Relativistic Quantum Fields 1

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

Autumn Term 2002
Tuesday 11:30am Arundel 1C
Friday 9:15pm Arundel 1C

PostScript and PDF versions of these notes available:

http://www.pact.cpes.susx.ac.uk/users/markh/RQF1/rqf1.ps
http://www.pact.cpes.susx.ac.uk/users/markh/RQF1/rqf1.pdf

Contents

0 Course information

Aim: to introduce 1st year physics graduate students to the theory of quantum fields.

Topics: canonical quantisation of scalar and vector fields; the path integral formalisms for scalar fields; Gaussian path integrals; the diagrammatic representation of Green's functions in real scalar field theory; to understand how field theory can be used to calculate physical processes including scattering.

Syllabus

- *Relativistic wave equations* (Ryder Ch 2, G & R Ch 4.1, 6.2). Relativistic notation; Klein-Gordon equation. Maxwell equations.
- *Lagrangian formulation* (Ryder Ch 3, G & R Ch 2). Lagragian particle mechanics; Euler-Lagrange for real scalar field; Noether's theorem for scalar fields. Electromagnetic field.
- *Canonical quantisation* (Ryder Ch 4, G & R Ch 4/7). Canonical commutation relations for real scalar field; Fock space; Number, energy and momentum operators. EM field in Lorentz gauge.
- *Path integrals* (Ryder Ch 5/6, G & R Ch 11/12). Path integrals in quantum mechanics; Functionals; Path integral quantisation of scalar field; Gaussian integration; Free particle Green's functions ; Vacuum-vacuum transition function Z[J].
- *Interacting fields* (Ryder Ch 5/6, G & R Ch 8). S-matrix and transition amplitudes; Time evolution operator; Wick's theorem; Matrix elements.

Teaching methods

There will be 2 lectures a week throughout the term. Problem sheets will be given every two weeks.

Assessment

MSc students Problem sheets will count for 40% of the total mark for the course, the remaining 60% will come from the 1.5 hour exam at the end of the year, which will have a choice of two questions from three.

DPhil students Problem sheets will count for 100% of the total mark.

Reading list

* *Quantum Field Theory*, L. Ryder (C.U.P., Cambridge, 1984).
* *Field Quantization*, W. Greiner and J. Reinhardt (Springer, Berlin, 1996),

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

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

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

  Fields, W. Siegel (`http://insti.physics.sunysb.edu/~siegel/plan.html`).

There are quite a few errors in Peskin and Schroeder which are supposed to be corrected in the next printing of the book. Meanwhile, a list can be found at
`http://www.slac.stanford.edu/~mpeskin/QFT.html`
The other books are complementary in some way. Kaku is very complete but rather rushed. Mandl and Shaw is pedagogical but adopts different Fourier transform and normalisation conventions which can confuse the unwary. Siegel's PDF lecture notes are the most complete exposition I have seen. They are very advanced and will cost you a lot to print at over 800 pages.

Prerequisites

The course assumes that you have already encountered

- 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).

It will help to have come across the following subjects, mostly covered in the Sussex undergraduate course Theoretical Physics II, as only a brief review is given here.

- Lagrangian formulation of classical mechanics (e.g. *Classical mechanics*, T.W.B. Kibble and F.H. Berkshire, London: Longman, 1996).
- 4-vector notation in Special Relativity (e.g. Chapter 2 in *General relativity and cosmology*, T.L. Chow, Winnipeg: Wuerz, 1994).

- Covariant formulation of Maxwell's equations.
- Natural Units (Mandl and Shaw, Section 6.1).

Course Lecturer

Mark Hindmarsh, CPES, Arundel 213.
Telephone: 8934
E-mail: m.b.hindmarsh@sussex.ac.uk
Course web: http://www.pact.cpes.susx.ac.uk/~markh/RQF1/
Office hour: Tuesday 2–3pm, or by arrangement (email is best).

1 Preliminaries

2 Relativistic wave equations

2.1 Special Relativity in relativistic notation

We start by quoting a formulation of the Principle of Special Relativity:

> 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.

2.1.1 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*

$$\begin{array}{rcl} t \to t' &=& \gamma\left(t - vx/c^2\right), \\ x \to x' &=& \gamma\left(x - vt\right), \\ y \to y' &=& y, \\ z \to z' &=& z, \end{array} \tag{2.1}$$

where

$$\gamma = 1/\sqrt{1 - v^2/c^2}. \tag{2.2}$$

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

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2. \tag{2.3}$$

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

2.1.2 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

$$x^i = (x^1, x^2, x^3) \equiv (x, y, z), \tag{2.4}$$

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

$$x^\mu = (x^0, x^1, x^2, x^3) \equiv (ct, x, y, z). \tag{2.5}$$

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. $\mathbf{x}$.

2.1.3 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 (2.1). This linear transformation can be represented by matrices Λ^μ_ν acting linearly on the coordinates:

$$x^\mu \to x'^\mu = \sum_{\nu=0}^{3} \Lambda^\mu{}_\nu x^\nu, \tag{2.6}$$

or in differential form

$$dx^\mu \to dx'^\mu = \sum_{\nu=0}^{3} \Lambda^\mu{}_\nu dx^\nu, \tag{2.7}$$

The matrix Λ has entries

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -\gamma v/c & \cdot & \cdot \\ -\gamma v/c & \gamma & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix}, \tag{2.8}$$

where μ labels the rows and ν the columns.

In order to save writing a large number of summation signs, we often use Einstein's convention that repeated indices in an expression are summed over automatically without the need for a summation sign. In this convention, Equation (2.7 becomes

$$dx^\mu \to dx'^\mu = \Lambda^\mu{}_\nu dx^\nu, \tag{2.9}$$

2.1.4 Space-time metric

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

$$ds^2 = \sum_{\mu,\nu=0}^{3} g_{\mu\nu} dx^\mu dx^\nu. \tag{2.10}$$

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

$$g_{\mu\nu} = \eta_{\mu\nu} \equiv \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & -1 & \cdot & \cdot \\ \cdot & \cdot & -1 & \cdot \\ \cdot & \cdot & \cdot & -1 \end{pmatrix}. \tag{2.11}$$

We often call $\eta_{\mu\nu}$ the *Minkowski metric*. In Einstein's repeated index convemtion, (2.10) can be written as

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu. \tag{2.12}$$

Note that the position of the indices is important. There is a difference between vectors with superscript indices and those with subscript indices, which shows up in their properties under Lorentz transformations. We may use the metric tensor to define an infinitesimal coordinate with a lower index:

$$dx_\mu = \eta_{\mu\nu} dx^\nu \tag{2.13}$$

Vectors with subscript indices are termed *covariant*, to distinguish them from their *contravariant* partners with superscript indices. They transform oppositely to contravariant 4-vectors, with the inverse of the Lorentz transformation matrix Λ:

$$dx_\mu \to dx'_\mu = (\Lambda^{-1})^\nu{}_\mu dx_\nu. \tag{2.14}$$

One can check that this is true by noting that the space-time interval can be written $ds^2 = dx_\mu dx^\mu$, and then substituting the transformation laws (2.7) and (2.13). One finds

$$ds^2 \to ds'^2 = dx'_\mu dx'^\mu = (\Lambda^{-1})^\nu{}_\mu dx_\nu \Lambda^\mu{}_\rho dx^\rho. \tag{2.15}$$

Note that we have not just blindly substituted (2.7): we have the repeated index from a ν to a ρ. The meaning is still the same: that index is to be summed over the values 0,1,2,3. However, if we have left the index as ν equation (2.15) would be ambiguous, as we would not know how to pair off indices in the summations. This is an important rule with index notation: *never use repeated indices twice on the same side of an equation.*

Continuing with equation (2.15), we note that the μ indices are paired and can be summed over. We can see that we are multiplying a matrix Λ by its inverse Λ^{-1} and therefore must obtain the identity, which expressed in index notation is

$$(\Lambda^{-1})^\nu{}_\mu \Lambda^\mu{}_\rho = \delta^\nu_\rho. \tag{2.16}$$

Here we introduce the *Kronecker delta*, defined by

$$\delta^\nu_\mu = \begin{cases} 1, & \mu = \nu, \\ 0, & \mu \neq \nu. \end{cases} \tag{2.17}$$

Note that 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):

$$x_0 = x^0, \quad x_i = -x^i. \tag{2.18}$$

Lastly, one can also define the metric tensor with raised indices as the matrix inverse of the covariant metric tensor

$$\eta_{\mu\nu}\eta^{\nu\rho} = \delta^\rho_\mu, \tag{2.19}$$

2.1.5 General Lorentz transformations: the Lorentz group

One can explicitly verify that the transformation law (2.1) leaves the space-time interval ds invariant. By choosing space coordinates so that the relative velocity of two inertial frames is along the x direction, it follows that all Lorentz transformations leave the interval invariant. Let us see what we can infer about the matrices Λ from this condition. The interval transforms as

$$ds^2 \to ds'^2 = \eta_{\mu\nu}\Lambda^\mu_\rho\Lambda^\nu_\sigma dx^\rho dx^\sigma = ds^2. \tag{2.20}$$

But ds^2 may be written $ds^2 = \eta_{\rho\sigma}dx^\rho dx^\sigma$: thus we may infer that

$$\eta_{\mu\nu}\Lambda^\mu_\rho\Lambda^\nu_\sigma = \eta_{\rho\sigma}. \tag{2.21}$$

Hence, any matrix Λ which leaves the metric $\eta_{\mu\nu}$ invariant under the transformation (2.21) represents a Lorentz transformation. These matrices form a group of transformations known as the *Lorentz group*. When combined with translation symmetry, $x^\mu \to x'^\mu = x^\mu + a^\mu$, with a^μ a constant 4-vector, it forms a larger group known as the Poincaré group.[1]

2.1.6 Derivatives

It is through studying derivatives that we discover the utility of the idea of covariant 4-vectors. Coordinate vectors are naturally contravariant, and it seems slightly perverse to introduce a covariant version with lowered indices. However, it turns out that there are 4-vectors which are naturally covariant, and the principal among these is the partial derivative 4-vector.

[1] A more detailed exploration of the properties of the Poincaré group, which is important for understanding the concepts of mass and spin in quantum field theory, can be found in Ryder, Chapter 2, or Peskin and Schroeder, Chapter 3.

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

$$\partial_i = \frac{\partial}{\partial x^i} = \left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3} \right). \tag{2.22}$$

The ∂_i are the components of the gradient operator $\boldsymbol{\nabla}$ in an orthonomal basis $\boldsymbol{e}^i$, or

$$\boldsymbol{\nabla} = \boldsymbol{e}^i \partial_i. \tag{2.23}$$

The relativistic generalisation is a covariant 4-vector ∂_μ, which is defined as

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x^i} \right). \tag{2.24}$$

The transformation law for ∂_μ may be found from the chain rule:

$$\frac{\partial}{\partial x^\mu} \rightarrow \frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu} \tag{2.25}$$

However, by differentiating

$$x^\nu = (\Lambda^{-1})^\nu{}_\mu x'^\mu, \tag{2.26}$$

which follows from equation (2.6), we find that

$$\frac{\partial}{\partial x'^\mu} = (\Lambda^{-1})^\nu{}_\mu \frac{\partial}{\partial x^\nu}. \tag{2.27}$$

Hence the 4-vector differential really is a covariant 4-vector.

2.1.7 4-vectors and the scalar product

To recap, a *contravariant* 4-vector can be defined as any collection of four quantities which transform the same way as dx^μ under a Lorentz transformation. Similarly, a *covariant* 4-vector can be defined as a collection of four quantities which transform the same way as ∂_μ. One can map any contravariant vector into a covariant one with the metric $\eta_{\mu\nu}$, an operation called lowering the indices, or one can go in the opposite direction with the inverse metric $\eta^{\mu\nu}$ (raising the indices).

There are many important quantities than can be assembled into 4-vectors: for example, energy E and momentum $\mathbf{p}$ of a particle belong together in a single momentum 4-vector p^μ, defined by

$$p^\mu = (E/c, \mathbf{p}). \tag{2.28}$$

We are used to the interval $ds^2 = dx_\mu dx^\mu$ being a Lorentz invariant quantity, but it is straightforward to see that one can form a Lorentz invariant quantity from any pair of 4-vectors a^μ, b^μ. This is the *scalar product*, written

$$a \cdot b \equiv p^2 = a^\mu b^\nu \eta_{\mu\nu} = a_\nu b^\nu = a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3. \tag{2.29}$$

We can take the scalar product of the momentum 4-vector with itself and recover the relativistic relation between energy, momentum and mass:

$$p^2 = E^2/c^2 - \mathbf{p}^2c^2 = m^2c^2, \tag{2.30}$$

Another invariant can be constructed by contracting the momentum 4-vector with the position 4-vector

$$p \cdot x = Et - \mathbf{p}\cdot\mathbf{x}. \tag{2.31}$$

Another important object is the velocity 4-vector. In Newtonian mechanics, the velocity is dx^i/dt, so it might be thought that dx^μ/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

$$v^\mu = \frac{dx^\mu}{d\tau}. \tag{2.32}$$

2.1.8 The d'Alembertian

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 $\Box$:

$$\Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2. \tag{2.33}$$

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.

2.1.9 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, $\mathbf{F} = d\mathbf{p}/dt$. 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

$$F^\mu = \frac{dp^\mu}{d\tau}. \tag{2.34}$$

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.

2.2 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,

$$i\hbar\frac{\partial}{\partial t}\psi = -\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2\psi. \tag{2.35}$$

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:

$$E \leftrightarrow i\hbar\frac{\partial}{\partial t}, \qquad \mathbf{p} \leftrightarrow -i\hbar\boldsymbol{\nabla}, \tag{2.36}$$

(where the double-headed arrow symbol means "is represented by"). Hence the Schrödinger equation (2.35) 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 Schrödinger equation ought to be

$$-\frac{\partial^2}{\partial t^2}\phi = -\boldsymbol{\nabla}^2\phi + m^2\phi, \tag{2.37}$$

or, in a manifestly Lorentz covariant form,

$$(\partial^2 + m^2)\phi = 0. \tag{2.38}$$

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

The solutions to this equation are

$$\phi(x) = Ae^{-iEt+\mathbf{p}\cdot\mathbf{x}} \tag{2.39}$$

with $E = \pm = \surd(\mathbf{p}^2 + m^2)$, and A an arbitrary complex constant.

