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} = (2\pi)^3\delta^3(\mathbf{k}-\mathbf{k}') \tag{3.18}$$

to obtain

$$\begin{aligned}
\phi(\mathbf{k}') &= \frac{1}{2\omega'_{\mathbf{k}}}(a(\mathbf{k}')e^{-i\omega_{\mathbf{k}}t} + a^*(-\mathbf{k}')e^{i\omega'_{\mathbf{k}}t}) & (3.19)\\
\pi(\mathbf{k}') &= \frac{1}{2\omega'_{\mathbf{k}}}(-i\omega'_{\mathbf{k}}a(\mathbf{k}')e^{-i\omega_{\mathbf{k}}t} + i\omega'_{\mathbf{k}}a^*(-\mathbf{k}')e^{i\omega'_{\mathbf{k}}t}). & (3.20)
\end{aligned}$$

Hence

$$a(\mathbf{k}') = \int d^3x(i\pi(x) + \omega'_{\mathbf{k}}\phi(x))e^{ik'\cdot x}. \tag{3.21}$$

By complex conjugation one obtains

$$\hat{a}^*(\mathbf{k}') = \int d^3x(-i\pi(x) + \omega'_{\mathbf{k}}\phi(x))e^{-ik'\cdot x}. \tag{3.22}$$

One is then well on the way to deriving the commutation relations

$$\begin{aligned}
[a(\mathbf{k}), a^*(\mathbf{k}')] &= 2\omega_{\mathbf{k}}(2\pi)^3\delta^3(\mathbf{k}-\mathbf{k}'), & (3.23)\\
[a(\mathbf{k}), a(\mathbf{k}')] &= 0, & (3.24)\\
[a^*(\mathbf{k}), a^*(\mathbf{k}')] &= 0. & (3.25)
\end{aligned}$$

These look very similar to the commutation relations for simple harmonic oscillators. We recall that the harmonic oscillator of unit mass is described by the Hamiltonian

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2x^2, \tag{3.26}$$

where ω is the angular frequency of oscillation. Upon quantisation, it is found that the energy eigenstates of the system are labelled by an integer n, and have energy $E_n = (n+\frac{1}{2})\omega$, and that the operators $a = (ip+\omega x)/\sqrt{2\omega}$ and $a^* = (-ip+\omega x)/\sqrt{2\omega}$ satisfy the commutation relations

$$[a, a^*] = 1, \qquad [a, a] = 0 = [a^*, a^*]. \tag{3.27}$$

If we denote the energy eigenstates by $|n\rangle$, then a^* and a act to change n by one unit – they are called *raising* and *lowering* operators for that reason. The lowest energy state or ground state $|0\rangle$ has the property that $a|0\rangle = 0$, by definition: if one could lower its energy any further it would not be the ground state.

So we have discovered that the field theory behaves very much like infinitely many harmonic oscillators, each with a different frequency $\omega_{\mathbf{k}} = \sqrt{(\mathbf{k}^2 + m^2)}$. We will see that the occupation number n of each oscillator actually corresponds to the number of *particles* with that particular momentum in the state considered. The raising and lowering operators $a^*(\mathbf{k})$ and $a(\mathbf{k})$ are also called *creation* and *annihilation* operators, because they change the number of particles with 3-momentum $\mathbf{k}$ in a state.

3.2 States of the scalar field; zero point energy

We saw in the previous section that the scalar field could be decomposed into a sum over operators obeying oscillator-like commutation relations, one pair for each wavevector $\mathbf{k}$. It will therefore be no surprise to see that the states of the quantum field theory are just towers of oscillator states, one for each $\mathbf{k}$. The ground state is defined as the state annihilated by all the annihilation operators:

$$a(\mathbf{k})|0\rangle = 0, \quad \forall\mathbf{k}. \tag{3.28}$$

We can use the creation operators $a^*(\mathbf{k})$ to construct excited states: for example, the first excited state in mode $\mathbf{k}$ may be written

$$|\mathbf{k}\rangle = a^*(\mathbf{k})|0\rangle. \tag{3.29}$$

The normalisation of these states follows from deciding that the vacuum state has unit norm, that is $\langle 0||0\rangle = 1$. Thus

$$
\begin{aligned}
\langle \mathbf{k}'|\mathbf{k}\rangle &= \langle 0|a(\mathbf{k}')a^*(\mathbf{k})|0\rangle &(3.30)\\
&= \langle 0|a^*(\mathbf{k})a(\mathbf{k}')|0\rangle + \langle 0|0\rangle 2\omega_{\mathbf{k}}\,\bar{\delta}(\mathbf{k}-\mathbf{k}'), &(3.31)
\end{aligned}
$$

where another peice of notation has been introduced: $\bar{\delta}(\mathbf{k}-\mathbf{k}') = (2\pi)^3\delta(\mathbf{k}-\mathbf{k}')$. The first term on the right hand side vanishes, and so we have that the first excited states satisfy the normalisation condition

$$\langle \mathbf{k}'|\mathbf{k}\rangle = 2\omega_{\mathbf{k}}\,\bar{\delta}(\mathbf{k}-\mathbf{k}'). \tag{3.32}$$

We call states $|\mathbf{k}\rangle$ single-particle states, as they describe states with a single particle of momentum $\mathbf{k}$. But we could continue applying creation operators $a^*(\mathbf{k}_1)$, $a^*(\mathbf{k}_2), \ldots$ to make states with any number of particles:

$$|\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2, \ldots\rangle = \ldots a^*(\mathbf{k}_2)a^*(\mathbf{k}_1)a^*(\mathbf{k})|0\rangle. \tag{3.33}$$

This is one of the special features of field theory: it is a quantum mechanical theory of many particles. Ordinary non-relativistic quantum mechanics describes only one particle at a time.

There is an operator which allows us to count the number of particles in a given state. Let us first introduce the set of operators $N(\mathbf{k})$ by

$$N(\mathbf{k}) = a^*(\mathbf{k})a(\mathbf{k}). \tag{3.34}$$

It is straightforward to see that they return zero when acting on the ground state,

$$N(\mathbf{k})|0\rangle = 0, \tag{3.35}$$

from the definition of the vacuum state (3.28). When acting on 1-particle states, we find

$$
\begin{aligned}
N(\mathbf{k})|\mathbf{k}'\rangle &= a^*(\mathbf{k})a(\mathbf{k})a^*(\mathbf{k}')|0\rangle &(3.36)\\
&= a^*(\mathbf{k})\left(a^*(\mathbf{k}')a(\mathbf{k}) + 2\omega_{\mathbf{k}}\,\bar{\delta}(\mathbf{k}-\mathbf{k}')\right)|0\rangle, &(3.37)
\end{aligned}
$$

where we have used the commutation relations (3.25) to change the order of $a(\mathbf{k})$ and $a^*(\mathbf{k}')$. Thus,

$$N(\mathbf{k})|\mathbf{k}'\rangle = 2\omega_{\mathbf{k}}\,\bar{\delta}(\mathbf{k}-\mathbf{k}')|\mathbf{k}\rangle. \tag{3.38}$$

The operators $N(\mathbf{k})$ are used to construct the total number operator

$$N_{\text{tot}} = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}N(\mathbf{k}), \tag{3.39}$$

which tells us the total number of particles in a particular state[1]. Thus, as expected, $N_{\text{tot}}|0\rangle = 0$, meaning that there are no particles in the ground state. In the 1-particle state $|\mathbf{k}'\rangle$, we find

$$N_{\text{tot}}|\mathbf{k}'\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}N(\mathbf{k})|\mathbf{k}'\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}2\omega_{\mathbf{k}}\,\bar{\delta}(\mathbf{k}-\mathbf{k}')|\mathbf{k}\rangle = |\mathbf{k}'\rangle. \tag{3.40}$$

The state $|\mathbf{k}'\rangle$ is indeed an eigenstate of the number operator, with eigenvalue 1 – there is 1 particle in state.

