 Michaelmas Term, 2006 and 2007

Quantum Field Theory

University of Cambridge Part III Mathematical Tripos

Dr David Tong

Department of Applied Mathematics and Theoretical Physics,
Centre for Mathematical Sciences,
Wilberforce Road,
Cambridge, CB3 OWA, UK

http://www.damtp.cam.ac.uk/user/tong/qft.html
d.tong@damtp.cam.ac.uk

Recommended Books and Resources

- M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*

This is a very clear and comprehensive book, covering everything in this course at the right level. It will also cover everything in the "Advanced Quantum Field Theory" course, much of the "Standard Model" course, and will serve you well if you go on to do research. To a large extent, our course will follow the first section of this book.

There is a vast array of further Quantum Field Theory texts, many of them with redeeming features. Here I mention a few very different ones.

- S. Weinberg, *The Quantum Theory of Fields, Vol 1*

This is the first in a three volume series by one of the masters of quantum field theory. It takes a unique route to through the subject, focussing initially on particles rather than fields. The second volume covers material lectured in "AQFT".

- L. Ryder, *Quantum Field Theory*

This elementary text has a nice discussion of much of the material in this course.

- A. Zee, *Quantum Field Theory in a Nutshell*

This is charming book, where emphasis is placed on physical understanding and the author isn't afraid to hide the ugly truth when necessary. It contains many gems.

- M Srednicki, *Quantum Field Theory*

A very clear and well written introduction to the subject. Both this book and Zee's focus on the path integral approach, rather than canonical quantization that we develop in this course.

There are also resources available on the web. Some particularly good ones are listed on the course webpage: http://www.damtp.cam.ac.uk/user/tong/qft.html

Contents

Acknowledgements

These lecture notes are far from original. My primary contribution has been to borrow, steal and assimilate the best discussions and explanations I could find from the vast literature on the subject. I inherited the course from Nick Manton, whose notes form the backbone of the lectures. I have also relied heavily on the sources listed at the beginning, most notably the book by Peskin and Schroeder. In several places, for example the discussion of scalar Yukawa theory, I followed the lectures of Sidney Coleman, using the notes written by Brian Hill and a beautiful abridged version of these notes due to Michael Luke.

My thanks to the many who helped in various ways during the preparation of this course, including Joe Conlon, Nick Dorey, Marie Ericsson, Eyo Ita, Ian Drummond, Jerome Gauntlett, Matt Headrick, Ron Horgan, Nick Manton, Hugh Osborn and Jenni Smillie. My thanks also to the students for their sharp questions and sharp eyes in spotting typos. I am supported by the Royal Society.

0. Introduction

> "There are no real one-particle systems in nature, not even few-particle systems. The existence of virtual pairs and of pair fluctuations shows that the days of fixed particle numbers are over."
>
> *Viki Weisskopf*

The concept of wave-particle duality tells us that the properties of electrons and photons are fundamentally very similar. Despite obvious differences in their mass and charge, under the right circumstances both suffer wave-like diffraction and both can pack a particle-like punch.

Yet the appearance of these objects in classical physics is very different. Electrons and other matter particles are postulated to be elementary constituents of Nature. In contrast, light is a derived concept: it arises as a ripple of the electromagnetic field. If photons and particles are truely to be placed on equal footing, how should we reconcile this difference in the quantum world? Should we view the particle as fundamental, with the electromagnetic field arising only in some classical limit from a collection of quantum photons? Or should we instead view the field as fundamental, with the photon appearing only when we correctly treat the field in a manner consistent with quantum theory? And, if this latter view is correct, should we also introduce an "electron field", whose ripples give rise to particles with mass and charge? But why then didn't Faraday, Maxwell and other classical physicists find it useful to introduce the concept of matter fields, analogous to the electromagnetic field?

The purpose of this course is to answer these questions. We shall see that the second viewpoint above is the most useful: the field is primary and particles are derived concepts, appearing only after quantization. We will show how photons arise from the quantization of the electromagnetic field and how massive, charged particles such as electrons arise from the quantization of matter fields. We will learn that in order to describe the fundamental laws of Nature, we must not only introduce electron fields, but also quark fields, neutrino fields, gluon fields, W and Z-boson fields, Higgs fields and a whole slew of others. There is a field associated to each type of fundamental particle that appears in Nature.

Why Quantum Field Theory?