In fact if ϕ is a solution to the Klein-Gordon equation it *cannot* be interpreted as a wave function as Schrödinger 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* $\mathbf{j} = -i(\psi^* \boldsymbol{\nabla}\psi - \boldsymbol{\nabla}\psi^*\psi)/2m$, in the sense that together they make up a probability conservation equation

$$\dot{\rho} + \boldsymbol{\nabla}\cdot\mathbf{j} = 0, \tag{2.40}$$

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

$$\rho = -\frac{i}{2}(\phi^*\partial_t\phi - \partial_t\phi^*\phi), \qquad \mathbf{j} = -\frac{i}{2}(\phi^*\boldsymbol{\nabla}\phi - \boldsymbol{\nabla}\phi^*\phi). \tag{2.41}$$

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. This was reasonable for his purposes, but the Klein-Gordon equation makes a comeback later, when we shall see that ρ can be interpreted as a charge density, which is allowed to take both positive and negative values.

2.3 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:

$$\begin{array}{cc} \textbf{Homogeneous} & \textbf{Inhomogeneous} \\ \boldsymbol{\nabla}\cdot\mathbf{B} = 0 & \boldsymbol{\nabla}\cdot\mathbf{E} = \rho \\ \dfrac{\partial}{\partial t}\mathbf{B} + \boldsymbol{\nabla}\wedge\mathbf{E} = 0 & -\dfrac{\partial}{\partial t}\mathbf{E} + \boldsymbol{\nabla}\wedge\mathbf{B} = \mathbf{j} \end{array} \tag{2.42}$$

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 measurable 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:

$$\begin{array}{cc} \textbf{Homogeneous} & \textbf{Inhomogeneous} \\ \text{potentials } \phi, \mathbf{A} & \text{current conservation} \\ \begin{aligned} \mathbf{B} &= \boldsymbol{\nabla}\wedge\mathbf{A} \\ \mathbf{E} &= -\dot{\mathbf{A}} - \boldsymbol{\nabla}\phi \\ \phi &\to \phi - \dot{\Lambda} \\ \mathbf{A} &\to \mathbf{A} + \boldsymbol{\nabla}\Lambda \end{aligned} & \dot{\rho} + \boldsymbol{\nabla}\cdot\mathbf{j} = 0 \end{array} \tag{2.43}$$

All these quantities can be assembled into explicitly Lorentz covariant objects. 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 *anti-symmetric tensor*, the field strength tensor $F^{\mu\nu}$, which is defined in terms of the gauge potential 4-vector A^μ:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{2.44}$$

The Lorentz transformation law for $F^{\mu\nu}$ follows from those of ∂^μ and A^μ:

$$F^{\mu\nu} \to F'^{\mu\nu} = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma F^{\rho\sigma}, \tag{2.45}$$

from which we explicitly see that $F^{\mu\nu}$ is a covariant tensor of rank 2. Let us examine the components where $\nu = 0$ and μ ranges over spatial indices i:

$$F^{i0} = \partial^i A^0 - \partial^0 A^i = \partial^i A^0 - \dot{A}^i = -\partial_i \phi - \dot{A}^i = E^i, \tag{2.46}$$

where we have used the expression for the electric field in terms of the gauge potentials in equation (2.43). The magnetic field is contained in the entries where both μ and ν take spatial values:

$$F^{ij} = \partial^i A^j - \partial^j A^i = -\partial_i A^j + \partial_j A^i = -\epsilon_{ijk} B^k, \tag{2.47}$$

where we have introduced the *Levi-Civita* symbol ϵ_{ijk}. The Levi-Civita symbol is defined by

$$\epsilon_{ijk} = \begin{cases} +1 & \text{if } i \neq j \neq k \text{ cyclic,} \\ -1 & \text{if } i \neq j \neq k \text{ anticyclic,} \\ 0 & \text{otherwise.} \end{cases} \tag{2.48}$$

Written in matrix form, the field strength tensor is:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}. \tag{2.49}$$

The reverse relations may be written

$$E^i = F^{i0}, \qquad B^i = -\frac{1}{2}\epsilon_{ijk} F^{jk}, \tag{2.50}$$

which can be verified with the help of the identity

$$\epsilon_{ijk}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}. \tag{2.51}$$

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

$$\partial_\mu j^\mu = 0. \tag{2.52}$$

Maxwell's equations become:

$$\begin{array}{cc} \textbf{Homogeneous} & \textbf{Inhomogeneous} \\ \partial^\lambda F^{\mu\nu} + \partial^\mu F^{\nu\lambda} + \partial^\nu F^{\lambda\mu} = 0, & \partial_\mu F^{\mu\nu} = j^\nu. \end{array} \tag{2.53}$$

The first of these expressions looks 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 ${}^4C_3$, which is equal to 4, precisely the number of equations we started with.

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

$$\epsilon_{\mu\nu\rho\sigma} = \begin{cases} +1 & \text{if } \mu \neq \nu \neq \rho \neq \sigma, \text{ symmetric,} \\ -1 & \text{if } \mu \neq \nu \neq \rho \neq \sigma, \text{ antisymmetric,} \\ 0 & \text{otherwise.} \end{cases} \tag{2.54}$$

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 raised:

$$\epsilon^{\mu\nu\rho\sigma} = \eta^{\mu\alpha}\eta^{\nu\beta}\eta^{\rho\gamma}\eta^{\sigma\delta}\epsilon_{\alpha\beta\gamma\delta}, \tag{2.55}$$

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

$$\epsilon_{\mu\nu\rho\sigma}\partial^\rho F^{\mu\nu} = 0. \tag{2.56}$$

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

2.3.1 Gauges

Recall that the gauge potential A^μ is specified only up to an arbitrary function of space and time Λ (see equation 2.43). The 4-vector version of a gauge transformation is written

$$A^\mu \to A^\mu - \partial^\mu \Lambda, \tag{2.57}$$

under which one can check that $F^{\mu\nu}$ is unchanged. Indeed, using (2.44) one finds

$$F^{\mu\nu} \to \partial^\mu (A^\nu - \partial^\nu \Lambda) - \partial^\nu (A^\mu - \partial^\mu \Lambda) = F^{\mu\nu} - (\partial^\mu \partial^\nu - \partial^\nu \partial^\mu)\Lambda = F^{\mu\nu}. \tag{2.58}$$

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

$$\boldsymbol{\nabla}\cdot\mathbf{A} = 0, \tag{2.59}$$

which implies that the first of the inhomogeneous equations in (2.42) becomes

$$\boldsymbol{\nabla}^2 \phi = -\rho. \tag{2.60}$$

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,

$$-\frac{\partial}{\partial t}(-\dot{\mathbf{A}} - \boldsymbol{\nabla}\phi) - \boldsymbol{\nabla}^2 \mathbf{A} = \mathbf{j}, \tag{2.61}$$

which we may write as

$$(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2)\mathbf{A} = \mathbf{j}_T, \tag{2.62}$$

where $\mathbf{j}_T = \mathbf{j} + \boldsymbol{\nabla}\dot{\phi}$. One can show, using the equation of current conservation and Eq. (2.60), that

$$\boldsymbol{\nabla}\cdot\mathbf{j}_T = 0. \tag{2.63}$$

This of course is necessary for the consistency of Eq. (2.62), 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

$$(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2)\mathbf{A} = 0. \tag{2.64}$$

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

$$\mathbf{A}(t,\mathbf{x}) = \mathbf{a}e^{-i\omega t + i\mathbf{k}\cdot\mathbf{x}} \tag{2.65}$$

where $\mathbf{a}$ is a constant 3-vector, and $\omega = \pm|\mathbf{k}|$.

The gauge condition (2.59) imposes a condition on the vector $\mathbf{a}$, for

$$\boldsymbol{\nabla}\cdot\mathbf{A} = i\mathbf{k}\cdot\mathbf{a}\, e^{-i\omega t + i\mathbf{k}\cdot\mathbf{x}} = 0, \tag{2.66}$$

which implies

$$\mathbf{k}\cdot\mathbf{a} = 0, \tag{2.67}$$

Hence the gauge potential $\mathbf{A}$ is orthogonal to the wave vector $\mathbf{k}$, which removes one of the three apparent degrees of freedom of the field. Thus the field oscillates in directions orthogonal to the direction of propagation: 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

$$\partial_\mu A^\mu = 0, \tag{2.68}$$

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

$$\partial_\mu F^{\mu\nu} = \partial_\mu\partial^\mu A^\nu - \partial_\mu\partial^\nu A^\mu = \partial^2 A^\nu = j^\nu. \tag{2.69}$$

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 (2.68), 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 (2.69)

$$\Box A^\mu \equiv (\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2)A^\mu = 0, \tag{2.70}$$

which looks very much like four copies of the Klein-Gordon equation for fields with zero mass. Plane wave solutions may be written

$$A^\mu(t,\mathbf{x}) = a^\mu e^{-ik\cdot x} \tag{2.71}$$

where we recall that $k\cdot x = k^0 t - \mathbf{k}\cdot\mathbf{x}$, and $k^0 = \pm|\mathbf{k}|$, Again, a^μ is a constant 4-vector.

The Lorentz gauge condition (2.68) are

$$\partial_\mu A^\mu = -ik_\mu a^\mu e^{-ik_\mu x^\mu} = 0, \tag{2.72}$$

which implies that

$$k \cdot a = 0. \tag{2.73}$$

Once again the field is orthogonal to the wave vector k^μ, but this time in the 4-vector sense.

Once the gauge condition is taken into account there are apparently three degrees of freedom, one more than in the Coulomb gauge. However, one of those is unphysical as a result of the remaining freedom to make harmonic gauge transformations.

3 Lagrangian formulation of field theory

3.1 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

$$\dot{\mathbf{p}}_A = \mathbf{F}_A, \tag{3.1}$$

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:

$$\mathbf{F}_A = -\boldsymbol{\nabla}_A V(\mathbf{x}_1, \ldots, \mathbf{x}_P). \tag{3.2}$$

There may also be *constraints* on the motion: for example, the particles may be forced to move on a sphere. 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):

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_P, t) = 0. \tag{3.3}$$

These are known as *holonomic* constraints. For example, the first particle may be constrained to move on a sphere of radius a, or $\mathbf{x}_1^2 = a^2$. Thus we lose one degree of freedon, the radial coordinate of the particle. If there are k such constraints, then the effective number of degrees of freedom is reduced by k: 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 solution 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

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{x}_1(q_1, \ldots, q_N, t) \\ \vdots &\quad \vdots \\ \mathbf{x}_P &= \mathbf{x}_P(q_1, \ldots, q_N, t) \end{aligned} \tag{3.4}$$

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

3.1.1 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

$$T = \frac{1}{2} \sum_a m_a \dot{q}_a^2, \qquad U = U(q_a, \dot{q}_a) \tag{3.5}$$

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

$$L = T - U. \tag{3.6}$$

The equations of motion of the system are then

$$\frac{\partial L}{\partial q_a} - \frac{\partial}{\partial t}\left(\frac{\partial L}{\partial \dot{q}_a}\right). \tag{3.7}$$

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

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

$$\frac{\partial L}{\partial q_a} = -\frac{\partial U}{\partial q_a}, \quad \frac{\partial L}{\partial \dot{q}_a} = m_a \dot{q}_a, \tag{3.8}$$

so that

$$\frac{\partial}{\partial t}(m_a \dot{q}_a) = -\frac{\partial U}{\partial q_a}. \tag{3.9}$$

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:

$$p_a = \frac{\partial L}{\partial \dot{q}_a}. \tag{3.10}$$

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

$$I = \int_{t_1}^{t_2} dt\, L(q_a, \dot{q}_a, t). \tag{3.11}$$

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 3.1). The action also changes, $I \to I' = I + \delta I$, where to first order in $\delta q_a(t)$,

$$\delta I = \int_{t_1}^{t_2} dt\, \delta L(q_a, \dot{q}_a, t) = \int_{t_1}^{t_2} dt\, \left(\frac{\partial L}{\partial q_a} \delta q_a + \frac{\partial L}{\partial \dot{q}_a} \delta \dot{q}_a \right). \tag{3.12}$$

Integrating by parts we find

$$\delta I = \int_{t_1}^{t_2} dt\, \left(\frac{\partial L}{\partial q_a} - \frac{\partial}{\partial t} \left(\frac{\partial L}{\partial \dot{q}_a} \right) \right) \delta q_a + \left[\frac{\partial L}{\partial \dot{q}_a} \delta q_a \right]_{t_1}^{t_2}. \tag{3.13}$$

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

$$\frac{\partial L}{\partial q_a} - \frac{\partial}{\partial t} \left(\frac{\partial L}{\partial \dot{q}_a} \right) = 0, \tag{3.14}$$

which are precisely the Euler-Lagrange equations (3.7).

3.1.2 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. (3.10). 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:

$$H(p_a, q_a, t) = \sum_a p_a \dot{q}_a - L(q_a, \dot{q}_a, t). \tag{3.15}$$

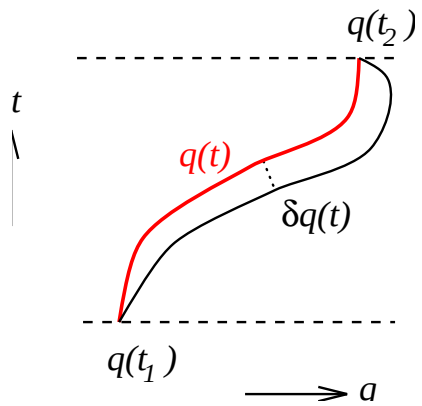


Figure 3.1: A path $q(t)$ for a system with one coordinate, and a small variation $\delta q(t)$ to that path. Classical paths are ones for which the action (see Equation 3.11) is stationary under such small variations.

The idea is to replace $\dot{q}_a$ by p_a, having solved equation (3.10). The first of Hamilton's equations follows from the Euler-Lagrange equations (3.7), after noting that $\partial L/\partial q_a = -\partial H/\partial q_a$. It is

$$\dot{p}_a = -\frac{\partial H}{\partial q_a}. \tag{3.16}$$

The second follows straightforwardly from the definition of the Hamiltonian (3.15) upon partial differentiation with respect to p_a:

$$\dot{q}_a = \frac{\partial H}{\partial p_a}. \tag{3.17}$$

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

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \sum_a \left(\frac{\partial H}{\partial p_a}\dot{p}_a + \frac{\partial H}{\partial q_a}\dot{q}_a \right). \tag{3.18}$$

Using Hamilton's equations (3.16,3.17) we see that

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \tag{3.19}$$

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

$$H = T + U, \tag{3.20}$$

where T and U are the kinetic and potential energies introduced in (3.5). 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

$$L = \frac{1}{2}\sum_a m\dot{q}_a^2, \tag{3.21}$$

from which we can easily derive the Euler-Lagrange equations

$$0 - \frac{\partial}{\partial t}(m\dot{q}_a) = 0, \tag{3.22}$$

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

$$H = \frac{1}{2}\sum_a \frac{p_a^2}{m}. \tag{3.23}$$

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

3.2 Lagrangian mechanics for the real scalar field

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 vector 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 $\phi(t, \mathbf{x})$, which takes complex values, is another example.