The energy of the states can be computed from the quantum Hamiltonian, which is constructed from the field operators in exactly the same way as the classical Hamiltonian is constructed from the classical fields:

$$\hat{H} = \int d^3x\left(\frac{1}{2}\hat{\pi}^2 + \frac{1}{2}(\boldsymbol{\nabla}\hat{\phi})^2 + \frac{1}{2}\mu^2\hat{\phi}^2\right). \tag{3.41}$$

Since the simple states we have seen so far have been constructed from creation operators $a^*(\mathbf{k})$ acting on the vacuum, it is more convenient to re-express $\hat{H}$ in terms of creation and annihilation operators. Recalling the expansions of $\hat{\phi}$ and $\hat{\pi}$ in terms of $\hat{a}(\mathbf{k})$ and $\hat{a}^*(\mathbf{k})$ in equations (3.15) and (3.16), we find

$$\hat{H} = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}\frac{\omega_{\mathbf{k}}}{2}\left(\hat{a}^*(\mathbf{k})\hat{a}(\mathbf{k}) + \hat{a}(\mathbf{k})\hat{a}^*(\mathbf{k})\right). \tag{3.42}$$

If we again try to change the order of the operators, we run into a problem, for

$$\hat{H} = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}\frac{\omega_{\mathbf{k}}}{2}\left(2\hat{a}^*(\mathbf{k})\hat{a}(\mathbf{k}) + \bar{\delta}(\mathbf{0})\omega_{\mathbf{k}}\right). \tag{3.43}$$

The second term in the brackets has the formally infinite term $\bar{\delta}(\mathbf{0})$. We regard $\bar{\delta}(\mathbf{0})$ as the spatial volume, the quantity it multiplies is then an energy *density*. However, even this energy density is infinite, in general. Suppose we try to find the energy of the ground state E_0, for which

$$\hat{H}|0\rangle = E_0|0\rangle = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}\omega_{\mathbf{k}}\hat{a}^*(\mathbf{k})\hat{a}(\mathbf{k})|0\rangle + \frac{1}{2}\int \bar{d}^3k\,\omega_{\mathbf{k}}\,\bar{\delta}(\mathbf{0})|0\rangle. \tag{3.44}$$

The first term on the right hand side is zero, but the second contains the integral $\frac{1}{2}\int \bar{d}^3k\,\omega_{\mathbf{k}}$ multiplying the volume factor $\bar{\delta}(\mathbf{0})$. This integral is divergent. If we integrate the momenta over the range $0 < |\mathbf{k}| < \Lambda$, with the upper cut-off in the integration range $\Lambda \gg \mu$, we find that the integral is proportional to Λ^4. As we take Λ to infinity, as we eventually must, the ground state energy diverges as the fourth power of the cut-off in the momentum – we say that the integral is *quartically divergent*. Thus the ground state energy of a field theory appears to

[1] if that state has a definite particle number – it may be that the state is not an eigenvalue of the number operator

be infinite. This is the first of many infinities in quantum field theory, and the way of dealing with this, and all others, is to *renormalise*, that is, to subtract off unobservable quantities. We say that the absolute value of the the energy of the ground state is unobservable: all we see is differences between that state and others. Thus we measure all energies relative to the ground state by subtracting the term which gives the infinite value for the ground state energy, $\frac{1}{2}\int d\!\!{}^-{}^3k\omega_{\mathbf{k}}\,\delta\!\!{}^-(\mathbf{0})$, and we define a new, *renormalised* Hamiltonian

$$\hat{H}' = \hat{H} - \frac{1}{2}\int d\!\!{}^-{}^3k\omega_{\mathbf{k}}\,\delta\!\!{}^-(\mathbf{0}) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}\omega_{\mathbf{k}}\hat{a}^*(\mathbf{k})\hat{a}(\mathbf{k}). \tag{3.45}$$

This has the sensible property that $\hat{H}'|0\rangle = 0$ – the ground state energy is zero. You might like to check that the energy of the state $|\mathbf{k}'\rangle$ is indeed $\omega'_{\mathbf{k}}$: a simple way to start is to note that

$$\hat{H}' = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}\omega_{\mathbf{k}}N(\mathbf{k}), \tag{3.46}$$

and use Equation (3.38).

We can subtract off ground state energies automatically by a procedure known as *normal ordering*. When we place an operator in normal order, we write it so that all the creation operators are to the left, and all the annihilation operators are to the right, so for example

$$:\hat{a}(\mathbf{k})\hat{a}(\mathbf{k}'): = \hat{a}^*(\mathbf{k}')\hat{a}(\mathbf{k}). \tag{3.47}$$

Hence,

$$:\hat{H}: = \hat{H}'. \tag{3.48}$$

3.3 The Electromagnetic Field

By now we should be getting the hand of how to quantise a given field theory. We find the dynamical coordinates, identify the conjugate momentum, and impose canonical commutation relations between them. Let us follow this procedure for the electromagnetic field, but we should be aware that there are extra complications not present for the real scalar field. Firstly, there is gauge invariance: recall that the transformation $A_\mu \to A_\mu - \partial_\mu\Lambda$ does not change any physical quantity associated with the field, which means we are in danger of encoutering unphysical degrees of freedom when we use A_μ as the dynamical coordinates. Secondly, the conjugate momentum to A_μ looks like it should be

$$\pi^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_0 A_\mu)} = F^{0\mu}. \tag{3.49}$$

But this means that A_0 does not have a canonical momentum, as $\pi^0 = F^{00}$ vanishes identically.

Our approach to dealing with these problems is to choose a gauge from the outset, trying to leave only physical degrees of freedom in the Lagrangian. In this section the gauge will be the Coulomb gauge, or $\partial_i A^i = 0$. In this gauge, A_0 is entirely determined at any time by the charge distribution, through the equation

$$-\nabla^2 A_0 = \rho, \tag{3.50}$$

and it is therefore not a true dynamical coordinate. In free space, that is, in the absence of charges or currents, we are entitled to set it to zero.

We recall the action of the electromagnetic field, in the absence of an external 4-current:

$$S = \int d^4x\,\mathcal{L} = -\int d^4x\,\frac{1}{4}F_{\mu\nu}F^{\mu\nu},$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Problem p(C.2) tells us that in the Coulomb gauge, this action is equivalent to

$$S_{\mathrm{Cg}} = \int d^4x\,\mathcal{L}_{\mathrm{Cg}} = \int d^4x\,\left(\frac{1}{2}\dot{A}^i\dot{A}^i - \frac{1}{2}\partial_j A^i\partial_j A^i\right). \tag{3.51}$$

Thus the momentum conjugate to A^i is

$$\pi_i = \frac{\partial\mathcal{L}_{\mathrm{Cg}}}{\partial\dot{A}^i} = \dot{A}^i, \tag{3.52}$$

and the Hamiltonian density is

$$\mathcal{H}_{\mathrm{Cg}} = \pi_i\dot{A}^i - \mathcal{L}_{\mathrm{Cg}} = \frac{1}{2}\dot{A}^i\dot{A}^i + \frac{1}{2}\partial_j A^i\partial_j A^i. \tag{3.53}$$

At this point it would seem natural to impose a commutation relation

$$[A^i(\mathbf{x},t), \pi_j(\mathbf{x}',t)] \stackrel{?}{=} i\delta^i_j\delta(\mathbf{x}-\mathbf{x}'), \tag{3.54}$$

based on our experience with scalar fields. This is *incorrect*, as differentiating both sides with respect to x^i quickly shows. The gauge condition implies that the left hand side vanishes. However,

$$[\partial_i A^i(\mathbf{x},t), \pi_j(\mathbf{x}',t)] = i\partial_j\delta(\mathbf{x}-\mathbf{x}') \neq 0. \tag{3.55}$$

So we need a more general function $\Delta^i{}_j(\mathbf{x},\mathbf{x}')$ for the right hand side of the CCR, which satisfies $\partial_i\Delta^i{}_j(\mathbf{x},\mathbf{x}') = 0$ (and also $\partial'_j\Delta^i{}_j(\mathbf{x},\mathbf{x}') = 0$). We also know that the dependence on $\mathbf{x}$ and $\mathbf{x}'$ must be in the combination $\mathbf{x}-\mathbf{x}'$ through translational invariance.