In classical physics, the primary reason for introducing the concept of the field is to construct laws of Nature that are *local*. The old laws of Coulomb and Newton involve "action at a distance". This means that the force felt by an electron (or planet) changes immediately if a distant proton (or star) moves. This situation is philosophically unsatisfactory. More importantly, it is also experimentally wrong. The field theories of Maxwell and Einstein remedy the situation, with all interactions mediated in a local fashion by the field.

The requirement of locality remains a strong motivation for studying field theories in the quantum world. However, there are further reasons for treating the quantum field as fundamental[1]. Here I'll give two answers to the question: Why quantum field theory?

Answer 1: Because the combination of quantum mechanics and special relativity implies that particle number is not conserved.

Particles are not indestructible objects, made at the beginning of the universe and here for good. They can be created and destroyed. They are, in fact, mostly ephemeral and fleeting. This experimentally verified fact was first predicted by Dirac who understood how relativity implies the necessity of anti-particles. An extreme demonstration of particle creation is shown in the picture, which comes from the Relativistic Heavy Ion Collider (RHIC) at Brookhaven, Long Island. This machine crashes gold nuclei together, each containing 197 nucleons. The resulting explosion contains up to 10,000 particles, captured here in all their beauty by the STAR detector.

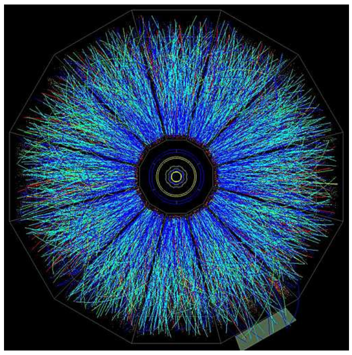

Figure 1:

We will review Dirac's argument for anti-particles later in this course, together with the better understanding that we get from viewing particles in the framework of quantum field theory. For now, we'll quickly sketch the circumstances in which we expect the number of particles to change. Consider a particle of mass m trapped in a box of size L. Heisenberg tells us that the uncertainty in the momentum is $\Delta p \geq \hbar/L$. In a relativistic setting, momentum and energy are on an equivalent footing, so we should also have an uncertainty in the energy of order $\Delta E \geq \hbar c/L$. However, when the uncertainty in the energy exceeds $\Delta E = 2mc^2$, then we cross the barrier to pop particle anti-particle pairs out of the vacuum. We learn that particle-anti-particle pairs are expected to be important when a particle of mass m is localized within a distance of order

$$\lambda = \frac{\hbar}{mc}$$

[1]A concise review of the underlying principles and major successes of quantum field theory can be found in the article by Frank Wilczek, http://arxiv.org/abs/hep-th/9803075

At distances shorter than this, there is a high probability that we will detect particle-anti-particle pairs swarming around the original particle that we put in. The distance λ is called the *Compton wavelength.* It is always smaller than the de Broglie wavelength $\lambda_{dB} = h/|\vec{p}|$. If you like, the de Broglie wavelength is the distance at which the wavelike nature of particles becomes apparent; the Compton wavelength is the distance at which the concept of a single pointlike particle breaks down completely.

The presence of a multitude of particles and antiparticles at short distances tells us that any attempt to write down a relativistic version of the one-particle Schrödinger equation (or, indeed, an equation for any *fixed* number of particles) is doomed to failure. There is no mechanism in standard non-relativistic quantum mechanics to deal with changes in the particle number. Indeed, any attempt to naively construct a relativistic version of the one-particle Schrödinger equation meets with serious problems. (Negative probabilities, infinite towers of negative energy states, or a breakdown in causality are the common issues that arise). In each case, this failure is telling us that once we enter the relativistic regime we need a new formalism in order to treat states with an unspecified number of particles. This formalism is quantum field theory (QFT).

Answer 2: Because all particles of the same type are the same

This sound rather dumb. But it's not! What I mean by this is that two electrons are identical in every way, regardless of where they came from and what they've been through. The same is true of every other fundamental particle. Let me illustrate this through a rather prosaic story. Suppose we capture a proton from a cosmic ray which we identify as coming from a supernova lying 8 billion lightyears away. We compare this proton with one freshly minted in a particle accelerator here on Earth. And the two are exactly the same! How is this possible? Why aren't there errors in proton production? How can two objects, manufactured so far apart in space and time, be identical in all respects? One explanation that might be offered is that there's a sea of proton "stuff" filling the universe and when we make a proton we somehow dip our hand into this stuff and from it mould a proton. Then it's not surprising that protons produced in different parts of the universe are identical: they're made of the same stuff. It turns out that this is roughly what happens. The "stuff" is the proton field or, if you look closely enough, the quark field.