A very powerful way of classifying fields is to organise them according to their properties under Lorentz transformations. The Klein-Gordon field is an example of a *scalar field*, which transforms as a scalar (a pure number):

$$\phi(x) \to \phi'(x') = \phi(x), \tag{3.24}$$

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

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:

$$L = \int d^3x\, \mathcal{L}(\phi, \dot{\phi}, \boldsymbol{\nabla}\phi, t). \tag{3.25}$$

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.

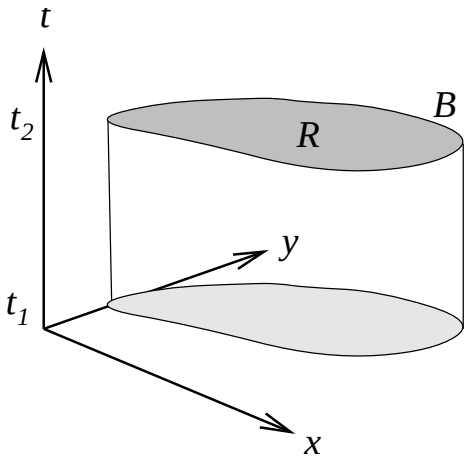


Figure 3.2: A space-time diagram (with the z coordinate suppressed) of the four-dimensional region in which the variation of the field $\delta\phi(x)$ is non-zero.

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

$$I = \int_{t_1}^{t_2} L dt = \int_{t_1}^{t_2} dt d^3x \; \mathcal{L}(\phi, \dot{\phi}, \boldsymbol{\nabla}\phi, t). \tag{3.26}$$

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 an arbitrary spatial region R with boundary B (see Figure 3.2). This variation causes a small change in the action:

$$I \to I' = I + \delta I. \tag{3.27}$$

Taylor expanding the Lagrangian density, we find

$$I' \simeq \int_{t_1}^{t_2} dt d^3x \; \left[\mathcal{L} + \frac{\partial \mathcal{L}}{\partial \phi}\delta\phi + \frac{\partial \mathcal{L}}{\partial \boldsymbol{\nabla}\phi} \cdot \boldsymbol{\nabla}\delta\phi + \frac{\partial \mathcal{L}}{\partial \dot{\phi}}\delta\dot{\phi} \right], \tag{3.28}$$

where

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\nabla}\phi} = \left(\frac{\partial \mathcal{L}}{\partial(\partial_x \phi)}, \frac{\partial \mathcal{L}}{\partial(\partial_y \phi)}, \frac{\partial \mathcal{L}}{\partial(\partial_z \phi)} \right). \tag{3.29}$$

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

$$\delta I = \int_{t_1}^{t_2} dt d^3x \ \left[\frac{\partial \mathcal{L}}{\partial \phi}\delta\phi + \frac{\partial \mathcal{L}}{\partial \boldsymbol{\nabla}\phi}\cdot\boldsymbol{\nabla}\delta\phi - \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\right)\delta\phi\right] + \left[\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\delta\phi\right]_{t_1}^{t_2}. \tag{3.30}$$

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

$$\delta I = \int_{t_1}^{t_2} dt d^3x \ \left[\frac{\partial \mathcal{L}}{\partial \phi}\delta\phi - \boldsymbol{\nabla}\cdot\frac{\partial \mathcal{L}}{\partial \boldsymbol{\nabla}\phi}\delta\phi - \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\right)\delta\phi\right] + \int_{t_1}^{t_2} dt \int_B \mathbf{dS}\cdot\left[\frac{\partial \mathcal{L}}{\partial \boldsymbol{\nabla}\phi}\delta\phi\right]. \tag{3.31}$$

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

$$\delta I = \int_{t_1}^{t_2} dt d^3x \ \left[\frac{\partial \mathcal{L}}{\partial \phi} - \boldsymbol{\nabla}\cdot\frac{\partial \mathcal{L}}{\partial \boldsymbol{\nabla}\phi} - \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\right)\right]\delta\phi. \tag{3.32}$$

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

$$\frac{\partial \mathcal{L}}{\partial \phi} - \boldsymbol{\nabla}\cdot\frac{\partial \mathcal{L}}{\partial \boldsymbol{\nabla}\phi} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{\phi}} = 0. \tag{3.33}$$

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

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_i\frac{\partial \mathcal{L}}{\partial(\partial_i\phi)} - \partial_0\frac{\partial \mathcal{L}}{\partial(\partial_0\phi)} = 0, \tag{3.34}$$

or

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)} = 0. \tag{3.35}$$

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

$$\frac{\partial \mathcal{L}}{\partial \phi_a} - \partial_\mu\frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi_a)} = 0. \tag{3.36}$$

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

$$(\partial^\mu\partial_\mu + m^2)\phi_R(x) = 0, \quad (\partial^\mu\partial_\mu + m^2)\phi_I(x) = 0. \tag{3.37}$$

Consider first the following Lagrangian density for the real part,

$$\mathcal{L}_R = \frac{1}{2}\partial^\mu\phi_R\partial_\mu\phi_R - \frac{1}{2}m^2\phi_R^2. \tag{3.38}$$

We can immediately derive

$$\frac{\partial \mathcal{L}}{\partial \phi_R} = -m^2 \phi_R, \quad \frac{\partial \mathcal{L}}{\partial(\partial_i \phi_R)} = \partial^i \phi_R, \quad \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi_R)} = \partial^0 \phi. \tag{3.39}$$

Hence,

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_R)} = (\partial^0 \phi_R, \partial^i \phi_R), \tag{3.40}$$

and so

$$\frac{\partial \mathcal{L}}{\partial \phi_R} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_R)} = -m^2 \phi_R - \partial_\mu \partial^\mu \phi_R = 0. \tag{3.41}$$

Thus the Lagrangian (3.38) 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

$$\begin{aligned} \mathcal{L} = \mathcal{L}_R + \mathcal{L}_L &= \frac{1}{2}\partial^\mu \phi_R \partial_\mu \phi_R - \frac{1}{2} m^2 \phi_R^2 + \frac{1}{2}\partial^\mu \phi_I \partial_\mu \phi_I - \frac{1}{2} m^2 \phi_I^2 \\ &= \frac{1}{2}\partial^\mu \phi^* \partial_\mu \phi - \frac{1}{2} m^2 \phi^* \phi. \end{aligned} \tag{3.42}$$

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

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(x)}. \tag{3.43}$$

Thus for our example (3.38), the conjugate momnetum $\pi(x) = \dot{\phi}(x)$. The Hamiltonian for the real scalar field is defined by

$$H = \int d^3x \; \pi(x)\dot{\phi}(x) - \int d^3x \mathcal{L}(\phi, \partial_i \phi, \dot{\phi}). \tag{3.44}$$

If the field obeys the Klein-Gordon equation, its Hamiltonian is

$$H = \int d^3x \left[\frac{1}{2}\pi^2 + \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 + \frac{1}{2} m^2 \phi^2 \right]. \tag{3.45}$$

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:

$$H = \int d^3x \left[\frac{1}{2}\pi^2 + \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 + V(\phi) \right]. \tag{3.46}$$

In classical field theory, V can be any function of ϕ, as long as it is bounded below, otherwise the energy of the field can become arbitrarily negative. Quantum field theory is much more restrictive about the terms it allows: it turns out that the theory does not make sense in four space-time dimensions unless $V(\phi)$ is restricted to be a polynomial of degree no more than four.

3.3 Lagrangian mechanics for the electromagnetic field

The electromagnetic gauge potential $A_\mu(x)$ is an example of a *4-vector field*, which is a field which transforms like a covariant 4-vector. Its behaviour under Lorentz transformations is

$$A^\mu(x) \to A'^\mu(x') = \Lambda^\mu_{\ \nu} A^\nu(x), \tag{3.47}$$

where $x'^\mu = \Lambda^\mu_{\ \nu} x^\nu$. In this section we will write down a Lagrangian density, and show that the resulting Euler-Lagrange equations are just the inhomogeneous Maxwell equations. The homogeneous ones will be shown to be an identity, rather than arising from any variational principle. The Lagrangian is

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - j^\mu A_\mu, \tag{3.48}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. The generalised coordinates of this system are the four components the gauge field A_μ. We can therefore derive Maxwell's equations by looking at the Euler-Lagrange equations:

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} - \frac{\partial \mathcal{L}}{\partial A_\nu} = 0. \tag{3.49}$$

Let us first consider the variation with respect to A_ν, which gives

$$\frac{\partial \mathcal{L}}{\partial A_\nu} = -j^\mu \frac{\partial A_\mu}{\partial A_\nu} = -j^\mu \delta^\nu_\mu = -j^\nu. \tag{3.50}$$

The variation with respect to $\partial_\mu A_\nu$ is a little more complex, and it will help to define $X_{\mu\nu} = \partial_\mu A_\nu$, so that $F_{\mu\nu} = X_{\mu\nu} - X_{\nu\mu}$. Hence

$$\begin{aligned}
\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} &= -\frac{1}{4} \frac{\partial F_{\rho\sigma}}{\partial X_{\mu\nu}} F^{\rho\sigma} - \frac{1}{4} F_{\rho\sigma} \frac{\partial F^{\rho\sigma}}{\partial X_{\mu\nu}}, \\
&= -\frac{1}{4} \left(\delta^\mu_\rho \delta^\nu_\sigma - \delta^\mu_\sigma \delta^\nu_\rho \right) F^{\rho\sigma} - \frac{1}{4} F_{\rho\sigma} \left(\eta^{\mu\rho} \eta^{\nu\sigma} - \eta^{\mu\sigma} \eta^{\nu\rho} \right), \\
&= -\frac{1}{4} (F^{\mu\nu} - F^{\nu\mu}) - \frac{1}{4} (F^{\mu\nu} - F^{\nu\mu}) \\
&= -F^{\mu\nu}.
\end{aligned} \tag{3.51}$$

Thus we can substitute (3.50) and (3.51) back into the Euler-Lagrange equation (3.49) to obtain

$$-\partial_\mu F^{\mu\nu} + j^\nu = 0, \tag{3.52}$$

which we recognise as the covariant form of the inhomogeneous Maxwell equations.

We now turn to the homogeneous Maxwell equations, and demonstrate that they are an identity, which is sometimes known as the *Bianchi identity*. This

can be seen by examining the 4-vector $k^\mu = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\partial^\nu F^{\rho\sigma} = 0$. Substituting the expression for the field strength tensor in terms of the gauge potential, we obtain

$$k^\mu = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\partial^\nu(\partial^\rho A^\sigma - \partial^\sigma A^\rho). \tag{3.53}$$

By exchanging the labels σ and ρ in the second term we find

$$\begin{aligned} k^\mu &= \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\partial^\nu\partial^\rho A^\sigma - \frac{1}{2}\epsilon_{\mu\nu\sigma\rho}\partial^\nu\partial^\rho A^\sigma, &(3.54)\\ &= \epsilon_{\mu\nu\rho\sigma}\partial^\nu\partial^\rho A^\sigma, &(3.55)\end{aligned}$$

with the last step following by the antisymmetry of the Levi-Civita tensor. Now we exchange the labels ν and ρ, from which we can deduce

$$\epsilon_{\mu\nu\rho\sigma}\partial^\nu\partial^\rho A^\sigma = \epsilon_{\mu\rho\nu\sigma}\partial^\rho\partial^\nu A^\sigma. \tag{3.56}$$

The antisymmetry of the Levi-Civita tensor then shows that

$$\epsilon_{\mu\nu\rho\sigma}\partial^\nu\partial^\rho A^\sigma = -\epsilon_{\mu\nu\rho\sigma}\partial^\nu\partial^\rho A^\sigma. \tag{3.57}$$

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

$$\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\partial^\nu F^{\rho\sigma} = 0. \tag{3.58}$$

Thus we have demonstrated that the Principle of Least Action applied to the Lagrangian (3.48), combined with the Bianchi identity (3.58), results in Maxwell's equations.

3.4 Noether's theorem and conservation laws

It is a profound feature of both classical and quantum dynamics that conservation laws, such as conservation of energy or momentum, follow from symmetries of the dynamical equations. In this section we shall show how this feature is realised in classical field theory.

A conservation law can be expressed in two equivalent ways. Firstly, and most obviously, one can say that there is a quantity Q which is constant in time, i.e.

$$\frac{dQ}{dt} = 0. \tag{3.59}$$

Entirely equivalently, one can also say that there is a 4-vector current $j^\mu(x)$ which satisfies a continuity equation,

$$\partial_\mu j^\mu(x) = 0. \tag{3.60}$$

If we integrate (3.60) over all space, we find

$$\int d^3x \, \frac{\partial}{\partial t} j^0(x) + \int d^3x \, \boldsymbol{\nabla} \cdot \mathbf{j} = 0. \tag{3.61}$$

Defining $Q = \int d^3x \, j^0(x)$, we find using Gauss's Law that

$$\frac{dQ}{dt} = -\int_\infty \mathbf{dS} \cdot \mathbf{j}, \tag{3.62}$$

where the surface integral is taken over a sphere at spatial infinity. With the assumption that all currents die off at infinity, we have

$$\frac{dQ}{dt} = 0. \tag{3.63}$$

Hence the existence of a 4-vector current satisfying a continuity equation $\partial_\mu j^\mu(x) = 0$ automatically implies that there is a conserved quantity, which is the integral over all space of the time component of the current.

Now consider a theory with N real fields $\phi_a(x)$, with $a = 1, \ldots, N$, and an action

$$S = \int d^4x \, \mathcal{L}(\phi_a, \partial\phi_a). \tag{3.64}$$

Noether's theorem for this theory can be stated as follows.

For every transformation $\phi_a(x) = \phi'_a(x')$ which leaves the action S invariant, there is a conserved current.

The proof of this theorem is quite involved, and we first need some definitions.

Let us first define the total variation in the field,

$$\delta\phi_a(x) = \phi'_a(x') - \phi_a(x). \tag{3.65}$$

Note that this variation can be split into two pieces, the first due to the fact that the functions ϕ_a are changing:

$$\tilde{\delta}\phi_a(x) = \phi'_a(x) - \phi_a(x), \tag{3.66}$$

which defines what we mean by the operator $\tilde{\delta}$. The remaining piece can be interpreted as the change in the field functions due to the change in coordinates,

$$\Delta\phi_a(x) = \phi_a(x') - \phi_a(x) \simeq \partial_\mu\phi_a(x)\delta x^\mu, \tag{3.67}$$

where $\delta x^\mu = x'^\mu - x^\mu$. Hence

$$\delta\phi_a(x) \simeq \tilde{\delta}\phi_a(x) + \Delta\phi'_a(x) \simeq \tilde{\delta}\phi_a(x) + \Delta\phi_a(x), \tag{3.68}$$

where we have dropped at term of second order in the variation, as $\Delta\phi'_a(x) = \Delta\phi_a(x) + O(\delta\phi\delta x)$.