It is easier to find the function $\Delta^i{}_j(\mathbf{x}-\mathbf{x}')$ in Fourier space: we write

$$\Delta^i{}_j(\mathbf{x}-\mathbf{x}') = \int \frac{d^3k}{(2\pi)^3}\tilde{\Delta}^i{}_j(\mathbf{k})e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}. \tag{3.56}$$

Then

$$\partial_i \Delta^i{}_j(\mathbf{x}-\mathbf{x}') = -i\int \frac{d^3k}{(2\pi)^3} k_i \tilde{\Delta}^i{}_j(\mathbf{k}) e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} = 0, \tag{3.57}$$

from which we infer that

$$k_i \tilde{\Delta}^i{}_j(\mathbf{k}) = 0. \tag{3.58}$$

The solution to this equation is

$$\Delta^{ij}(\mathbf{k}) = \left(\delta^{ij} - \frac{k^i k^j}{\mathbf{k}^2}\right) F(\mathbf{k}), \tag{3.59}$$

where $F(\mathbf{k})$ is some function of $\mathbf{k}$. One has to do a little more work to show that the correct function is just $F = 1$: for now we will just have to take it on trust. Thus we may write the canonical commutation relations in the Coulomb gauge as

$$[A^i(\mathbf{x}, t), E^j(\mathbf{x}', t)] = -i\int \frac{d^3k}{(2\pi)^3}\left(\delta^{ij} - \frac{k^i k^j}{\mathbf{k}^2}\right) e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}. \tag{3.60}$$

This is often written

$$[A^i(\mathbf{x}, t), E^j(\mathbf{x}', t)] = \left(\delta^{ij} - \frac{\partial^i \partial^j}{\nabla^2}\right)\delta(\mathbf{x}-\mathbf{x}'), \tag{3.61}$$

where ∇^{-2} is *defined* in terms of its Fourier transform. Given any function $f(\mathbf{x})$ with Fourier transform $\tilde{f}(\mathbf{k})$, $\nabla^{-2} f(\mathbf{x})$ is defined as that function whose Fourier transform is $\mathbf{k}^{-2}\tilde{f}(\mathbf{k})$.

As we saw in Section 2.2.1, the free space Coulomb gauge equations of motion for the electromagnetic field are

$$(\partial_0^2 - \partial_j^2) A^i(t, \mathbf{x}) = 0,$$

which have the general solution

$$A^i(t, \mathbf{x}) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}\left(a^i e^{-ik\cdot x} + a^{*i} e^{ik\cdot x}\right),$$

where $k_i a^i = 0$. Thus the amplitudes a^i are not independent: they satisfy a constraint arising from the Coulomb gauge condition. It is much more convenient to have two freely chosen functions of $\mathbf{k}$ rather than three functions subject to a constraint, so we instead write the general solution

$$A^i(t, \mathbf{x}) = \sum_{\lambda=1,2}\int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}\left(a_{(\lambda)}\epsilon^i_{(\lambda)} e^{-ik\cdot x} + a^*_{(\lambda)}\epsilon^*_{(\lambda)}{}^i e^{ik\cdot x}\right). \tag{3.62}$$

The $\epsilon^i_{(\lambda)}$ are called *polarisation vectors*, and have the properties

$$k_i \epsilon^i_{(\lambda)} = 0, \quad \epsilon^i_{(\lambda)}\epsilon^i_{(\lambda')} = \delta_{\lambda\lambda'}. \tag{3.63}$$

These conditions mean that, along with $\hat{k}^i = k^i/|\mathbf{k}|$, the two polarisation vectors form an orthonormal basis for 3-vectors. The completeness property therefore follows:

$$\epsilon^i_{(1)}\epsilon^j_{(1)} + \epsilon^i_{(2)}\epsilon^j_{(2)} + \hat{k}^i\hat{k}^j = \delta^{ij}, \tag{3.64}$$

or $\sum_\lambda \epsilon^i_{(\lambda)}\epsilon^j_{(\lambda)} = \delta^{ij} - \hat{k}^i\hat{k}^j$.

Upon quantisation, the independent amplitudes $a_{(\lambda)}(\mathbf{k})$ become operators, and by substituting the expansions for $A^i(t, \mathbf{k})$ and $E^j(t, \mathbf{x}')$ into the canonical commutation relations, and using the properties of the polarisation vectors, one discovers the commutation relations for $a_{(\lambda)}(\mathbf{k})$ and $a^*_{(\lambda)}(\mathbf{k})$:

$$[a_{(\lambda)}(\mathbf{k}), a^*_{(\lambda')}(\mathbf{k}')] = 2\omega_{\mathbf{k}}\delta_{\lambda\lambda'}(2\pi)^3\delta(\mathbf{k}-\mathbf{k}'). \tag{3.65}$$

Thus each polarisation behaves like an independent scalar field.

As before, we can define a vacuum state $|0\rangle$ satisfying

$$a_{(\lambda)}(\mathbf{k})|0\rangle = 0, \quad \forall \mathbf{k}, \lambda. \tag{3.66}$$

Other states are derived from the vacuum by acting with raising operators. For example, there is a set of "first excited states"

$$|\mathbf{k}, \lambda\rangle = a^*_{(\lambda)}(\mathbf{k})|0\rangle. \tag{3.67}$$

Note that there is an extra label, λ, which indicates the polarisation of the electromagnetic field in the state. The quantised states of the electromagnetic field are of course the particles known as photons.

Last in this section, we write down the Hamiltonian for the electromagnetic field in terms of the ladder operators in its normal-ordered form:

$$:H: = \int d^3x :\mathcal{H}: = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}\sum_\lambda \omega_{\mathbf{k}} a^*_{(\lambda)}(\mathbf{k}) a_{(\lambda)}(\mathbf{k}). \tag{3.68}$$

This is just two copies of the scalar field Hamiltonian ?, one for each polarisation.

4 The Casimir Effect

The Casimir effect is a fairly simple demonstration of one of the unexpected properties of the vacuum quantum field theory: that its properties, in particular its energy, depend on the boundary conditions. Briefly stated, the Casimir effect is the existence of a force between conducting plates in a vacuum, whose origin lies in the nature of the vacuum of the electromagnetic field. The "nothing" between the plates is a different kind of "nothing" to that outside: its energy density is lower, and becomes even lower as the plates approach one another.

The existence of this phenomenon was pointed out by Henryk Casimir in 1948,[2] although amazingly enough conclusive experimental proof was not obtained until very recently by Lamoureux[3].

