Part II Classical Field Theory

William Royce

September 14, 2024

Part II Physics, The University of Cambridge

Preface

Classical field theory deals with the general idea of a quantity that is a function of time and space, which can be used to describe wave-like physical phenomena such as sound and light, or other continuous phenomena such as fluid flow. The word 'classical' is here used in the sense 'not quantum mechanical'. We shall define a field to be classical if it satisfies the following criteria:

1. The state of the field at a given time is represented by furnishing, for each point in space, a finite set of numbers (e.g. a single real number or a tensor or a spinor, depending on the type of field).
2. The field can in principle be observed without disturbing it.

By contrast, a quantum field would be described by furnishing at each point in space a set of operators not numbers, and it could not in general be observed without disturbing it. It is important to maintain a tight grip on terminology here, because in many textbooks the equations described in this chapter are first introduced in the context of quantum mechanics. However I believe it is better to become acquainted with these fields in their classical guise first, and then quantise them afterwards.

In the whole of this book up to and including this chapter, all fields mentioned are classical fields. The electromagnetic field has been shown to be a tensor field. We have sometimes invoked scalar fields for illustrative purposes. In this chapter the idea of a classical spinor field will be introduced. We shall allow that high relative speeds may be involved; to make sure our results satisfy the Postulates of Relativity we shall only write Lorentz covariant field equations. The language of tensors and spinors makes this easy to accomplish.

A general discussion of classical field theory would require a text book in its own right. In this chapter the aim is to introduce some field equations that are close cousins of the wave equation. This will allow the basic principles of quantum field theory and particle physics to be opened up in the next chapter.

Contents

1 Basic Lagrangian Mechanics **1**
- 1.1 Hamilton's Principle . . . 1
- 1.2 Derivation of the Equations of Motion . . . 2
- 1.3 Symmetry and Conservation Laws; Canonical Momenta . . . 6
  - 1.3.1 Translational Invariance ↔ Conservation of Linear Momentum . . . 6
  - 1.3.2 Rotational Invariance ↔ Conservation of Angular Momentum . . . 7
  - 1.3.3 Time Invariance ↔ Conservation of Energy . . . 7

2 Hamilton's Equations of Motion **9**
- 2.1 Liouville's theorem . . . 10
- 2.2 Poisson Brackets and the Analogy with Quantum Commutators . . . 11
  - 2.2.1 Quantum Variational Principle . . . 12
  - 2.2.2 Canonical Transformations . . . 13
- 2.3 Lagrangian Dynamics of a Charged Particle . . . 14
  - 2.3.1 Gauge Invariance . . . 16
- 2.4 Relativistic Particle Dynamics . . . 17
  - 2.4.1 Relativistic Lagrangians . . . 18
  - 2.4.2 Relativistic Particle in Electromagnetic Field . . . 19
  - 2.4.3 Lagrangian vs. Hamiltonian Methods . . . 19

3 Classical Fields **21**
- 3.1 Waves in One Dimension . . . 21
- 3.2 Multidimensional Space . . . 23
- 3.3 Relativistic Scalar Field . . . 24
- 3.4 Natural units . . . 25
- 3.5 Fourier Analysis . . . 26
- 3.6 Multi-Component Fields . . . 27
  - 3.6.1 Transverse Waves on a String . . . 28
- 3.7 Complex Scalar Field . . . 28
- 3.8 Electromagnetic Field . . . 30
  - 3.8.1 Gauge Invariance . . . 32

4 Symmetries and Conservation Laws **33**
- 4.1 Noether's Theorem . . . 33
- 4.2 Global Phase Symmetry . . . 34
- 4.3 Local Phase (Gauge) Symmetry . . . 35
- 4.4 Electromagnetic Interaction . . . 36
- 4.5 Stress-Energy (-Momentum) Tensor . . . 37
  - 4.5.1 Longitudinal waves in 1D . . . 38
  - 4.5.2 Relativistic Scalar Field . . . 38
  - 4.5.3 Electromagnetic Field . . . 38
  - 4.5.4 General Relativity . . . 39
- 4.6 Angular Momentum and Spin . . . 40
- 4.7 Quantum Fields . . . 41

5 Broken Symmetry **45**
5.1 Self-Interacting Scalar Field 45
5.2 Spontaneously Broken Global Symmetry 46
5.3 Spontaneously Broken Local Symmetry 48
5.4 Higgs Mechanism 48

6 Dirac Field [*not examinable*] **51**
6.1 Dirac Lagrangian and Hamiltonian 53
6.2 Global and Local Phase Symmetry 54
6.3 Stress-Energy Tensor, Angular Momentum and Spin 54

7 Phase Transitions and Critical Phenomena **55**

A Appendix **A.1**

List of Tables

List of Figures

1.1 An example of a constraint, restricting the motion of a particle (which may be subject to external forces, such as gravity) along a specific path: the bead on a wire. 3
1.2 Examples of dynamical systems with holonomic constraints (in each case their expression is in the frame). In each case the number of coordinates is reduced (for the pulley and rolling cylinder Cartesians are sufficient, the circular wire requires plane polar coordinates and rolling on the inside of a cone, spherical polar coordinates; the last case has two free generalised coordinates left). 5

5.1 The form of the 'potential' $V(\varphi)$ that leads to spontaneous symmetry breaking. 46
5.2 Summary of interactions between particles described by the Standard Model. 49

Chapter 1

Basic Lagrangian Mechanics

The initial purpose of Lagrangian mechanics is to express the relevant equations of motion, essentially Newton's laws, in a form involving a set $q_1, q_2, \cdots, q_n$ of generalised position coordinates, and their first time-derivatives $\dot{q}_1, \dot{q}_2, \cdots, \dot{q}_n$. The n-component vector $\{\mathbf{q}\}$ can represent any physical system or process, as long as this set of numbers completely describes the state of the system (n is the number of degrees of freedom). In the most straightforward case, these can be the 3 Cartesian coordinates of a material point (a "particle"), but, say, the spherical polar set (r, θ, ϕ) is just as good. Just as $\{\mathbf{q}\}$ is the generalised coordinate vector, so is $\{\dot{\mathbf{q}}\}$ the generalised velocity, both explicit functions of time. It is assumed that simultaneous knowledge of $\{\mathbf{q}(t)\}$ and $\{\dot{\mathbf{q}}(t)\}$ completely defines the mechanical state of the system. Mathematically, this means that the complete set of $\{\mathbf{q}(t)\}$ and $\{\dot{\mathbf{q}}(t)\}$ also determines the accelerations $\{\ddot{\mathbf{q}}(t)\}$. The mathematical relations that relate accelerations with coordinates and velocities are what one calls the equations of motion.

In many cases, not all n degrees of freedom are completely free. A system may have *constraints*; for example $q_1 = \text{const.}$, $q_2 = \text{const.}, \cdots, q_r = \text{const.}$ could represent the r constraints and $q_{r+1}, \cdots, q_n$ the remaining independent coordinates. Most often the choice of generalised coordinates $\{\mathbf{q}\}$ is dictated by the nature of the constraints. For instance, if a particle is constrained to move on the surface of an expanding balloon of radius $R = a\sqrt{t}$, we might use spherical polar coordinates, scaled such that $q_1 = r/a\sqrt{t}, q_2 = \theta, q_3 = \phi$; in that case the single constraint is expressed as $q_1 = 1$ (it would look a lot more complicated if we tried to express it in Cartesians). The Lagrangian formalism is developed, partially, to enable one to deal efficiently with the sometimes complicated constraints imposed on the evolution of physical systems. Constraints are called **holonomic** if they are of the form $g(q_1, q_2, \cdots, q_n, t) = 0$. We shall shortly return to their treatment, but first, let us revise some basic starting points.

1.1 Hamilton's Principle

A very general formulation of the equations of motion of mechanical (and many other) systems is given by Hamilton's *Principle of Least Action*. It states that every mechanical system can be characterised by a certain function

$$L(q_1, q_2, \cdots, q_n; \dot{q}_1, \dot{q}_2, \cdots, \dot{q}_n; t) \equiv L(q, \dot{q}; t). \tag{1.1}$$

Theorem 1 (Hamilton's Principle)**.** *The actual motion of a system from A to B is that which makes the integral $S = \int_A^B L \, \mathrm{d}t$ a minimum.*

The function $L(q, \dot{q}; t)$ is called the Lagrangian of the given system and the integral S defined in Theorem 1 is called the **action**. Later in this course we shall have some deeper

insights into what this object, the action functional $S[\mathbf{q}(t)]$, represents and why it has to be minimal. For the time being let us take this as an axiom.

The principle of least action implies that, with a sufficient command of mathematics, in particular the calculus of variations, the solution of any mechanical problem is achieved by the following recipe:

'Minimise $S = \int_A^B L\,\mathrm{d}t$ for fixed starting and finishing (representative) points, $A = (\mathbf{q}_A, t_A)$ to $B = (\mathbf{q}_B, t_B)$, taking proper account of all the constraints.'

Remark. *There is no worldline whose action is a true maximum, that is, for which* $\delta^2 S < 0$ *or more generally for which* $S - S_0 < 0$ *for every adjacent curve.*

Proof. Here follows an intuitive proof by contradiction for the Lagrangian $L = T - V$ with T positive definite. Consider an actual worldline for which it is claimed that S is a true maximum. Now modify this worldline by adding wiggles somewhere in the middle. These wiggles are to be of very high frequency and very small amplitude so that they increase the kinetic energy T compared to that along the original worldline with only a small change in the corresponding potential energy V. The Lagrangian $L = T - V$ for the region of wiggles is larger for the new curve and so is the overall time integral S. The new worldline has greater action than the original worldline, which we claimed to have maximum action. Therefore S cannot be a true maximum for any actual worldline. □

1.2 Derivation of the Equations of Motion

First, let's examine the "standard derivation" based on d'Alembert's principle: consider a particle that is subject to the total force $\mathbf{F}$ and has momentum $\mathbf{p}$. Then if we construct a vector $(\mathbf{F} - \dot{\mathbf{p}})$, this vector will always be perpendicular to the instantaneous line of motion. In other words, the scalar product is zero:

$$\sum_i (F_i - \dot{p}_i)\delta x_i = 0. \tag{1.2}$$

That's almost trivially true for an arbitrary set of coordinate variations δx_i because Newton's second law ($\mathbf{F}^{\text{total}} = m\ddot{\mathbf{r}}$ for each particle) makes each $(F_i - \dot{p}_i) = 0$. However, we shall only be interested in sets of displacements δx_i consistent with the constraints. Constraints exert their own forces on each particle, which we call internal: see the reaction force R exerted by the wire in Fig. 1.1. By definition of the constraint, these internal forces are perpendicular to the line of motion, that is $\sum_i F_i^{\text{internal}} \delta x_i = 0$. Therefore, d'Alembert's principle states

$$\sum_i \left(F_i^{\text{internal}} - \dot{p}_i\right)\delta x_i = 0. \tag{1.3}$$

Let us try rewriting this in an arbitrary set of generalised coordinates $\{\mathbf{q}\}$ to which the Cartesians $\{\mathbf{r}\}$ could be transformed via matrices $\partial q_i/\partial x_j$. The aim is to present Eq. (1.3) as a generalised scalar product $\sum_j (\text{something})\delta q_j = 0$, so that we can say this is true for arbitrary sets of variations δq_i of the reduced number $(n - r)$ of generalised coordinate that are not subject to the constraints.

Fig. 1.1: An example of a constraint, restricting the motion of a particle (which may be subject to external forces, such as gravity) along a specific path: the bead on a wire.

The coordinate transformation in the first term, involving the external force, is easy:

$$\sum_i F_i \delta x_i = \sum_{i,j} F_i \frac{\partial x_i}{\partial q_j} \delta q_j \equiv \sum_j Q_j \delta q_j. \tag{1.4}$$

One must take great care over precisely what partial differentials mean. In the following, $\partial/\partial q_j$ means evaluating $(\partial/\partial q_j)$ with the other components $q_{i\neq j}$, all velocities $\dot{q}_i$ and time t held constant.

It is clear that $\partial x_i/\partial q_j$ should mean $(\partial x_i/\partial q_j)_{\text{all other } q,t}$; holding the $\dot{q}_i$ constant only becomes relevant when we differentiate a velocity w.r.t. q_j – a velocity component changes with $\mathbf{q}$ for fixed $\dot{\mathbf{q}}$ because the conversion factors from the $\dot{q}_j$ to the $\dot{r}_i$ change with position. Similarly, $\partial/\partial t$ means $(\partial/\partial t)_{\mathbf{q},\dot{\mathbf{q}}}$, e.g. $\partial x_i/\partial t$ refers to the change in position, for *fixed* $\mathbf{q}$ and $\dot{\mathbf{q}}$, due to the prescribed motion of the q-coordinate system.

Dealing with the second term, involving the rate of change of momentum, is a bit harder – it takes a certain amount of algebra to manipulate it into the required form. First, by definition of momentum in Cartesians:

$$\sum_i \dot{p}_i \delta x_i = \sum_{i,j} m_i \dot{v}_i \frac{\partial x_i}{\partial q_j} \delta q_j. \tag{1.5}$$

We shall need

$$v_i \equiv \dot{x}_i = \sum_j \frac{\partial x_i}{\partial q_j} \dot{q}_j + \frac{\partial x_i}{\partial t}, \quad \text{whence} \quad \frac{\partial v_i}{\partial \dot{q}_j} = \frac{\partial x_i}{\partial q_j}. \tag{1.6}$$

Now we are in a position to start work on the second term. The relevant product is

$$\dot{v}_i \frac{\partial x_i}{\partial q_j} = \frac{\mathrm{d}}{\mathrm{d}t}\left(v_i \frac{\partial x_i}{\partial q_j} \right) - v_i \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial x_i}{\partial q_j} \right). \tag{1.7}$$

Further transforming the second term in (1.7):

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial x_i}{\partial q_j} \right) &= \frac{\partial}{\partial q_k}\left(\frac{\partial x_i}{\partial q_j} \right) \dot{q}_k + \frac{\partial}{\partial t}\left(\frac{\partial x_i}{\partial q_j} \right) \quad \text{(summed over } k\text{)} \\ &= \frac{\partial}{\partial q_j}\left[\left(\frac{\partial x_i}{\partial q_k} \right) \dot{q}_k + \frac{\partial x_i}{\partial t} \right] \\ &= \left(\frac{\partial v_i}{\partial q_j} \right)_{\text{other } q,\dot{q},t} \end{aligned} \tag{1.8}$$

Using (1.6) on the first term and (1.8) on the second term of (1.7) we finally get

$$\begin{aligned} \sum_i \dot{p}_i \delta x_i &= \sum_{i,j} \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(m_i v_i \frac{\partial v_i}{\partial \dot{q}_j} \right) - m_i v_i \frac{\partial v_i}{\partial q_j} \right] \delta q_j \\ &= \sum_j \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} \right] \delta q_j, \end{aligned} \tag{1.9}$$

which is equal to $\sum_j Q_j \delta q_j$, from Eq. (1.4). Here the total kinetic energy of the system has been defined from the Cartesian representation $T = \sum_i \frac{1}{2} m_i v_i^2$. The last equation is a consequence of D'Alembert's principle. Since the components δq_j allowed by the constraints are all independent, it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = Q_j. \tag{1.10}$$

In many systems the external forces are gradients of a scalar potential; in Cartesians:

$$F_i = -\frac{\partial V}{\partial x_i}, \quad \text{where} \quad V = V(\mathbf{x}, t), \tag{1.11}$$

(i.e. V is independent of the particle velocities), so that

$$Q_j = -\sum_i \frac{\partial V}{\partial x_i}\frac{\partial x_i}{\partial q_j} = -\frac{\partial V}{\partial q_j}. \tag{1.12}$$

Therefore, substituting this vector $\mathbf{Q}$ into the r.h.s. of Eq. (1.10) and using the fact that it does not depend on $\dot{\mathbf{q}}$, we can write

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0, \tag{1.13}$$

where $L \equiv T - V$, but by construction it is a function of the generalised coordinates q, $\dot{q}$ and not the original Cartesians. We discover that this function is exactly the Lagrangian of the system, the one used in the definition of the action in Hamilton's principle. Indeed, (1.13) is the differential equation that one obtains by the calculus of variations from the condition $\delta S = 0$ for the minimum of the action.

Now let's reverse the argument. Define a function $L = T - V$. In Cartesian coordinates, an N-particle system moving in a potential $V(x_1, x_2, \cdots, x_{3N})$ has

$$p_i \equiv m_i v_i = \frac{\partial}{\partial v_i}\left(\frac{1}{2} m_i v_i^2\right) = \frac{\partial T}{\partial v_i} \equiv \frac{\partial L}{\partial v_i}, \tag{1.14}$$

and

$$F_i = -\frac{\partial V}{\partial x_i} = \frac{\partial L}{\partial x_i}, \tag{1.15}$$

provided that V is independent of velocities, so the equations of motion $\dot{\mathbf{p}} = \mathbf{F}$ are

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}_i}\right) - \frac{\partial L}{\partial x_i} = 0. \tag{1.16}$$

Therefore the motion obeys Hamilton's principle of least action. But Hamilton's principle is a statement independent of any particular coordinate system; it is true in any coordinate system. Therefore write down the Euler-Lagrange equations for Hamilton's principle in our new q coordinate system (in which constraints are of the kind $q_j = \text{const.}$):

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0, \tag{1.17}$$

a much more memorable way to derive Lagrange's equations.

But what about the constraints? We must invent some local potentials near the actual path of the system, whose gradients are perpendicular to the actual path and just right to keep the system on the constrained trajectory, see Fig. 1.1. The important thing is that they will have no gradient along the local 'directions' of the q_is, allowed by the constraints, and will therefore not affect the dynamics for those q_is.

1.2.0.1 What is Lagrangian Mechanics Good For?

Lagrangian mechanics will do nothing that Newtonian mechanics won't do. It's just a reformulation of the same physics. In fact, it will do slightly less, because some problems (notoriously, the motion of a bicycle) have what is called non-holonomic constraints. One thing that one can say for the Lagrangian formulation is that it involves scalars (T and V) instead of vectors (forces, couples) which makes it less confusing to use in messy problems. Several examples will be treated in the lectures and, more particularly, in the examples classes. Figure 1.2 gives a few simple examples of dynamical systems with constraints. Use them to practise: in each case first write the full potential (in all cases it's gravity) and kinetic energies (don't forget the moment of inertia for the rotating cylinder), then implement the constraint (for this you need to have the appropriate choice of coordinates) and write the resulting Lagrangian, as well as the dynamical equation(s).

Fig. 1.2: Examples of dynamical systems with holonomic constraints (in each case their expression is in the frame). In each case the number of coordinates is reduced (for the pulley and rolling cylinder Cartesians are sufficient, the circular wire requires plane polar coordinates and rolling on the inside of a cone, spherical polar coordinates; the last case has two free generalised coordinates left).

1.2.0.2 A Few More Remarks

So far we have assumed forces depend only on position: $F_i = -\,\partial V(\mathbf{x},t)/\partial x_i$. For *some* **velocity-dependent** forces one can still use Lagrange's equations; it's possible if you can find a $V(\mathbf{x},\dot{\mathbf{x}},t)$ such that

$$F_i = -\frac{\partial V(\mathbf{x},\dot{\mathbf{x}},t)}{\partial x_i} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial V(\mathbf{x},\dot{\mathbf{x}},t)}{\partial \dot{x}_i}\right). \tag{1.18}$$

To derive this condition you should repeat the steps between Eqs.(1.4) and (1.13), only now allowing a contribution from V in the $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right)$ terms of the Euler-Lagrange equations. The hard bit is to find a V that satisfies such a condition. A bit later we shall examine the case of magnetic forces, which is a good example of such a potential.