One can show straightforwardly from the definition of $\tilde{\delta}$ that

$$\tilde{\delta}(\partial_\mu \phi_a(x)) = \partial_\mu(\tilde{\delta}\phi_a(x)), \tag{3.69}$$

but that

$$\delta(\partial_\mu \phi_a(x)) = \partial_\mu(\delta\phi_a(x)) - \partial_\nu \phi_a(x)\partial_\mu \delta x^\nu. \tag{3.70}$$

Hence the partial derivative operator commutes with $\tilde{\delta}$, but not with δ.

Now consider the effect of the variation in Eq. 3.65 on the action:

$$\delta S = \int_\Omega d^4x'\, \mathcal{L}'(x') - \int_\Omega d^4x\, \mathcal{L}(x), \tag{3.71}$$

where $\mathcal{L}'(x') = \mathcal{L}(\phi'_a(x'), \partial\phi'_a(x'))$, and Ω is an arbitrary region in spacetime. By adding and subtracting $\mathcal{L}(x)$ to the first integrand, we see

$$\delta S = \int_\Omega d^4x'\, \delta\mathcal{L}'(x) + \int_\Omega d^4x'\, \mathcal{L}(x) - \int_\Omega d^4x\, \mathcal{L}(x). \tag{3.72}$$

We now need to express the integration measure d^4x' in terms of d^4x, using the Jacobian of the coordinate transformation:

$$d^4x' = d^4x \left| \frac{\partial(x'^0, \ldots, x'^3)}{\partial(x^0, \ldots, x^3)} \right| = d^4x |\det \mathbf{M}|, \tag{3.73}$$

where the components of the matrix $\mathbf{M}$ are given by

$$M^\mu{}_\nu = \frac{\partial x'^\mu}{\partial x^\nu} \simeq \delta^\mu_\nu + \partial_\nu \delta x^\mu. \tag{3.74}$$

In matrix notation, we may write $\mathbf{M} = 1 + \delta\mathbf{M}$, and using the standard matrix identity

$$\ln\det \mathbf{M} = \mathrm{tr} \ln \mathbf{M}, \tag{3.75}$$

we see that

$$\det \mathbf{M} \simeq 1 + \mathrm{tr}\, \delta\mathbf{M} = 1 + \partial_\mu \delta x^\mu. \tag{3.76}$$

Substituting into the varation of the action Eq. (3.71), we find

$$\delta S = \int_\Omega d^4x\, (1 + \partial_\mu \delta x^\mu)\delta\mathcal{L}'(x) + \int_\Omega d^4x\, (1 + \partial_\mu \delta x^\mu)\mathcal{L}(x) - \int_\Omega d^4x\, \mathcal{L}(x). \tag{3.77}$$

Dropping terms which are second order in the variation we arrive at

$$\delta S = \int_\Omega d^4x\, \delta\mathcal{L}(x) + \int_\Omega d^4x\, \partial_\mu \delta x^\mu \mathcal{L}(x) \tag{3.78}$$

Now we can again split the variation in the Lagrangian function into a piece arising from the change in the function itself, and that due to the change of coordinates:

$$\delta\mathcal{L}(x) = \tilde{\delta}\mathcal{L}(x) + \partial_\mu \mathcal{L}(x)\delta x^\mu. \tag{3.79}$$

The change in the Lagrangian function arises because of the change in the functional form of the fields, so

$$\tilde{\delta}\mathcal{L}(x) = \frac{\partial \mathcal{L}}{\partial \phi_a}(x)\tilde{\delta}\phi_a(x) + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}(x)\tilde{\delta}(\partial_\mu \phi_a(x)). \tag{3.80}$$

At this point, the commutativity of $\tilde{\delta}$ and ∂_μ stated earlier means that

$$\begin{aligned} \delta S &= \int_\Omega d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi_a}(x)\tilde{\delta}\phi_a(x) + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}(x)\partial_\mu(\tilde{\delta}\phi_a(x)) \right. \\ &+ \left. \int_\Omega d^4x \, \left(\partial_\mu \mathcal{L}(x)\delta x^\mu + \mathcal{L}(x)\partial_\mu \delta x^\mu\right) \right). \end{aligned} \tag{3.81}$$

Some simple algebra shows that

$$\delta S = \int_\Omega d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi_a} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \right) \tilde{\delta}\phi_a(x) + \int_\Omega d^4x \, \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \tilde{\delta}\phi_a + \mathcal{L}(x)\delta x^\mu \right). \tag{3.82}$$

Note that the first term is precisely the variation in the action we found when deriving the Euler-Lagrange equation, where we considered variations in the fields alone, vanishing on the boundary of Ω. Suppose we consider a particular field configuration $\phi(x)$ satisfying the Euler-Lagrange equation. If we then make an infinitesimal transformation on it which leaves the action invariant, i.e. which results in no first-order change in the action, we must have that

$$\partial_\mu f^\mu = 0, \tag{3.83}$$

where

$$f^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \tilde{\delta}\phi_a(x) + \mathcal{L}\delta x^\mu, \tag{3.84}$$

or substituting back for $\tilde{\delta}\phi(x)$,

$$f^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \delta\phi_a(x) - \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \partial_\nu \phi_a(x) - \mathcal{L}\delta^\mu_\nu \right) \delta x^\nu. \tag{3.85}$$

Thus we see the appearance of an infinitesimal 4-vector, f^μ, which satisfies a continuity equation.

3.4.1 Conservation of energy-momentum

Consider a translation on the coordinates, with the fields unchanged:

$$\delta x^\mu = \epsilon^\mu, \qquad \delta\phi_a = 0, \tag{3.86}$$

where ϵ^μ is a constant 4-vector. This clearly includes both space and time translations. Hence we can write

$$f^\mu = \theta^\mu{}_\nu \epsilon^\nu, \tag{3.87}$$

where we have introduced the *canonical energy-momentum tensor*

$$\theta^{\mu}{}_{\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \partial_\nu \phi_a(x) - \mathcal{L}\delta^{\mu}_{\nu}. \tag{3.88}$$

Because the translations ϵ^μ are arbitrary, translation invariance of the Lagrangian implies that there must be four continuity relations,

$$\partial_\mu \theta^{\mu}{}_{\nu} = 0. \tag{3.89}$$

The conserved "charge" associated with these four continuity relations is the total 4-momentum,

$$P_\mu = \int d^3x \theta^0{}_{\mu}(x). \tag{3.90}$$

4 Canonical quantisation

We have seen that the Klein-Gordon field $\phi(x)$ does not have the interpretation as a single-particle wave-function, as it does not have a conserved probability current, and it possesses negative energy states which cannot form a "Dirac sea". However, as we shall see in this section, the theory of the quantum field operator $\hat{\phi}(x)$ is in fact a relativisitic many-particle quantum theory. The process of developing the theory of a quantum field is sometimes misleadingly called "second quantisation", perhaps because of some idea that one is quantising a wavefunction, itself a quantum object. It should be emphasised at the outset that the scalar field, real or complex, is not a wavefunction.

In this section we 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. In practice this means that the Lagrangian can be written as a quadratic function of the fields. Free fields may sound somewhat irrelevant, but nearly all quantum field theories have to be treated as small perturbations away from free field theories, so it is an appropriate place to start. There are non-trivial effects in free field theory which come from boundary conditions: one of these effects is the Casimir effect in the free electromagnetic field.

4.1 Quantisation of finite systems

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

$$L = \frac{1}{2}\sum_{a=1}^{N} m\dot{q}_a^2 - U(q_a) \tag{4.1}$$

(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

$$H = \sum_{a=1}^{N} \frac{p_a^2}{2m} + U(q_a). \tag{4.2}$$

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

$$[\hat{q}_a, \hat{p}_b] = i\delta_{ab}, \quad [\hat{q}_a, \hat{q}_b] = 0 = [\hat{p}_a, \hat{p}_b]. \tag{4.3}$$

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*

$$i\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle. \tag{4.4}$$

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

$$i\frac{d}{dt}O_H(t) = [O_H(t), H]. \tag{4.5}$$

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 Hermitian conjugate). Thus, if we suppose that the two sets of operators and states coincide at $t = 0$, they are related at all times by

$$\begin{aligned} O_H(t) &= e^{iHt} O_S e^{-iHt} \\ |\psi\rangle_H &= e^{iHt}|\psi(t)\rangle_S. \end{aligned} \tag{4.6}$$

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:

$$\begin{aligned} {}_H\langle\psi|O_H(t)|\psi'\rangle_H &= {}_S\langle\psi(t)|e^{-iHt}e^{iHt}O_S e^{-iHt}e^{iHt}|\psi'(t)\rangle_S \\ &= {}_S\langle\psi(t)|O_S|\psi'(t)\rangle_S. \end{aligned} \tag{4.7}$$

This demonstrates that the Heisenberg picture matrix elements are the same as those in the Schrödinger one, which means there is no difference in the physical predictions one makes.

4.2 The real scalar field

We begin with the Lagrangian density

$$\mathcal{L} = \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}m^2\phi^2, \tag{4.8}$$

from which we can define the conjugate momentum

$$\pi(x) = \frac{\partial\mathcal{L}}{\partial\dot{\phi}(x)}. \tag{4.9}$$

The Hamiltonian density is then found from

$$\mathcal{H} = \pi(x)\dot{\phi}(x) - \mathcal{L} = \frac{1}{2}\pi^2 + \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 + \frac{1}{2}m^2\phi^2. \tag{4.10}$$

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

$$\begin{aligned}
[\phi(t,\mathbf{x}),\pi(t,\mathbf{x}')] &= i\hbar\delta^3(\mathbf{x}-\mathbf{x}'), &(4.11)\\
[\phi(t,\mathbf{x}),\phi(t,\mathbf{x}')] &= 0, &(4.12)\\
[\pi(t,\mathbf{x}),\pi(t,\mathbf{x}')] &= 0, &(4.13)
\end{aligned}$$

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 (4.3) 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*, obeying the Heisenberg equations

$$\dot{\hat{\phi}}(x) = \frac{1}{i}\left[\hat{\phi}(x),\hat{H}\right], \qquad \dot{\hat{\pi}}(x) = \frac{1}{i}\left[\hat{\pi}(x),\hat{H}\right], \tag{4.14}$$

where $\hat{H}$ is the quantum Hamiltonian, constructed by replacing the classical fields by operators in the original:

$$\hat{H} = \int d^3x\left(\frac{1}{2}\hat{\pi}^2 + \frac{1}{2}(\boldsymbol{\nabla}\hat{\phi})^2 + \frac{1}{2}\mu^2\hat{\phi}^2\right). \tag{4.15}$$

It is straightforward to show for a free field that

$$\dot{\hat{\phi}}(x) = \hat{\pi}(x), \qquad \dot{\hat{\pi}}(x) = (\nabla^2 - m^2)\hat{\phi}(x). \tag{4.16}$$

Hence the field operator obeys the same equation of motion as the classical field. It is convenient to expand the field operator in terms of eigenfunctions of the operator $(-\nabla^2 + m^2)$,

$$(-\nabla^2 + m^2)e^{i\mathbf{k}\bullet\mathbf{x}} = \omega_{\mathbf{k}}^2 e^{i\mathbf{k}\bullet\mathbf{x}}. \tag{4.17}$$

Writing

$$\hat{\phi}(x) = \int \frac{d^3k}{(2\pi)^3} \left(f_{\mathbf{k}}(x)\hat{a}(\mathbf{k}) + f_{\mathbf{k}}^*(x)\hat{a}^*(\mathbf{k}) \right). \tag{4.18}$$

with $f_{\mathbf{k}}(x) = A_{\mathbf{k}}(t)e^{i\mathbf{k}\bullet\mathbf{x}}$, we see that $f_{\mathbf{k}}(x)$ obeys the field equation $(\partial^2 + m^2)f_{\mathbf{k}}(x) = 0$ if

$$A_{\mathbf{k}}(t) = \mathcal{N}_{\mathbf{k}} e^{-i\omega_{\mathbf{k}}t} \tag{4.19}$$

The normalisation factor $\mathcal{N}_{\mathbf{k}}$ is essentially arbitrary, although almost everyone uses one of two conventions outlined below. The properties of the operators $\hat{a}(\mathbf{k})$ and $\hat{a}^*(\mathbf{k})$ can be deduced from the commutation relations of $\hat{\phi}$ and $\hat{\pi}$. We therefore need expressions for $\hat{a}(\mathbf{k})$ and $\hat{a}^*(\mathbf{k})$ in terms of $\hat{\phi}$ and $\hat{\pi}$, where

$$\hat{\pi}(x) = \dot{\hat{\phi}}(x) = \int \frac{d^3k}{(2\pi)^3} \left(-i\omega_{\mathbf{k}}\mathcal{N}_{\mathbf{k}}\hat{a}(\mathbf{k})e^{-ik\cdot x} + i\omega_{\mathbf{k}}\mathcal{N}_{\mathbf{k}}^*\hat{a}^*(\mathbf{k})e^{ik\cdot x} \right). \tag{4.20}$$

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

$$\hat{\phi}(\mathbf{k}', t) = \int d^3x \hat{\phi}(x)e^{-i\mathbf{k}'\bullet\mathbf{x}}, \quad \hat{\pi}(\mathbf{k}', t) = \int d^3x \hat{\pi}(x)e^{-i\mathbf{k}'\bullet\mathbf{x}}, \tag{4.21}$$

and use the relation

$$\int d^3x e^{i(\mathbf{k}-\mathbf{k}')\bullet\mathbf{x}} = (2\pi)^3\delta^3(\mathbf{k} - \mathbf{k}') \tag{4.22}$$

to obtain

$$\hat{\phi}(\mathbf{k}', t) = \mathcal{N}_{\mathbf{k}}\hat{a}(\mathbf{k}')e^{-i\omega_{\mathbf{k}'}t} + \mathcal{N}_{\mathbf{k}}^*\hat{a}^*(-\mathbf{k}')e^{i\omega_{\mathbf{k}'}t} \tag{4.23}$$

$$\hat{\pi}(\mathbf{k}', t) = -i\omega_{\mathbf{k}'}\mathcal{N}_{\mathbf{k}'}\hat{a}(\mathbf{k}')e^{-i\omega_{\mathbf{k}'}t} + i\omega_{\mathbf{k}'}\mathcal{N}_{\mathbf{k}'}^* a(\mathbf{k}')\hat{a}^*(-\mathbf{k}')e^{i\omega_{\mathbf{k}'}t}. \tag{4.24}$$