4.1 Boundary conditions

The secret of the Casimir effect is how the boundary conditions imposed by the conducting plates change the allowed modes of the electromagnetic field between them. We first study the configuration considered by Casimir: two square parallel conducting plates of size L and area $A = L^2$ separated by a distance $a \ll L$ (see Figure 4.1). The exact shape of the plates is not important, as their closeness

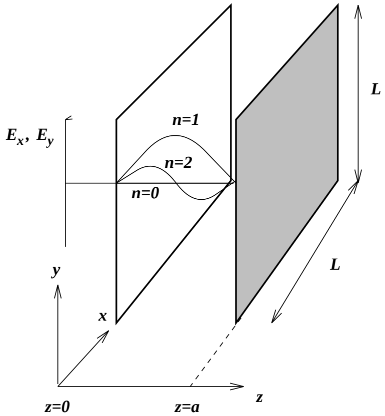

Figure 4.1: The geometry of the Casimir effect. Two conducting plates of size L are placed parallel at a distance a apart, with $L \gg a$. The component of the electric field parallel to the plates are forced to vanish at the plates, which quantises the allowed modes.

enables us to neglect any effects associated with their edges. We suppose that they are perfect conductors, so that the parallel component of the electric field vanishes at the plates. If we choose a Cartesian coordinate system in which the plates are at $z = 0$ and $z = a$, we can therefore say

$$\mathbf{E}_\parallel(x, y, 0, t) = 0 = \mathbf{E}_\parallel(x, y, a, t), \tag{4.1}$$

[2] H.B.G. Casimir, Proc. Kon. Ned. Akad. Wetenschap., ser. B, 51, 793 (1948)
[3] S.K. Lamoureux, Phys. Rev. Lett. 78, 5 (1997)

where $\mathbf{E}_\parallel = \mathbf{E} - (\hat{\mathbf{z}}\cdot\mathbf{E})\hat{\mathbf{z}}$ is the component of the electric field parallel to the plates. As $\mathbf{E} = -\dot{\mathbf{A}}$, This implies that $\mathbf{A}_\parallel(x, y, 0, t)$ and $\mathbf{A}_\parallel(x, y, a, t)$ are (different) constants. We shall see that we can choose both these constants to be zero, provided we exclude the possibility of a constant background magnetic field. The solution to the field equations

$$(\partial_0^2 - \nabla^2)\mathbf{A} = 0$$

subject to these boundary conditions should look like

$$A^i(x, y, z, t) = \frac{1}{a}\left(\sum_{n=0}^{\infty} \alpha_n^i(x, y, t)\sin(k_n z) + \alpha_0^i(x, y, t)\right) + \beta^i(x, y, t)z, \tag{4.2}$$

with $k_n = \pi n/a$. The boundary conditions imply that $\boldsymbol{\alpha}_\parallel$ and $\boldsymbol{\alpha}_\parallel$ are constants. However, if $\boldsymbol{\beta} \neq 0$ there is a component of the magnetic field in the x and y directions at all values of z, equal to $-\beta^2$ and β^1 respectively. This we can exclude by considering only fields which vanish at infinity. The gauge transformation $\mathbf{A} \to \mathbf{A} + \boldsymbol{\nabla}\Lambda$, with $\Lambda = -\boldsymbol{\alpha}_\parallel\cdot\mathbf{x}$, eliminates $\boldsymbol{\alpha}_\parallel$ from consideration. Hence the solution reduces to

$$A^i(x, y, z, t) = \frac{1}{a}\left(\sum_{n=0}^{\infty} \alpha_n^i(x, y, t)\sin(k_n z) + \alpha_{0\perp}^i(x, y, t)\right). \tag{4.3}$$

Only the polarisation perpendicular to the plates survives in the component of the field which is independent of z. This component corresponds to modes which have no momentum in the z direction, that is, modes which travel straight between the plates. The dependence of the functions α_n on x, y and t is fixed by the field equations, which give

$$\begin{aligned} (\partial_0^2 - \partial_1^2 - \partial_2^2 + k_n^2)\alpha_n^i(x, y, t) &= 0, & (4.4)\\ (\partial_0^2 - \partial_1^2 - \partial_2^2)\alpha_{0\perp}^i(x, y, t) &= 0. & (4.5) \end{aligned}$$

The general solutions of these equations are

$$\begin{aligned} \alpha_n^i(x, y, t) &= \int \frac{d^2k}{(2\pi)^2}\frac{1}{2\omega_{\mathbf{k}}}\left(\epsilon_{(\lambda)}^i a_{(\lambda)}^n(k^1, k^2)e^{ik^1x^1 + ik^2x^2 - i\omega_{\mathbf{k}}t} + \text{c.c.}\right), & (4.6)\\ \alpha_{0\perp}^i(x, y, t) &= \int \frac{d^2k}{(2\pi)^2}\frac{1}{2\omega_{\mathbf{k}}}\left(\epsilon_\perp^i a^0(k^1, k^2)e^{ik^1x^1 + ik^2x^2 - i\omega_{\mathbf{k}}t} + \text{c.c}\right), & (4.7) \end{aligned}$$

with $\omega_{\mathbf{k}} = +[(k^1)^2 + (k^2)^2 + (k_n)^2]^{\frac{1}{2}}$, and where $\epsilon_\perp^i$ is the polarisation vector perpendicular to the plates. The notation d^2k in the integration measure denotes an integration over k^1 and k^2 only, and the abbreviation "c.c." stands for "complex conjugate".

4.2 Quantisation

The new boundary conditions also modify the commutation relations of the ladder operators $a^n_{(\lambda)}(k^1,k^2)$. Substituting the expansion (4.3) into the canonical commutation relations, and using the expressions (4.6,4.7) one can derive

$$[a^n_{(\lambda)}(k^1,k^2), a^m_{(\lambda)}(k'^1,k'^2)] = 2\omega_{\mathbf{k}}(2\pi)^2\delta(k^1-k'^1)\delta(k^2-k'^2)a\delta_{nm}. \tag{4.8}$$

The Hamiltonian is

$$H = \frac{1}{a}\int\frac{d^2k}{(2\pi)^2}\frac{1}{2\omega_{\mathbf{k}}}\left[\sum_{n=1}^{\infty}\sum_{\lambda}\frac{1}{2}\hbar\omega_{\mathbf{k}}\left(a^{*n}_{(\lambda)}a^n_{(\lambda)} + a^n_{(\lambda)}a^{*n}_{(\lambda)}\right) + \frac{1}{2}\hbar\omega_{\mathbf{k}}(a^{*0}a^0 + a^0a^{*0})\right], \tag{4.9}$$

where the dependence of the operators $a^n_{(\lambda)}$ and a^0 on the momentum has not been shown in order to keep the length of the expression under control. Planck's constant $\hbar$ has also returned, as we are close to deriving a number which can be compared with experiment.

The vacuum state is defined in the usual way as the state annihilated by all lowering operators:

$$a^n_{(\lambda)}(k^1,k^2)|0\rangle = 0 \quad \forall \quad \lambda, k^1, k^2, n, \tag{4.10}$$
$$a^0(k^1,k^2)|0\rangle = 0 \quad \forall \quad k^1, k^2. \tag{4.11}$$

4.3 The vacuum energy

The energy in the vacuum E_0 is defined as the expectation value of the Hamiltonian in the vacuum state, or

$$E_0 = \langle 0|H|0\rangle. \tag{4.12}$$

We saw for the scalar field in an ordinary vacuum state that this was infinite, but that the energy difference between the vacuum state and others was a meaningful and finite quantity, and obtained by considering the normal-ordered Hamiltonian. We are interested in the energy difference between the vacuum state in the presence of the plates and the vacuum state without the plates, so normal ordering does not help us here: the normal ordered Hamiltonian just gives us energy differences between states with the same boundary conditions. We must adopt a different approach.