There is also an issue of uniqueness of the Lagrangian function. Clearly the condition of zero variation $\delta[S] = 0$ can be maintained if we:

- add any constant to V;
- multiply L by any constant;
- add a total time derivative, $f = \frac{\mathrm{d}}{\mathrm{d}t}g(q,\dot{q},t)$;

etc. Usually the convention is to take $L = T - V$, with no "additives", which then corresponds to the "classical action S".

1.3 Symmetry and Conservation Laws; Canonical Momenta

Symmetry is one of the most powerful tools used in theoretical physics. In this section we will show how symmetries of L correspond to important conservation laws. This theme will be taken further later in this course, and in subsequent solid state and particle physics courses.

When L does not depend explicitly on one of the q_i then Lagrange's equations show that the corresponding $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right)$ is zero, directly from Eq. (1.13). Hence we can define an object, $p_i \equiv \frac{\partial L}{\partial \dot{q}_i}$, which is conserved for such a system and q_i is called an ignorable coordinate. The object p_i is generally referred to as the *canonical momentum* corresponding to the coordinate q_i. For a particle moving in Cartesian coordinates this would be the "ordinary" momentum component: for each x_i this is just $p_i = m\dot{x}_i$. However, in generalised coordinates, the physical meaning of each component of the canonical momentum may be very different. In particular, since a generalised coordinate can have any dimensions, the dimensions of the corresponding canonical momentum need not be those of ordinary momentum.

1.3.1 Translational Invariance $\leftrightarrow$ Conservation of Linear Momentum

Suppose L does not depend on the position of the system as a whole, i.e. we can move the position of every particle by the *same* vector ϵ without changing L. Suppose we move the whole system by δx in the x direction (in Cartesian coordinates!). Then

$$L \to L + \sum_n \frac{\partial L}{\partial x_n}\delta x. \tag{1.19}$$

(N.B. In this section and the next we sum over particles, assuming only 1-dimensional motion for simplicity: n is an index of summation over the particles. Also, V in $L = T - V$ now includes the mutual potential energies that impose e.g. the constraints that keep rigid bodies rigid.) Therefore if L does not change

$$\sum_n \frac{\partial L}{\partial x_n} = 0, \quad \text{and so, by Lagrange's equations,} \quad \sum_n \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}_n}\right) = 0, \tag{1.20}$$

i.e. if L is independent of the position of the system $\sum \partial L/\partial \dot{x}_n$ is constant.

Now if V depends on particle positions only, then

$$\sum \frac{\partial L}{\partial \dot{x}_n} = \sum m_n \dot{x}_n \tag{1.21}$$

is just the x-component of the total momentum of the system, and we conclude

$$\text{Homogeneity of space} \implies \text{conservation of linear momentum.} \tag{1.22}$$

N.B.: This is clearly not true for velocity-dependent potentials, e.g. for charged ions moving in magnetic fields.

1.3.2 Rotational Invariance ↔ Conservation of Angular Momentum

Suppose L is independent of the orientation of the system. In particular, suppose L is invariant under rotation of the whole system about the z-axis; then, proceeding as before, we can assume a rotation by $\delta\theta$ and use cylindrical coordinates (r, θ, z) to obtain

$$\sum_n \frac{\partial L}{\partial \theta_n}\delta\theta = 0 \implies \sum_n \frac{\partial L}{\partial \theta_n} = 0 \tag{1.23}$$

$$\implies \sum_n \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\theta}_n}\right) = 0 \implies \sum_n \frac{\partial L}{\partial \dot{\theta}_n} = \text{const.} \tag{1.24}$$

Again, when V depends on particle positions only,

$$\begin{aligned}\frac{\partial L}{\partial \dot{\theta}_n} &= \frac{\partial}{\partial \dot{\theta}_n}\left(\frac{1}{2}m_n\dot{r}_n^2 + \frac{1}{2}m_n r_n^2 + \dot{\theta}_n^2 + \frac{1}{2}m_n\dot{z}_n^2\right)\\ &= m_n r_n^2\dot{\theta}_n \qquad (1.25)\\ &= \text{ angular momentum of } n^{\text{th}} \text{ particle about } z\text{-axis} \qquad (1.26)\end{aligned}$$

Thus the total angular momentum about the z axis is conserved.

$$\text{Isotropy of space} \implies \text{conservation of angular momentum.} \tag{1.27}$$

1.3.3 Time Invariance ↔ Conservation of Energy

This is a bit more complicated and it also gives us a chance to explore a very useful mathematical result called Euler's homogeneous function theorem. If L does not depend explicitly on time, i.e.

$$L = L(q_i\dot{q}_i) \implies \frac{\partial L}{\partial t} = 0, \tag{1.28}$$

then the total time derivative of the Lagrangian is

$$\begin{aligned}\frac{\mathrm{d}L}{\mathrm{d}t} &= \sum_i \frac{\partial L}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_n}\ddot{q}_n \qquad (1.29)\\ &= \sum_i \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial q_i}\right)\dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_n \quad \text{(using the E-L equation)}\\ &= \sum_i \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\dot{q}_i\right) \quad \text{(assembling the total derivative from both terms),} \qquad (1.30)\end{aligned}$$

or, combining the total time derivatives from l.h.s. and r.h.s. into one expression, we have

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}H, \quad \text{where} \quad H = \sum_i \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L = \text{a constant, } E \tag{1.31}$$

We shall now identify the new function H, the **Hamiltonian**, as the total energy of the system. Assume the generalised kinetic energy, T, is given by:

$$T = \sum_{i,j}\frac{1}{2}c_{ij}\dot{q}_i\dot{q}_j, \tag{1.32}$$

where we can generically choose $c_{ij} = c_{ji}$ (by symmetry of dummy indices of summation). The coefficients c_{ij} might be functions of $q_1, \cdots, q_n$ but not $\dot{q}_1, \cdots, \dot{q}_n$ or time. This quadratic form is also called a homogeneous second-order polynomial function of $\dot{\mathbf{q}}$. Assume also that the potential energy, V, is given by:

$$V = V(q_1, \cdots, q_n), \tag{1.33}$$

(velocity-independent) and $L = T - V$ as usual (with no explicit time-dependence as we agreed). Then:

$$\begin{aligned}
\sum_i \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i &= \sum_i \dot{q}_i \frac{\partial}{\partial \dot{q}_i}\left(\sum_{j,k}\frac{1}{2}c_{jk}\dot{q}_j\dot{q}_k\right) \quad \text{since} \quad \frac{\partial V}{\partial \dot{q}_i} = 0 \qquad (1.34)\\
&= \sum_i \dot{q}_i \left(\sum_{j,k}\left[\frac{1}{2}c_{jk}\delta_{jk}\dot{q}_k + \frac{1}{2}c_{jk}\dot{q}_j\delta_{ik}\right]\right) \quad \text{since} \quad \frac{\partial \dot{q}_j}{\partial \dot{q}_i} = \delta_{ij}\\
&= \sum_i \dot{q}_i \left(\sum_k \frac{1}{2}c_{ik}\dot{q}_k + \sum_j \frac{1}{2}c_{ji}\dot{q}_j\right)\\
&= \sum_i \dot{q}_i \left(\sum_j c_{ij}\dot{q}_j\right) \quad \text{renaming the dummy index } k \to j\\
&= \sum_{i,j} c_{ij}\dot{q}_i\dot{q}_j\\
&= 2T. \qquad (1.35)
\end{aligned}$$

In effect, what we have just proven is that for any homogeneous quadratic function $T = T(q_i)$, the following property holds:

$$\sum_i q_i \frac{\partial T}{\partial q_i} = 2T \tag{1.36}$$

(an aspect of Euler's more general theorem; guess how this would change for linear, or cubic functions). Returning to our Hamiltonian, we have $H = 2T - L = 2T - (T - V) = T + V \equiv$ total energy $= E$, a constant from Eq. (1.31).

Chapter 2

Hamilton's Equations of Motion

We have already defined the i^{th} component of generalised (canonical) momentum as

$$p_i \equiv \left(\frac{\partial L}{\partial \dot{q}_i}\right)_{\text{other } \dot{q},\mathbf{q},t} \tag{2.1}$$

In Cartesians and for $V = V(\mathbf{r})$, $p_i = m_i v_i$.

Now try to re-write the equations of motion in terms of $\mathbf{q}$ and $\mathbf{p}$ instead of $\mathbf{q}$ and $\dot{\mathbf{q}}$. This operation is fully analogous to what is called the Legendre transformation in thermodynamics, when we change from one potential depending on a given variable to another, depending on its conjugate (like $T\,\mathrm{d}S \to -S\,\mathrm{d}T$ or $-p\,\mathrm{d}V \to V\,\mathrm{d}p$).

The Euler-Lagrange equations say

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) \equiv \dot{p}_i = \left(\frac{\partial L}{\partial q_i}\right)_{\text{other } q,\dot{\mathbf{q}},t}, \tag{2.2}$$

but that's not quite what we want, for it refers to $L(\mathbf{q}, \dot{\mathbf{q}}, t)$, not $L(\mathbf{q}, \mathbf{p}, t)$. We proceed thus:

$$\begin{aligned} \delta L &= \frac{\partial L}{\partial t}\delta t + \sum_i \frac{\partial L}{\partial q_i}\delta q_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i \\ &= \frac{\partial L}{\partial t}\delta t + \sum_i \dot{p}_i \delta q_i + \sum_i p_i \delta \dot{q}_i. \end{aligned} \tag{2.3}$$

Hence

$$\delta\left(L - \sum_i p_i \dot{q}_i\right) = \frac{\partial L}{\partial t}\delta t + \sum_i \dot{p}_i \delta q_i - \sum_i \dot{q}_i \delta p_i. \tag{2.4}$$

This last equation suggests that the expression in brackets on the l.h.s. is a function of (t, q, p). It is clear that every pair of a generalised velocity $\dot{q}_i$ and its canonical momentum p_i have the same status as conjugate variables in thermodynamics. (Of course, historically, people developed the underlying maths behind Lagrangian and Hamiltonian dynamics first; the physical concepts of thermodynamics were then easy to formalise.)

Defining the Hamiltonian as in Eq. (1.31):

$$H \equiv \left(\sum_i p_i \dot{q}_i - L\right)\Bigg|_{\dot{q}_i = \dot{q}_i(\mathbf{q},\mathbf{p},t)}, \tag{2.5}$$

then (2.4) implies

$$\dot{p}_i = -\left(\frac{\partial H}{\partial q_i}\right)_{\text{other } q,\mathbf{p},t} \qquad \dot{q}_i = \left(\frac{\partial H}{\partial p_i}\right)_{\text{other } p,\mathbf{q},t}, \tag{2.6}$$

which are **Hamilton's equations of motion**. Notice the importance of completing the transformation in Eq. (2.5) and writing H as a function of $\mathbf{q}, \mathbf{p}, t$ only. None of the $\dot{q}_i$

should appear explicitly; rather, they should be replaced by inverting Eq.(2.1) that defines p_i.

We have already shown that, provided L does not depend explicitly on time, T has the form (1.32) and V does not depend on velocities, then $H = T + V$, the total energy, and it is a constant of the motion. This does not imply that the right-hand sides of Hamilton's equations are zero! They are determined by the functional form of the dependence of H on the p_i and q_i.

Note also that p_i and q_i are now on an equal footing. In Hamilton's equations q_i can be anything, not necessarily a position coordinate. For example we could interchange the physical meaning of what we regard as p_i with q_i, and q_i with $-p_i$, and Hamilton's equations would still work.

2.1 Liouville's theorem

Phase space is defined as the space spanned by the canonical coordinates and conjugate momenta, e.g. (x, y, z, p_x, p_y, p_z) for a single particle (a 6D space) and $\{q_1, q_2, \cdots, q_n, p_1, p_2, \cdots, p_n\}$ for a system of n particles. This defines phase space to be a $6n$-dimensional space. A single point in phase space represents the state of the whole system, i.e. the positions and velocities of all its particles. It is called a **representative point**. If there are constraints acting on the system, the representative points are confined to some lower dimensional subspace. The representative points move with velocities $\mathbf{v}$ where:

$$\mathbf{v} = \{\dot{q}_1, \dot{q}_2, \cdots, \dot{q}_n, \dot{p}_1, \dot{p}_2, \cdots, \dot{p}_n\} = \left\{\frac{\partial H}{\partial p_1}, \frac{\partial H}{\partial p_2}, \cdots, \frac{\partial H}{\partial p_n}, -\frac{\partial H}{\partial q_1}, -\frac{\partial H}{\partial q_2}, \cdots, -\frac{\partial H}{\partial q_n}\right\}. \tag{2.7}$$

Liouville's theorem is a very powerful result concerning the evolution in time of ensembles of systems. We can regard the initial state of the ensemble of systems as corresponding to a distribution or density of representative points in phase space. Then Liouville's theorem states that:

Theorem 2 (Liouville's theorem)**.** *The density in phase space evolves as an incompressible fluid.*

Proof. The proof is a simple application of Hamilton's equations and the n-dimensional divergence theorem. The n-dimensional divergence theorem states that for an n-dimensional vector function of n variables $\mathbf{V}(x_1, \cdots, x_n)$:

$$\int_{\mathcal{V}} \sum_i \frac{\partial V_i}{\partial x_i} \, \mathrm{d}\tau = \int_{\mathcal{S}} \sum_i V_i \, \mathrm{d}S_i \,, \tag{2.8}$$

where $\mathrm{d}\tau$ is an n-dimensional volume element and $\mathrm{d}\mathbf{S}$ is an $(n-1)$-dimensional element of surface area. This theorem and its proof are the obvious generalisation the divergence theorem (the Gauss theorem) in 3D:

$$\int_{\mathcal{V}} \boldsymbol{\nabla} \cdot \mathbf{V} \, \mathrm{d}\tau = \int_{\mathcal{S}} \mathbf{V} \cdot \mathrm{d}\mathbf{S} \,. \tag{2.9}$$

To prove Liouville's theorem, suppose that the representative points are initially confined to some (n-dimensional) volume $\mathcal{V}$ with surface $\mathcal{S}$. The points move with velocity $\mathbf{v}$ given by Eq. (2.7). Therefore, at the surface, the volume occupied by the points is changing at a rate $\mathrm{d}V = \mathbf{v} \cdot \mathrm{d}\mathbf{S}$. Hence:

$$\begin{aligned} \Delta V &= \int_{\mathcal{S}} \mathbf{v} \cdot \mathrm{d}\mathbf{S} \\ &= \int_{\mathcal{V}} \boldsymbol{\nabla} \cdot \mathbf{v} \, \mathrm{d}\tau \,, \end{aligned} \tag{2.10}$$

by (2.9). However,

$$\begin{aligned} \boldsymbol{\nabla} \cdot \mathbf{v} &= \sum_i \frac{\partial}{\partial q_i} \dot{q}_i + \frac{\partial}{\partial p_i} \dot{p}_i \\ &= \sum_i \frac{\partial^2 H}{\partial q_i \partial p_i} - \frac{\partial^2 H}{\partial p_i \partial q_i} = 0. \end{aligned} \tag{2.11}$$

So the volume occupied by the ensemble's representative points does not change. □

2.2 Poisson Brackets and the Analogy with Quantum Commutators

Suppose $f = f(q_i, p_i, t)$, i.e. f is a function of the dynamical variables $\mathbf{p}$ and $\mathbf{q}$. Then

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial q_i} \dot{q}_i + \frac{\partial f}{\partial p_i} \dot{p}_i + \frac{\partial f}{\partial t} \tag{2.12}$$

$$= \frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} + \frac{\partial f}{\partial t}, \tag{2.13}$$

which we can rewrite for notational convenience as

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \{f, H\} + \frac{\partial f}{\partial t}, \tag{2.14}$$

where

$$\{f, g\} \equiv \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i}, \tag{2.15}$$

is called the *Poisson Bracket* of the functions f and g and was first introduced into mechanics by Simon Poisson in 1809 (Note the use of summation convention in Eq. (2.15)).

Eq. (2.14) is remarkably similar to the Ehrenfest theorem in Quantum Mechanics

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\langle \hat{O} \right\rangle = \frac{1}{i\hbar} \left\langle \left[\hat{O}, \hat{H} \right] \right\rangle + \left\langle \frac{\partial \hat{O}}{\partial t} \right\rangle, \tag{2.16}$$

for the variation of expectation values of (Hermitian) operators $\hat{O}$, where $\hat{H}$ is of course the quantum mechanical Hamiltonian operator and $\left[\hat{O}, \hat{H}\right]$ represents the commutator $\hat{O}\hat{H} - \hat{H}\hat{O}$.

It is easy to check that $\{f, g\}$ has many of the properties of the commutator:

$$\{f, g\} = -\{g, f\}, \quad \{f, f\} = 0 \quad \text{etc.} \tag{2.17}$$

Also, if in Eq. (2.12) $\partial f/\partial t = 0$ (no explicit t dependence) and $\{f, H\} = 0$ then $\mathrm{d}f/\mathrm{d}t = 0$, i.e. f is a constant of the motion.

This suggests we can relate classical and quantum mechanics by formulating classical mechanics in terms of Poisson Brackets and then associating these with the corresponding Quantum Mechanical commutator

$$\begin{gathered} \{A, B\} \leftrightarrow \frac{1}{i\hbar}\left[\hat{A}, \hat{B}\right] \\ \text{classical} \leftrightarrow \text{quantum} \\ H \leftrightarrow \hat{H}. \end{gathered} \tag{2.18}$$

Classically

$$\{q, p\} = \frac{\partial q}{\partial p}\frac{\partial p}{\partial p} - \frac{\partial q}{\partial p}\frac{\partial p}{\partial q} = 1. \tag{2.19}$$

Quantum Mechanically,

$$\left[q, -i\hbar\frac{\partial}{\partial q}\right]\Psi = -qi\hbar\frac{\partial}{\partial q}\Psi + i\hbar\frac{\partial}{\partial q}(q\Psi) = i\hbar\Psi, \tag{2.20}$$

i.e.

$$\frac{1}{i\hbar}\left[q, -i\hbar\frac{\partial}{\partial q}\right] = 1 = \{q, p\}. \tag{2.21}$$

Therefore we confirm the identification of $-i\hbar\,\partial/\partial q$ with the canonical **momentum operator** $\hat{p}$ corresponding to q.

2.2.1 Quantum Variational Principle

In QM we associate a wave vector $\mathbf{k} = \mathbf{p}/\hbar$ with a particle of momentum p (de Broglie relation), and frequency $\omega = H/\hbar$ with its total energy $E = H$ (Einstein relation). Thus we can write Hamilton's principle for the classical motion of the particle as

$$\frac{1}{\hbar}\int L\,\mathrm{d}t = \int\left(\frac{1}{\hbar}\mathbf{p}\cdot\dot{\mathbf{q}} - \frac{H}{\hbar}\right)\mathrm{d}t = \int(\mathbf{k}\cdot\mathrm{d}\mathbf{q} - \omega\,\mathrm{d}t) = \text{stationary}, \tag{2.22}$$

i.e. the wave-mechanical phase shall be stationary (because multiplication by a constant factor does not alter the condition for the minimum of S). This is the condition for constructive interference of waves; what the Hamilton principle really says is that the particle goes where the relevant de Broglie waves reinforce. If we move a little away from the classical path, the waves do not reinforce so much and the particle is less likely to be found there. If we imagine taking the limit $\hbar \to 0$, the wavefunction falls off so rapidly away from the classical path that the particle will *never* deviate from it.

We see that classical mechanics is the "geometrical optics" limit of QM: the "rays" correspond to the classical paths and quantum effects (like diffraction) are due the finite frequency and wave number of waves of given energy and momentum, i.e the finite value of $\hbar$.