Hence

$$2\omega_{\mathbf{k}}\mathcal{N}_{\omega_{\mathbf{k}}}\hat{a}(\mathbf{k}') = \int d^3x (i\hat{\pi}(x) + \omega_{\mathbf{k}'}\hat{\phi}(x))e^{ik'\cdot x}. \tag{4.25}$$

By complex conjugation one obtains

$$2\omega_{\mathbf{k}}\mathcal{N}_{\omega_{\mathbf{k}'}}^*\hat{a}^*(\mathbf{k}') = \int d^3x (-i\hat{\pi}(x) + \omega_{\mathbf{k}'}\hat{\phi}(x))e^{-ik'\cdot x}. \tag{4.26}$$

One can then derive the commutation relation

$$[\hat{a}(\mathbf{k}), \hat{a}^*(\mathbf{k}')] = (2\pi)^3\delta^3(\mathbf{k} - \mathbf{k}')/2\omega_{\mathbf{k}}|\mathcal{N}_{\mathbf{k}}|^2, \tag{4.27}$$

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

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 x^2, \tag{4.28}$$

where ω is the angular frequency of oscillation. Upon quantisation, the operators $\hat{a} = (i\hat{p} + \omega\hat{x})/\sqrt{2\omega}$ and $\hat{a}^* = (-i\hat{p} + \omega\hat{x})/\sqrt{2\omega}$ satisfy the commutation relations

$$[\hat{a}, \hat{a}^*] = 1. \tag{4.29}$$

This correspondence suggests that one choice for the normalisation factor:

$$\mathcal{N}_{\mathbf{k}} = \frac{1}{\sqrt{2\omega_{\mathbf{k}}}}, \tag{4.30}$$

in which case we find for the complete set of commutation relations

$$\begin{aligned} [a(\mathbf{k}), a^*(\mathbf{k}')] &= (2\pi)^3\delta^3(\mathbf{k} - \mathbf{k}'), & (4.31)\\ [a(\mathbf{k}), a(\mathbf{k}')] &= 0, & (4.32)\\ [a^*(\mathbf{k}), a^*(\mathbf{k}')] &= 0. & (4.33) \end{aligned}$$

The right hand side of the first equation, with its δ-fucntion, is the closest we can get to unity when we are dealing with functions of a continuous variable like $a(\mathbf{k})$. However, we shall adopt a convention which keeps the Fourier expansions of the field operators Lorentz invariant,

$$\mathcal{N}_{\mathbf{k}} = \frac{1}{2\omega_{\mathbf{k}}}, \tag{4.34}$$

which means that the first commutation relation becomes

$$[\hat{a}(\mathbf{k}), \hat{a}^*(\mathbf{k}')] = 2\omega_{\mathbf{k}}(2\pi)^3\delta^3(\mathbf{k} - \mathbf{k}'). \tag{4.35}$$

It is also quite common to work in a finite volume V, which means that instead of integrals over $\mathbf{k}$ we have sums over allowed wavevectors $\mathbf{k} = 2\pi(n_1, n_2, n_3)/L$, which go over to integrals in the infinite volume limit:

$$\frac{1}{V}\sum_{\mathbf{k}} \to \int \frac{d^3k}{(2\pi)^3}. \tag{4.36}$$

All wavefunctions are multiplied by a factor $1/\sqrt{V}$: hence the plane wave expansion of the field operator is written

$$\phi(x) = \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_{\mathbf{k}}V}} \left(a(\mathbf{k})e^{-ik\cdot x} + a^*(\mathbf{k})e^{ik\cdot x} \right). \tag{4.37}$$

The non-trivial equal time canonical commutation relation then becomes

$$[\hat{a}(\mathbf{k}), \hat{a}^*(\mathbf{k}')] = \delta_{\mathbf{k},\mathbf{k}'}, \tag{4.38}$$

where $\delta_{\mathbf{k},\mathbf{k}'} = \delta_{n_1 n_1'}\delta_{n_2 n_2'}\delta_{n_3 n_3'}$.

To summarise, we have discovered that the field theory behaves very much like infinitely many harmonic oscillators, labelled by their wavenumber $\mathbf{k}$, and 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 $\hat{a}^*(\mathbf{k})$ and $\hat{a}(\mathbf{k})$ are also called *creation* and *annihilation* operators, because they change the number of particles with 3-momentum $\mathbf{k}$ in a state.

4.3 States of the scalar field; zero point energy

In this section we will drop the hat notation for operators, unless there could be ambiguity, as all fields will be assumed to be operators unless otherwise specified. 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:

$$a(\mathbf{k})|0\rangle = 0, \quad \forall \mathbf{k}. \tag{4.39}$$

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

$$|\mathbf{k}\rangle = a^*(\mathbf{k})|0\rangle. \tag{4.40}$$

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

$$\begin{aligned} \langle \mathbf{k}'|\mathbf{k}\rangle &= \langle 0|a(\mathbf{k}')a^*(\mathbf{k})|0\rangle & (4.41)\\ &= \langle 0|a^*(\mathbf{k})a(\mathbf{k}')|0\rangle + \langle 0|0\rangle(2\pi)^3\delta(\mathbf{k}-\mathbf{k}'). & (4.42) \end{aligned}$$

The first term on the right hand side vanishes, and so we have that the first excited states satisfy the normalisation condition

$$\langle \mathbf{k}'|\mathbf{k}\rangle = (2\pi)^3\delta(\mathbf{k}-\mathbf{k}'). \tag{4.43}$$

The corresponding completeness relation is

$$\int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{k}}}|\mathbf{k}\rangle\langle\mathbf{k}| = 1. \tag{4.44}$$

The states $|\mathbf{k}\rangle$ are interpreted in quantum field theory as single-particle states, an interpretation discussed further below. We could continue applying creation operators $a^*(\mathbf{k}_1)$, $a^*(\mathbf{k}_2), \ldots$ to make states

$$|\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2, \ldots\rangle = \ldots a^*(\mathbf{k}_2)a^*(\mathbf{k}_1)a^*(\mathbf{k})|0\rangle. \tag{4.45}$$

The general state could have multiple applications of creation operators with the same momenta, in which case it is conventional to divide by a symmetry factor, to obtain

$$|\{n_\mathbf{k}\}\rangle = \prod_\mathbf{k} \frac{(a^*_\mathbf{k})^{n_\mathbf{k}}}{\sqrt{n_\mathbf{k}!}}|0\rangle, \tag{4.46}$$

where $n_\mathbf{k}$ is the number of times the creation operator with label $\mathbf{k}$ is applied. The Hilbert space spanned by the set of all states $|n_\mathbf{k}\rangle$ is called *Fock space*. The basis states are interpreted as multiparticle states, which demonstrates 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 $\hat{n}(\mathbf{k})$ by

$$\hat{n}(\mathbf{k}) = a^*(\mathbf{k})a(\mathbf{k}). \tag{4.47}$$

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

$$\hat{n}(\mathbf{k})|0\rangle = 0, \tag{4.48}$$

from the definition of the vacuum state (4.39). When acting on 1-particle states, we find

$$\begin{aligned} \hat{n}(\mathbf{k})|\mathbf{k}'\rangle &= a^*(\mathbf{k})a(\mathbf{k})a^*(\mathbf{k}')|0\rangle & (4.49)\\ &= a^*(\mathbf{k})\left(a^*(\mathbf{k}')a(\mathbf{k}) + 2\omega_\mathbf{k}(2\pi)^3\delta(\mathbf{k}-\mathbf{k}')\right)|0\rangle, & (4.50) \end{aligned}$$

where we have used the commutation relations (4.33) to change the order of $a(\mathbf{k})$ and $a^*(\mathbf{k}')$. Thus,

$$\hat{n}(\mathbf{k})|\mathbf{k}'\rangle = 2\omega_\mathbf{k}(2\pi)^3\delta(\mathbf{k}-\mathbf{k}')|\mathbf{k}\rangle. \tag{4.51}$$

The operators $\hat{n}(\mathbf{k})$ are used to construct the *number operator*

$$N = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_\mathbf{k}}\hat{n}(\mathbf{k}), \tag{4.52}$$

which tells us the total number of particles in a particular state[2]. Thus, as expected, $N|0\rangle = 0$, meaning that there are no particles in the ground state. In the 1-particle state $|\mathbf{k}'\rangle$, we find

$$N_{\rm tot}|\mathbf{k}'\rangle = \int \frac{d^3k}{(2\pi)^3}N(\mathbf{k})|\mathbf{k}'\rangle = \int \frac{d^3k}{(2\pi)^3}(2\pi)^3\delta(\mathbf{k}-\mathbf{k}')|\mathbf{k}\rangle = |\mathbf{k}'\rangle. \tag{4.53}$$

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

The state $|\mathbf{k}'\rangle$ is indeed an eigenstate of the number operator, with eigenvalue 1 – there is 1 particle in state. When applied to the general state (4.46), it can be shown that

$$N|\{n_{\mathbf{k}}\}\rangle = \int \frac{d^3k}{(2\pi)^3} n_{\mathbf{k}} |\{n_{\mathbf{k}}\}\rangle. \tag{4.54}$$

The energy of the states can be computed from the quantum energy operator

$$\hat{E} = \int d^3x \theta^0{}_0(x) = \int d^3x \left(\frac{1}{2}\hat{\pi}^2 + \frac{1}{2}(\boldsymbol{\nabla}\hat{\phi})^2 + \frac{1}{2}\mu^2\hat{\phi}^2 \right), \tag{4.55}$$

which is immediately seen to be identical to the Hamiltonian. 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 the Hamiltonian operator H in terms of creation and annihilation operators. Recalling the expansions of ϕ and π in terms of $a(\mathbf{k})$ and $a^*(\mathbf{k})$ in equations (4.18) and (4.20), we find

$$H = \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{k}}} \frac{\omega_{\mathbf{k}}}{2} \left(a^*(\mathbf{k})a(\mathbf{k}) + a(\mathbf{k})a^*(\mathbf{k}) \right). \tag{4.56}$$

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

$$\hat{H} = \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{k}}} \frac{\omega_{\mathbf{k}}}{2} \left(2a^*(\mathbf{k})a(\mathbf{k}) + (2\pi)^3\delta(\mathbf{0})\omega_{\mathbf{k}} \right). \tag{4.57}$$

The second term in the brackets has the formally infinite term $(2\pi)^3\delta(\mathbf{0})$. The source of this infinity is the fact that we are working in infinite spatial volume. Let us revert to a finite cubic region with sides of length L, in which case we can write

$$(2\pi)^3\delta(\mathbf{0}) = \lim_{L\to\infty} \lim_{\mathbf{k}'\to\mathbf{k}} \int_{-L/2}^{L/2} d^3x e^{i\mathbf{x}\cdot(\mathbf{k}-\mathbf{k}')} = \lim_{L\to\infty} \int_{-L/2}^{L/2} d^3x = V, \tag{4.58}$$

where V is a formal volume factor, and the limit of infinite volume should strictly be taken at the end of the calculation of any physical quantity. Thus we can interpret the quantity multiplied by V as an energy *density*. However, even this energy density is infinite, in general. Suppose we try to find the energy density of the ground state ρ_0, for which

$$H|0\rangle = \int d^3x \rho_0 |0\rangle = \int \frac{d^3k}{(2\pi)^3} \omega_{\mathbf{k}} a^*(\mathbf{k})a(\mathbf{k})|0\rangle + \frac{1}{2}\int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{k}}} \omega_{\mathbf{k}} (2\pi)^3\delta(\mathbf{0})|0\rangle. \tag{4.59}$$

Hence,

$$\rho_0 = \frac{1}{2}\int \frac{d^3k}{(2\pi)^3} \omega_{\mathbf{k}}. \tag{4.60}$$

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 dominant behaviour in the integral is

$$\rho_0 \sim \int^{\Lambda} dk k^3 \sim \Lambda^4. \tag{4.61}$$

that As we take Λ to infinity, 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 density of a field theory appears to 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

$$H_{\rm ren} = H - \frac{1}{2}\int d^3k \omega_{\mathbf{k}} \delta(\mathbf{0}) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{k}}} \omega_{\mathbf{k}} a^*(\mathbf{k}) a(\mathbf{k}). \tag{4.62}$$

This has the sensible property that $\hat{H}_{\rm ren}|0\rangle = 0$ – 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

$$H_{\rm ren} = \int \frac{d^3k}{(2\pi)^3} \omega_{\mathbf{k}} n(\mathbf{k}), \tag{4.63}$$

and use Eq. (4.51).

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

$$:a(\mathbf{k})a^*(\mathbf{k}'): = a^*(\mathbf{k}')a(\mathbf{k}). \tag{4.64}$$

Hence,

$$:H: = H_{\rm ren}. \tag{4.65}$$

4.4 Particle interpretation of states

We shall finish this section by returning to the justification for interpreting the states of the field as multiparticle states. Let us first compute the energy of a one-particle state:

$$H_{\rm ren}|\mathbf{k}\rangle = \int \frac{d^3k'}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{k}'}} \omega_{k'} \hat{n}(\mathbf{k}')|\mathbf{k}\rangle = \omega_{\mathbf{k}}|\mathbf{k}\rangle. \tag{4.66}$$

Hence the energy of the state we have decided is a one-particle state is $\omega_{\mathbf{k}}$, which we recall is equal to $\sqrt{(\mathbf{k}^2 + m^2)}$, the energy of a relativistic particle of mass m and momentum $\mathbf{k}$. Thus in order to strengthen our case we should check that the momentum of the state $|\mathbf{k}\rangle$ is indeed $\mathbf{k}$.

The momentum operator in field theory is

$$\mathbf{P} = \int d^3x\, \dot{\hat{\phi}}(x) \boldsymbol{\nabla} \phi(x) = \int \frac{d^3k'}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{k}'}} \mathbf{k} \hat{n}(\mathbf{k}). \tag{4.67}$$

There is in fact no need to normal order the momentum operator, as the momentum of the vacuum is zero as a result of its translational invariance. Thus we may evaluate the momentum of the state $|\mathbf{k}\rangle$ through

$$\mathbf{P}|\mathbf{k}\rangle = \int \frac{d^3k'}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{k}'}} \mathbf{k}' \hat{n}(\mathbf{k}')|\mathbf{k}\rangle = \mathbf{k}|\mathbf{k}\rangle. \tag{4.68}$$

Thus the state $|\mathbf{k}\rangle$ is an eigenstate of the momentum operator, with eigenvalue $\mathbf{k}$, properties we expect from a one-particle state.