Firstly, we derive a formal expression for the vacuum energy, ignoring for the moment the fact that it is infinite. In calculating E_0 we need to find $\langle 0|a^n_{(\lambda)}a^{*n}_{(\lambda)}|0\rangle$ and $\langle 0|a^0a^{*0}|0\rangle$. The commutation relations for the ladder operators can be used to show that

$$\langle 0|a^n_{(\lambda)}(k^1,k^2)a^{*n}_{(\lambda)}(k^1,k^2)|0\rangle = \langle 0|a^0(k^1,k^2)a^{*0}(k^1,k^2)|0\rangle = 2\omega_{\mathbf{k}}(2\pi)^2\delta^2(0)\cdot a\,. \tag{4.13}$$

The δ-functions evaluated at zero are interpreted as factors of L, the size of the plates, because

$$\delta(0) = \lim_{L\to\infty}\int_{-L/2}^{L/2} dx e^{ikx}\Big|_{k=0} = \lim_{L\to\infty} L. \tag{4.14}$$

Hence

$$E_0 = \lim_{L\to\infty}\frac{1}{2}\hbar L^2\int\frac{d^2k}{(2\pi)^2}\left(2\sum_{n=1}^{\infty}\omega_{\mathbf{k}} + \omega_{\mathbf{k}}|_{n=0}\right). \tag{4.15}$$

The presence of the factor of L^2 on the right hand side shows that we really should be discussing the vacuum energy *per unit area* of the plates, we shall denote by σ. Using the fact that $\omega_{\mathbf{k}} = c|\mathbf{k}| = c[(k^1)^2+(k^2)^2+\pi^2n^2/a^2]^{\frac{1}{2}}$, we have

$$\sigma = \frac{1}{2}\hbar c\int\frac{d^2k}{(2\pi)^2}\left(2\sum_{n=1}^{\infty}[(k^1)^2+(k^2)^2+\pi^2n^2/a^2]^{\frac{1}{2}} + [(k^1)^2+(k^2)^2]^{\frac{1}{2}}\right). \tag{4.16}$$

This energy per unit area is to be compared with that in the absence of the plates, for which the Hamiltonian is given by

$$H = \int\frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}\sum_{\lambda}\frac{1}{2}\hbar\omega_{\mathbf{k}}\left(a^*_{(\lambda)}(\mathbf{k})a_{(\lambda)}(\mathbf{k}) + a_{(\lambda)}(\mathbf{k})a^*_{(\lambda)}(\mathbf{k})\right) \tag{4.17}$$

The vacuum energy *density* without plates is

$$\rho_0 = \frac{1}{2}\hbar c\int\frac{d^3k}{(2\pi)^3}2[(k^1)^2+(k^2)^2+(k^3)^2]^{\frac{1}{2}}, \tag{4.18}$$

The energy per unit area in the volume where the plates are should therefore be compared with

$$\sigma_0 = a\rho_0 = \frac{1}{2}\hbar ca\int\frac{d^3k}{(2\pi)^3}2[(k^1)^2+(k^2)^2+(k^3)^2]^{\frac{1}{2}}, \tag{4.19}$$

which is again a formally infinite expression. Thus, in order to compare the two surface energies, we need a way of subtracting two infinite quantities. This is done by rendering the infinite quantities finite by some consistent set of rules, a procedure known as *regularisation.*

4.4 Regularisation

The regularisation procedure we will adopt is to restrict the domain of integrations to a momenta $|\mathbf{k}| < \Lambda$, or $[(k^1)^2+(k^2)^2+\pi^2n^2/a^2]^{\frac{1}{2}} < \Lambda$. We can then subtract σ_0 from σ meaningfully. We then take Λ to infinity, to recover the full range of integration, and we shall find that the result is in fact finite.

Let us use a function $f(|\mathbf{k}|)$ to help us with the regularisation, with the following properties:

$$f(x)=1, \quad \frac{d^n}{dx^n}f(x)\bigg|_{x=0}=0, \quad f(x)=0 \quad \text{for } x \gg \Lambda. \tag{4.20}$$

The essential point of this function is that it rapidly approaches zero for $x > \Lambda$, so that we can multiply it into the integrands with the effect of limiting the range of integration.

We can thus write the regularised surface energies

$$\begin{aligned} \sigma_{\rm reg} &= \frac{1}{2}\hbar c \int \frac{d^2k}{(2\pi)^2}\left(2\sum_{n=1}^{\infty} f([k_\parallel^2+(\pi n/a)^2]^{\frac{1}{2}})[k_\parallel^2+(\pi n/a)^2]^{\frac{1}{2}} + f(k_\parallel)k_\parallel\right), & (4.21)\\ \sigma_{0,\rm reg} &= \frac{1}{2}\hbar c a\int \frac{d^3k}{(2\pi)^3} 2[k_\parallel^2+(k^3)^2]^{\frac{1}{2}} f([k_\parallel^2+(k^3)^2]^{\frac{1}{2}}), & (4.22)\end{aligned}$$

where $k_\parallel^2=(k^1)^2+(k^2)^2$.

We can simplify a little by a couple of changes of the integration variables. Firstly, $u=a^2k_\parallel^2/\pi^2$, from which we infer

$$\frac{d^2k}{(2\pi)^2}=\frac{k_\parallel dk_\parallel}{2\pi}=\frac{\pi}{4a^2}du. \tag{4.23}$$

Secondly, in the expression for $\sigma_{0,\rm reg}$ we replace k^3 by a *continuous* variable $n = k^3a/2\pi$. Thus

$$\begin{aligned} \sigma_{\rm reg} &= \hbar c\frac{\pi^2}{4a^2}\int du\left[\frac{1}{2}u^{\frac{1}{2}}f(\frac{\pi}{a}u^{\frac{1}{2}})+\sum_{n=1}^{\infty}(u+n^2)^{\frac{1}{2}}f(\frac{\pi}{a}(u+n^2)^{\frac{1}{2}})\right], & (4.24)\\ \sigma_{0,\rm reg} &= \hbar c\frac{\pi^2}{4a^2}\int du\int dn(u+n^2)^{\frac{1}{2}}f(\frac{\pi}{a}(u+n^2)^{\frac{1}{2}}). & (4.25)\end{aligned}$$

If we define the following function of n,

$$F(n)=\int du(u+n^2)^{\frac{1}{2}}f(\frac{\pi}{a}(u+n^2)^{\frac{1}{2}}), \tag{4.26}$$

we can write the difference in the two regularised surface energies in the compact form

$$\Delta\sigma=\sigma_{\rm reg}-si_{0,\rm reg}=\hbar c\frac{\pi^2}{4a^2}\left[\frac{1}{2}F(0)+\sum_{n=1}^{\infty}F(n)-\int dn\,F(n)\right]. \tag{4.27}$$

Miraculously, there is a ready-made formula from the theory of complex analysis which we can turn to our purposes, called the Euler-Maclaurin formula, which can be paraphrased as

$$\frac{1}{2}F(0)+\sum_{n=1}^{\infty}F(n)-\int dn\,F(n)=-\frac{1}{2!}B_2F'(0)-\frac{1}{4!}B_4F'''(0)+\ldots, \tag{4.28}$$

where $B_2=1/6$, $B_4=-1/30$, and the ellipsis indicates an expansion in higher derivatives of the function F evaluated at $n=0$. The constants B_2 and B_4 have a wider significance: the are two of the Bernoulli numbers, which appear in many other places in analysis.

In order to evaluate the derivatives of F, let us first shift the integration variable by defining $v=u+n^2$. Hence

$$F(n)=\int_{n^2}^{\infty}dv\sqrt{v}f(\frac{\pi}{a}\sqrt{v}). \tag{4.29}$$

Therefore,

$$\begin{aligned} F'(n) &= -2n^2f(\frac{\pi}{a}n), & (4.30)\\ F''(n) &= -4nf(\frac{\pi}{a}n)-2n^2\frac{a}{\pi}f'(\frac{\pi}{a}n), & (4.31)\\ F'''(n) &= -4f(\frac{\pi}{a}n)-8n\frac{a}{\pi}f'(\frac{\pi}{a}n)-2n^2\frac{a^2}{\pi^2}f''(\frac{\pi}{a}n). & (4.32)\end{aligned}$$

Thus, using the properties of the regularising function f, we find

$$F'(0)=0, \quad F''(0)=0, \quad F'''(0)=-4. \tag{4.33}$$

Higher derivatives of $F(n)$ all involve a sum of derivatives of f, and so vanish at $n=0$. We therefore need not consider any higher order terms, and we arrive at the result

$$\frac{1}{2}F(0)+\sum_{n=1}^{\infty}F(n)-\int dn\,F(n)=\frac{1}{4!}\cdot\frac{1}{30}\cdot(-4)=-\frac{1}{180}. \tag{4.34}$$

This result is rather extraordinary, as it is independent of the regularisation parameter Λ. We may therefore take it to infinity with impunity, and write down the final result for the surface energy of the conducting plates measured relative to the vacuum without plates (which we conventionally call the zero of energy):

$$\Delta\sigma=-\frac{\pi^2}{720}\frac{\hbar c}{a^3}. \tag{4.35}$$

By differentiation we can derive a force per unit area:

$$\text{Force/area}=-\frac{d\Delta\sigma}{da}=-\frac{\pi^2}{240}\frac{\hbar c}{a^4}\simeq -0.016\,a^{-4}\ \text{dyn}(\mu\text{m})^4\text{cm}^{-2}. \tag{4.36}$$

Thus plates 1 cm across separated by a distance of 1 μm experience an attractive force of about 10^{-8} N, which is the weight of 1 μg.