In fact, there's more to this tale. Being the "same" in the quantum world is not like being the "same" in the classical world: quantum particles that are the same are truely indistinguishable. Swapping two particles around leaves the state completely unchanged — apart from a possible minus sign. This minus sign determines the statistics of the particle. In quantum mechanics you have to put these statistics in by hand

and, to agree with experiment, should choose Bose statistics (no minus sign) for integer spin particles, and Fermi statistics (yes minus sign) for half-integer spin particles. In quantum field theory, this relationship between spin and statistics is not something that you have to put in by hand. Rather, it is a consequence of the framework.

What is Quantum Field Theory?

Having told you why QFT is necessary, I should really tell you what it is. The clue is in the name: it is the quantization of a classical field, the most familiar example of which is the electromagnetic field. In standard quantum mechanics, we're taught to take the classical degrees of freedom and promote them to operators acting on a Hilbert space. The rules for quantizing a field are no different. Thus the basic degrees of freedom in quantum field theory are *operator valued functions of space and time.* This means that we are dealing with an infinite number of degrees of freedom — at least one for every point in space. This infinity will come back to bite on several occasions.

It will turn out that the possible interactions in quantum field theory are governed by a few basic principles: locality, symmetry and renormalization group flow (the decoupling of short distance phenomena from physics at larger scales). These ideas make QFT a very robust framework: given a set of fields there is very often an almost unique way to couple them together.

What is Quantum Field Theory Good For?

The answer is: almost everything. As I have stressed above, for any relativistic system it is a necessity. But it is also a very useful tool in non-relativistic systems with many particles. Quantum field theory has had a major impact in condensed matter, high-energy physics, cosmology, quantum gravity and pure mathematics. It is literally the language in which the laws of Nature are written.

0.1 Units and Scales

Nature presents us with three fundamental dimensionful constants; the speed of light c, Planck's constant (divided by 2π) $\hbar$ and Newton's constant G. They have dimensions

$$\begin{aligned}[c] &= LT^{-1} \\ [\hbar] &= L^2MT^{-1} \\ [G] &= L^3M^{-1}T^{-2}\end{aligned}$$

Throughout this course we will work with "natural" units, defined by

$$c = \hbar = 1 \tag{0.1}$$

which allows us to express all dimensionful quantities in terms of a single scale which we choose to be mass or, equivalently, energy (since $E = mc^2$ has become $E = m$). The usual choice of energy unit is eV, the electron volt or, more often $GeV = 10^9 eV$ or $TeV = 10^{12} eV$. To convert the unit of energy back to a unit of length or time, we need to insert the relevant powers of c and $\hbar$. For example, the length scale λ associated to a mass m is the Compton wavelength

$$\lambda = \frac{\hbar}{mc}$$

With this conversion factor, the electron mass $m_e = 10^6 eV$ translates to a length scale $\lambda_e = 2 \times 10^{-12} m$.

Throughout this course we will refer to the *dimension* of a quantity, meaning the mass dimension. If X has dimensions of $(\text{mass})^d$ we will write $[X] = d$. In particular, the surviving natural quantity G has dimensions $[G] = -2$ and defines a mass scale,

$$G = \frac{\hbar c}{M_p^2} = \frac{1}{M_p^2} \tag{0.2}$$

where $M_p \approx 10^{19} GeV$ is the *Planck scale*. It corresponds to a length $l_p \approx 10^{-33} cm$. The Planck scale is thought to be the smallest length scale that makes sense: beyond this quantum gravity effects become important and it's no longer clear that the concept of spacetime makes sense. The largest length scale we can talk of is the size of the cosmological horizon, roughly $10^{60} l_p$.