The state is also rotationally invariant (a property following from the rotational invariance of the scalar field), and hence describes a spinless particle. Lastly, multiparticle states automatically have the correct Bose symmetry. Consider a state of two particles with momenta $\mathbf{k}_1$ and $\mathbf{k}_2$. Then

$$|\mathbf{k}_1\, \mathbf{k}_2\rangle = a^*(\mathbf{k}_1) a^*(\mathbf{k}_2)|0\rangle a^*(\mathbf{k}_2) a^*(\mathbf{k}_1)|0\rangle = |\mathbf{k}_2\, \mathbf{k}_1\rangle, \tag{4.69}$$

which shows that the quantum state is symmetric under the exchange of particles. This property stems from the canonical commutation relations of the field operators.

4.5 The Electromagnetic Field

In order to quantise the electromagnetic field we shall follow exactly the same procedure as we did for the real scalar field. We shall find the dynamical coordinates, identify the conjugate momentum, and impose canonical commutation relations between them. However, 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

$$\pi^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_\mu)} = F^{0\mu}. \tag{4.70}$$

But this means that A_0 does not have a canonical momentum, as π^0 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

$$-\nabla^2 A_0 = \rho, \tag{4.71}$$

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 $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. In the Coulomb gauge, this action is equivalent to

$$S_{\text{Cg}} = \int d^4x\, \mathcal{L}_{\text{Cg}} = \int d^4x\, \left(\frac{1}{2} \dot{A}^i \dot{A}^i - \frac{1}{2} \partial_j A^i \partial_j A^i \right). \tag{4.72}$$

Thus the momentum conjugate to A^i is

$$\pi_i = \frac{\partial \mathcal{L}_{\text{Cg}}}{\partial \dot{A}^i} = \dot{A}^i = -E^i, \tag{4.73}$$

and the Hamiltonian density is

$$\mathcal{H}_{\text{Cg}} = \pi_i \dot{A}^i - \mathcal{L}_{\text{Cg}} = \frac{1}{2} \dot{A}^i \dot{A}^i + \frac{1}{2} \partial_j A^i \partial_j A^i. \tag{4.74}$$

At this point it would seem natural to impose a commutation relation [3]

$$[A^i(\mathbf{x}, t), \pi_j(\mathbf{x}', t)] \stackrel{?}{=} i \delta_{ij} \delta(\mathbf{x} - \mathbf{x}'), \tag{4.75}$$

based on our experience with scalar fields. This is *incorrect*, as differentiating both sides with respect to x^i quickly shows. The gauge condition $\partial_i A^i = 0$ implies that the left hand side vanishes. However,

$$[\partial_i A^i(\mathbf{x}, t), \pi_j(\mathbf{x}', t)] = i \partial_j \delta(\mathbf{x} - \mathbf{x}') \neq 0. \tag{4.76}$$

The derivative of a δ-function is not zero. So we need a more general function $\Delta_{ij}(\mathbf{x}, \mathbf{x}')$ for the right hand side of the CCR, which satisfies $\partial_i \Delta_{ij}(\mathbf{x}, \mathbf{x}') = 0$ (and also $\partial'_j \Delta_{ij}(\mathbf{x}, \mathbf{x}') = 0$, as one can easily check by substituting for π_j in the commutation relations). We also know that the dependence on $\mathbf{x}$ and $\mathbf{x}'$ must be in the combination $\mathbf{x} - \mathbf{x}'$ through translational invariance.

[3] Note that as this is not a covariant equation, there is no need for covariant and contravariant indices to balance on both sides of the equation.

It is easier to find the function $\Delta_{ij}(\mathbf{x}-\mathbf{x}')$ in Fourier space: we write

$$\Delta_{ij}(\mathbf{x}-\mathbf{x}') = \int \frac{d^3k}{(2\pi)^3}\tilde{\Delta}_{ij}(\mathbf{k})e^{i\mathbf{k}\bullet(\mathbf{x}-\mathbf{x}')}. \tag{4.77}$$

Then

$$\partial_i\Delta_{ij}(\mathbf{x}-\mathbf{x}') = i\int \frac{d^3k}{(2\pi)^3}k^i\tilde{\Delta}_{ij}(\mathbf{k})e^{i\mathbf{k}\bullet(\mathbf{x}-\mathbf{x}')} = 0, \tag{4.78}$$

from which we infer that

$$k^i\tilde{\Delta}_{ij}(\mathbf{k}) = 0. \tag{4.79}$$

The solution to this equation is

$$\tilde{\Delta}_{ij}(\mathbf{k}) = \left(\delta_{ij} - \frac{k^ik^j}{\mathbf{k}^2}\right)\tilde{F}(\mathbf{k}), \tag{4.80}$$

where $F(\mathbf{k})$ is some function of $\mathbf{k}$. The correct function is just $F = 1$: this keeps the function $\Delta_{ij}(\mathbf{x}-\mathbf{x}')$ as close as possible to the δ-function. Thus the canonical commutation relations in the Coulomb gauge are

$$[A^i(\mathbf{x},t), E^j(\mathbf{x}',t)] = -i\int \frac{d^3k}{(2\pi)^3}\left(\delta_{ij} - \frac{k^ik^j}{\mathbf{k}^2}\right)\tilde{e}^{i\mathbf{k}\bullet(\mathbf{x}-\mathbf{x}')}. \tag{4.81}$$

This we can write

$$[A^i(\mathbf{x},t), E^j(\mathbf{x}',t)] = -i\left(\delta_{ij} - \frac{\partial_i\partial_j}{\nabla^2}\right)\delta(\mathbf{x}-\mathbf{x}'), \tag{4.82}$$

where ∇^{-2} is the inverse of $\nabla^2 = \partial_i\partial_i$, and can be thought of as being *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 that function whose Fourier transform is $\mathbf{k}^{-2}\tilde{f}(\mathbf{k})$. One can now verify that the Coulomb gauge commutation relations 4.82 are now consistent with the gauge condition $\partial_iA^i = 0$.

As we saw in Section 2.3.1, 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

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

where $k_ia^i = 0$. Thus the three components of the vector amplitude $a^i(\mathbf{k})$ are not independent: they satisfy a constraint arising from the Coulomb gauge condition.

Rather than worry about this constraint all the time, it is much more convenient to have two freely chosen functions of $\mathbf{k}$, so we instead write the general solution

$$A^i(t,\mathbf{x}) = \sum_{\lambda=1,2} \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{k}}}} \left(a_{(\lambda)} \epsilon^i_{(\lambda)} e^{-ik\cdot x} + a^*_{(\lambda)} \epsilon^*_{(\lambda)}{}^i e^{ik\cdot x} \right). \tag{4.83}$$

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

$$k^i \epsilon^i_{(\lambda)} = 0, \quad \epsilon^i_{(\lambda)} \epsilon^i_{(\lambda')} = \delta_{\lambda\lambda'}. \tag{4.84}$$

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:

$$\epsilon^i_{(1)} \epsilon^j_{(1)} + \epsilon^i_{(2)} \epsilon^j_{(2)} + \hat{k}^i \hat{k}^j = \delta_{ij}, \tag{4.85}$$

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 $a^*_{(\lambda)}(\mathbf{k})$:

$$[a_{(\lambda)}(\mathbf{k}), a^*_{(\lambda')}(\mathbf{k}')] = \delta_{\lambda\lambda'} (2\pi)^3 \delta(\mathbf{k} - \mathbf{k}'). \tag{4.86}$$

Thus each polarisation behaves like an independent scalar field.

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

$$a_{(\lambda)}(\mathbf{k})|0\rangle = 0, \quad \forall \mathbf{k}, \lambda. \tag{4.87}$$

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

$$|\mathbf{k}, \lambda\rangle = a^*_{(\lambda)}(\mathbf{k})|0\rangle. \tag{4.88}$$

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:

$$:H: = \int d^3x \, :\mathcal{H}: = \int \frac{d^3k}{(2\pi)^3} \sum_\lambda \omega_{\mathbf{k}} a^*_{(\lambda)}(\mathbf{k}) a_{(\lambda)}(\mathbf{k}). \tag{4.89}$$

This is just two copies of the renormalised scalar field Hamiltonian (4.62), one for each polarisation.

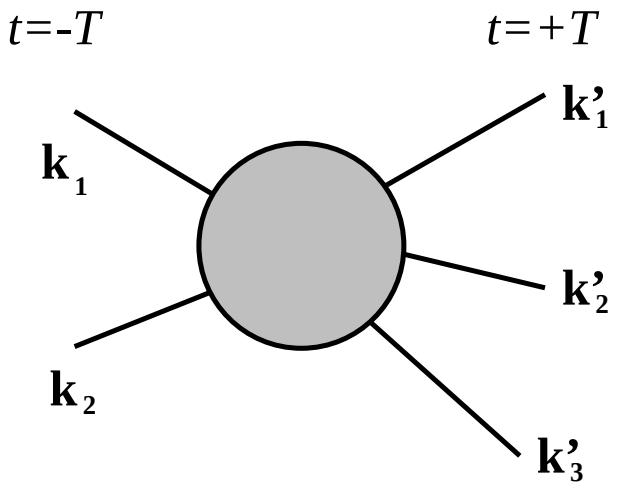


Figure 6.1: A scattering experiment. A set of particles with momenta $\mathbf{k}_1$, $\mathbf{k}_2$ come in from a large separation a long time in the past, interact, and head out towards infinity again in a different state, where they have different momenta. There may also be a different number of particles of various species in the final state. infinity, interact,

5 Path integrals

6 Interacting fields

So far we have been dealing with *free* field theories, which we have been able to solve exactly. Thus we have found the equations of motion for the field operators, solved them in terms of a sum of operator-valued Fourier modes ($a(\mathbf{k})$, $b(\mathbf{k})$ and their complex conjugates) and found all the possible quantum states. The dynamics of a free field theory are therefore rather trivial: one decides which state (or superposition of states) the field is in at some initial time, and the field then remains in this state for all subsequent times.

We would like to able to describe more realistic situations in which the field changes its state. Field theory was developed to describe the result of scattering experiments in which a few particles head towards each other from a long distance, interact, and the resulting particles (there may be more than the incoming set) then head back to infinity again (see Figure 6.1). The goal is to calculate the quantum mechanical amplitude for the initial state to change in to the final state. From this we can calculate the transition probability, which is usually expressed as a *cross-section*.

Unfortunately, it is almost always impossible to calculate these probabilities

exactly, and one must use a perturbation expansion in some small number which parametrises the strength of the interaction – a *coupling constant*. For example, q, the charge of the complex scalar field of the previous section, is such an expansion parameter, which we must assume to be small. Most of the technical complexity of quantum field theory is due to the difficulty of writing this expansion in a manageable way, and in then in making sense of the expressions once they have been written down. The breakthrough that made field theory calculations feasible was Feyman's realisation that the perturbation expansion could be written down in a graphical way with *Feynman diagrams*. The procedure of making sense of the result, which is formally infinite, is *renormalisation*.

6.1 The interaction picture

In order to develop this expansion we must first recall some results from ordinary time-dependent perturbation theory in quantum mechanics. There are three ways of looking at the time evolution of a state in quantum mechanics, which are equivalent in the sense that all operators (which correspond to observable quantities) have the same matrix elements. These ways are called "pictures".

Schrödinger picture This is probably the most familiar picture. The time dependence of the system is carried in the states, while the operators O_S are time-independent. The time-dependent states $|\psi(t)\rangle_S$ obey the Schrödinger equation

$$i\frac{d}{dt}|\psi(t)\rangle_S = H|\psi(t)\rangle_S, \tag{6.1}$$

where H is the Hamiltonian.

Heisenberg picture In this picture, it is 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

$$i\frac{d}{dt}O_H(t) = [O_H(t), H]. \tag{6.2}$$

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

$$\begin{aligned} O_H(t) &= e^{iHt}O_S e^{-iHt} \\ |\psi\rangle_H &= = e^{iHt}|\psi(t)\rangle_S. \end{aligned} \tag{6.3}$$

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:

$$\begin{aligned} {}_H\langle\psi|O_H(t)|\psi'\rangle_H &= {}_S\langle\psi(t)|e^{-iHt}e^{iHt}O_S e^{-iHt}e^{iHt}|\psi'(t)\rangle_S \\ &= {}_S\langle\psi(t)|O_S|\psi'(t)\rangle_S. \end{aligned} \tag{6.4}$$

This demonstrates the point.

Interaction picture This picture is the one suited for situations in which we can't solve the Schrödinger (or Heisenberg) equation exactly for our Hamiltonian H, but we can split it into two parts, one solvable with Hamiltonian H_0, and the other H' which is small compared to H_0:

$$H = H_0 + H'. \tag{6.5}$$

In the interaction picture, the time dependence is divided between the operators and the states. The operators evolve according to the Heisenberg equations for the solvable Hamiltonian H_0,

$$O_I(t) = e^{iH_0t}O_S e^{-iH_0t}, \tag{6.6}$$

which means that the states must be related to the Schrödinger picture states by

$$i\frac{d}{dt}|\psi(t)\rangle_I = e^{iH_0t}|\psi(t)\rangle_S. \tag{6.7}$$

Thus by differentiating both sides with respect to time,

$$i\frac{d}{dt}|\psi(t)\rangle_I = e^{iH_0t}(-H_0 + H)|\psi(t)\rangle_S = (e^{iH_0t}H'e^{-iH_0t})e^{iH_0t}ket\psi(t)_S. \tag{6.8}$$

The piece is brackets o the right hand side is just the interaction picture representation of the interaction Hamiltonian, which we write $H'(t)_I$. Thus

$$i\frac{d}{dt}|\psi(t)\rangle_I = H'(t)_I|\psi(t)\rangle_I. \tag{6.9}$$

Thus the interaction picture states evolve according to the non-trivial part of the Hamiltonian only. In the limit $H' \to 0$, the interaction picture reduces to the Hesinberg picture.