Conversely, if we have a classical theory for a physical system, which works for macroscopic systems of that kind, we can get a wave-mechanical description that reduces to this classical theory as $\hbar \to 0$ by making the Hamiltonian operator $\hat{H}$ the same function of $-i\hbar\,\partial/\partial q_i$ and q_i as the classical Hamiltonian H is of p_i and q_i.

N.B.: This is not necessarily the only or the correct QM description! There may be other bits of physics (terms in H) which vanish as $\hbar \to 0$ but are important, e.g., the electron spin.

2.2.2 Canonical Transformations

Another advantage of the Hamiltonian formulation of dynamics is that we have considerable freedom to redefine the generalised coordinates and momenta, which can be useful for solving the equations of motion. For example, as we already saw, we can redefine p_i as q_i and q_i as $-p_i$. This is an example of a much more general change of variables known as a *canonical transformation*. This is a transformation of the form

$$Q_j = Q_j(\{q_i\},\{p_i\}), \quad P_j = P_j(\{q_i\},\{p_i\}), \tag{2.23}$$

that preserves the form of Hamilton's equations of motion:

$$\dot{P}_j = -\left(\frac{\partial H}{\partial Q_j}\right)_{\text{other } Q,\mathbf{P},t}, \quad \dot{Q}_j = -\left(\frac{\partial H}{\partial P_j}\right)_{\text{other } P,\mathbf{Q},t}. \tag{2.24}$$

The condition for a transformation to be canonical is very simple: the transformed variables have to satisfy the canonical Poisson bracket relations

$$\{Q_j,P_j\} = 1, \quad \{Q_j,Q_k\} = \{P_j,P_k\} = \{Q_j,P_k\} = 0 \quad \text{for} \quad j \neq k, \tag{2.25}$$

with respect to the original generalised coordinates and momenta. We prove this for a single coordinate and momentum; the generalization to many variables is straightforward. For any function $Q(q,p)$, not explicitly time-dependent, we have

$$\dot{Q} = \{Q,H\} = \frac{\partial Q}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial H}{\partial q}. \tag{2.26}$$

Expressing H in terms of Q and some other function $P(q,p)$,

$$\frac{\partial H}{\partial p} = \frac{\partial H}{\partial Q}\frac{\partial Q}{\partial p} + \frac{\partial H}{\partial P} + \frac{\partial P}{\partial p} \tag{2.27}$$

$$\frac{\partial H}{\partial q} = \frac{\partial H}{\partial Q}\frac{\partial Q}{\partial q} + \frac{\partial H}{\partial P} + \frac{\partial P}{\partial q}. \tag{2.28}$$

Inserting these in Eq. (2.26) and rearranging terms,

$$\begin{aligned}\dot{Q} &= \frac{\partial H}{\partial P}\left(\frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q}\right)\\ &= \frac{\partial H}{\partial P}\{Q,P\}.\end{aligned} \tag{2.29}$$

Similarly for P we find

$$\begin{aligned}\dot{P} &= \frac{\partial H}{\partial Q}\left(\frac{\partial P}{\partial q}\frac{\partial Q}{\partial p} - \frac{\partial P}{\partial p}\frac{\partial Q}{\partial q}\right) \\ &= -\frac{\partial H}{\partial Q}\{Q, P\}. \end{aligned} \tag{2.30}$$

Hence the necessary and sufficient condition to preserve Hamilton's equations is $\{Q, P\} = 1$.

2.3 Lagrangian Dynamics of a Charged Particle

The Lorentz force is an example of a velocity dependent force. Another example is the 'fictitious' Coriolis force found in rotating (non-inertial) frames. A deeper treatment of these forces leads to special relativity in the case of electromagnetism and general relativity in the case of inertial forces.

In this section we examine how the Lorentz force and basic electromagnetism can be incorporated into the Lagrangian formalism without explicit mention of special relativity. In the following sections we sketch the much more powerful ideas involved in the relativistic approach.

The derivation of Lagrange's equations of motion is valid provided the external forces satisfy

$$F_i = -\frac{\partial V}{\partial x_i} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial V}{\partial \dot{x}_i}\right). \tag{2.31}$$

The second term (with the derivative w.r.t. velocity) is *not* usually present for conventional (i.e. potential, $\mathbf{F} = -\boldsymbol{\nabla} V$) forces. For the Lorentz force problem, a particle of charge e in fields $\mathbf{E}$ and $\mathbf{B}$ experiences a velocity dependent force $\mathbf{F}$,

$$\mathbf{F} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \tag{2.32}$$

and we can in fact take

$$V = e(\phi - \mathbf{v} \cdot \mathbf{A}), \tag{2.33}$$

where $\mathbf{A}$ is the magnetic vector potential such that $\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}$ and $\mathbf{E} = -\boldsymbol{\nabla}\phi - \partial \mathbf{A}/\partial t$.

The potential (2.33), when plugged into (2.31), gives the correct expression for the force (2.32). To verify this, we need to perform a calculation:

$$F_i = e(\mathbf{E} + \mathbf{v} \times \mathbf{B})_i \stackrel{?}{=} -\frac{\partial V}{\partial x_i} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial V}{\partial \dot{x}_i}\right). \tag{2.34}$$

We will need the result of vector analysis

$$[\mathbf{v} \times (\boldsymbol{\nabla} \times \mathbf{A})]_i = v_j \frac{\partial A_j}{\partial x_i} - v_j \frac{\partial A_i}{\partial x_j}, \tag{2.35}$$

which follows from

$$[\mathbf{v}\times(\boldsymbol{\nabla}\times\mathbf{A})]_i = \epsilon_{ijk}v_j(\boldsymbol{\nabla}\times\mathbf{A})_k \tag{2.36}$$

$$= \epsilon_{ijk}v_j\epsilon_{kpq}\frac{\partial A_q}{\partial x_p} \tag{2.37}$$

$$\epsilon_{ijk}\epsilon_{pqk} \equiv (\delta_{ip}\delta_{jq}-\delta_{iq}\delta_{jp}) \quad\Longrightarrow\quad = (\delta_{ip}\delta_{jq}-\delta_{iq}\delta_{jp})v_j\frac{\partial A_q}{\partial x_p}$$

$$= v_j\frac{\partial A_j}{\partial x_i} - v_j\frac{\partial A_i}{\partial x_j}. \tag{2.38}$$

The rest is a simple manipulation

$$F_i = -\frac{\partial V}{\partial x_i} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial V}{\partial v_i}\right) \tag{2.39}$$

$$= -\frac{\partial}{\partial x_i}e(\phi - v_jA_j) + \frac{\mathrm{d}}{\mathrm{d}t}(-eA_i)$$

$$= -e\frac{\partial\phi}{\partial x_i} + ev_j\frac{\partial A_j}{\partial x_i} - e\frac{\partial A_i}{\partial x_j}v_j - e\frac{\partial A_i}{\partial t}$$

$$= e(\mathbf{E}+[\mathbf{v}\times(\boldsymbol{\nabla}\times\mathbf{A})])_i. \tag{2.40}$$

Having satisfied our sense of caution to some extent, we can now write the Lagrangian, as usual,

$$L = T - V = \frac{1}{2}mv^2 - e(\phi - \mathbf{v}\cdot\mathbf{A}). \tag{2.41}$$

The components of the canonical momentum $\mathbf{p}$ are obtained by the familiar

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = \frac{\partial L}{\partial v_i} = mv_i + eA_i, \tag{2.42}$$

or, for a charged particle in an electromagnetic field,

$$\text{canonical momentum} = \text{mechanical momentum} + e\mathbf{A}. \tag{2.43}$$

Knowing p_i we can write down the Hamiltonian H, formally following our previous definitions,

$$H = \mathbf{p}\cdot\dot{\mathbf{q}} - L \tag{2.44}$$

$$= (m\mathbf{v}+e\mathbf{A})\cdot\mathbf{v} - \frac{1}{2}mv^2 + e(phi - \mathbf{v}\cdot\mathbf{A})$$

$$= \frac{1}{2}mv^2 + e\phi \quad \text{(total energy)}$$

$$= \frac{1}{2m}(\mathbf{p}-e\mathbf{A})^2 + e\phi, \tag{2.45}$$

where $\mathbf{p}$ is the canonical momentum (2.42).

Suppose we reverse the argument and formally start from the Lagrangian (2.41), looking for the equations of motion by minimisation of the corresponding action:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \mathbf{v}}\right) = \frac{\partial L}{\partial \mathbf{x}} \equiv \boldsymbol{\nabla}L = e\boldsymbol{\nabla}(\mathbf{v}\cdot\mathbf{A}) - e\boldsymbol{\nabla}\phi. \tag{2.46}$$

Another useful formula from vector analysis says:

$$\boldsymbol{\nabla}(\mathbf{a}\cdot\mathbf{b}) = (\mathbf{a}\cdot\boldsymbol{\nabla})\mathbf{b} + (\mathbf{b}\cdot\boldsymbol{\nabla})\mathbf{a} + \mathbf{a}\times(\boldsymbol{\nabla}\times\mathbf{b}) + \mathbf{b}\times(\boldsymbol{\nabla}\times\mathbf{a}), \tag{2.47}$$

for any two vectors $\mathbf{a}$ and $\mathbf{b}$. Remembering that $\boldsymbol{\nabla}L$ in (2.46) is evaluated at constant $\mathbf{v}$, we find for its r.h.s.

$$\boldsymbol{\nabla}L = e(\mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{A} + e[\mathbf{v}\times(\boldsymbol{\nabla}\times\mathbf{A})] - e\boldsymbol{\nabla}\phi. \tag{2.48}$$

The l.h.s. of (2.46) is the total time-derivative of the canonical momentum $\mathbf{p} = m\mathbf{v} + e\mathbf{A}$. The total time-derivative of $\mathbf{A}$, which may be a function of time and position, is given by

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} + \frac{\partial\mathbf{A}}{\partial t} + (\mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{A}. \tag{2.49}$$

Substituting this and (2.48) into (2.46) we find that the awkward $(\mathbf{v}\cdot\boldsymbol{\nabla})$ term cancels and the equation of motion becomes

$$\frac{\mathrm{d}(m\mathbf{v})}{\mathrm{d}t} = -e\frac{\partial\mathbf{A}}{\partial t} - e\boldsymbol{\nabla}\phi + e[\mathbf{v}\times(\boldsymbol{\nabla}\times\mathbf{A})]. \tag{2.50}$$

The force on the r.h.s. is thus made up of two parts. The first (the first two terms) does not depend on the particle velocity; the second part is proportional to $|\mathbf{v}|$ and is perpendicular to it. The first force, per unit of particle charge e, is *defined* as the **electric field**,

$$\mathbf{E} = -\boldsymbol{\nabla}\phi - \frac{\partial\mathbf{A}}{\partial t}, \tag{2.51}$$

and the force proportional to the velocity, per unit charge, is *defined* as the **magnetic flux density**

$$\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}, \tag{2.52}$$

which returns the familiar expression for the full Lorentz force (2.32).

2.3.1 Gauge Invariance

The equation of motion of a physical particle is determined by the physically observable fields $\mathbf{E}$ and $\mathbf{B}$. How unique are the potentials ϕ and $\mathbf{A}$ which determine these fields and contribute to the Lagrangian function? It turns out that they are not unique at all ...

If we add the gradient of an arbitrary scalar function $f(\mathbf{x}, t)$ to the vector potential $\mathbf{A}$, i.e.

$$A_i' = A_i + \frac{\partial f}{\partial x_i}, \tag{2.53}$$

the magnetic flux density $\mathbf{B}$ will not change, because $\boldsymbol{\nabla}\times\boldsymbol{\nabla}f \equiv 0$. To have the electric field unchanged as well, we must simultaneously subtract the time-derivative of f from the scalar potential:

$$\phi' = \phi - \frac{\partial f}{\partial t}. \tag{2.54}$$

The invariance of all electromagnetic processes with respect to the above transformation of the potentials by an arbitrary function f is called **gauge invariance**. As always, the

discovery of an additional symmetry is an indication of much deeper underlying physics and you will meet gauge invariance, and its consequences, many times in the future.

But how are we to deal with such non-uniqueness of the electromagnetic potentials and, accordingly, the Lagrangian? Because an arbitrary scalar function is governing the invariance transformation, one is free to choose *any* additional condition, an equation relating the potentials ϕ and $\mathbf{A}$ – but only one such condition. For instance, we may choose to formulate electrodynamics with no scalar electric potential, $\phi = 0$. However we cannot have $\mathbf{B} = \mathbf{0}$, since this represents three conditions for its components, instead of the allowed one. We can at most choose $\mathbf{n} \cdot \mathbf{A} = 0$ for some constant vector $\mathbf{n}$. Vector potentials satisfying such a condition are said to be in an *axial gauge*, with gauge vector $\mathbf{n}$.

Alternatively , since one can add an arbitrary gradient to $\mathbf{A}$, we could enforce the condition $\boldsymbol{\nabla} \cdot \mathbf{A} = 0$. Potentials satisfying this condition are said to be in the *Coulomb gauge*. Such a gauge leads to a convenient form of wave equation for $\mathbf{A}$, used in the theory of electromagnetic waves.

In relativistic dynamics, a commonly used condition is

$$\frac{\partial \phi}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{A} = 0, \tag{2.55}$$

which defines the *Lorenz gauge*. Notice that in this gauge there remains a residual ambiguity: we can still vary the electromagnetic potentials using any function f that satisfies the wave equation

$$\frac{\partial^2 f}{\partial t^2} - \nabla^2 f = 0. \tag{2.56}$$

2.4 Relativistic Particle Dynamics

The 4-index or **covariant notation** is widely used in theoretical physics; this subsection contains a brief (and not very rigorous) introduction. Consider:

$$x^\mu : (x^0, x^1, x^2, x^3) = (ct, x, y, z) \quad \text{a } \textbf{contravariant} \text{ 4-vector and} \tag{2.57}$$

$$x_\mu : (x_0, x_1, x_2, x_3) = (ct, -x, -y, -z) \quad \text{a } \textbf{covariant} \text{ 4-vector.} \tag{2.58}$$

Only (implicit) summations involving one raised and one lowered suffix are allowed, thus:

$$x^\mu x_\mu = c^2 t^2 - r^2, \tag{2.59}$$

is valid (and is of course Lorentz invariant), but neither $x^\mu x^\mu$ nor $x_\mu x_\mu$ is allowed.

If $\phi = \phi(x^\mu)$ then $\mathrm{d}\phi = \frac{\partial \phi}{\partial x^\mu}\, \mathrm{d}x^\mu$ is invariant, hence:

$$\frac{\partial \phi}{\partial x^\mu} \quad \text{is a covariant 4-vector and} \tag{2.60}$$

$$\frac{\partial}{\partial x^\mu} \quad \text{is a covariant operator.} \tag{2.61}$$

The operator $\frac{\partial}{\partial x^\mu}$ is often simply written as ∂_μ. Similarly the contravariant operator $\frac{\partial}{\partial x_\mu} = \partial^\mu$. The "metric tensor":

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{2.62}$$

can be used to raise or lower indices, for example:

$$x_\mu = g_{\mu\nu} x^\nu, \quad \text{etc.} \tag{2.63}$$

Finally we remark that in General Relativity $g^{\mu\nu}$ becomes a function of the mass distribution.

2.4.1 Relativistic Lagrangians

To derive the Lagrangian for a relativistic particle, we will begin from the equation of motion

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -\boldsymbol{\nabla} V, \tag{2.64}$$

but of course the relativistic momentum is more complicated and the simple recipe "$L = T - V$" can be readily checked *not* to work.

Two possible routes to the relativistic Lagrangian are instructive. First let us try guessing that $\gamma(v)m\mathbf{v}$ is the canonical momentum, which then requires

$$\frac{\partial L}{\partial \mathbf{v}} = \gamma(v) m \mathbf{v}. \tag{2.65}$$

Integrating this with respect to v (using explicitly that $\gamma = \left[1 - v^2/c^2\right]^{-1/2}$) then gives

$$L(\mathbf{r}, \mathbf{v}) = -\frac{1}{\gamma(v)} mc^2 - f(\mathbf{r}). \tag{2.66}$$

The 'constant of integration' $f(\mathbf{r})$ is readily determined as $V(\mathbf{r})$ by requiring that the Euler-Lagrange equations get the r.h.s. of the equation of motion correctly. You can recover the familiar non-relativistic limit by expanding in powers of v/c, and discarding the constant $-mc^2$.

The second method is more elegant but specialised to relativistic principles: let us analyse the motion in a frame of reference where it is non-relativistic, and then rewrite the analysis in a manner which is evidently frame-independent. This should then apply even when the motion appears highly relativistic. We saw that one can write the Lagrangian Action as

$$S = \int L \, \mathrm{d}t = \int (\mathbf{p} \cdot \mathrm{d}\mathbf{r} - H \, \mathrm{d}t). \tag{2.67}$$

Now $(H, c, \mathbf{p}) = p^\mu$, the (contravariant) 4-momentum, and hence we can write this in terms of four-vectors as

$$S = -\int p^\mu \, \mathrm{d}x_\mu = -\int p_\mu \, \mathrm{d}x^\mu \,. \tag{2.68}$$

Now $p_\mu \, \mathrm{d}x^\mu$ is frame-independent, and in the comoving frame of a free particle it evaluates to $mc^2 \, \mathrm{d}\tau$, where τ is the proper time. All observers thus agree on this form for the action, and for a free particle it obviously matches $mc^2\gamma(v) \, \mathrm{d}t$ which we got before.

The result is that for a free relativistic particle, the trajectory from one point in space-time (=event) to another fixed event is that which maximises the elapsed proper time. Because the equations of motion are only obeyed *after* maximising, it should be clarified that the proper time is to be evaluated on the basis of $\mathrm{d}\tau = \mathrm{d}t\,/\gamma(v)$ – that is the form for which we showed that the Euler-Lagrange equations gave the right results.

2.4.2 Relativistic Particle in Electromagnetic Field

To cope with EM interactions relativistically, we generalise the non-relativistic contribution to the action, $-\int V \, \mathrm{d}t$, to a four-vector EM potential interaction contribution, $-e\int(\phi \, \mathrm{d}t - \mathbf{A}\cdot\mathrm{d}\mathbf{r}) = -e\int A_\mu \, \mathrm{d}x^\mu$, where the scalar electric potential ϕ has been absorbed as the timelike component $\phi/c = A_t$ of the EM 4-potential and a factor of charge e has been introduced. Then the action becomes

$$S = -\int mc^2 \, \mathrm{d}\tau - \int eA_\mu \, \mathrm{d}x^\mu \, . \tag{2.69}$$

$$\text{free particle} + \text{interaction with field.} \tag{2.70}$$

Now we have to be very careful to distinguish between the mechanical momentum $\mathbf{p}^{\text{mechanical}}$ $= \gamma m\mathbf{v}$ and the canonical momentum $\mathbf{p}^{\text{canonical}} = \partial L/\partial \mathbf{v}$. Writing out carefully the Lagrangian corresponding to the expression for the action S gives:

$$L = -\frac{1}{\gamma(v)}mc^2 - e(\phi - \mathbf{v}\cdot\mathbf{A}), \tag{2.71}$$

from which

$$\mathbf{p}^{\text{canonical}} = \gamma m\mathbf{v} + e\mathbf{A}. \tag{2.72}$$

This gives an elegant form for the action of a particle interacting with an electromagnetic field,

$$S = -\int p_\mu^{\text{canonical}} \, \mathrm{d}x^\mu \, , \tag{2.73}$$

exactly as in the free particle case, except that the canonical momentum has to be written in terms of the velocity and potential using

$$p_\mu^{\text{canonical}} = \gamma(v)m\frac{\mathrm{d}x_\mu}{\mathrm{d}t} + eA_\mu. \tag{2.74}$$

2.4.3 Lagrangian vs. Hamiltonian Methods

Something you might like to check is how the Hamiltonian comes out from the relativistic Lagrangian above; it is of course a time-like quantity and not in any sense frame invariant. Hamilton's equations, because they are equations of *motion*, involve time and hence the particular frame of reference quite explicitly.