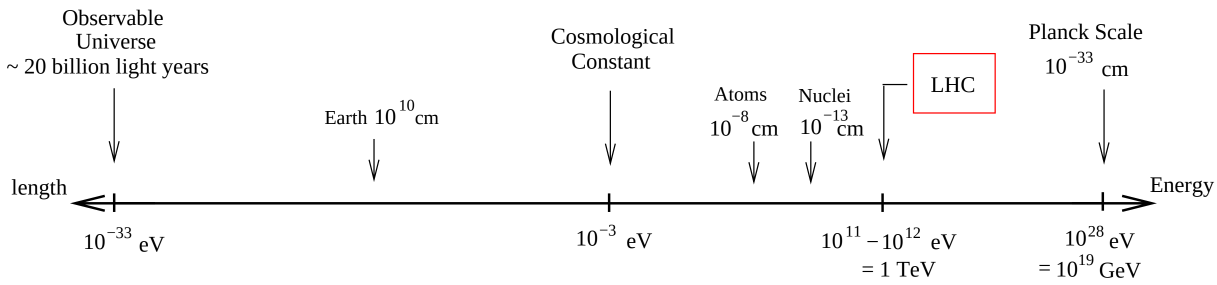

Figure 2: Energy and Distance Scales in the Universe

Some useful scales in the universe are shown in the figure. This is a logarithmic plot, with energy increasing to the right and, correspondingly, length increasing to the left. The smallest and largest scales known are shown on the figure, together with other relevant energy scales. The standard model of particle physics is expected to hold up to about the TeV. This is precisely the regime that is currently being probed by the Large Hadron Collider (LHC) at CERN. There is a general belief that the framework of quantum field theory will continue to hold to energy scales only slightly below the Planck scale — for example, there are experimental hints that the coupling constants of electromagnetism, and the weak and strong forces unify at around 10^{18} GeV.

For comparison, the rough masses of some elementary (and not so elementary) particles are shown in the table,

Particle	Mass
neutrinos	$\sim 10^{-2}$ eV
electron	0.5 MeV
Muon	100 MeV
Pions	140 MeV
Proton, Neutron	1 GeV
Tau	2 GeV
W,Z Bosons	80-90 GeV
Higgs Boson	125 GeV

1. Classical Field Theory

In this first section we will discuss various aspects of classical fields. We will cover only the bare minimum ground necessary before turning to the quantum theory, and will return to classical field theory at several later stages in the course when we need to introduce new ideas.

1.1 The Dynamics of Fields

A *field* is a quantity defined at every point of space and time $(\vec{x}, t)$. While classical particle mechanics deals with a finite number of generalized coordinates $q_a(t)$, indexed by a label a, in field theory we are interested in the dynamics of fields

$$\phi_a(\vec{x}, t) \tag{1.1}$$

where both a and $\vec{x}$ are considered as labels. Thus we are dealing with a system with an infinite number of degrees of freedom — at least one for each point $\vec{x}$ in space. Notice that the concept of position has been relegated from a dynamical variable in particle mechanics to a mere label in field theory.

An Example: The Electromagnetic Field

The most familiar examples of fields from classical physics are the electric and magnetic fields, $\vec{E}(\vec{x}, t)$ and $\vec{B}(\vec{x}, t)$. Both of these fields are spatial 3-vectors. In a more sophisticated treatement of electromagnetism, we derive these two 3-vectors from a single 4-component field $A^\mu(\vec{x}, t) = (\phi, \vec{A})$ where $\mu = 0, 1, 2, 3$ shows that this field is a vector in *spacetime*. The electric and magnetic fields are given by

$$\vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t} \quad \text{and} \quad \vec{B} = \nabla \times \vec{A} \tag{1.2}$$

which ensure that two of Maxwell's equations, $\nabla \cdot \vec{B} = 0$ and $d\vec{B}/dt = -\nabla \times \vec{E}$, hold immediately as identities.

The Lagrangian

The dynamics of the field is governed by a Lagrangian which is a function of $\phi(\vec{x}, t)$, $\dot{\phi}(\vec{x}, t)$ and $\nabla\phi(\vec{x}, t)$. In all the systems we study in this course, the Lagrangian is of the form,

$$L(t) = \int d^3x \; \mathcal{L}(\phi_a, \partial_\mu \phi_a) \tag{1.3}$$

where the official name for $\mathcal{L}$ is the *Lagrangian density*, although everyone simply calls it the Lagrangian. The action is,

$$S = \int_{t_1}^{t_2} dt \int d^3x \; \mathcal{L} \; = \int d^4x \; \mathcal{L} \tag{1.4}$$

Recall that in particle mechanics L depends on q and $\dot{q}$, but not $\ddot{q}$. In field theory we similarly restrict to Lagrangians $\mathcal{L}$ depending on ϕ and $\dot{\phi}$, and not $\ddot{\phi}$. In principle, there's nothing to stop $\mathcal{L}$ depending on $\nabla\phi$, $\nabla^2\phi$, $\nabla^3\phi$, etc. However, with an eye to later Lorentz invariance, we will only consider Lagrangians depending on $\nabla\phi$ and not higher derivatives. Also we will not consider Lagrangians with explicit dependence on x^μ; all such dependence only comes through ϕ and its derivatives.