The interaction picture state evolution equation has a formal solution which is

$$|\psi(t)\rangle_I = U(t,t_0)|\psi(t_0)\rangle_I, \tag{6.10}$$

where $U(t,t_0)$ is an operator which satisfies the equation

$$i\frac{\partial}{\partial t}U(t,t_0) = H'(t)_I U(t,t_0), \tag{6.11}$$

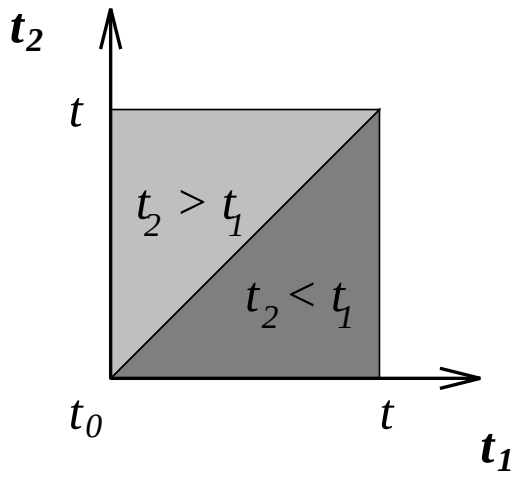


Figure 6.2: The simplification of the integration region in Equation (6.14). The orginal integration is over the darker shaded area, but by using the symmetry of the integrand and careful time ordering the region can be extended over the entire shaded area.

with the boundary condition $U(t,t_0) = 1$, the identity operator. We can convert this into an integral equation

$$U(t,t_0) = 1 - i\int_{t_0}^{t} dt_1 H'_I(t_1)U(t_1,t_0), \tag{6.12}$$

which we solve by iteration. The starting point is to take $U(t,t_0) = 1$. We substitute this back into the integral equation, from which we obtain

$$U^{(1)}(t,t_0) = 1 - i\int_{t_0}^{t} dt_1 H'_I(t_1). \tag{6.13}$$

Substituting again, we get

$$\begin{aligned} U^{(2)}(t,t_0) &= 1 - i\int_{t_0}^{t} dt_1 H'_I(t_1)U^{(1)}(t_1,t_0) \\ &= 1 + (-i)^2\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 H'_I(t_1)H'_I(t_2). \end{aligned} \tag{6.14}$$

We can simplify the integration region by exploiting the symmetry of the integrand, and by using the time ordered product introduced in Section (**??**). Note first that we are integrating over the darker shaded area in Figure (6.2). Let us now rewrite the second order term in Equation (6.14) as

$$\frac{(-i)^2}{2}\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 H'_I(t_1)H'_I(t_2) + \frac{(-i)^2}{2}\int_{t_0}^{t} dt_2 \int_{t_0}^{t_2} dt_1 H'_I(t_2)H'_I(t_1). \tag{6.15}$$

All we have done here is to divide the integration into two and interchange the integration variables in the second term. In doing so, we can see that the second term is effectively an integration over the lighter shaded area. In both terms the earlier time appears to the right. Thus we may write this as

$$\frac{(-i)^2}{2}\int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 T\left[H_I'(t_1)H_I'(t_2)\right]. \tag{6.16}$$

It is not too hard to satsify oneself that the nth term in this iteration involves the time ordered product of n copies of the Hamiltonian, with a symmetry factor of $1/n!$. Hence

$$U(t,t_0) = \sum_n \frac{(-i)^2}{n!}\int_{t_0}^{t} dt_1 \ldots \int_{t_0}^{t} dt_n T\left[H_I'(t_1)\ldots H_I'(t_n)\right]. \tag{6.17}$$

This can be written more compactly as

$$U(t,t_0) = T\left[\exp\left(-i\int_{t_0}^{t} dt' H_I'(t')\right)\right], \tag{6.18}$$

where the operation of time ordering on an exponential is defined by its operation on the individual terms in the Taylor series.

6.2 The S-matrix and transition amplitudes

In principle, the matrix $U(t,t_0)$, whose formal solution we found in the previous section, completely determines the time evolution of the system. In scattering processes, we are generally not interested in the intermediate times when the particles are interacting, only in how the states in the distant past change into other states in the future. We assume that interactions are localised in space and time, which amounts to assuming that $H_I'(t) \to 0$ as $t \to \pm\infty$. Hence

$$|\Psi(t)\rangle_I \to |\Psi(\pm\infty)\rangle_I, \tag{6.19}$$

which are states of the free Hamiltonian H_0. We define the *S-matrix* to be

$$S = \lim_{t\to+\infty t_0\to-\infty} U(t,t_0), \tag{6.20}$$

so that

$$|\Psi(+\infty)\rangle_I = S|\Psi(-\infty)\rangle_I. \tag{6.21}$$

We call states $|\Psi(-\infty)\rangle_I$ *in states*, and $|\Psi(+\infty)\rangle_I$ *out states*. In order to reduce the amount of typing, they will be denoted $|i\rangle$ and $|f\rangle$ respectively.

In a scattering experiment, we prepare the system in a state $|\Psi(t)\rangle_I$ at $t \to -\infty$, and we would like to be able to compute the amplitude for a quantum

mechanical transition to any other state $|\Phi(t)\rangle$ as $t \to +\infty$. The probability amplitude for this transition is

$$A_{\Phi\Psi} = \lim_{t\to+\infty} {}_I\langle\Phi(t)|\Psi(t)\rangle_I = {}_I\langle\Phi(+\infty)|S|\Psi(-\infty)\rangle_I. \tag{6.22}$$

Simplifying the notation, the probabilty amplitude for the transition from an in state $|i\rangle$ to an out-state $|f\rangle$ is

$$A_{fi} = \langle f|S|i\rangle \equiv S_{fi}. \tag{6.23}$$

That is, the matrix elements of the S-matrix give the transition probability amplitudes between initial and final states.

The S-matrix has a number of very important properties, the first of which is that probability must be conserved. That is, if we prepare an in-state with unit probability, the sum of all the probabilities of the out-states must also be unity. Thus we are assuming that $\langle i|i\rangle = 1$, and that this normalisation is preserved throughout the evolution: that is, ${}_I\langle\Psi(+\infty)|\Psi(+\infty)\rangle_I = 1$ also. Hence

$$\langle i|S^\dagger S|i\rangle = 1. \tag{6.24}$$

Let us insert a complete set of out-states:

$$1 = \sum_f |f\rangle\langle f|, \tag{6.25}$$

so that

$$\sum_f \langle i|S^\dagger|f\rangle\langle f|S|i\rangle = 1, \tag{6.26}$$

or

$$\sum_f S^*_{if} S_{fi} = 1. \tag{6.27}$$

Thus the conservation of probability demands that the S-matrix be unitary ($S^\dagger S = 1$).

Finally for this section it is important to point out a major problem with this approach to dealing with interactin field theories. The interaction picture Hamiltonian $H'_I(t)$ does *not* become negligible as $t \to \pm\infty$, because particles are always interacting with themselves. The result of these self-interactions is to shift the mass and charge of the particle by an infinite amount, although this only happens at second order in the perturbation expansion for S. These infinite shifts must be dealt with by the procedure of *renormalisation*.

A Problem Sheets

A.1 Problem Sheet 1

1. In the following, x^μ is a 4-vector, and $\partial_\mu \equiv \frac{\partial}{\partial x^\mu}$.

   (a) Show that

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

   and show that $\partial \cdot x = 4$.

   (b) Show that, if $\phi(x) = Ae^{-ip.x}$, with p a constant 4-vector, then $(\partial^2 + m^2)\phi(x) = (-p^2 + m^2)\phi(x)$.

   (c) Given the function $f(x) = e^{-\frac{1}{2}ax^2}$, where a is a constant scalar, calculate $\partial_\mu \partial_\nu f(x)$ and show that $\partial^2 f(x)$ vanishes on the curve $x^2 = 4/a$.

2. (a) Let $\psi(\mathbf{x}, t)$ be a wave function satisfying the Schrödinger equation

   $$i\hbar \frac{\partial}{\partial t}\psi = -\frac{\hbar^2}{2m}\nabla^2\psi + V(\mathbf{x})\psi.$$

   Show that the probability density $\rho = |\psi|^2$ and the probability current $\mathbf{j} = -i(\psi^* \boldsymbol{\nabla}\psi - \psi \boldsymbol{\nabla}\psi^*)/2m$ satisfy a continuity equation

   $$\dot{\rho} + \boldsymbol{\nabla}\cdot\mathbf{j} = 0.$$

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

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

   Show that the 4-vector current

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

   satisfies a covariant continuity equation $\partial \cdot J = 0$. Give a reason why J^μ can not be a probability current.

3. 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

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

   Note that λ and μ^2 are both real and positive.

(a) Write down the Euler-Lagrange equation for a real scalar field, and hence derive the field equation obeyed by ϕ.

(b) Find the solutions with ϕ constant, and show also that

$$\phi(z) = v \tanh(mz)$$

is a solution, provided $v = \sqrt{\mu^2/\lambda}$ and $m = \mu/\sqrt{2}$.

(c) By considering the properties of ϕ under a Lorentz transformation, or by explicit evaluation of the field equation, show that

$$\phi(t, z) = v \tanh[m\gamma(z - vt)],$$

with $\gamma = \sqrt{(1 - v^2)}$, is also a solution,.

4. (a) Show that $\epsilon_{ijk}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$, and use this relation to show that $\nabla \times \mathbf{B} = -\nabla^2\mathbf{A} + \nabla(\nabla \cdot \mathbf{A})$. (The components of $\nabla \times \mathbf{B}$ in terms of B_i, the components of $\mathbf{B}$, are $(\nabla \times \mathbf{B})_i = \epsilon_{ijk}\partial_j B_k$).

(b) We define the dual electromagnetic field strength tensor as

$$\tilde{F}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}F^{\rho\sigma},$$

where $\epsilon_{\mu\nu\rho\sigma}$ is the 4-dimensional Levi-Civita tensor. Show that $\tilde{F}_{0j} = -B_j$ and $\tilde{F}_{ij} = \epsilon_{ijk}E_k$, and hence verify that the 4-vector equation

$$\partial^\nu \tilde{F}_{\mu\nu} = 0$$

correctly reproduces the two homogeneous Maxwell equations.

A.2 Problem Sheet 2

1. A real scalar field has Lagrangian density

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

where $V(\phi)$ is bounded from below.

   (a) Calculate the canonical energy-momentum tensor $\theta^{\mu\nu}$ from this Lagrangian density, using the formula derived in the lectures.

   (b) Calculate the Hamiltonian density $\mathcal{H}$, and verify that it is equal to the energy density $\theta^{00}(x)$.

   (c) Verify, using the field equation $\partial^2\phi + V'(\phi) = 0$, that the energy-momentum tensor is conserved (satisfies $\partial_\mu\theta^{\mu\nu} = 0$).

2. (a) Consider a real scalar field whose Lagrangian is of the same form as the Problem 1, with $V(\phi) = \frac{1}{2}m^2\phi^2$. Find the Hamiltonian operator $\hat{H}$, and show that the Heisenberg equations of motion

$$\dot{\hat{\phi}}(x) = \frac{1}{i}\left[\hat{\phi}(x), \hat{H}\right], \qquad \dot{\hat{\pi}}(x) = \frac{1}{i}\left[\hat{\pi}(x), \hat{H}\right],$$

   imply that $(\partial^2 + m^2)\hat{\phi}(x) = 0$.

   (b) Write down the Lorentz transformation which takes the 4-vector $k^\mu = (\omega_{\mathbf{k}}, k^i)$, where $\omega_{\mathbf{k}} = |(\mathbf{k}^2 + m^2)^{\frac{1}{2}}|$, to a frame moving with velocity v in the x direction relative to the original.

   (c) By applying the Lorentz transformation of the previous part, or otherwise, show that the momentum integration measure $d^3k/(2\pi)^3 2\omega_{\mathbf{k}}$ is Lorentz invariant,

3. The real scalar field operator is expanded in terms of ladder operators $\hat{a}_{\mathbf{k}}$ and $\hat{a}^*_{\mathbf{k}}$ as

$$\hat{\phi}(x) = \int \frac{đ^3k}{2\omega_{\mathbf{k}}} \left(\hat{a}_{\mathbf{k}} f_{\mathbf{k}}(x) + \hat{a}^*_{\mathbf{k}} f^*_{\mathbf{k}}(x)\right),$$

where $f_{\mathbf{k}}(x) = e^{-ik\cdot x}$, with $k^0 = \omega_{\mathbf{k}} \equiv |(\mathbf{k}^2 + m^2)^{\frac{1}{2}}|$.

(a) Show that the functions $f_{\mathbf{k}}(x)$ obey orthogonality requirements

$$\begin{aligned} i\int d^3x \ (f^*_{\mathbf{k}}(x)\partial_0 f_{\mathbf{k}'}(x) - \partial_0 f^*_{\mathbf{k}}(x)\, f_{\mathbf{k}'}(x)) &= 2\omega_{\mathbf{k}}\, \delta^3(\mathbf{k}-\mathbf{k}'), \\ i\int d^3x \ (f_{\mathbf{k}}(x)\partial_0 f_{\mathbf{k}'}(x) - \partial_0 f_{\mathbf{k}}(x)\, f_{\mathbf{k}'}(x)) &= 0. \end{aligned}$$

(b) Show that

$$\hat{a}_{\mathbf{k}} = i\int d^3x \ \left(f^*_{\mathbf{k}}(x)\hat{\pi}(x) - \partial_0 f^*_{\mathbf{k}}(x)\hat{\phi}(x)\right),$$

where $\hat{\pi}(x) = \partial_0\hat{\phi}(x)$ is the canonical momentum operator.

(c) Show that the non-trivial commutation relation for the ladder operators is

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

4. The number operator for a real scalar field is defined as

$$\hat{N} = \int \frac{đ^3k}{2\omega_{\mathbf{k}}} \hat{a}^*_{\mathbf{k}}\hat{a}_{\mathbf{k}},$$

where $\hat{a}^*_{\mathbf{k}}$ and $\hat{a}_{\mathbf{k}}$ are the ladder operators, which obey the commutation relation of Problem 3.

(a) Show that $[\hat{N}, (\hat{a}^*_{\mathbf{k}})^n] = n(\hat{a}^*_{\mathbf{k}})^n$.

(b) Hence, or otherwise, show that the state

$$|\{n_{\mathbf{k}}\}\rangle = \prod_{\mathbf{k}} \frac{(\hat{a}^*_{\mathbf{k}})^{n_{\mathbf{k}}}}{\sqrt{n_{\mathbf{k}}!}}|0\rangle$$

is an eigenstate of the number operator with eigenvalue

$$N = V\int \frac{d^3k}{(2\pi)^3} n_{\mathbf{k}},$$

where V is a formal volume factor.

NB The product over the continuous variable $\mathbf{k}$ is defined as

$$\prod_{\mathbf{k}} = \lim_{L\to\infty} \prod_{\mathbf{m}},$$

where $\mathbf{k} = 2\pi\mathbf{m}/L$, and $\mathbf{m} = (m^1, m^2, m^3)$, with m^1, m^2, m^3 integers.

A.3 Problem Sheet 3

1. Consider the one-parameter family of Lagrangians

$$\mathcal{L}_\zeta = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{\zeta}{2}(\partial \cdot A)^2.$$

   (a) Find the field equations for the gauge potential A^μ which result from this family.

   (b) Show that they all give the same field equations in the Lorentz gauge. What is special about the field equations when $\zeta = 1$?