4.5 Lamoreaux's experiment

The problem in trying to experimentally observe the force we have just calculated is that in practice it is very difficult to keep the plates parallel to the required accuracy, which is something like 1 μm over 1 cm. Instead, Lamoreaux used one conducting plate and a spherical lens instead of the other plate, for which one does not have to be especially careful with the orientation. Both were made of quartz coated in copper and gold. Unfortunately, this means doing the calculation all over again for the new geometry. In this case it is quite a lot harder, and no exact solution has been found: instead there is an approximate one for the case where the radius of curvature of the lens is much greater than the separation of the surfaces. The prediction is that the attractive force between the lens and the plate should be

$$F(a) = 2\pi R \left(\frac{1}{3} \frac{\pi^2}{240} \frac{\hbar c}{a^3} \right). \tag{4.37}$$

In Lamoreaux's experiment, the radius of curvature of the lens was 2m, which is very large compared to the separation, which was between 0.6 and 11 μm. The formula (4.37) is therefore a very good approximation.

The results can be found in Lamoreaux's paper: a beautiful agreement to within about 5% is found over the range of separations in the experiment.

A Constants and Conversion Factors

Newton's constant	G	$6.672\,59(85) \times 10^{-11}$	$\mathrm{m^3\,kg^{-1}\,sec^{-2}}$
Speed of light in vacuum	c	$2.99\,792\,458 \times 10^{8}$	$\mathrm{m\,sec^{-1}}$
permittivity of vacuum	ϵ_0	$8.854\,187\,817 \times 10^{-12}$	$\mathrm{F\ m^{-1}}$
permeability of vacuum	μ_0	$4\pi \times 10^{-7}$	$\mathrm{N\ A^{-2}}$
Planck constant	h	$6.626\,075\,5(40) \times 10^{-34}$	J s
Reduced Planck constant	$\hbar = h/2\pi$	$1.054\,572\,66(63) \times 10^{-34}$	J s
elementary charge	e	$1.602\,177\,33(49) \times 10^{-19}$	C
fine-structure constant $e^2/4\pi\epsilon_0\hbar c$	α	7.297×10^{-3}	
inverse fine-structure constant	α^{-1}	137.036	
Boltzmann constant	k_{B}	$1.380\,658(12) \times 10^{-23}$	$\mathrm{J\,K^{-1}}$

Table A.1: Some fundamental constants.

1 eV	=	$1.602\,177\,33(49) \times 10^{-19}$ J
1 J	=	10^{7} erg
1 C	=	10^{-1} e.m.u.
	=	2.998×10^{9} e.s.u.
1 T	=	10^{4} G
1 Å	=	10^{-10} m
1 fm	=	10^{-15} m
1 b	=	10^{-28} $\mathrm{m^2}$
	=	100 $\mathrm{fm^2}$

Table A.2: Some conversion factors.

B Wave function normalisation

In this section we shall see how to normalise eigenfunctions with continuous eigenvalues using the Dirac δ-function. Our model will be the momentum eigenfunctions of a particle moving in 3 dimensions.

Recall that the momentum operator is given by $\hat{\mathbf{p}} = -i\hbar\boldsymbol{\nabla}$, which has eigenfunctions

$$v_{\mathbf{k}}(\mathbf{x}) = N_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{B.1}$$

where $\mathbf{k} = \mathbf{p}/\hbar$ is the wave vector and $N_{\mathbf{k}}$ is a normalisation constant, which in principle might depend on the wave vector. There are two ways of finding $N_{\mathbf{k}}$.

B.1 Box normalisation

We restrict the domain of the wave function to a large cubical box of side L, centred on the origin, and impose periodic boundary conditions. Hence $\mathbf{k}$ takes a resticted set of values

$$\mathbf{k} = \frac{2\pi}{L}\mathbf{n}, \tag{B.2}$$

where $\mathbf{n}$ is a vector of integers, $\mathbf{n} = (n_1, n_2, n_3)$. Consider the inner product of two eigenfunctions with wave vectors $\mathbf{k}$ and $\mathbf{k}'$:

$$\int d^3x\, v^*_{\mathbf{k}'}(\mathbf{x}) v_{\mathbf{k}}(\mathbf{x}) = N^*_{\mathbf{k}'} N_{\mathbf{k}} \left[\int_{-L/2}^{L/2} dx^1 e^{i2\pi(n_1-n_1')x^1/L}\right][\,1 \leftrightarrow 2\,][\,1 \leftrightarrow 3\,]. \tag{B.3}$$

This notation means that the right hand side consists of a product of identical factors, which may be obtained by replacing the label "1" in the first set of square brackets by "2" and "3" in the second and third set respectively. Performing the integrations, we find that they vanish unless $\mathbf{n}' = \mathbf{n}$, in which case we get a factor of L for each integration. Hence

$$\int d^3x v^*_{\mathbf{k}'}(\mathbf{x}) v_{\mathbf{k}}(\mathbf{x}) = N^*_{\mathbf{k}'} N_{\mathbf{k}} L^3 \delta_{n_1 n_1'} \delta_{n_2 n_2'} \delta_{n_3 n_3'}, \tag{B.4}$$

where δ_{mn} is the Kronecker delta, defined by

$$\delta_{mn} = \begin{cases} 1 & \text{if } m = n, \\ 0 & \text{otherwise.} \end{cases} \tag{B.5}$$

Thus with box normalisation, $N_{\mathbf{k}} = L^{-3/2}$ and the momentum eigenfunctions are

$$v_{\mathbf{k}}(\mathbf{x}) = L^{-3/2} e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{B.6}$$

B.2 δ-function normalisation

With box normalisation, one has to be careful about how one takes the limit $L \to \infty$ if one is interested in the infinite volume limit. Clearly the eigenfunctions themselves are badly defined. The way around this problem is to revise the definition of normalisation itself. In infinite volume, the inner product of the two momentum eigenfunctions is

$$\int d^3x v^*_{\mathbf{k}'}(\mathbf{x}) v_{\mathbf{k}}(\mathbf{x}) = N^*_{\mathbf{k}'} N_{\mathbf{k}} \int d^3x e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}}. \tag{B.7}$$

The right hand side may be written as the limit of box normalisation when L is taken to infinity:

$$\begin{aligned} \int d^3x\, v^*_{\mathbf{k}'}(\mathbf{x}) v_{\mathbf{k}}(\mathbf{x}) &= \lim_{L\to\infty} N^*_{\mathbf{k}'} N_{\mathbf{k}} \left[\int_{-L/2}^{L/2} dx^1 e^{i(k_1-k_1')x^1/L}\right][\,1 \leftrightarrow 2\,][\,1 \leftrightarrow 3\,], \\ &= N^*_{\mathbf{k}'} N_{\mathbf{k}} \left[\frac{2\sin L(k_1-k_1')/2}{k_1-k_1'}\right][\,1 \leftrightarrow 2\,][\,1 \leftrightarrow 3\,], \end{aligned}$$

Consider the sequence of functions $\sin(Lk/2)/\pi k$ as $L \to \infty$. They have value $L/2\pi$ at $k = 0$, and oscillate away from the origin with zeroes spaced roughly at intervals of $2\pi/L$, and the amplitude of the peaks decreasing as k^{-1}. One way of defining the δ-function is as the limit

$$\delta(k) = \lim_{L\to\infty} \sin(Lk/2)/\pi k. \tag{B.8}$$

It is possible (but difficult) to show that so defined, the δ-function has the following properties:

(1) $\delta(k) = 0$ if $k \neq 0$.