Although the Lagrangian is not itself frame independent either, the Lagrangian formulation is frame-invariant. The quantity

$$S[x^\mu(t)] = \int L \, \mathrm{d}t = -\int p_\mu \, \mathrm{d}x^\mu \,, \tag{2.75}$$

is a functional of the path $x^\mu(t)$ which is frame-invariant, as is the variational condition $\delta S = 0$.

Chapter 3

Classical Fields

Much of modern theoretical physics is, one way or another, field theory, the first example of which is the Maxwell approach to electromagnetism. Our next step is into Lagrangians depending on 'fields' rather than 'particle coordinates'. For simplicity we will start with a non-relativistic case, picking up electromagnetism as a relativistic example towards the end.

3.1 Waves in One Dimension

The basic idea is a very simple adaptation of the standard Lagrangian problem. Consider for example the longitudinal modes of an elastic rod (i.e. sound waves in one dimension). Each material point x has a displacement $\varphi(x,t)$; the *dynamical variables* are the φ, one for each value of x, the coordinate values x playing the role of labels on these (infinite number of) physical degrees of freedom.

We can write the kinetic energy as

$$T = \int \frac{1}{2}\rho\left(\frac{\partial\varphi}{\partial t}\right)^2 \mathrm{d}x\,, \tag{3.1}$$

where ρ is the mass per unit length, and the (elastic) potential energy as

$$V = \int \kappa\left(\frac{\partial\varphi}{\partial x}\right)^2 \mathrm{d}x\,, \tag{3.2}$$

where κ is (Young's Modulus)$\times$(cross-sectional area). Indeed, if we modelled this rod as a set of point masses connected by springs, each of potential energy $\frac{1}{2}K(\Delta x)^2$, we would express the total potential energy as a sum of Hookean contributions for each spring, stretched by the relative amount measured by the local displacements $\varphi(x)$:

$$V = \sum_{\{x\}} \frac{1}{2}K[\varphi(x+\delta x)-\varphi(x)]^2 \quad \Longrightarrow \quad \int \frac{1}{2}\kappa\left(\frac{\partial\varphi}{\partial x}\right)^2 \mathrm{d}x\,, \tag{3.3}$$

after transforming the discrete sum into a continuum integral and setting $\kappa = \lim_{\delta x\to 0} K\cdot\delta x$. We can now write down the Lagrangian and action, both as functionals of the **field** $\varphi(x,t)$, respectively

$$L = T - V = \int \left[\frac{1}{2}\rho\left(\frac{\partial\varphi}{\partial t}\right)^2 - \frac{1}{2}\kappa\left(\frac{\partial\varphi}{\partial x}\right)^2\right]\mathrm{d}x \equiv \int \mathcal{L}\,\mathrm{d}x\,, \tag{3.4}$$

$$\text{and,}\quad S = \int L\,\mathrm{d}t = \int \mathcal{L}\,\mathrm{d}x\,\mathrm{d}t\,, \tag{3.5}$$

where $\mathcal{L}$ is the **Lagrangian density**. We use the term 'field' here in the general sense of a function of space and time. The Lagrangian density is a function of the field φ and its derivatives:

$$\mathcal{L}\left(\varphi, \frac{\partial\varphi}{\partial t}, \frac{\partial\varphi}{\partial x}\right) = \frac{1}{2}\rho\left(\frac{\partial\varphi}{\partial t}\right)^2 - \frac{1}{2}\kappa\left(\frac{\partial\varphi}{\partial x}\right)^2. \tag{3.6}$$

The Euler-Lagrange equations from the condition of minimal action $\delta S = 0$ for this type of problem are a straightforward generalisation of the case where we had variables depending on t only. For brevity, write

$$\frac{\partial\varphi}{\partial t} = \dot{\varphi}, \quad \frac{\partial\varphi}{\partial x} = \varphi'. \tag{3.7}$$

For a small variation of the field, $\delta\varphi$, we have

$$\delta S = \int \left(\frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \frac{\partial\mathcal{L}}{\partial\varphi'}\delta\varphi' + \frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\delta\dot{\varphi} \right) \mathrm{d}x\,\mathrm{d}t\,. \tag{3.8}$$

But

$$\int \frac{\partial\mathcal{L}}{\partial\varphi'}\delta\varphi'\,\mathrm{d}x = \int \frac{\partial\mathcal{L}}{\partial\varphi'}\frac{\partial}{\partial x}\delta\varphi\,\mathrm{d}x = \left[\frac{\partial\mathcal{L}}{\partial\varphi'}\delta\varphi\right]_{-\infty}^{+\infty} - \int \frac{\partial}{\partial x}\left(\frac{\partial\mathcal{L}}{\partial\varphi'}\right)\delta\varphi\,\mathrm{d}x\,. \tag{3.9}$$

Just as in the ordinary Lagrangian problem, there are conditions which require the integrated term to vanish. In this case, for the action integral to exist we require the displacement φ, and hence also $\delta\varphi$, to vanish at $x = \pm\infty$. Similarly, for the motion in the time interval $[t_1, t_2]$,

$$\int \frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\delta\dot{\varphi})\,\mathrm{d}t = \left[\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\delta\varphi\right]_{t_1}^{t_2} - \int \frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\right)\delta\varphi\,\mathrm{d}t\,. \tag{3.10}$$

We are interested in minimising the action for given initial and final configurations $\varphi(x, t_1)$ and $\varphi(x, t_2)$, so $\delta\varphi(x, t_{1,2}) = 0$ and again the integrated term vanishes, giving

$$\delta S = \int \left[\frac{\partial\mathcal{L}}{\partial\varphi} - \frac{\partial}{\partial x}\left(\frac{\partial\mathcal{L}}{\partial\varphi'}\right) - \frac{\partial}{\partial x}\left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\right)\right]\delta\varphi\,\mathrm{d}x\,\mathrm{d}t\,. \tag{3.11}$$

This has to vanish for any $\delta\varphi(x, t) = 0$ satisfying the boundary conditions, so we obtain the Euler-Lagrange equation of motion for the field $\varphi(x, t)$:

$$\frac{\partial\mathcal{L}}{\partial\varphi} - \frac{\partial}{\partial x}\left(\frac{\partial\mathcal{L}}{\partial\varphi'}\right) - \frac{\partial}{\partial x}\left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\right) = 0. \tag{3.12}$$

Applying this to our example, we have

$$\mathcal{L} = \frac{1}{2}\rho\dot{\varphi}^2 - \frac{1}{2}\kappa\varphi'^2, \tag{3.13}$$

and so we obtain

$$0 + \frac{\partial}{\partial x}\kappa\varphi' - \frac{\partial}{\partial t}\rho\dot{\varphi} = 0, \tag{3.14}$$

which is just the one dimensional wave equation, as we should have expected. Indeed, writing it in a more familiar format, we recognise both the equation and its solution:

$$\frac{\partial^2\varphi}{\partial t^2} = \frac{\kappa}{\rho}\frac{\partial^2\varphi}{\partial x^2}, \quad \varphi \propto e^{i(\omega t - kx)}, \tag{3.15}$$

with dispersion relation

$$\omega = \sqrt{\frac{\kappa}{\rho}}k. \tag{3.16}$$

We can define a **canonical momentum density**, by analogy with $p = \partial L/\partial v$, as

$$\pi(x, t) = \frac{\partial\mathcal{L}}{\partial\dot{\varphi}} = \rho\dot{\varphi}, \tag{3.17}$$

in our example. This is sensibly analogous to our previous ideas about momentum, in particular, $p = \int \pi \, \mathrm{d}x$. In a system obeying translational invariance, when L does not explicitly depend on q_i (that is, on φ in our example), we would by analogy expect to find the momentum conservation law, although it now involves a more complicated quantity. For $\partial\mathcal{L}/\partial\varphi = 0$, our generalised Euler-Lagrange equation reads

$$\frac{\partial}{\partial t}\pi(x,t) + \frac{\partial}{\partial x}J(x,t) = 0, \tag{3.18}$$

where $J(x,t) = \partial\mathcal{L}/\partial\varphi'$ can be interpreted as the current of canonical momentum. Since $\pi(x,t)$ was the density of canonical momentum, this is just the statement that canonical momentum overall is conserved. In fact, we see that the Euler-Lagrange equation (3.12) or (3.18) for the equilibrium trajectory $\varphi(x,t)$ is the equation for momentum conservation, or the balance of local forces. If you think about balls and springs, we have arrived at the obvious result that the springs cause exchange of momentum between particles (i.e. current of momentum) but conserve momentum overall.

Again in close analogy with particle mechanics, we can define the **Hamiltonian density** $\mathcal{H}$,

$$\mathcal{H}(\varphi, \varphi', \pi) = \pi\dot{\varphi} - \mathcal{L}, \tag{3.19}$$

where $\dot{\varphi}$ is replaced by π as an independent variable. In the case of the elastic rod, this gives

$$\mathcal{H} = \frac{\pi^2}{2\rho} + \frac{1}{2}\kappa\varphi'^2, \tag{3.20}$$

which (since the kinetic energy is a homogeneous quadratic function of $\dot{\varphi}$) is just the total energy density.

3.2 Multidimensional Space

Consider now the extension of this to several dimensions of space, but keeping the physical field $\varphi(\mathbf{x},t)$ as a scalar for the present. We have

$$S = \iint \cdots \int \mathcal{L}\left(\varphi, \frac{\partial\varphi}{\partial t}, \boldsymbol{\nabla}\varphi\right) \mathrm{d}t \, \mathrm{d}x_1 \cdots \mathrm{d}x_d \,, \tag{3.21}$$

and the Euler-Lagrange equation for $\delta S = 0$ gives us

$$\frac{\partial\mathcal{L}}{\partial\varphi} = \frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial(\partial\varphi/\partial t)} + \frac{\partial}{\partial x_1}\frac{\partial\mathcal{L}}{\partial(\partial\varphi/\partial x_1)} + \cdots + \frac{\partial}{\partial x_d}\frac{\partial\mathcal{L}}{\partial(\partial\varphi/\partial x_d)}, \tag{3.22}$$

or if one wants to be more succinct about the spatial derivatives,

$$\frac{\partial\mathcal{L}}{\partial\varphi} = \frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}}{\partial(\partial\varphi/\partial t)}\right) + \boldsymbol{\nabla}\cdot\left(\frac{\partial\mathcal{L}}{\partial(\boldsymbol{\nabla}\varphi)}\right). \tag{3.23}$$

Note that we still have just one such equation, for the single physical field $\varphi(\mathbf{x},t)$ – the result of having several spatial coordinates is the multicomponent gradient on the r.h.s. The momentum density is also a scalar function, the definition (3.17) remains valid.

The condition of momentum conservation in the case when no external forces are applied, $\partial\mathcal{L}/\partial\varphi = 0$, now resembles the so-called continuity equation:

$$\dot{\pi}(\mathbf{x},t) + \boldsymbol{\nabla}\cdot\mathbf{J}(\mathbf{x},t) = 0, \quad \text{with the vector} \quad \mathbf{J} = \frac{\partial\mathcal{L}}{\partial(\boldsymbol{\nabla}\varphi)}. \tag{3.24}$$

Now it should be fairly obvious that we have in fact put time t and space $\mathbf{x}$ on the same footing, and we can simply regard time (strictly speaking, ct) as one of the coordinate variables x^μ to give

$$\frac{\partial\mathcal{L}}{\partial\varphi} = \frac{\partial}{\partial x^\mu}\left(\frac{\partial\mathcal{L}}{\partial(\,\partial\varphi/\partial x^\mu\,)}\right) \equiv \partial_\mu \frac{\partial\mathcal{L}}{\partial[\partial_\mu\varphi]}, \tag{3.25}$$

(recall that ∂_μ is shorthand for $\partial/\partial x^\mu$) as the Euler-Lagrange equation for the minimal-action condition. Here we are assuming Greek indices to run over time and space and repeated indices in the same expression are summed, just as in relativity. However it should be stressed that our equations are in no way particular to Special Relativity - though of course they very naturally encompass it, as we now explore in more detail.

3.3 Relativistic Scalar Field

In Special Relativity the action S is a Lorentz invariant quantity and Eq. (3.21) involves an integration with respect to the invariant space-time volume element $\mathrm{d}^4x = c\,\mathrm{d}t\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$. It follows that the Lagrangian density $\mathcal{L}$ is also a Lorentz invariant (scalar) function. If we require that the Euler-Lagrange equation of motion for the field be linear and at most a second-order differential equation, this limits $\mathcal{L}$ to the general form

$$\mathcal{L} = \alpha(\partial^\mu\varphi)(\partial_\mu\varphi) + \beta\partial^\mu\partial_\mu\varphi + \gamma\varphi\partial^\mu\partial_\mu\varphi + \delta\varphi + \epsilon\varphi^2, \tag{3.26}$$

where α, β, γ, δ and ϵ are *constants*. Writing this as

$$\mathcal{L} = (\alpha-\gamma)(\partial^\mu\varphi)(\partial_\mu\varphi) + \partial^\mu(\beta\partial_\mu\varphi + \gamma\varphi\partial_\mu\varphi) + \delta\varphi + \epsilon\varphi^2, \tag{3.27}$$

we note that the total derivative term $\partial^\mu(\cdots)$ can be integrated to give a (4D) surface contribution to the action, which does not affect the equation of motion since the field vanishes at infinite distance and is fixed in the distant past and future. Furthermore the equation of motion is unaffected by an overall rescaling of the action, so we may as well choose $\alpha - \gamma = \frac{1}{2}$. Thus the most general physically significant form is

$$\mathcal{L} = \frac{1}{2}(\partial^\mu\varphi)(\partial_\mu\varphi) + \delta\gamma + \epsilon\gamma^2. \tag{3.28}$$

which leads to the equation of motion

$$\partial^\mu\partial_\mu\varphi - \delta - 2\epsilon\varphi = 0. \tag{3.29}$$

(To get the first term, write $(\partial^\mu\varphi)(\partial_\mu\varphi) = g^{\mu\nu}(\partial_\nu\varphi)(\partial_\mu\varphi)$ and note that both derivatives contribute to the r.h.s. of Eq. (3.25) because of the summation convention.)

According to the boundary conditions, $\varphi = 0$ at infinity and therefore we must have $\delta = 0$. Finally, we shall see shortly that ϵ must be negative, so it is convenient to redefine

$\epsilon = -m^2/2$. In summary, the most general acceptable Lagrangian density for a real scalar field with linear dynamics is

$$\mathcal{L} = \frac{1}{2}(\partial^\mu\varphi)(\partial_\mu\varphi) - \frac{1}{2}m^2\varphi^2, \tag{3.30}$$

with the **Klein-Gordon** equation of motion,

$$\partial^\mu\partial_\mu\varphi + m^2\varphi = 0. \tag{3.31}$$

Writing out the Lagrangian density (3.30) in more detail,

$$\mathcal{L} = \frac{1}{2c^2}\left(\frac{\partial\varphi}{\partial t}\right)^2 - \frac{1}{2}(\boldsymbol{\nabla}\varphi)^2 - \frac{1}{2}m^2\varphi^2, \tag{3.32}$$

we see that the momentum density is

$$\pi = \frac{\partial\mathcal{L}}{\partial\dot{\varphi}} = \frac{1}{c^2}\frac{\partial\varphi}{\partial t}, \tag{3.33}$$

and so the Klein-Gordon Hamiltonian density is

$$\mathcal{H} = \frac{1}{2}c^2\pi^2 + \frac{1}{2}(\boldsymbol{\nabla}\varphi)^2 + \frac{1}{2}m^2\varphi^2. \tag{3.34}$$

This quantity is positive-definite if and only if the coefficient of φ^2 is positive. If the coefficient were negative, there would be field configurations with arbitrarily large negative energy, and the system would have no stable ground state. This justifies our decision to write the coefficient $(-\epsilon)$ as $m^2/2$.

3.4 Natural units

In dealing with relativistic systems it is convenient to use units such that $c = 1$. Then lengths are measured in the same units as times (the time it takes light to travel that distance), and mass in the same units as energy (the energy released by annihilating that mass). In these units Eqs. (3.32) and (3.34) become

$$\mathcal{L} = \frac{1}{2}\left(\frac{\partial\varphi}{\partial t}\right)^2 - \frac{1}{2}(\boldsymbol{\nabla}\varphi)^2 - \frac{1}{2}m^2\varphi^2, \tag{3.35}$$

$$\mathcal{H} = \frac{1}{2}\pi^2 + \frac{1}{2}(\boldsymbol{\nabla}\varphi)^2 + \frac{1}{2}m^2\varphi^2, \tag{3.36}$$

since now $\pi = \partial\varphi/\partial t$.

However, these units are still not optimal since $\mathcal{L}$ and $\mathcal{H}$ are supposed to have the dimensions of energy density, which in $c = 1$ units ($[L] = [T]$) are $[M][T]^{-3}$. That would mean that φ has to have dimensions $[M]^{1/2}[T]^{-1/2}$. Since in practice we are dealing with fields that represent phenomena on the subatomic scale, it is simplest to couple the dimensions of energy/mass and time as well, by using units in which $\hbar = c = 1$. Since $[\hbar] = [E][T]$ this means that time is measured in units of inverse energy or mass (the energy of a quantum whose angular frequency is the inverse of that time). These are called *natural units*, at least by particle physicists. For them the natural scale of energy is

measured in giga-electron-volts, GeV. Then the magic formula for converting to everyday units is

$$\hbar c = 0.2 \text{ GeV fm}, \tag{3.37}$$

where 1 fm (femtometre) is 10^{-15} m.

In natural units, $[T] = [E]^{-1} = [M]^{-1}$ and the dimensions of $\mathcal{L}$ and $\mathcal{H}$ are $[M]^4$, so φ has simply dimensions of mass. You can check that every term in Eq. (3.36) has dimension $[M]^4$, provided the constant m is itself interpreted as a mass.

We shall usually employ natural units in this section from now on. With a little practice, it is straightforward to reinsert the correct number of factors of $\hbar$ and c to convert any given expression into SI units.