   (c) Show that the Lagrangian

$$\mathcal{L} = -\frac{1}{2}\partial_\mu A^\nu \partial^\mu A_\nu$$

   gives the same field equations as $\mathcal{L}_1$, and find $\mathcal{L} - \mathcal{L}_1$. Why is it that $\mathcal{L}$ and $\mathcal{L}_1$ give the same field equations?

2. (a) Starting with the Lagrangian $\mathcal{L} = -\frac{1}{2}\partial_\mu A^\nu \partial^\mu A_\nu$, show that the associated Hamiltonian density $\mathcal{H}$ is

$$\mathcal{H} = -\frac{1}{2}\dot{A}_\mu \dot{A}^\mu - \frac{1}{2}\nabla A_\mu \nabla A^\mu.$$

   (b) What is strange about this Hamiltonian density?

   (c) Given the expansion of the electromagnetic field operator in a plane wave basis

$$\hat{A}^\mu(x) = \int \frac{đ^3k}{2\omega_{\mathbf{k}}} \left(\hat{a}^\mu_{\mathbf{k}} e^{-ik\cdot x} + \hat{a}^{\dagger\mu}_{\mathbf{k}} e^{ik\cdot x}\right),$$

   with $k^0 = \omega_{\mathbf{k}} = (\mathbf{k}^2 + m^2)^{\frac{1}{2}}$, show that

$$:H: = \int d^3x :\mathcal{H}: = -\int \frac{đ^3k}{2\omega_{\mathbf{k}}} \omega_{\mathbf{k}} \hat{a}^{\dagger\mu}_{\mathbf{k}} \hat{a}_{\mathbf{k}\mu}.$$

   (d) Re-expressing the ladder operators in a polarisation basis $\hat{a}^\mu_{\mathbf{k}} = \epsilon^A_{\mathbf{k}} \hat{a}_{\mathbf{k}A}$, show that if $|\phi\rangle$ is a physical state, then

$$:H:|\phi\rangle = \int \frac{đ^3k}{2\omega_{\mathbf{k}}} \omega_{\mathbf{k}} \left(a_{\mathbf{k}1} a_{\mathbf{k}1} + a_{\mathbf{k}2} a_{\mathbf{k}2}\right) |\phi\rangle.$$

   Comment on this result in the light of Question 2b.

3. The 4-momentum operator for a real scalar field obeying the Klein-Gordon equation is $\hat{P}^\mu$, with

$$\hat{P}^0 = \frac{1}{2}\int d^3x \left(\pi^2 + \nabla\phi^2 + m^2\phi^2\right), \qquad \hat{P}^i = \int d^3x \pi \partial^i \phi.$$

(a) Show that $[\hat{P}^\mu, \hat{\phi}(x)] = -i\partial^\mu \hat{\phi}(x)$

(b) Using the Baker-Campbell-Hausdorff formula

$$e^{\hat{A}}\hat{B}e^{-\hat{A}} = [\hat{A}, \hat{B}] + \frac{1}{2}[\hat{A}, [\hat{A}, \hat{B}]] + \frac{1}{3!}[\hat{A}, [\hat{A}, [\hat{A}, \hat{B}]]] + \ldots$$

show that, if a^μ is a constant 4-vector,

$$e^{ia\cdot\hat{P}}\hat{\phi}(x)e^{-ia\cdot\hat{P}} = \hat{\phi}(x+a).$$

What does this tell us about the 4-momentum operator $\hat{P}^\mu$?

4. The scalar Feynman propagator is defined as the vacuum expectation value of the time-ordered product of two fields

$$\begin{aligned} i\Delta_F(x-y) &= \langle 0|T[\hat{\phi}(x)\hat{\phi}(y)]|0\rangle \\ &= \langle 0|\hat{\phi}(x)\hat{\phi}(y)|0\rangle\theta(x^0-y^0) + \langle 0|\hat{\phi}(y)\hat{\phi}(x)|0\rangle\theta(y^0-x^0). \end{aligned}$$

We split the field into its positive and negative frequency components $\hat{\phi}(x) = \hat{\phi}^+(x) + \hat{\phi}^-(x)$, where

$$\hat{\phi}^+(x) = \int \frac{đ^3k}{2\omega_{\mathbf{k}}} a_{\mathbf{k}} e^{-ik\cdot x}, \qquad \hat{\phi}^-(x) = \int \frac{đ^3k}{2\omega_{\mathbf{k}}} a^\dagger_{\mathbf{k}} e^{ik\cdot x}$$

(a) Defining $\Delta^+(x) = \langle 0|\hat{\phi}^+(x)\hat{\phi}^-(0)|0\rangle$, $\Delta^-(x) = \langle 0|\hat{\phi}^+(0)\hat{\phi}^-(x)|0\rangle$, find the Feynman propagator $i\Delta_F(x)$ in terms of $\Delta^+(x)$ amd $\Delta^-(x)$.

(b) Express the functions $\Delta^+(x)$ and $\Delta^-(x)$ in terms of integrals over wave vectors $\mathbf{k}$ and show that they are Lorentz invariant.

(c) Show that $\Delta(x) = \langle 0|[\hat{\phi}(x), \hat{\phi}(0)]|0\rangle$ vanishes for $x^2 < 0$. *[Hint: take $x^0 = 0$]*

(d) Discuss the interpretation of the Feynman propagator $i\Delta_F(x-y)$ as the probability amplitude for a particle to propagate from y to x.

A.4 Problem Sheet 4

1. In the following, $A(x)$, $B(x)$, $C(x)$, and $D(x)$ are field operators.

   (a) Show that $:A(x_1)B(x_2):\, = \,:B(x_2)A(x_1):$.

   (b) Show that $:A(x_1)B(x_2):\, = 0$.

   (c) Use Wick's theorem to expand $T[A(x_1)B(x_2)C(x_3)D(x_4)]$ in terms of contractions and normal ordered products.

   (d) How many ways are there of taking c contractions of n operators?

2. Consider a real scalar field theory with potential $V(\phi) = \frac{1}{2}\phi^2 + \frac{1!}{4!}\lambda\phi^4$. A particular scattering process has four particles in the in-state, with 3-momenta $\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}_3$, and $\mathbf{k}_4$, and two particles in the out-state, with 3-momenta $\mathbf{k}'_1$ and $\mathbf{k}'_2$. It is assumed that none of the momenta are equal.

   The S-matrix element is written

   $$\langle \mathbf{k}'_1, \mathbf{k}'_2 | S | \mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4 \rangle .$$

   (a) Write down the Dyson-Wick expansion for the S operator, and show that the $\mathrm{O}(\lambda^0)$ and $\mathrm{O}(\lambda^1)$ terms in the corresponding expansion of the matrix element vanish.

   (b) Apply Wick's theorem to $T[:\phi^4(x_1)::\phi^4(x_2):]$ and hence calculate the $\mathrm{O}(\lambda^2)$ contribution to the matrix element. You may leave your answer in the form of a double integral over x_1 and x_2.

3. (a) Write down the definition of the functional derivative of a functional $F[f]$, defined on functions of a single real variable $f(x)$. Show directly from this definition that

$$\frac{\delta f(x)}{\delta f(y)} = \delta(x-y)$$

(b) For the functional

$$F[f] = \exp\left(\int dx\, b(x) f(x)\right)$$

find $\frac{\delta F[f]}{\delta f(x)}$, and show that

$$\left.\frac{\delta^n F[f]}{\delta f(x_1)\delta f(x_2)\ldots\delta f(x_n)}\right|_{f=0} = b(x_1)b(x_2)\ldots b(x_n),$$

where the notation means that we evaluate the result of taking the functional derivatives at $f(x) = 0$.

(c) Consider the functional

$$D[f] = \exp\left(\int dx\, b(x) f'(x)\right),$$

where $f'(x) = df(x)/dx$. Show that $\left.\frac{\delta D[f]}{\delta f(x)}\right|_{f=0} = -b'(x)$.

4. Suppose the action for a particle of unit mass in the presence of a force $\hbar j(t)$ is

$$S_j[q] = \int dt\, \left(\frac{1}{2}\dot{q}^2 - V(q) + \hbar j q\right).$$

(a) Find the functional derivative $\frac{\delta S[q]}{\delta q(t)}$, and show that

$$\frac{\delta^2 S[q]}{\delta q(t)\delta q(t')} = -\left(\frac{\partial^2}{\partial t^2} + V''(q)\right)\delta(t-t').$$

(b) Let $\bar{q}(t)$ be an extremum of the action. If $\eta(t) = q(t) - \bar{q}(t)$ is the deviation of the path from the extremal path, assumed small, show that we may approximate the action by

$$S_j[q] \simeq S_j[\bar{q}] + \frac{1}{2}\int dt\, \left(\dot{\eta}^2 - V''(\bar{q})\eta^2\right).$$

Under what circumstances is this expression exact?

A.5 Problem Sheet 5

1. (a) Write down the completeness relation for single-particle momentum eigenstates in one dimension $|p\rangle$, and use it to show that

$$\langle q|q'\rangle = \int \frac{dp}{2\pi\hbar} e^{ip(q-q')/\hbar},$$

where $|q\rangle$ is a one-dimensional position eigenstate. Comment on your result.

(b) The transition amplitude between two free particle position eigenstates can be written

$$\langle q,t|q',t'\rangle = \langle q|e^{-i\hat{H}(t-t')/\hbar}|q'\rangle,$$

where $\hat{H} = \frac{1}{2}\hat{p}^2$. By inserting a complete set of momentum eigenstates, show

$$K(q,t;q',t') \equiv \langle q,t|q',t'\rangle = \left(\frac{1}{2\pi i\hbar T}\right)^{\frac{1}{2}} \exp\left(\frac{i}{2\hbar}\frac{(q-q')^2}{T}\right),$$

where $T = t - t'$. Show also that $K(q,t;q',t')$ solves the free particle Schrödinger equation $(i\hbar\partial_t + \frac{1}{2}\hbar^2\partial_q^2)K(q,t;q',t') = 0$.

(c) Write down the path integral expression for $K(q,t;q',t')$ and show when the time interval is split into segments $\Delta t = T/n$ that it can be written in the form

$$K(q,t;q',t') = \lim_{n\to\infty} \int dq_1 \ldots dq_{n-1} \left(\frac{\alpha}{2\pi}\right)^{\frac{n-1}{2}} \exp\left(-\frac{\alpha}{2}\mathbf{q}^T\mathbf{K}_{n-1}\mathbf{q} + \alpha\mathbf{b}^T\mathbf{q} - c\right),$$

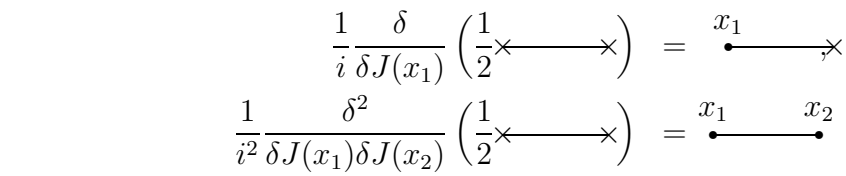
where $\mathbf{q}^T = (q_1,\ldots,q_{n-1})$, $\mathbf{b}^T = (q',0,\ldots,0,q)$, $\mathbf{K}_{n-1}$ is a tridiagonal $(n-1)\times(n-1)$ matrix, and c and $\alpha = 1/i\hbar\Delta t$ are constants.

Using the formula for a multidimensional Gaussian integral given in the lectures, show that on performing the integral and taking the limit, the same result as part (b) is obtained.

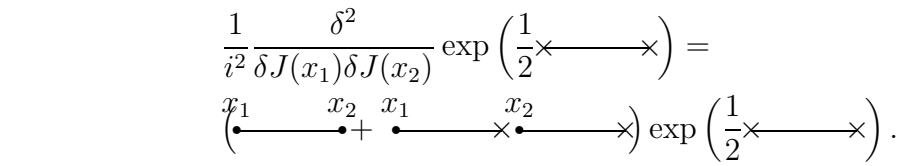
Hint 1: $\mathbf{K}_{n-1}^{-1}$ is symmetric, with $(\mathbf{K}_{n-1}^{-1})_{ij} = i(n-j)/n$, $(i<j)$.

Hint 2: To find the determinant of $\mathbf{K}$, prove and use the recursion relation $\det\mathbf{K}_n = 2\det\mathbf{K}_{n-1} - \det\mathbf{K}_{n-2}$

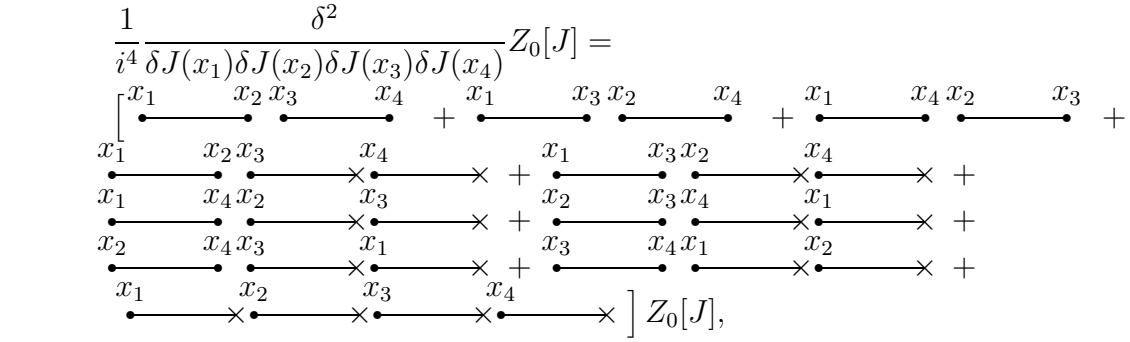
2. (a) By rewriting the following equation in terms of sources $J(x)$ and propagators $\Delta_F(x)$, verify that



(b) Show, by using the results of (2a) that



(c) The normalised generating functional for a free real scalar field is $Z[J] = \exp\left(\frac{1}{2}\times\!\!\longrightarrow\!\!\times\right)$. Show that



and hence that the 4-point function for this theory is

$$G^4(x_1,x_2,x_3,x_4) = -\Delta_F(x_1-x_2)\Delta_F(x_3-x_4)$$
$$-\Delta_F(x_1-x_3)\Delta_F(x_2-x_4) - \Delta_F(x_2-x_3)\Delta_F(x_1-x_4),$$

where $\Delta_F(x_1-x_2)$ is the Feynman propagator.

(d) Show that the odd functions $G^{2n+1}(x_1,\ldots,x_{2n+1}) = \langle 0|T[\hat{\phi}(x_1)\ldots\hat{\phi}(x_{2n+1})]|0\rangle$ all vanish.

(e) Show that

$$Z[J] = \langle 0|T\left[\exp\left(i\int d^4x J(x)\hat{\phi}(x)\right)\right]|0\rangle.$$