(2) $\int dk\, \delta(k) f(k) = f(0)$ if the region of integration includes $k = 0$, and $f(k)$ is continuous there.

Returning to Eq. (B.8), we can use the δ-function to write it as

$$\begin{aligned} \int d^3x v^*_{\mathbf{k}'}(\mathbf{x}) v_{\mathbf{k}}(\mathbf{x}) &= N^*_{\mathbf{k}'} N_{\mathbf{k}} (2\pi)^3 \delta(k_1-k_1')\delta(k_2-k_2')\delta(k_3-k_3') \\ &= N^*_{\mathbf{k}'} N_{\mathbf{k}} (2\pi)^3 \delta(\mathbf{k}-\mathbf{k}'). \end{aligned} \tag{B.9}$$

Thus if we choose the normalisation constants to be unity, we obtain the following normalisation condition for the eigenstates:

$$\int d^3x v^*_{\mathbf{k}'}(\mathbf{x}) v_{\mathbf{k}}(\mathbf{x}) = (2\pi)^3 \delta(\mathbf{k}-\mathbf{k}'). \tag{B.10}$$

In many text books you will see the normalisation constants chosen to be $1/(2\pi)^{3/2}$, which removes the factor of $(2\pi)^3$ from the right hand side of (B.10).

C Problem Sheets

C.1 Problem Sheet 1

1.1 (a) Two nucleons collide with equal and opposite momenta. What threshold energy and velocity do they need for the reaction

$$N + N \to N + N + N + \bar{N}$$

to occur? Ignore the mass difference between protons and neutrons, and take the nucleon mass to be $m_N = 940$ MeV.

(b) Find the threshold energy in a fixed target experiment, where one of the nucleons is at rest.

(c) The Cosmotron, built in 1952, was an accelerator at Brookhaven, which could accelerate protons to a total energy of 3 GeV. Did it discover the antiproton?

1.2 (a) Show that

$$\partial_\mu x^\nu = \delta_\mu^\nu,$$

and hence that

$$\partial_\mu x_\nu = \eta_{\mu\nu}, \quad \partial^\mu x_\nu = \delta^\mu_\nu, \quad \partial^\mu x^\nu = \eta^{\mu\nu}.$$

(b) Calculate $i\partial^\mu e^{-ik\cdot x}$, where k^μ is a constant 4-vector.

(c) Calculate

$$\partial_\mu \partial_\nu e^{-\frac{1}{2}x^2}.$$

(Recall that for 4-vectors, $x^2 = x \cdot x$.)

1.3 (a) Newton's constant G has the value 6.67×10^{-11} m^{3}kg^{-1}s^{-2}. Express G in natural units, using the relations 1 GeV $= 1.6 \times 10^{-10}$ Joule, $\hbar = 6.6 \times 10^{-25}$ GeV s, and $\hbar c = 2.0 \times 10^{-16}$ GeV m.

(b) The average density of the Universe today ρ_0 is thought to be roughly 10^{-26} kg m^{-3}. Using this order-of-magnitude estimate, express ρ_0 in natural units, and then form a dimensionless combination of ρ_0 and a power of G. Give its value. (The smallness of this number if the source of a major cosmological puzzle).

1.4 A harmonic oscillator with unit mass has Hamiltonian operator

$$H = \tfrac{1}{2}p^2 + \tfrac{1}{2}\omega^2 x^2,$$

where x is the displacement operator and p its conjugate momentum, satisfying

$$[\hat{x}, \hat{p}] = i.$$

The ladder operators a and a^* are defined as

$$a = \frac{\omega\hat{x} + i\hat{p}}{\sqrt{2\omega}}, \qquad a^* = \frac{\omega\hat{x} - i\hat{p}}{\sqrt{2\omega}}.$$

(a) Verify the commutation relation for a and a^*,

$$[a, a^*] = 1.$$

(b) Show that H may be written in terms of the ladder operators as

$$H = \tfrac{1}{2}\omega(aa^* + a^*a).$$

(c) Show that

$$[H, a] = -\omega a, \qquad [H, a^*] = \omega a^*.$$

Hence show that, if $|\alpha\rangle$ is an eigenstate of the Hamiltonian with energy E_α, then $a|E_\alpha\rangle$ has energy $E_\alpha - \omega$, and $a^*|E_\alpha\rangle$ has energy $E_\alpha + \omega$.

(d) We postulate the existence of a lowest energy or ground state $|0\rangle$, which satisfies $a|0\rangle = 0$. What is the energy of the ground state?

1.5 Prove the following identities for the δ-function by multiplying by an arbitrary "good" function $f(x)$ and integrating.

(a) $\delta(ax) = |a|^{-1}\delta(x)$ ($a \neq 0$, and real).

(b) $\dfrac{d}{dx}\theta(x) = \delta(x)$, where $\theta(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0. \end{cases}$

(c) If $g(x) = 0$ has n simple roots $x = a_i$, then

$$\delta(g(x)) = \sum_{i=1}^{n} |g'(a_i)|^{-1}\delta(x - a_i).$$

[Hint: split the range of integration into segments around each root a_i.]

C.2 Problem Sheet 2

2.1 We define the dual electromagnetic field strength tensor as

$$\tilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}.$$

Write down its components in matrix form. Hence verify that the 4-vector equation

$$\partial_\nu \tilde{F}^{\mu\nu} = 0$$

correctly reproduces the two homogeneous Maxwell equations.

(*Something to think about – no marks for this:* What would the physical significance be if the right hand side of this equation were not zero, but equal to a "current" k_μ?)

2.2 A general Lorentz transformation on 4-vectors and tensors is represented by the matrix $\Lambda^\mu_{\ \nu}$.

(a) What property must $\Lambda^\mu_{\ \nu}$ satisfy in order for it to be a Lorentz transformation?

(b) Write down the Lorentz transformation laws for the partial derivative ∂_μ, the 4-current $j^\mu(x)$, and the field strength tensor $F^{\mu\nu}(x)$.

(c) Show that the relativistic for of the inhomogeneous Maxwell equations

$$\partial_\mu F^{\mu\nu}(x) = j^\nu(x)$$

keep their form under a Lorentz transformation, i.e. that the transformed quantities satisfy

$$\partial'_\mu F'^{\mu\nu}(x') = j'^\nu(x').$$

(We say in this case that the equation is Lorentz covariant.)

2.3 Let $\phi(x)$ be a complex-valued field satisfying the Klein-Gordon equation

$$\partial^2\phi + m^2\phi = 0.$$

We define a 4-vector current

$$J^\mu = -\frac{i}{2}\phi^*\partial^\mu\phi + \frac{i}{2}\phi\partial^\mu\phi^*.$$

(a) Show that J^μ is conserved, i.e. that $\partial \cdot J = 0$. (You will need to use the equations of motion.)

(b) Suppose $\phi(x) = Ae^{ik\cdot x}$, with A a complex constant. What condition must k satisfy in order for $\phi(x)$ to be a solution? And what is the current J^μ when the field is in this state?

(2.4) Let $\phi(x)$ be a real scalar field, with Lagrangian density

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

where the potential energy density is given by

$$V(\phi) = \frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4.$$

Note that λ and μ^2 are both real, and $\lambda > 0$.