3.5 Fourier Analysis

Consider first, for simplicity, solutions of the Klein-Gordon equation that depend only on x and t. They satisfy the 1-dimensional version of Eq. (3.31), i.e. (in natural units)

$$\frac{\partial^2\varphi}{\partial t^2} - \frac{\partial^2\varphi}{\partial x^2} + m^2\varphi = 0. \tag{3.38}$$

We can express any real field $\varphi(x,t)$ as a Fourier integral:

$$\varphi(x,t) = \int \mathrm{d}k\, N(k)\Big[a(k)e^{i(kx-\omega t)} + a^*(k)e^{-i(kx-\omega t)}\Big], \tag{3.39}$$

where $N(k)$ is a convenient normalizing factor for the Fourier transform $a(k)$, which we choose to satisfy $N(-k) = N(k)$ for later convenience. The frequency $\omega(k)$ is obtained by solving the equation of motion: the Klein-Gordon equation gives $\omega^2 = k^2 + m^2$ and therefore

$$\omega = +\sqrt{k^2+m^2}, \tag{3.40}$$

where we choose the positive root because Eq. (3.39) includes $+\omega$ and $-\omega$ explicitly. The Hamiltonian

$$H = \int \left(\frac{1}{2}\pi^2 + \frac{1}{2}\varphi'^2 + \frac{1}{2}m^2\varphi^2\right)\mathrm{d}x \tag{3.41}$$

takes a simpler form in terms of the Fourier amplitudes $a(k)$. We can write e.g.

$$\varphi^2 = \int \mathrm{d}k\, N(k)[\cdots] \int \mathrm{d}k'\, N(k')[\cdots], \tag{3.42}$$

and use

$$\int \mathrm{d}x\, e^{i(k\pm k')x} = 2\pi\delta(k \pm k') \tag{3.43}$$

to show that

$$\int \varphi^2\,\mathrm{d}x = 2\pi \int \mathrm{d}k\,\mathrm{d}k'\, N(k)N(k')\Big[a(k)a(k')\delta(k+k')e^{-i(\omega+\omega')t} + a^*(k)a^*(k')\delta(k+k')e^{i(\omega+\omega')t} + a(k)a^*(k')\delta(k-k')e^{-i(\omega-\omega')t} + a^*(k)a(k')\delta(k-k')e^{i(\omega-\omega')t}\Big] \tag{3.44}$$

Noting that $\omega(-k) = \omega(k)$ and with the choice $N(-k) = N(k)$, this gives

$$\int \varphi^2 \, \mathrm{d}x = 2\pi \int \mathrm{d}k \, [N(k)]^2 \Big[a(k)a(-k)e^{-2i\omega t} + a^*(k)a^*(-k)e^{2i\omega t} + a(k)a^*(k) + a^*(k)a(k) \Big]. \tag{3.45}$$

Similarly,

$$\int \varphi'^2 \, \mathrm{d}x = 2\pi \int \mathrm{d}k \, [kN(k)]^2 \Big[a(k)a(-k)e^{-2i\omega t} + a^*(k)a^*(-k)e^{2i\omega t} + a(k)a^*(k) + a^*(k)a(k) \Big], \tag{3.46}$$

while,

$$\int \dot{\varphi}^2 \, \mathrm{d}x = 2\pi \int \mathrm{d}k \, [\omega(k)N(k)]^2 \Big[-a(k)a(-k)e^{-2i\omega t} - a^*(k)a^*(-k)e^{2i\omega t} + a(k)a^*(k) + a^*(k)a(k) \Big]. \tag{3.47}$$

and hence, using $k^2 = \omega^2 - m^2$,

$$H = 2\pi \int \mathrm{d}k \, [N(k)\omega(k)]^2 \Big[a(k)a^*(k) + a^*(k)a(k) \Big], \tag{3.48}$$

or, choosing

$$N(k) = \frac{1}{2\pi \cdot 2\omega(k)}, \tag{3.49}$$

$$H = \int \mathrm{d}k \, N(k) \frac{1}{2} \omega(k) \Big[a(k)a^*(k) + a^*(k)a(k) \Big], \tag{3.50}$$

i.e. the integrated *density of modes* $N(k)$ times the *energy per mode* $\omega(k)|a(k)|^2$.

Hence each normal mode of the system behaves like an independent harmonic oscillator with amplitude $a(k)$.

In 3 spatial dimensions we write

$$\varphi(\mathbf{r}, t) = \int \mathrm{d}^3\mathbf{k} \, N(\mathbf{k}) \Big[a(\mathbf{k}) e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} + a^*(\mathbf{k}) e^{-i(\mathbf{k}\cdot\mathbf{r} - \omega t)} \Big], \tag{3.51}$$

and use

$$\int \mathrm{d}^3\mathbf{r} \, e^{i(\mathbf{k} \pm \mathbf{k}')\cdot\mathbf{r}} = (2\pi)^3 \delta^3(\mathbf{k} \pm \mathbf{k}'). \tag{3.52}$$

Therefore we should choose

$$N(\mathbf{k}) = \frac{1}{(2\pi)^3 2\omega(\mathbf{k})} \tag{3.53}$$

to obtain an integral with the usual relativistic phase space (density of states) factor:

$$H = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3 2\omega(\mathbf{k})} \omega(\mathbf{k}) |a(\mathbf{k})|^2. \tag{3.54}$$

3.6 Multi-Component Fields

We now look at examples where the field itself has a more complicated structure. In principle, there is no significant difference and the analogies we have been observing will continue to hold, if each component of the physical field is regarded as an independent scalar variable as was studied above. Let us consider first an example with an intuitively clear 2-dimensional vector field.

3.6.1 Transverse Waves on a String

Instead of longitudinal modes of a rod, consider small transverse displacements $\boldsymbol{\varphi} = (\varphi_y, \varphi_z)$ of a flexible elastic string stretched along the x-axis at constant tension F. Then the kinetic energy is

$$T = \frac{1}{2}\rho \int \left[\left(\frac{\partial \varphi_y}{\partial t} \right)^2 + \left(\frac{\partial \varphi_z}{\partial t} \right)^2 \right] \mathrm{d}x \,, \tag{3.55}$$

and the elastic potential energy is

$$\begin{aligned} V = F\left[\int \mathrm{d}s - \int \mathrm{d}x \right] &= F \int \left[\sqrt{1 + \left(\frac{\partial \varphi_y}{\partial x} \right)^2 + \left(\frac{\partial \varphi_z}{\partial x} \right)^2} - 1 \right] \mathrm{d}x \\ &= \frac{1}{2} F \int \left[\left(\frac{\partial \varphi_y}{\partial x} \right)^2 + \left(\frac{\partial \varphi_z}{\partial x} \right)^2 \right] \mathrm{d}x \,, \end{aligned} \tag{3.56}$$

for small displacements (and small displacement gradients). Therefore the Lagrangian density just becomes

$$\mathcal{L} = \frac{1}{2} \sum_{j-y,z} \left[\rho \left(\frac{\partial \varphi_j}{\partial t} \right)^2 - F \left(\frac{\partial \varphi_j}{\partial x} \right)^2 \right]. \tag{3.57}$$

The action is to be mimimised with respect to variations in both φ_y and φ_z, so we now get Euler-Lagrange conditions for each component:

$$\frac{\partial \mathcal{L}}{\partial \varphi_j} = \frac{\partial}{\partial t} \left(\frac{\partial \mathcal{L}}{\partial (\partial \varphi_j / \partial t)} \right) + \frac{\partial}{\partial x} \left(\frac{\partial \mathcal{L}}{\partial (\partial \varphi_j / \partial x)} \right), \tag{3.58}$$

which in this case give identical wave equations for $j = y$ and z:

$$0 = \frac{\partial}{\partial t} \rho \left(\frac{\partial \varphi_j}{\partial t} \right) - \frac{\partial}{\partial x} F \left(\frac{\partial \varphi_j}{\partial x} \right). \tag{3.59}$$

Thus transverse waves propagate with velocity $\sqrt{F/p}$, independent of the direction of the displacement vector φ, i.e. independent of their *polarisation*.

For a multi-component field in multidimensional space we again have to regard each component of the vector field φ as a separate field giving us an Euler-Lagrange condition for each of these components:

$$\frac{\partial \mathcal{L}}{\partial \varphi_j} = \frac{\partial}{\partial t} \left(\frac{\partial \mathcal{L}}{\partial (\partial \varphi_j / \partial t)} \right) + \boldsymbol{\nabla} \cdot \left(\frac{\partial \mathcal{L}}{\partial (\boldsymbol{\nabla} \varphi_j)} \right). \tag{3.60}$$

For example, in the case of elastic waves in a three-dimensional medium, $\boldsymbol{\varphi}$ represents the displacement of a material point in the medium and Eq. (3.60) leads to the **acoustic wave equation**. For a general medium, the wave velocity now depends on the polarization, but the details of this would take us too far into the theory of elasticity.

3.7 Complex Scalar Field

Suppose φ is a complex scalar field, i.e. $\varphi^* \neq \varphi$. We can always decompose it into

$$\varphi = \frac{1}{\sqrt{2}} (\varphi_1 + i\varphi_2), \tag{3.61}$$

where φ_1 and φ_2 are real. Then, writing φ_1 and φ_2 as Fourier integrals as in Eq. (3.39),

$$\varphi(x,t) = \int \mathrm{d}k\, N(k)\Big[a(k)e^{i(kx-\omega t)} + b^*(k)e^{-i(kx-\omega t)}\Big], \tag{3.62}$$

where

$$a = \frac{1}{\sqrt{2}}(a_1 + ia_2), \quad b^* = \frac{1}{\sqrt{2}}(a_1^* + ia_2^*) \neq a^*. \tag{3.63}$$

The Lagrangian density

$$\mathcal{L} = \mathcal{L}[\varphi_1] + \mathcal{L}[\varphi], \tag{3.64}$$

can be written as

$$\mathcal{L} = \frac{\partial\varphi^*}{\partial t}\frac{\partial\varphi}{\partial t} - \frac{\partial\varphi^*}{\partial x}\frac{\partial\varphi}{\partial x} - m^2\varphi^*\varphi. \tag{3.65}$$

The canonical momentum densities conjugate to φ and φ^* are thus

$$\pi = \frac{\partial\mathcal{L}}{\partial\dot{\varphi}} = \frac{\partial\varphi^*}{\partial t}, \quad \pi^* = \frac{\partial\mathcal{L}}{\partial\dot{\varphi}^*} = \frac{\partial\varphi}{\partial t}, \tag{3.66}$$

and the Hamiltonian density is

$$\mathcal{H} = \pi\dot{\varphi} + \pi^*\dot{\varphi}^* - \mathcal{L} = \pi^*\pi + \frac{\partial\varphi^*}{\partial x}\frac{\partial\varphi}{\partial x} + m^2\varphi^*\varphi. \tag{3.67}$$

Using the Fourier expansion of φ and integrating over all space (left as an exercise!), we find

$$\begin{aligned} H = \int \mathrm{d}x\,\mathcal{H} &= \frac{1}{2}\int \mathrm{d}k\, N(k)\omega(k)\Big[a(k)a^*(k) + a^*(k)a(k) + b(k)b^*(k) + b^*(k)b(k)\Big] \\ &= \int \mathrm{d}k\, N(k)\omega(k)\Big[|a(k)|^2 + |b(k)|^2\Big]. \end{aligned} \tag{3.68}$$

Therefore the positive and negative frequency Fourier components of the field contribute to the energy with the same (positive) sign.

In 3 spatial dimensions Eq. (3.65) becomes

$$\begin{aligned} \mathcal{L} &= \frac{\partial\varphi^*}{\partial t}\frac{\partial\varphi}{\partial t} - \boldsymbol{\nabla}\varphi^*\cdot\boldsymbol{\nabla}\varphi - m^2\varphi^*\varphi \\ &= \partial_\mu\varphi^*\partial^\mu\varphi - m^2\varphi^*\varphi, \end{aligned} \tag{3.69}$$

and the Fourier decomposition of the field is

$$\varphi = \int \mathrm{d}^3\mathbf{k}\, N(\mathbf{k})\Big[a(\mathbf{k})e^{-ik\cdot x} + b^*(\mathbf{k})e^{ik\cdot x}\Big] \tag{3.70}$$

where $N(\mathbf{k})$ is given by Eq. (3.53) and for brevity we have introduced the wave 4-vector $k^\mu = (\omega/c, \mathbf{k})$ in the exponents, so that

$$k\cdot x = k^\mu x_\mu = \omega t - \mathbf{k}\cdot\mathbf{r}. \tag{3.71}$$

Then in place of (3.68) we have

$$H = \int \mathrm{d}^3\mathbf{k}\, N(\mathbf{k})\omega(\mathbf{k})\Big[|a(\mathbf{k})|^2 + |b(\mathbf{k})|^2\Big]. \tag{3.72}$$

3.8 Electromagnetic Field

Finally, let us have a look at the electromagnetic field, although this is a much more difficult and involved subject! The starting point has to be the four-potential A_μ in terms of which the physical fields $\mathbf{E}$ and $\mathbf{B}$ can be found (with a bit of effort) amongst the components of the electromagnetic field strength tensor

$$F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha \quad \text{or} \quad F_{\alpha\beta} = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix}. \tag{3.73}$$

Note that $F_{\alpha\beta}$ is antisymmetric (and therefore has zero diagonal elements). This is an example of a physical variable expressed by a second-rank tensor field. Now one needs to construct the appropriate Lagrangian density from it.

As we already discussed for the relativistic scalar field, given that the Lagrangian action S is a frame invariant scalar, it follows we should expect the Lagrangian density $\mathcal{L}$ to be a scalar also. This rather limits the possibilities of how we could construct $\mathcal{L}$ from the traceless second-rank tensor $F_{\alpha\beta}$.

We want $\mathcal{L}$ to be a scalar, and to give us Euler-Lagrange equations which are linear in the physical fields $\mathbf{E}$ and $\mathbf{B}$ (the components of $F_{\alpha\beta}$). Therefore, we need $\mathcal{L}$ to be at most quadratic in these components and the only possible form is $\mathcal{L}_0 = aF_{\alpha\beta}F^{\alpha\beta}$ (a linear term, if it existed, would have to be the trace of $F^{\alpha\beta}$, which is zero by construction). This is the first level of approximation, corresponding to Maxwell electromagnetism. Extensions could arise, for instance, by bringing in spatial gradients: squared powers of $\partial_\mu F^{\alpha\beta}$ would lead to a variety of effects such as spatial dispersion, optical rotation, etc. Let us stay on the most basic level here.

In view of what we found for the single relativistic particle, it is evidently prudent to anticipate the coupling of the free field $F^{\alpha\beta}$ to the electric current distribution, characterised by the four-current J^μ, defined as $\rho(\mathrm{d}x^\mu/\mathrm{d}t)$ with ρ the charge density. The timelike component of J^μ is just the charge density (with a factor c), while the spacelike components are $\rho\mathbf{v} \equiv \mathbf{J}$, the density of electric current. This corresponds to the external force term in the potential energy of the elastic Lagrangians in the preceding examples and, in general, leads to a non-zero l.h.s. in the corresponding Euler-Lagrange equations, such as (3.12) and (3.60). In the electromagnetic field case this then leads us to consider the Lagrangian density

$$\mathcal{L} = aF_{\alpha\beta}F^{\alpha\beta} - J^\mu A_\mu. \tag{3.74}$$

Another important constraint in electromagnetism, which we have already discussed in Section 2.3, is gauge invariance: we can let $a_\mu \to A_\mu + \partial_\mu f$ (in 4-vector notation), where f is any scalar function, without altering the physical fields $F_{\alpha\beta}$. Evidently our Lagrangian density with the coupling to an external current is not gauge invariant, but if you compute the corresponding change in the action S you find

$$\Delta S = -\int J^\mu(\partial_\mu f)\,\mathrm{d}^4x = +\int f(\partial_\mu J^\mu)\,\mathrm{d}^4x + (\text{boundary current terms}), \tag{3.75}$$

and the r.h.s. vanishes if the current J^μ is conserved, that is if the 4-gradient $\partial_\mu J^\mu = 0$ (and if the current does not flow out through the boundaries). Therefore coupling the

vector potential to conserved currents leaves the Lagrangian Action (but not $\mathcal{L}$) gauge invariant.

Now we have to check that all this really works, that is, it leads to real electromagnetism. For instance, let us check that we do get Maxwell's equations for the field components! Clearly, they have to satisfy the minimal-action condition, i.e. to be the relevant Euler-Lagrange equations. All we need to do is to rewrite the canonical form in an appropriate way.

Let us take the four-potential A_μ as the basic field variable of the Euler-Lagrange equations for $\delta S = 0$. We have then

$$\frac{\partial \mathcal{L}}{\partial A_\alpha} = \frac{\partial}{\partial x^\mu}\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\alpha)}\right), \tag{3.76}$$

and the l.h.s. immediately gives $-J^\alpha$, the external "force". For the r.h.s. we need to calculate the derivative

$$\frac{\partial}{\partial(\partial_\mu A_\alpha)} a F_{\delta\gamma} F^{\delta\gamma} = a F^{\delta\gamma} \frac{\partial}{\partial(\partial_\mu A_\alpha)} F_{\delta\gamma} + a F_{\delta\gamma} \frac{\partial}{\partial(\partial_\mu A_\alpha)} F^{\delta\gamma}. \tag{3.77}$$

It is not too hard to convince yourself that the two terms are in fact equal, and that by permuting indices each of these is equal to

$$2a F^{\delta\gamma} \frac{\partial}{\partial(\partial_\mu A_\alpha)} \partial_\delta A_\gamma = 2a F^{\mu\alpha}. \tag{3.78}$$

The Euler-Lagrange equations therefore reduce to the 4-vector relation

$$J^\alpha + 4a \partial_\mu F^{\mu\alpha} = 0, \tag{3.79}$$

or if you prefer

$$J^\alpha + 4a(\partial_\mu \partial^\mu A^\alpha - \partial^\alpha \partial_\mu A^\mu) = 0. \tag{3.80}$$

With a suitable choice of the constant, $a = -1/4\mu_0$, these are just the inhomogeneous pair of Maxwell equations (recall here that $c^2 = 1/\epsilon_0\mu_0$)

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad \boldsymbol{\nabla} \times \mathbf{B} = \epsilon_0\mu_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{J}. \tag{3.81}$$

The important continuity equation is then obtained by covariant differentiation of (3.79):

$$\partial_\mu \partial_\nu F^{\mu\nu} = 0 = \partial_\nu J^\nu = \mu_0 \left(\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{J}\right), \tag{3.82}$$

meaning that the charge is conserved (recall that $F^{\mu\nu}$ is antisymmetric and therefore vanishes whenever both indices are contracted with a symmetric tensor, $\partial_\mu \partial_\nu$ in this case). The other pair of Maxwell equations is actually contained in the definition of the antisymmetric tensor $F^{\mu\nu}$:

$$\partial^\lambda F^{\mu\nu} + \partial^\nu F^{\lambda\mu} + \partial^\mu F^{\nu\lambda} = 0, \tag{3.83}$$

(the so-called *Bianchi identity*). When written out explicitly, this gives $\boldsymbol{\nabla} \cdot \mathbf{B} = \boldsymbol{\nabla} \times \mathbf{E} + \dot{\mathbf{B}} = 0$.

To sum up much of what we have covered regarding the electromagnetic field, the action for the EM field plus charged relativistic particles is given by

$$S = \sum_{\text{particles}} \left[-\int mc^2 \, \mathrm{d}\tau \quad - \int eA_\mu \, \mathrm{d}x^\mu \, (t) \right] - \frac{1}{4\mu_0} \int F_{\alpha\beta} F^{\alpha\beta} \, \mathrm{d}^4 x \, , \tag{3.84}$$

free particles	coupling to EM field	free EM field
Section 2.4	Section 2.3	here

and the condition $\delta S = 0$ gives both the motion of the particles in the field and the dynamics of the field due to the particles. This means that the full relativistic electromagnetic interactions between the particles (retardation, radiation and all) are included, excluding of course quantum mechanical effects.

3.8.1 Gauge Invariance

The gauge transformations (2.53) and (2.54) of the EM potentials are simply expressed in covariant form as

$$A'_\mu = A_\mu + \partial_\mu f. \tag{3.85}$$

The invariance of electromagnetism with respect to this transformation allows us to impose one constraint on A_μ, for example the axial gauge condition $n^\mu A_\mu = 0$, which now includes the choice $\phi = 0$ or $\mathbf{n} \cdot \mathbf{A} = 0$, according to the choice of the arbitrary 4-vector n^μ.

We also already mentioned the Lorenz gauge, where the condition (2.55) in covariant notation becomes $\partial_\mu A^\mu = 0$, which is manifestly Lorentz invariant. Furthermore this choice leaves Eq. (3.80) as simply the wave equation $\partial_\mu \partial^\mu A^\alpha$ in the absence of charges. The residual ambiguity (2.56) in this gauge similarly takes the form $\partial_\mu \partial^\mu f = 0$..