(a) Sketch the function $V(\phi)$, both for $\mu^2 > 0$ and $\mu^2 < 0$ (i.e. imaginary μ).

(b) Derive the equations obeyed by ϕ, starting from the Euler-Lagrange equations.

(c) Show that $\phi = 0$ is a solution, for both positive and negative μ^2.

(d) Consider values of $\phi(x)$ for which $|\phi(x)| \ll |\mu^2/\lambda|^{\frac{1}{2}}$. What is the linear equation which ϕ approximately satisfies for these "small" values?

(2.5) Show that the Lagrangian density

$$\mathcal{L} = \frac{1}{2}\dot{A}^j\dot{A}^j - \frac{1}{2}\partial_j A^i\partial_j A^i.$$

gives the correct field equations for the electromagnetic field in the Coulomb gauge, in the absence of charges or currents. The canonical momentum conjugate to A^j is

$$\pi_j = \frac{\partial\mathcal{L}}{\partial\dot{A}^j}.$$

(a) To what physical quantity does π_j correspond?

(b) Derive the Hamiltonian density $\mathcal{H}$, and show that it may be written

$$\mathcal{H} = \frac{1}{2}\mathbf{E}^2 + \frac{1}{2}\mathbf{B}^2.$$

C.3 Problem Sheet 3

In Questions 3.3 and 3.4, you will need the relation

$$\int_{-\infty}^{\infty} dx e^{i(k+k')x} = (2\pi)\delta(k+k')$$

and its 3-dimensional generalisation

$$\int d^3x\ e^{i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x}} = (2\pi)^3\delta^3(\mathbf{k}+\mathbf{k}'),$$

where $\delta^3(\mathbf{k}+\mathbf{k}') = \delta(k^1+k'^1)\delta(k^2+k'^2)\delta(k^3+k'^3)$.

3.1 This question is designed to show that one can give meaning to the differential of a δ-function. Suppose $f(x)$ is continuous and differentiable at $x = 0$.

(a) Evaluate

$$\int_{-\infty}^{\infty} dx\ f(x)\frac{d}{dx}\delta(x)$$

through integration by parts.

(b) Evaluate also the 3-dimensional version

$$\int d^3x\ f(\mathbf{x})\boldsymbol{\nabla}\delta^3(\mathbf{x}),$$

where $f(\mathbf{x})$ is again continuous and differentiable at $\mathbf{x} = 0$, and the integration is taken over all $\mathbf{x}$. (Reminder: $\delta^3(\mathbf{x}) = \delta(x^1)\delta(x^2)\delta(x^3)$.)

3.2 (a) Write down the Lorentz transformation which takes the 4-vector $k^\mu = (\omega_{\mathbf{k}}, k^i)$ to a frame moving with velocity v in the x direction relative to the original. Write down also the transformation law for the differentials $d\omega_{\mathbf{k}}$ and dk^i.

(b) Given that $\omega_{\mathbf{k}}^2 = \mathbf{k}^2 + m^2$, with m^2 constant, find the relation between dk^1 and $d\omega_{\mathbf{k}}$ at constant k^2 and k^3.

(c) Hence show that the momentum integration measure $\frac{d^3k}{2\omega_{\mathbf{k}}}$ is Lorentz invariant, by applying the Lorentz transformation of Part (a).

3.3 The real scalar field operator is expanded in terms of ladder operators $a(\mathbf{k})$ and $a^*(\mathbf{k})$ as

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}\left(a(\mathbf{k})e^{-ik\cdot x} + a^*(\mathbf{k})e^{ik\cdot x}\right),$$

with $\omega_{\mathbf{k}} = +(\mathbf{k}^2 + m^2)^{\frac{1}{2}}$.

(a) We define the Fourier transform of the field operator as $\phi(\mathbf{k}) = \int d^3x\phi(x)e^{-i\mathbf{k}\cdot\mathbf{x}}$. Show that

$$\phi(\mathbf{k}) = \frac{1}{2\omega_{\mathbf{k}}}\left(a(\mathbf{k})e^{-i\omega_{\mathbf{k}}t} + a^*(-\mathbf{k})e^{i\omega_{\mathbf{k}}t}\right)$$

(b) Find the analogous expression for the conjugate momentum $\pi(\mathbf{k})$, assuming that $\pi(\mathbf{x}) = \dot{\phi}(\mathbf{x})$.

(c) Hence show that

$$a(\mathbf{k}) = \omega_{\mathbf{k}}\int d^3x\ \phi(x)e^{ik\cdot x} + i\int d^3x\pi(x)e^{ik\cdot x}.$$

3.4 From Question 3.3, we can infer (by taking the complex conjugate, and renaming variables) that

$$a^*(\mathbf{k}') = \omega_{\mathbf{k}}\int d^3x'\ \phi(x')e^{-ik'\cdot x'} - i\int d^3x'\pi(x')e^{-ik'\cdot x'}.$$

(a) Write down the equal time canonical commutation relations of ϕ and π.

(b) By taking $t = t'$, show from the expressions for $a(\mathbf{k})$ and $a^*(\mathbf{k}')$ above that

$$[a(\mathbf{k}), a^*(\mathbf{k}')] = 2\omega_{\mathbf{k}}(2\pi)^3\delta^3(\mathbf{k} - \mathbf{k}').$$

C.4 Problem Sheet 4

4.1 (a) Write down the Hamiltonian for a free real scalar field $\phi(x)$, giving the expression for the conjugate momentum $\pi(x)$ in terms of $\phi(x)$.

Let the expansion for the field in terms of the ladder operators $a_{\mathbf{k}}$ and $a_{\mathbf{k}}^*$ be

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}\left(a_{\mathbf{k}}e^{-ik\cdot x} + a_{\mathbf{k}}^*e^{ik\cdot x}\right).$$

(I have written $a_{\mathbf{k}}$ for $a(\mathbf{k})$ to save space.)

(b) Show that

$$\int d^3x\ m^2\phi^2(x) =$$
$$\int \frac{d^3k}{(2\pi)^3}\frac{1}{(2\omega_{\mathbf{k}})^2}m^2\left(a_{\mathbf{k}}^*a_{\mathbf{k}} + a_{\mathbf{k}}a_{\mathbf{k}}^* + a_{\mathbf{k}}a_{-\mathbf{k}}e^{-2i\omega_{\mathbf{k}}t} + a_{\mathbf{k}}^*a_{-\mathbf{k}}^*e^{2i\omega_{\mathbf{k}}t}\right).$$

(Note that $a_{-\mathbf{k}} \equiv a(-\mathbf{k})$.)

(c) Show also that

$$\boldsymbol{\nabla}\phi(x) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}(i\mathbf{k})\left(a_{\mathbf{k}}e^{-ik\cdot x} - a_{\mathbf{k}}^*e^{ik\cdot x}\right),$$

and hence that

$$\int d^3x\ (\boldsymbol{\nabla}\phi(x))^2 =$$
$$\int \frac{d^3k}{(2\pi)^3}\frac{1}{(2\omega_{\mathbf{k}})^2}\mathbf{k}^2\left(a_{\mathbf{k}}^*a_{\mathbf{k}} + a_{\mathbf{k}}a_{\mathbf{k}}^* - a_{\mathbf{k}}a_{-\mathbf{k}}e^{-2i\omega_{\mathbf{k}}t} - a_{\mathbf{k}}^*a_{-\mathbf{k}}^*e^{2i\omega_{\mathbf{k}}t}\right).$$

(d) Finally, calculate $\int d^3x\ \dot{\phi}^2(x)$, and using your result, and the results of parts (b) and (c), show that

$$H = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}\frac{\omega_{\mathbf{k}}}{2}\left[a_{\mathbf{k}}^*a_{\mathbf{k}} + a_{\mathbf{k}}a_{\mathbf{k}}^*\right].$$

For parts (b), (c) and (d) you will need the relation:

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