We shall have a good deal more to say about gauge invariance after we have considered in more detail the relationship between symmetries and conservation laws in the next chapter.

Chapter 4

Symmetries and Conservation Laws

The relationship between symmetries and conserved quantities, and the effects of symmetry breaking, are amongst the most important in theoretical physics. We start with the simplest case of the scalar (Klein-Gordon) field, then add electromagnetism. Finally we introduce the transition from classical to quantum fields, which clarifies the interpretation of conserved quantities such as energy and charge.

4.1 Noether's Theorem

Let us try to find a current and a density that satisfy the continuity equation for the complex Klein-Gordon field. We use an important general result called *Noether's theorem* (Emmy Noether, 1918), which tells us that there is a *conserved current* associated with every continuous *symmetry* of the Lagrangian, i.e. with symmetry under a transformation of the form

$$\varphi \to \varphi + \delta\varphi, \tag{4.1}$$

where $\delta\varphi$ is infinitesimal. Symmetry means that $\mathcal{L}$ doesn't change:

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \frac{\partial\mathcal{L}}{\partial\varphi'}\delta\varphi' + \frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\delta\dot{\varphi} = 0, \tag{4.2}$$

where

$$\delta\varphi' = \delta\left(\frac{\partial\varphi}{\partial x}\right) = \frac{\partial}{\partial x}\delta\varphi \tag{4.3}$$

$$\delta\dot{\varphi} = \delta\left(\frac{\partial\varphi}{\partial t}\right) = \frac{\partial}{\partial t}\delta\varphi \tag{4.4}$$

(easily generalised to 3 spatial dimensions).

The Euler-Lagrange equation of motion

$$\frac{\partial\mathcal{L}}{\partial\varphi} - \frac{\partial}{\partial x}\left(\frac{\partial\mathcal{L}}{\partial\varphi'}\right) - \frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\right) = 0, \tag{4.5}$$

then implies that

$$\begin{aligned}\delta\mathcal{L} = \frac{\partial}{\partial x}\left(\frac{\partial\mathcal{L}}{\partial\varphi'}\right)\delta\varphi + \frac{\partial\mathcal{L}}{\partial\varphi'}\frac{\partial}{\partial x}(\delta\varphi) + \frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\right)\delta\varphi + \frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial}{\partial t}(\delta\varphi) = 0 \\ \implies \quad \frac{\partial}{\partial x}\left(\frac{\partial\mathcal{L}}{\partial\varphi'}\delta\varphi\right) + \frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\delta\varphi\right) = 0.\end{aligned} \tag{4.6}$$

Comparing with the conservation/continuity equation (in 1 dimension)

$$\frac{\partial}{\partial x}(J_x) + \frac{\partial\rho}{\partial t} = 0, \tag{4.7}$$

we see that the conserved density and current are (proportional to)

$$\rho = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} \delta\varphi, \quad J_x = \frac{\partial \mathcal{L}}{\partial \varphi'} \delta\varphi. \tag{4.8}$$

In 3 spatial dimensions

$$J_x = \frac{\partial \mathcal{L}}{\partial(\,\partial\varphi/\partial x\,)} \delta\varphi, \quad J_y \frac{\partial \mathcal{L}}{\partial(\,\partial\varphi/\partial y\,)} \delta\varphi, \quad \cdots \quad , \tag{4.9}$$

and hence in covariant notation

$$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \delta\varphi. \tag{4.10}$$

If the Lagrangian involves several fields $\varphi_1, \varphi_2, \cdots$, the symmetry may involve changing them all: invariance w.r.t. $\varphi_j \to \varphi_j + \delta\varphi_j$ then implies the existence of a conserved *Noether current*

$$J^\mu = \sum_j \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_j)} \delta\varphi_j. \tag{4.11}$$

In general the transformation may mix the different fields, so that

$$\delta\varphi_j = \varepsilon \sum_k t_{jk} \varphi_k, \tag{4.12}$$

where ε is a small parameter and t_{ik} are constants. Then, dividing out ε, the Noether current is

$$J^\mu = \sum_{j,k} t_{jk} \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_j)} \delta\varphi_k. \tag{4.13}$$

4.2 Global Phase Symmetry

As an important example, consider the Klein-Gordon Lagrangian density for a complex field, Eq. (3.69):

$$\mathcal{L} = \partial_\mu \varphi^* \partial^\mu \varphi - m^2 \varphi^* \varphi. \tag{4.14}$$

This is invariant under a *global phase change* in φ:

$$\varphi \to e^{-i\varepsilon} \varphi \simeq \varphi - i\varepsilon\varphi \tag{4.15}$$

$$\varphi^* \to e^{+i\varepsilon} \varphi^* \simeq \varphi^* + i\varepsilon\varphi^*, \tag{4.16}$$

i.e. $\delta\varphi \propto -i\varphi$, $\delta\varphi^* \propto +i\varphi^*$. By 'global' we mean that the phase change ε is the same at all points in space-time.

The corresponding conserved Noether current is

$$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \delta\varphi + \delta\varphi^* \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi^*)} = -i(\partial^\mu \varphi^*)\varphi + i\varphi^*(\partial^\mu \varphi). \tag{4.17}$$

We can define an associated *conserved charge*, which is the integral of ρ over all space:

$$Q = \int \rho \, \mathrm{d}^3\mathbf{r} \,. \tag{4.18}$$

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \int \frac{\partial \rho}{\partial t} \, \mathrm{d}^3\mathbf{r} = -\int \boldsymbol{\nabla} \cdot \mathbf{J} \, \mathrm{d}^3\mathbf{r} = -\int_{\infty \text{ sphere}} \mathbf{J} \cdot \mathrm{d}\mathbf{S} = 0. \tag{4.19}$$

In this case

$$Q = -i \int \left(\frac{\partial \varphi^*}{\partial t} \varphi - \varphi^* \frac{\partial \varphi}{\partial t} \right) \mathrm{d}^3\mathbf{r} \,. \tag{4.20}$$

Inserting the Fourier decomposition (3.70) for the field,

$$\varphi = \int \mathrm{d}^3\mathbf{r} \, N(\mathbf{k}) \Big[a(\mathbf{k}) e^{-ik\cdot x} + b^*(\mathbf{k}) e^{ik\cdot x} \Big], \tag{4.21}$$

we find (another exercise!)

$$Q = \int \mathrm{d}^3\mathbf{k} \, N(\mathbf{k}) \Big[|a(\mathbf{k})|^2 - |b(\mathbf{k})|^2 \Big]. \tag{4.22}$$

Therefore the positive and negative frequency Fourier components of the field contribute to the charge with *opposite* signs, in contrast to their contributions to the energy. We shall see that, when a complex classical field is quantised, the quanta can be either *particles* (quanta with positive energy and positive charge) or *antiparticles* (quanta with positive energy and negative charge).

4.3 Local Phase (Gauge) Symmetry

Suppose now that we make a *local* phase change in a complex scalar field φ, i.e. we allow the phase ε to be a function of the space-time coordinates x^μ. This is not a symmetry of the free-field Klein-Gordon Lagrangian $\mathcal{L}_{KG}$ because

$$\varphi \to e^{-i\varepsilon(x)} \varphi \implies \partial^\mu \varphi \to e^{-i\varepsilon(x)} [(\partial^\mu \varphi) - i(\partial^\mu \varepsilon)\varphi], \tag{4.23}$$

and therefore

$$\begin{aligned} \mathcal{L}_{KG} &\to [\partial_\mu \varphi^* + i(\partial_\mu \varepsilon)\varphi^*][\partial^\mu \varphi - i(\partial^\mu \varepsilon)\varphi] - m^2 \varphi^* \varphi \\ &= \mathcal{L}_{KG} - i(\partial_\mu \varepsilon)[(\partial^\mu \varphi^*)\varphi - \varphi^*(\partial^\mu \varphi)] + (\partial_\mu \varepsilon)(\partial^\mu \varepsilon)\varphi^* \varphi. \end{aligned} \tag{4.24}$$

Notice that the second term on the r.h.s. is proportional to the current $J^\mu = -i[(\partial^\mu \varphi^*)\varphi - \varphi^*(\partial^\mu \varphi)]$. In the presence of an electromagnetic field this term will cancel with the change in the interaction term $-eJ^\mu A_\mu$ if we make a simultaneous gauge transformation

$$A_\mu \to A_\mu + \partial_\mu \varepsilon / e, \tag{4.25}$$

i.e. we choose the scalar function f in Eq. (3.85) to be ε/e. And in fact the last term will also cancel if we introduce the electromagnetic interaction through the so-called *covariant derivative*

$$\partial_\mu \to D_\mu = \partial_\mu + ieA_\mu, \tag{4.26}$$

for then the combined effect of the phase change in φ and the gauge change in A_μ will be such that

$$D_\mu\varphi \to [\partial_\mu + ieA_\mu + i(\partial_\mu\varepsilon)]e^{-i\varepsilon}\varphi = e^{-i\varepsilon}D_\mu\varphi. \tag{4.27}$$

In other words, the covariant derivative of φ transforms in the same way as φ itself, so that

$$\mathcal{L}_{KG} = (D_\mu\varphi)^*(D^\mu\varphi) - m^2\varphi^*\varphi \tag{4.28}$$

remains unchanged.

This is a profound result: we have found that the electromagnetic field is an essential requirement if the theory is to remain invariant under local phase transformations of a complex (i.e. charged) field. Furthermore the interaction between the fields must be of the form prescribed by the covariant derivative (4.26). Bearing in mind that in quantum mechanics the *canonical* 4-momentum is obtained from the wave function using the operator $i\partial^\mu$ (in units where $\hbar = 1$), we see that the covariant derivative operator D^μ represents the *mechanical* 4-momentum.

4.4 Electromagnetic Interaction

Expanding Eq. (4.28), we find

$$\mathcal{L}_{KG} = \partial_\mu\varphi^*\partial^\mu\varphi - m^2\varphi^*\varphi + ieA_\mu[(\partial^\mu\varphi^*)\varphi - \varphi^*(\partial^\mu\varphi)] + e^2A_\mu A^\mu\varphi^*\varphi. \tag{4.29}$$

The first two terms on the r.h.s. are those of the free-field Klein-Gordon equation. The third is the expected interaction term $eA_\mu J^\mu$ where J^μ is the free-field current (4.17). In addition, we now have a surprising extra term, quadratic in the electromagnetic potential, which is required to preserve gauge invariance.

Let us now apply Noether's theorem to deduce the conserved current associated with invariance of the combined charged scalar + electromagnetic field Lagrangian density

$$\mathcal{L} = \mathcal{L}_{em} + \mathcal{L}_{KG} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\varphi)^*(D^\mu\varphi) - m^2\varphi^*\varphi, \tag{4.30}$$

under the infinitesimal local phase + gauge transformation

$$\varphi \to \varphi - ie\varepsilon\varphi, \quad A_\mu \to a_\mu + \partial_\mu\varepsilon. \tag{4.31}$$

We have

$$J^\mu \propto -ie\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\varepsilon\varphi + ie\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi^*)}\varepsilon\varphi^* + \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)}\partial_\nu\varepsilon. \tag{4.32}$$

The first two terms get contributions only from $\mathcal{L}_{KG}$, proportional to

$$J^\mu_{KG} = ie[\varphi^*(D^\mu\varphi) - (D^\mu\varphi)^*\varphi] = ie[\varphi^*(\partial^\mu\varphi) - (\partial^\mu\varphi)^*\varphi] - 2e^2A^\mu\varphi^*\varphi. \tag{4.33}$$

Thus the Klein-Gordon current is modified in the presence of the electromagnetic field, and the interaction of the extra piece gives rise to the final term in Eq. (4.29).

The third term in Eq. (4.32) gets a contribution only from $\mathcal{L}_{em}$:

$$J^\mu_{em} \propto -F^{\mu\nu}\partial_\nu\varepsilon = -\partial_\nu(F^{\mu\nu}\varepsilon) + (\partial_\nu F^{\mu\nu})\varepsilon. \tag{4.34}$$

As usual, we drop the first term on the r.h.s. since a total derivative cannot contribute to the charge as long as the fields vanish on the surface of the integration region. This leaves us with the current

$$J^{\mu}_{em} = \partial_{\nu} F^{\mu\nu}, \tag{4.35}$$

which is indeed conserved since

$$\partial_{\mu} J^{\mu}_{em} = \partial_{\mu}\partial_{\nu} F^{\mu\nu} = 0, \tag{4.36}$$

by the antisymmetry of the field strength tensor $F^{\mu\nu}$.

4.5 Stress-Energy (-Momentum) Tensor

If the symmetry involves a space-time transformation instead of (or in addition to) a redefinition of the fields at each space-time point, the situation is a little more complicated, because after the transformation the Lagrangian density is defined at the transformed point. Let us consider in particular a small space-time displacement, $x^{\mu} \to x^{\mu} + \varepsilon^{\mu}$. The corresponding field transformation is (always working to first order in ε)

$$\varphi(x^{\mu}) \to \tilde{\varphi}(x^{\mu}) = \varphi(x^{\mu} + \varepsilon^{\mu}) = \varphi + \varepsilon^{\mu}\partial_{\mu}\varphi = \varphi + \varepsilon_{\mu}\partial^{\mu}\varphi. \tag{4.37}$$

Provided the Lagrangian does not depend *explicitly* on the space-time coordinates, for this to be a symmetry transformation we require

$$\mathcal{L}(\tilde{\varphi}, \partial^{\mu}\tilde{\varphi}) = \mathcal{L}(x^{\mu} + \varepsilon^{\mu}) = \mathcal{L} + \varepsilon_{\mu}\partial^{\mu}\mathcal{L}. \tag{4.38}$$

Combining this condition with the expansion

$$\mathcal{L}(\tilde{\varphi}, \partial^{\mu}\tilde{\varphi}) = \mathcal{L}(\varphi, \partial^{\mu}\varphi) + \varepsilon_{\mu}\frac{\partial\mathcal{L}}{\partial\varphi}\partial^{\mu}\varphi + \varepsilon_{\mu}\frac{\partial\mathcal{L}}{\partial(\partial^{\nu}\varphi)}\partial^{\mu}\partial^{\nu}\varphi, \tag{4.39}$$

and recalling that ε_{μ} is an arbitrary constant, we obtain

$$\frac{\partial\mathcal{L}}{\partial\varphi}\partial^{\mu}\varphi + \frac{\partial\mathcal{L}}{\partial(\partial^{\nu}\varphi)}\partial^{\mu}\partial^{\nu}\varphi = \partial^{\mu}\mathcal{L}. \tag{4.40}$$

As in the derivation of the Noether current, we may now use the equation of motion to write this as

$$\partial^{\nu}\left(\frac{\partial\mathcal{L}}{\partial(\partial^{\nu}\varphi)}\partial^{\mu}\varphi\right) = \partial^{\mu}\mathcal{L}, \tag{4.41}$$

or in other words

$$\partial^{\nu}\left(\frac{\partial\mathcal{L}}{\partial(\partial^{\nu}\varphi)}\partial^{\mu}\varphi - \delta^{\mu}_{\nu}\mathcal{L}\right) = 0. \tag{4.42}$$

Relabelling and rearranging indices, we see that this implies that the **stress-energy tensor**,

$$T^{\mu\nu} = \frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\varphi)}\partial^{\nu}\varphi - g^{\mu\nu}\mathcal{L}, \tag{4.43}$$

(sometimes called the energy-momentum tensor) is conserved:

$$\partial_{\mu} T^{\mu\nu} = 0. \tag{4.44}$$

As in the case of the Noether current, for multi-component fields we have simply to add the contributions of the components:

$$T^{\mu\nu} = \sum_{j} \frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\varphi_j)}\partial^{\nu}\varphi_j - g^{\mu\nu}\mathcal{L}. \tag{4.45}$$

4.5.1 Longitudinal waves in 1D

Although we have used relativistic notation, these results are not limited to covariant systems, as long as the Lagrangian density is invariant under translations in space and time. For longitudinal waves on an elastic rod, for example, we saw that

$$\mathcal{L} = \frac{1}{2}\rho(\dot{\varphi})^2 - \frac{1}{2}\kappa(\varphi')^2, \tag{4.46}$$

and therefore

$$\begin{aligned} T_{tt} &= \rho(\dot{\varphi})^2 - \mathcal{L} = \mathcal{H}, & T_{tx} &= -\rho\dot{\varphi}\varphi', \\ T_{xx} &= \kappa(\varphi')^2 + \mathcal{L} = \mathcal{H}, & T_{xt} &= -\kappa\dot{\varphi}\varphi'. \end{aligned} \tag{4.47}$$

As expected, T_{tt} is the Hamiltonian density, i.e. the wave energy per unit length in the rod. Furthermore from Eq. (4.44)

$$\frac{\partial T_{tt}}{\partial t} = -\frac{\partial T_{xt}}{\partial x}, \tag{4.48}$$

as can easily be verified using the equation of motion, so that $T_{xt} = -\kappa\dot{\varphi}\varphi'$ must represent the flow of wave energy along the rod. Similarly

$$\frac{\partial T_{tx}}{\partial t} = -\frac{\partial T_{xx}}{\partial x}, \tag{4.49}$$

where T_{tx} is the momentum density in the wave and T_{xx} is the associated flow of momentum (which in this case coincides again with the energy (Hamiltonian) of the system, as per (4.47)).

4.5.2 Relativistic Scalar Field

In the case of a relativistic scalar field, we saw that the field must satisfy the Klein-Gordon equation, with Lagrangian density (3.30):

$$\mathcal{L} = \frac{1}{2}(\partial^\mu\varphi)(\partial_\mu\varphi) - \frac{1}{2}m^2\varphi^2. \tag{4.50}$$

Thus the stress-energy tensor is

$$T^{\mu\nu} = (\partial^\mu\varphi)(\partial^\nu\varphi) - g^{\mu\nu}\mathcal{L}. \tag{4.51}$$

4.5.3 Electromagnetic Field

For the free electromagnetic field (in units where $\mu_0 = \epsilon_0 = c = 1$) we have

$$\mathcal{L} = -\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} = -\frac{1}{4}g^{\alpha\gamma}g^{\beta\delta}(\partial_\alpha A_\beta - \partial_\beta A_\alpha)(\partial_\gamma A_\delta - \partial_\delta A_\gamma), \tag{4.52}$$

and the stress-energy tensor is

$$\begin{aligned} T^{\mu\nu} &= \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\lambda)}\partial^\nu A_\lambda - g^{\mu\nu}\mathcal{L} \\ &= -F^{\mu\lambda}\partial^\nu A_\lambda + \frac{1}{4}g^{\mu\nu}F_{\alpha\beta}F^{\alpha\beta} \\ &= -F^\mu_\lambda\partial^\nu A^\lambda + \frac{1}{4}g^{\mu\nu}F_{\alpha\beta}F^{\alpha\beta}. \end{aligned} \tag{4.53}$$

While this is indeed a conserved tensor, it is not in a convenient form since it is not gaugeinvariant and cannot be expressed in terms of the field strengths $\mathbf{E}$ and $\mathbf{B}$. Notice, however, that we are free to redefine $T^{\mu\nu}$ by adding any tensor of the form $\partial_\lambda \Omega^{\lambda\mu\nu}$ where $\Omega^{\lambda\mu\nu}$ is antisymmetric with respect to the indices λ and μ, for then $\partial_\mu\partial_\lambda\Omega^{\lambda\mu\nu} = 0$ (this is true for any stress-energy tensor, not only in the case of an electromagnetic field). Let us choose

$$\Omega^{\lambda\mu\nu} = -F^{\lambda\mu}A^\nu. \tag{4.54}$$

so that the added terms are

$$\partial_\lambda \Omega^{\lambda\mu\nu} = -(\partial_\lambda F^{\lambda\mu})A^\nu - F^{\lambda\mu}\partial_\lambda A^\nu. \tag{4.55}$$

By virtue of the free-field Maxwell equations (3.79), the first term on the r.h.s. vanishes (in the absence of sources), and therefore

$$\partial_\lambda \Omega^{\lambda\mu\nu} = -F^{\lambda\mu}\partial_\lambda A^\nu = F^{\mu\lambda}\partial_\lambda A^\nu = F^\mu{}_\lambda\, \partial^\lambda A^\nu. \tag{4.56}$$

Our redefined stress-energy tensor is thus

$$T^{\mu\nu} = -F^\mu{}_\lambda F^{\nu\lambda} + \frac{1}{4} g^{\mu\nu} F_{\alpha\beta}F^{\alpha\beta}, \tag{4.57}$$

which is now expressed in terms of the field strengths and therefore gauge invariant. Notice that it is also now a symmetric tensor, which means for example that $T^{01} = T^{10}$, i.e. the density of the x-component of the field momentum is equal to the flow of energy in the x-direction. In terms of the field strengths, we have explicitly

$$\mathcal{L} = \frac{1}{2}\Big(\mathbf{E}^2 - \mathbf{B}^2\Big), \quad T^{00} = \mathcal{H} = \frac{1}{2}\Big(\mathbf{E}^2 + \mathbf{B}^2\Big), \quad T^{0j} = (\mathbf{E}\times\mathbf{B})_j. \tag{4.58}$$

You may recognise that the term on the right is none other than the j-component of the Poynting vector, which is indeed the directional energy flux density of an electromagnetic field.

4.5.4 General Relativity

In general relativity, the element of space-time that is invariant under general coordinate transformations is $\mathrm{d}^4x\,\sqrt{-g}$ where g is the determinant of the metric tensor, $g = \det(g_{\mu\nu})$, which is of course -1 for the Minkowski metric (2.62). The invariant action integral therefore becomes

$$S = \int \mathrm{d}^4x\,\sqrt{-g}\mathcal{L}. \tag{4.59}$$

There is then a very general and powerful way of defining the stress-energy tensor, which is to say that it measures the response of the action to small changes in the metric, according to the equation

$$\delta S = \frac{1}{2}\int \mathrm{d}^4x\,\sqrt{-g}T_{\mu\nu}\delta g^{\mu\nu}, \tag{4.60}$$

that is,

$$T_{\mu\nu} = \frac{2}{\sqrt{-g}}\frac{\partial(\sqrt{-g}\mathcal{L})}{\partial g^{\mu\nu}} = 2\frac{\partial\mathcal{L}}{\partial g^{\mu\nu}} + \frac{1}{g}\frac{\partial g}{\partial g^{\mu\nu}}\mathcal{L}. \tag{4.61}$$

Now from the properties of determinants and the inverse matrix,

$$\frac{\partial(\det M)}{\partial M_{jk}} = \det M (M^{-1})_{kj}, \tag{4.62}$$

and (even in general relativity)

$$g_{\mu\lambda} g^{\nu\lambda} = \delta^{\nu}_{\mu}, \tag{4.63}$$

so $(g^{-1})^{\nu\mu} = g^{\mu\nu}$ and

$$\frac{\partial g}{\partial g_{\mu\nu}} = g g^{\mu\nu}. \tag{4.64}$$

This is almost what we need: differentiating (4.63) gives $g_{\mu\lambda}\, \mathrm{d}g^{\nu\lambda} = -\,\mathrm{d}g_{\mu\lambda}\, g^{\nu\lambda}$, so $\mathrm{d}g_{\mu\nu} = -g_{\mu\lambda} g_{\sigma\nu}\, \mathrm{d}g^{\sigma\lambda}$ and

$$\frac{\partial g}{\partial g^{\mu\nu}} = -g g_{\mu\nu}. \tag{4.65}$$

Thus we finally obtain for the stress-energy tensor the general expression

$$T_{\mu\nu} = 2\frac{\partial \mathcal{L}}{\partial g^{\mu\nu}} - g_{\mu\nu}\mathcal{L}. \tag{4.66}$$

You can check that (4.66) agrees with the results we obtained earlier for the Klein-Gordon and electromagnetic fields. Notice that this tensor is manifestly symmetric (since $g_{\mu\nu}$ is) and we automatically obtain the symmetric, gauge-invariant form (4.57) in the electromagnetic case.

4.6 Angular Momentum and Spin

Having found that the stress-energy tensor $T^{\mu\nu}$ is conserved and (properly defined) symmetric, we can construct a conserved tensor of higher rank as follows:

$$M^{\lambda\mu\nu} = x^{\mu} T^{\lambda\nu} - x^{\nu} T^{\lambda\mu}, \tag{4.67}$$

for then,

$$\partial_{\lambda} M^{\lambda\mu\nu} = T^{\mu\nu} - T^{\nu\mu} = 0, \tag{4.68}$$

since $T^{\mu\nu}$ is symmetric. Recalling that T^{0j} is the density of the j^{th} component of momentum, we see that for example $M^{012} = xT^{02} - yT^{01}$ is the density of the z-component of angular momentum, and so we define the *total angular momentum tensor* of the field as

$$J^{\mu\nu} = \int \mathrm{d}^3\mathbf{r}\, M^{0\mu\nu}. \tag{4.69}$$

Let us open a parenthesis here to remark how we have arrived at a conserved angular momentum. At the beginning of Section 4.5 we introduced the conserved stress-energy tensor as a consequence of invariance to small space-time displacements. In general, this symmetry alone does not warrant conservation of *angular* momentum. However, we considered systems that have a larger symmetry: invariance with respect to general coordinate transformations. These include rotations and thus lead to conservation of the total angular momentum. In this section we see that this additional conservation is

directly related to whether the stress-energy tensor is symmetric or not in (some of) its space-time indices.

The components of the more familiar total angular momentum vector $\mathbf{J}$ are then given by

$$J_i = \frac{1}{2}\varepsilon_{ijk}J^{jk}. \tag{4.70}$$

The other non-zero components of $J^{\mu\nu}$ are of the form

$$J^{0j} = -J^{j0} = \int \mathrm{d}^3\mathbf{r}\, M^{00j} = t\int \mathrm{d}^3\mathbf{r}\, T^{0j} - \int \mathrm{d}^3\mathbf{r}\, x^j T^{00} = tP_j - R_j E, \tag{4.71}$$

where $\mathbf{P}$ is the total momentum of the field, E the total energy, and $\mathbf{R}$ the position of the centre-of-energy. Thus the conservation of this quantity implies that

$$\mathbf{R} = \mathbf{V}t + \text{const.}, \tag{4.72}$$

where $\mathbf{V} = \mathbf{P}/E$ is the velocity of the centre-of-energy, i.e. the velocity of the zero-momentum frame (in units where $c = 1$). This is just Newton's first law: the centre-of-energy of an isolated system moves with constant velocity.

Since in general the centre-of-energy is moving, the total angular momentum includes both "orbital" and "intrinsic" parts, where the intrinsic or *spin* angular momentum is defined in the zero-momentum frame. We can make a covariant definition of the spin by using the 4-velocity $U^\mu P^\mu/\sqrt{P_\nu P^\nu} = \gamma(1, \mathbf{V})$, where $\gamma = 1/\sqrt{1-V^2}$. Then we define the *spin 4-vector* as

$$S^\mu = -\frac{1}{2}\varepsilon^{\mu\nu\alpha\beta}U_\nu J_{\alpha\beta}, \tag{4.73}$$

where $\varepsilon^{\mu\nu\alpha\beta}$ is the totally antisymmetric Levi-Civita tensor ($\varepsilon = +1$ for any even permutation of indices 0123, -1 for any odd permutation, 0 otherwise). Then in the zero-momentum frame $U_\nu = (1, \mathbf{0})$, $S^0 = 0$ and

$$S^i = -\frac{1}{2}\varepsilon^{i0jk}J_{jk} = J^i. \tag{4.74}$$

Notice that S^μ is always orthogonal to the 4-velocity, $S^\mu U_\mu = 0$, so in fact the spin only has 3 independent components in any frame, whereas $J^{\mu\nu}$ has the extra components (4.71), which are mixed with those of $\mathbf{J}$ in different frames. There is no covariant way of separating the total angular momentum into orbital and spin contributions. J is a 4-tensor and S is a 4-vector: to take their difference "$L = J - S$" does not make sense, any more than subtracting apples from oranges.

4.7 Quantum Fields

Although this is a course on classical field theory, it is worth making a short excursion into the quantum mechanics of fields, which in fact clarifies many of their properties, especially in the relativistic case. The transition from classical to quantum fields is through a procedure called 'second quantisation'.

First quantisation was the procedure of replacing the classical dynamical variables q and p by quantum operators $\hat{q}$ and $\hat{p}$ such that

$$[\hat{q}, \hat{p}] = i \qquad (\hbar = 1). \tag{4.75}$$

Second quantisation means replacing the field variable $\varphi(x,t)$ and its conjugate momentum density $\pi(x,t)$ by operators such that

$$[\hat{\varphi}(x,t), \hat{\pi}(x',t)] = i\delta(x - x'). \tag{4.76}$$

N.B.: x and x' are not dynamical variables but labels for the field values at different points. Compare (and contrast)

$$[\hat{q}_j, \hat{p}_k] = i\delta_{jk} \qquad (j,k = x,y,z). \tag{4.77}$$

The field (wavefunction) φ satisfying the Klein-Gordon equation is replaced by a *field operator* $\hat{\varphi}$, satisfying the same equation. The Fourier representation of a real field becomes

$$\hat{\varphi}(x,t) = \int \mathrm{d}k\, N(k)\Big[\hat{a}(k)e^{i(kx-\omega t)} + \hat{a}^\dagger(k)e^{-i(kx-\omega t)}\Big]. \tag{4.78}$$

Note that $\hat{\varphi}$ is hermitian but the Fourier conjugate operator $\hat{a}$ is not. Keeping track of the order of operators, the Hamiltonian operator is

$$\hat{H} = \int \mathrm{d}k\, N(k)\frac{1}{2}\omega(k)\Big[\hat{a}(k)\hat{a}^\dagger(k) + \hat{a}^\dagger(k)\hat{a}(k)\Big]. \tag{4.79}$$

Comparing this with the simple harmonic oscillator,

$$\hat{H}_{\mathrm{SHO}} = \frac{1}{2}\omega\Big(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\Big), \tag{4.80}$$

we see that $\hat{a}^\dagger(k)$ and $\hat{a}(k)$ must be the ladder operators for the mode of wave number k. They add/remove one quantum of excitation of the mode. These quanta are the particles corresponding to that field

$$\Longrightarrow \quad \hat{a}^\dagger(k) = \text{the } \textit{creation} \text{ operator}, \quad \hat{a}(k) = \text{the } \textit{annihilation} \text{ operator} \tag{4.81}$$

for Klein-Gordon particles.

The ladder operators of the simple harmonic oscillator satisfy

$$\Big[\hat{a}, \hat{a}^\dagger\Big] = 1. \tag{4.82}$$

The analogous commutation relation for the creation and annihilation operators is

$$\begin{aligned} &N(k)\Big[\hat{a}(k), \hat{a}^\dagger(k')\Big] = \delta(k - k') \\ &\Longrightarrow \ \Big[\hat{a}(k), \hat{a}^\dagger(k')\Big] = 2\pi \cdot 2\omega(k)\delta(k - k'), \end{aligned} \tag{4.83}$$

or in 3 spatial dimensions

$$\Big[\hat{a}(\mathbf{k}), \hat{a}^\dagger(\mathbf{k}')\Big] = (2\pi)^3 \cdot 2\omega(\mathbf{k})\delta^3(\mathbf{k} - \mathbf{k}'). \tag{4.84}$$

On the other hand

$$[\hat{a}(\mathbf{k}), \hat{a}(\mathbf{k}')] = \left[\hat{a}^\dagger(\mathbf{k}), \hat{a}^\dagger(\mathbf{k}')\right] = 0. \tag{4.85}$$

The commutators of the creation and annihilation operators correspond to the field commutation relation

$$\begin{aligned}[\hat{\varphi}(\mathbf{r}, t), \hat{\pi}(\mathbf{r}', t)] &= \int \mathrm{d}^3\mathbf{k}\, \mathrm{d}^3\mathbf{k}'\, N(\mathbf{k}) N(\mathbf{k}') [-i\omega(\mathbf{k}')] \\ &\quad \times \left[\hat{a}(\mathbf{k}) e^{-ik\cdot x} + \hat{a}^\dagger(\mathbf{k}) e^{ik\cdot x}, \hat{a}(\mathbf{k}') e^{-ik'\cdot x'} - \hat{a}^\dagger(\mathbf{k}') e^{ik'\cdot x'}\right] \\ &= i \int \mathrm{d}^3\mathbf{k}\, N(\mathbf{k}) \omega(\mathbf{k}) \left[e^{-ik\cdot(x-x')} + e^{ik\cdot(x-x')}\right], \end{aligned} \tag{4.86}$$

where $x^\mu = (ct, \mathbf{r})$ and $x'^\mu = (ct, \mathbf{r}')$. Hence

$$[\hat{\varphi}(\mathbf{r}, t), \hat{\pi}(\mathbf{r}', t)] = i\delta^3(\mathbf{r} - \mathbf{r}'), \tag{4.87}$$

as expected. On the other hand

$$[\hat{\varphi}(\mathbf{r}, t), \hat{\varphi}(\mathbf{r}', t)] = [\hat{\pi}(\mathbf{r}, t), \hat{\pi}(\mathbf{r}', t)] = 0. \tag{4.88}$$

The fact that the field operator has positive- and negative-frequency parts now appears quite natural:

- The *positive* frequency part $\hat{a}(\mathbf{k}) e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}$ *annihilates* particles;
- The *negative* frequency part $\hat{a}^\dagger(\mathbf{k}) e^{-i(\mathbf{k}\cdot\mathbf{r} - \omega t)}$ *creates* particles;

and $\pm\hbar\omega$ is the energy released/absorbed in the annihilation/creation process.

The hermitian field describes particles that are identical to their antiparticles, e.g. π^0 mesons. For a complex (non-hermitian) field, the negative-frequency part of $\hat{\varphi}$ *creates antiparticles*. To see this, consider the second-quantised version of Eq. (4.22) for a complex field $\hat{\varphi}$:

$$\hat{Q} = \int \mathrm{d}^3\mathbf{k}\, N(\mathbf{k}) \left[\hat{a}^\dagger(\mathbf{k}) \hat{a}(\mathbf{k}) - \hat{b}^\dagger(\mathbf{k}) \hat{b}(\mathbf{k})\right]. \tag{4.89}$$

Comparing with the energy

$$\hat{H} = \int \mathrm{d}^3\mathbf{k}\, N(\mathbf{k}) \omega(\mathbf{k}) \left[\hat{a}^\dagger(\mathbf{k}) \hat{a}(\mathbf{k}) + \hat{b}^\dagger(\mathbf{k}) \hat{b}(\mathbf{k})\right], \tag{4.90}$$

we see that the particles created by $\hat{a}^\dagger$ and $\hat{b}^\dagger$ have the same energy $\hbar\omega$ but *opposite charge*: for example, they could be π^+ and π^- mesons.

To summarise, in quantum field theory:

- The object that satisfies the Klein-Gordon equation is the field operator $\hat{\varphi}$;
- The Fourier decomposition of $\hat{\varphi}$ has a *positive frequency* part that *annihilates a particle* (with energy $\hbar\omega$ and charge $+1$) AND a *negative frequency* part that *creates an antiparticle* (with energy $\hbar\omega$ and charge -1);
- Similarly, $\hat{\varphi}^\dagger$ creates a particle or annihilates an antiparticle.

Chapter 5

Broken Symmetry

We have seen that symmetries can have profound implications (e.g., conserved currents), but sometime the breaking of symmetry can be even more interesting. We introduce here the important phenomenon of *spontaneous* symmetry breaking, in the simple context of the scalar Klein-Gordon field. We consider only the classical case, but all the features of interest remain valid after second quantisation.

5.1 Self-Interacting Scalar Field

Let us consider the effect of adding terms of higher order in the field φ to the Klein-Gordon Lagrangian density (3.69). The simplest addition that preserves the global phase symmetry is a quartic term $-\frac{1}{2}\lambda(\varphi^*\varphi)^2$:

$$\mathcal{L} = (\partial^\mu\varphi^*)(\partial_\mu\varphi) - m^2\varphi^*\varphi - \frac{1}{2}\lambda(\varphi^*\varphi)^2. \tag{5.1}$$

As was argued earlier for the quadratic term $-m^2\varphi^*\varphi$, the coefficient λ must be positive in order for the Hamiltonian to be positive-definite, otherwise no state of lowest energy will exist.

Notice that λ is a dimensionless quantity: this is easy to see in natural units, since then $\mathcal{L}$ and $(\varphi^*\varphi)^2$ both have dimensions $[M]^4$. Any higher powers, $(\varphi^*\varphi)^p$ with $p > 2$, would have coefficients with dimensions of inverse powers of mass, $[M]^{4-2p}$. One can argue on rather general grounds that such terms should be negligible since the relevant mass scale should be large.

The equation for the conjugate momentum density, $\pi = \partial\varphi^*/\partial t$, remains unchanged and the Hamiltonian density is

$$\mathcal{H} = \pi^*\pi + \boldsymbol{\nabla}\varphi^* \cdot \boldsymbol{\nabla}\varphi + V(\varphi), \tag{5.2}$$

where the 'potential' is

$$V(\varphi) = m^2\varphi^*\varphi + \frac{1}{2}\lambda(\varphi^*\varphi)^2. \tag{5.3}$$

The equation of motion becomes

$$\partial^\mu\partial_\mu\varphi + m^2\varphi + \lambda(\varphi^*\varphi)\varphi = 0. \tag{5.4}$$

We can think of the extra term as a self-interaction of the field, with strength λ.

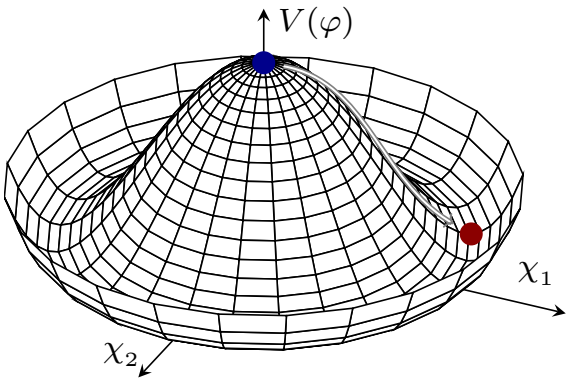

Fig. 5.1: The form of the 'potential' $V(\varphi)$ that leads to spontaneous symmetry breaking.

5.2 Spontaneously Broken Global Symmetry

The Hamiltonian density (5.2) with the potential (5.3) implies that the state of the field with minimum energy is that in which $\varphi = 0$ everywhere (assuming that the coefficient m^2 of the quadratic term is non-negative). Note, however, that in the presence of the quartic term in the potential our previous argument for the positive sign of the quadratic term (see Section 3.3) is no longer valid: the Hamiltonian is always bounded from below as long as $\lambda > 0$. Let us therefore study the effect of a negative quadratic term in the potential, i.e.

$$V(\varphi) = -m^2\varphi^*\varphi + \frac{1}{2}\lambda(\varphi^*\varphi)^2. \tag{5.5}$$

The state of minimum energy will now be one in which φ has a constant value φ_0 such that $V(\varphi_0)$ is a minimum, i.e.

$$\varphi_0^*\varphi_0 = \frac{m^2}{\lambda}, \tag{5.6}$$

which describes a circle in the complex φ plane,

$$\varphi_0 = \frac{m}{\sqrt{\lambda}}e^{i\theta}, \tag{5.7}$$

where θ is arbitrary.

Thus the system does not have a unique state of minimum energy but an infinite number of equivalent ones corresponding to different values of θ: it is said to be *degenerate*. However, if we take any particular example of the system and reduce its energy somehow to the minimum value, it will be in a state with a particular value of θ. The situation is like that of a thin rod initially balanced vertically on its tip on a horizontal plane: when it falls under gravity, it will lie at a particular angle on the plane, although all angles have equal energy. The dynamics and the initial state are symmetrical with respect to rotations about the vertical axis, but the final minimum-energy state is not: the rotational symmetry has been *spontaneously broken*.

Similarly, in the presence of the quartic interaction, the Klein-Gordon field will undergo spontaneous symmetry breaking by choosing some particular minimum-energy state, with a particular global value of θ. And since the dynamics has phase symmetry we may as well choose to label that state as $\theta = 0$ (like measuring angles with respect to the fallen

rod). Thus we define $\varphi_0 = m/\sqrt{\lambda}$ and measure variations of the field with respect to this 'ground-state' value: $\varphi = \varphi_0 + \chi$ (for instance, if we are interested in fluctuations of the system about the symmetry broken ground state φ_0). In terms of the 'dynamical' field χ, the potential is

$$\begin{aligned} V &= V(\varphi_0) + \frac{1}{2}\lambda[\varphi_0(\chi^* + \chi) + \chi^*\chi]^2 \\ &= V(\varphi_0) + \frac{1}{2}m^2(\chi^* + \chi)^2 + \mathcal{O}(\chi^3) \end{aligned} \tag{5.8}$$

(where we have used the fact that $\lambda\varphi_0^2 = m^2$ to cancel a lot of terms). Resolving χ into its real and imaginary parts as in Eq. (3.61):

$$\chi = \frac{1}{\sqrt{2}}(\chi_1 + i\chi_2), \tag{5.9}$$

where χ_1 and χ_2 are real, we have

$$V = V(\varphi_0) + m^2\chi_1^2 + \mathcal{O}(\chi^3), \tag{5.10}$$

and the Lagrangian and Hamiltonian reduce to, respectively:

$$\mathcal{L} = \frac{1}{2}(\partial^\mu\chi_1)(\partial_\mu\chi_1) + \frac{1}{2}(\partial^\mu\chi_2)(\partial_\mu\chi_2) - V(\varphi_0) - m^2\chi_1^2 + \mathcal{O}(\chi^3) \tag{5.11}$$

$$\begin{aligned} \mathcal{H} = &\frac{1}{2}\left(\frac{\partial\chi_1}{\partial t}\right)^2 + \frac{1}{2}\boldsymbol{\nabla}\chi_1 \cdot \boldsymbol{\nabla}\chi_1 + m^2\chi_1^2 \\ &+ \frac{1}{2}\left(\frac{\partial\chi_2}{\partial t}\right)^2 + \frac{1}{2}\boldsymbol{\nabla}\chi_2 \cdot \boldsymbol{\nabla}\chi_2 + V(\varphi_0) + \mathcal{O}(\chi^3), \end{aligned} \tag{5.12}$$

where we immediately recognise the separate contributions pertaining to the fields χ_1 and χ_2.

Equation (5.10) displays many interesting features:

1. The phase symmetry breaking is now explicit: there is a quadratic term for the real part of χ but not for the imaginary part;
2. The quadratic term for the real part φ_1 is positive even though we started off with a negative quadratic term for φ.
3. Comparing with the original Klein-Gordon equation (3.30), we see that the quadratic term corresponds to the dispersion relation

   $$\omega = \sqrt{k^2 + 2m^2} \tag{5.13}$$

   for the field χ_1. This means that, after second-quantisation, the quanta of the field will be particles of mass (in natural units) $\sqrt{2}m$;
4. On the other hand the field χ_2 has no quadratic term and therefore its dispersion relation is $\omega = k$, i.e. its quanta have zero mass, like those of the electromagnetic field (photons), except that their spin is zero instead of one.

This last point is an example of *Goldstone's theorem*: for every spontaneously broken (global) continuous symmetry there is a field with massless quanta, which is therefore called a Goldstone field.

5.3 Spontaneously Broken Local Symmetry

Recall that the symmetry of the Klein-Gordon Lagrangian with respect to global (space-time independent) phase changes can be promoted to local (space-time dependent) phase symmetry by introducing a vector field A^μ that undergoes a compensating gauge transformation. The full Lagrangian density then takes the form

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\varphi)^*(D^\mu\varphi) - V(\varphi), \tag{5.14}$$

where $D_\mu = \partial_\mu + ieA_\mu$. As in the previous subsection, the presence of negative quadratic and positive quartic terms in V will trigger spontaneous symmetry breaking and a non-zero value of φ satisfying Eq. (5.6) in the state of minimum energy. However, in such a state (i.e., $\varphi = \varphi_0$) we now obtain a non-zero contribution to $\mathcal{L}$ from the second term on the r.h.s., namely

$$(ie\varphi_0 A_\mu)^*(ie\varphi_0 A^\mu) = \frac{e^2m^2}{\lambda}A_\mu A^\mu, \tag{5.15}$$

and the corresponding equation of motion for the field A_μ (in Lorenz gauge, $\partial_\mu A^\mu = 0$) is

$$\partial_\nu\partial^\nu A_\mu + 2\frac{e^2m^2}{\lambda}A_\mu = 0. \tag{5.16}$$

This means that the dispersion relation for the vector field has become

$$\omega = \sqrt{k^2 + 2e^2m^2/\lambda}, \tag{5.17}$$

and correspondingly the quanta of the vector field have acquired a non-zero mass equal to $em\sqrt{2/\lambda}$. In the following section we shall see in detail how this generation of mass has come to pass, through a phenomenon known as the *Higgs mechanism*.

5.4 Higgs Mechanism

Let us examine what happens to the dynamical part of the scalar field when the local phase symmetry is spontaneously broken. As in the case of global symmetry breaking, we write

$$\varphi = \varphi_0 + \chi = \varphi_0 + \frac{1}{\sqrt{2}}(\chi_1 + i\chi_2). \tag{5.18}$$

When χ is non-zero, the covariant derivative of φ becomes

$$D_\mu\varphi = \frac{1}{\sqrt{2}}(\partial_\mu\chi_1 + i\partial_\mu\chi_2) + ie\varphi_0 A_\mu + \cdots, \tag{5.19}$$

where the dots represent terms of higher order in the fields. We see that the term involving χ_2 can be removed by a redefinition of the vector field

$$A_\mu \to A_\mu - \frac{1}{\sqrt{2}e\varphi_0}\partial_\mu\chi_2, \tag{5.20}$$

Fig. 5.2: Summary of interactions between particles described by the Standard Model.

which is just a gauge transformation, so it leaves the field strength tensor $F_{\mu\nu}$ unchanged. The Lagrangian density then takes the form

$$\begin{aligned}\mathcal{L} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + e^2\varphi_0^2 A_\mu A^\mu + \frac{1}{2}(\partial_\mu\chi_1)(\partial^\mu\chi_1) - V(\varphi) + \cdots \\ &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{e^2m^2}{\lambda}A_\mu A^\mu + \frac{1}{2}(\partial_\mu\chi_1)(\partial^\mu\chi_1) - m^2\chi_1^2 - V(\varphi_0) + \cdots . \end{aligned} \tag{5.21}$$

As we saw before, there is now a term quadratic in the vector field, which corresponds to its quanta having a mass, and a similar term for the dynamical scalar field χ_1. However, unlike the case of the broken *global* phase symmetry, there are no terms involving the field χ_2: it has been absorbed in a redefinition of the vector field. In more colourful terms, the vector field has *eaten* the Goldstone field and thereby acquired a mass. This is called the *Higgs mechanism*.

Another way to see that the field χ_2 can be removed by a gauge transformation is to recall that such a transformation is equivalent to a change of phase of the scalar field. Since χ_2 is the imaginary part of the field φ, it can be removed by a phase change

$$\varphi \to e^{-i\theta} \quad \text{where} \quad \theta = \sin^{-1}(\chi_2/|\varphi|), \tag{5.22}$$

which makes φ real everywhere. Local phase/gauge symmetry means we are allowed to make a different phase change at every space-time point, so χ_2 can always be removed in this way. In the absence of a vector field, on the other hand, we can only make a constant global phase change, so χ_2 cannot be removed and remains a physical field.

The Higgs mechanism can also be applied more generally when the Lagrangian density of a system has several local gauge symmetries, some of which are broken spontaneously by nonzero values of scalar fields in the state of minimum energy. For each gauge symmetry there is a vector field whose quanta are massless if the symmetry is unbroken but massive if it is broken. The masses are acquired by 'eating' the Goldstone components of the scalar fields, which would otherwise have massless quanta themselves. The remaining, uneaten components of the scalar fields have quanta that should be observable as massive *Higgs bosons*.

In the Standard Model of particle physics, the electromagnetic and weak nuclear interactions are described by a unified theory with four distinct gauge symmetries, three of which are broken spontaneously by components of a pair of complex scalar fields. Thus three of the four associated vector fields have massive quanta (the W^+, W^- and Z^0 bosons) while one remains massless (the photon). Correspondingly, three of the four (real and imaginary) components of the scalar fields are eaten and one physical scalar Higgs boson remains. In the Standard Model, the Higgs field also couples to the fundamental fermions (i.e. the quarks and leptons), and through these interactions the fermions also acquire mass (see Figure 5.2).

Chapter 6

Dirac Field [*not examinable*]

You may have already encountered the relativistic wave equation for the electron, namely the so-called Dirac equation

$$i\hbar\frac{\partial\psi}{\partial t} = -i\hbar c\,\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\psi + \beta mc^2\psi, \tag{6.1}$$

where ψ is a 4-component wave function called a *Dirac spinor* and α_j, β are 4×4 hermitian *Dirac matrices* with the properties

$$\alpha_j^2 + \beta^2 = 1, \quad \alpha_j\beta + \beta\alpha_j = \alpha_j\alpha_k + \alpha_k\alpha_j = 0 \quad (i,j=1,2,3; j\neq k). \tag{6.2}$$

To make the covariant form of the Dirac equation more apparent, we multiply through by β and rearrange terms to give

$$i\hbar(\beta\partial_0\psi + c\beta\alpha_j\partial_j\psi) = mc^2\psi. \tag{6.3}$$

Defining

$$\gamma^0 = \beta, \quad \gamma^j = \beta\alpha_j, \tag{6.4}$$

and dividing by the speed of light c, we can write this as

$$i\hbar\gamma^\mu\partial_\mu\psi = mc\psi, \tag{6.5}$$

where one needs to recall that $\partial_0\ \partial/\partial(ct)$. In natural units:

$$(i\gamma^\mu\partial_\mu - m)\psi = 0. \tag{6.6}$$

From the properties of the α_j and β matrices we find (check it!) that

$$\{\gamma^\mu,\gamma^\nu\} = \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}I_4, \tag{6.7}$$

where the curly brackets $\{\,,\,\}$ denote the so called anti-commutator, and ${}_I4$ is the 4×4 identity matrix. Whereas the notation may already look covariant, it is important to appreciate that γ^μ is NOT a 4-vector like x^μ or ∂_μ: it is a set of 4 4×4 matrices, which do not transform under Lorentz transformations.

In the so-called standard representation, the γ matrices take the form

$$\gamma^0 = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \tag{6.8}$$

where I_2 is the 2×2 identity matrix and σ_i are Pauli matrices.

To make a 4-vector, we have to sandwich γ^μ between some spinors. Recall that the probability density associated with the Dirac spinor ψ is

$$\rho = \psi^\dagger\psi, \tag{6.9}$$

and the probability current is

$$\mathbf{J} = c\psi^\dagger \boldsymbol{\alpha}\psi. \tag{6.10}$$

Introducing the notation

$$\overline{\psi} = \psi^\dagger\beta = \psi^\dagger\gamma^0, \tag{6.11}$$

we have

$$\rho = \overline{\psi}\gamma^0\psi, \quad J_j = c\overline{\psi}\gamma^j\psi. \tag{6.12}$$

However, we know that the density and current together form a 4-vector $J^\mu = (c\rho, \mathbf{J})$ and so we may write this as

$$J^\mu = c\overline{\psi}\gamma^\mu\psi, \tag{6.13}$$

which shows that $\overline{\psi}\gamma^\mu\psi$ is a 4-vector.

The density and current satisfy the conservation (continuity) equation

$$\frac{\partial\rho}{\partial t} = -\boldsymbol{\nabla}\cdot\mathbf{J}, \tag{6.14}$$

which becomes simply

$$\partial_\mu J^\mu = c\partial_\mu\left(\overline{\psi}\gamma^\mu\psi\right) = 0. \tag{6.15}$$

Let us verify this directly from the Dirac equation. (For simplicity, from now on we use natural units.) We have

$$\partial_\mu\left(\overline{\psi}\gamma^\mu\psi\right) = \left(\partial_\mu\overline{\psi}\right)\gamma^\mu\psi + \overline{\psi}\gamma^\mu\partial_\mu\psi. \tag{6.16}$$

Now the Dirac equation gives $\gamma^\mu\partial_\mu\psi = -im\psi$, so

$$\overline{\psi}\gamma^\mu\partial_\mu\psi = -im\overline{\psi}\psi. \tag{6.17}$$

Taking the adjoint of the Dirac equation gives

$$(\partial_\mu\psi^\dagger)(\gamma^\mu)^\dagger = +im\psi^\dagger. \tag{6.18}$$

From the definition of the γ^μ matrices in terms of the hermitian α_j and β matrices, one may readily verify that

$$(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0, \tag{6.19}$$

and hence

$$\left(\partial_\mu\psi^\dagger\right)(\gamma^\mu)^\dagger = \left(\partial_\mu\overline{\psi}\right)\gamma^\mu\gamma^0 = +im\psi^\dagger. \tag{6.20}$$

Thus, multiplying both sides on the right by γ^0 and using $(\gamma^0)^2 = 1$,

$$\left(\partial_\mu\overline{\psi}\right)\gamma^\mu = +im\overline{\psi}, \tag{6.21}$$

which, together with Eq. (6.17), shows that

$$\partial_\mu J^\mu = \partial_\mu\left(\overline{\psi}\gamma^\mu\psi\right) = +im\overline{\psi}\psi - im\overline{\psi}\psi = 0. \tag{6.22}$$

6.1 Dirac Lagrangian and Hamiltonian

Now we want to formulate the Lagrangian density $\mathcal{L}$ which has the Dirac equation as its equation of motion. A suitable choice is

$$\mathcal{L} = i\overline{\psi}\gamma^\mu\partial_\mu\psi - m\overline{\psi}\psi. \tag{6.23}$$

This is a choice in the sense that, as usual, we are free to add a total derivative. For example we could add $-\frac{i}{2}\partial_\mu\left(\overline{\psi}\gamma^\mu\psi\right)$ and write $\mathcal{L}$ in a more symmetrically pleasing (yet equivalent) form:

$$\mathcal{L} = \frac{i}{2}\left[\overline{\psi}\gamma^\mu\partial_\mu\psi - \left(\partial_\mu\overline{\psi}\right)\gamma^\mu\psi\right] = m\overline{\psi}\psi. \tag{6.24}$$

Just as we must treat φ and φ^* as independent fields in the Klein-Gordon case, we have to treat ψ and $\overline{\psi}$ as independent here. Using the Lagrangian density (6.23), the equation of motion for $\overline{\psi}$ gives the Dirac equation for ψ directly:

$$\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\overline{\psi})} = 0 = \frac{\partial\mathcal{L}}{\partial\overline{\psi}} = i\gamma^\mu\partial_\mu\psi - m\psi, \tag{6.25}$$

while that for ψ gives the corresponding equation (6.21) for $\overline{\psi}$:

$$\partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\psi} = i\left(\partial_\mu\overline{\psi}\gamma^\mu\right) = \frac{\partial\mathcal{L}}{\partial\psi} = -m\overline{\psi}. \tag{6.26}$$

You should check that the Lagrangian density (6.24) gives the same results. Note that when ψ and $\overline{\psi}$ satisfy their equations of motion, $\mathcal{L} = 0$. In other words, the minimum value of the action is zero.

The Hamiltonian density is

$$\mathcal{H} = \pi\frac{\partial\psi}{\partial t} + \frac{\partial\overline{\psi}}{\partial t}\overline{\pi} - \mathcal{L}, \tag{6.27}$$

where π and $\overline{\pi}$ are the canonical momentum densities conjugate to ψ and $\overline{\psi}$, respectively. Note that we have been careful with the order of factors: if ψ is represented by a column vector then pi is a row vector, and conversely for $\overline{\psi}$ and $\overline{\pi}$. Sticking with the simpler form (6.23) of the Dirac Lagrangian density, we have

$$\pi = \frac{\partial\mathcal{L}}{\partial(\partial\psi/\partial t)} = -\overline{\psi}\gamma^0 = i\psi^\dagger, \quad \overline{\pi} = \frac{\partial\mathcal{L}}{\partial(\partial\overline{\psi}/\partial t)} = 0, \tag{6.28}$$

and so (in natural units)

$$\begin{aligned}\mathcal{H} &= i\psi^\dagger\frac{\partial\psi}{\partial t} - \mathcal{L} \\ &= \psi^\dagger(-i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta m)\psi,\end{aligned} \tag{6.29}$$

which is just what we would expect for the energy density, given the Dirac equation (6.1).

6.2 Global and Local Phase Symmetry

Like the Klein-Gordon Lagrangian, The Dirac Lagrangian is invariant with respect to a global (space-time independent) phase change in ψ:

$$\psi \to e^{-i\varepsilon}\psi, \quad \overline{\psi} \to e^{+i\varepsilon}\overline{\psi}, \tag{6.30}$$

or, for an infinitesimal change,

$$\psi \to \psi - i\varepsilon\psi, \quad \overline{\psi} \to \overline{\psi} + i\varepsilon\overline{\psi}. \tag{6.31}$$

The associated Noether current is the one we found already:

$$J^\mu = -i\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\psi + i\overline{\psi}\frac{\partial\mathcal{L}}{\partial(\partial_\mu\overline{\psi})} = \overline{\psi}\gamma^\mu\psi + 0. \tag{6.32}$$

In the presence of an electromagnetic 4-potential A_μ we expect an interaction term $-eJ^\mu A_\mu$, so that the Lagrangian density becomes

$$\mathcal{L} = i\overline{\psi}\gamma^\mu(\partial_\mu + ieA_\mu)\psi \equiv i\overline{\psi}\gamma^\mu D_\mu\psi, \tag{6.33}$$

where $D_\mu = \partial_\mu + ieA_\mu$ is the covariant derivative introduced in Section 4.3. Then, just as in the case of the Klein-Gordon field, we see that $\mathcal{L}$ is invariant under a *local* (space-time dependent) phase change $\varepsilon(x)$ combined with a gauge change $A_\mu \to A_\mu + \partial_\mu\varepsilon/e$.

6.3 Stress-Energy Tensor, Angular Momentum and Spin

The stress-energy tensor for the Dirac field is obtained by following the steps in Section 4.5. Using the Lagrangian density (6.23),

Chapter 7

Phase Transitions and Critical Phenomena

Many of the ideas developed so far in this course (Hamiltonian formalism, classical field theory, spontaneous symmetry breaking, etc) have application in another very important area of physics: critical phenomena and phase transitions. In this section we introduce some of the basic ideas, then discuss the simple Ising model for ferromagnetism and finally describe the Ginzburg-Landau Theory of second order phase transitions. This has some similarities with the Higgs mechanism discussed in the previous section.

Appendix A

Appendix