We can determine the equations of motion by the principle of least action. We vary the path, keeping the end points fixed and require $\delta S = 0$,

$$\begin{aligned} \delta S &= \int d^4x \; \left[\frac{\partial \mathcal{L}}{\partial \phi_a}\,\delta\phi_a + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\,\delta(\partial_\mu\phi_a)\right] \\ &= \int d^4x \; \left[\frac{\partial \mathcal{L}}{\partial \phi_a} - \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\right)\right]\,\delta\phi_a + \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\,\delta\phi_a\right) \end{aligned} \tag{1.5}$$

The last term is a total derivative and vanishes for any $\delta\phi_a(\vec{x},t)$ that decays at spatial infinity and obeys $\delta\phi_a(\vec{x},t_1) = \delta\phi_a(\vec{x},t_2) = 0$. Requiring $\delta S = 0$ for all such paths yields the Euler-Lagrange equations of motion for the fields ϕ_a,

$$\partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\right) - \frac{\partial \mathcal{L}}{\partial \phi_a} = 0 \tag{1.6}$$

1.1.1 An Example: The Klein-Gordon Equation

Consider the Lagrangian for a real scalar field $\phi(\vec{x},t)$,

$$\begin{aligned} \mathcal{L} &= \tfrac{1}{2}\,\eta^{\mu\nu}\,\partial_\mu\phi\partial_\nu\phi - \tfrac{1}{2}m^2\phi^2 \\ &= \tfrac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\nabla\phi)^2 - \tfrac{1}{2}m^2\phi^2 \end{aligned} \tag{1.7}$$

where we are using the Minkowski space metric

$$\eta^{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} +1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{1.8}$$

Comparing (1.7) to the usual expression for the Lagrangian $L = T - V$, we identify the kinetic energy of the field as

$$T = \int d^3x \; \tfrac{1}{2}\dot{\phi}^2 \tag{1.9}$$

and the potential energy of the field as

$$V = \int d^3x \; \tfrac{1}{2}(\nabla\phi)^2 + \tfrac{1}{2}m^2\phi^2 \tag{1.10}$$

The first term in this expression is called the gradient energy, while the phrase "potential energy", or just "potential", is usually reserved for the last term.

To determine the equations of motion arising from (1.7), we compute

$$\frac{\partial\mathcal{L}}{\partial\phi} = -m^2\phi \quad \text{and} \quad \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} = \partial^\mu\phi \equiv (\dot{\phi}, -\nabla\phi) \tag{1.11}$$

The Euler-Lagrange equation is then

$$\ddot{\phi} - \nabla^2\phi + m^2\phi = 0 \tag{1.12}$$

which we can write in relativistic form as

$$\partial_\mu\partial^\mu\phi + m^2\phi = 0 \tag{1.13}$$

This is the *Klein-Gordon Equation.* The Laplacian in Minkowski space is sometimes denoted by $\Box$. In this notation, the Klein-Gordon equation reads $\Box\phi + m^2\phi = 0$.

An obvious generalization of the Klein-Gordon equation comes from considering the Lagrangian with arbitrary potential $V(\phi)$,

$$\mathcal{L} = \tfrac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi) \quad \Rightarrow \quad \partial_\mu\partial^\mu\phi + \frac{\partial V}{\partial\phi} = 0 \tag{1.14}$$

1.1.2 Another Example: First Order Lagrangians

We could also consider a Lagrangian that is linear in time derivatives, rather than quadratic. Take a complex scalar field ψ whose dynamics is defined by the real Lagrangian

$$\mathcal{L} = \frac{i}{2}(\psi^\star\dot{\psi} - \dot{\psi}^\star\psi) - \nabla\psi^\star\cdot\nabla\psi - m\psi^\star\psi \tag{1.15}$$

We can determine the equations of motion by treating ψ and $\psi^\star$ as independent objects, so that

$$\frac{\partial\mathcal{L}}{\partial\psi^\star} = \frac{i}{2}\dot{\psi} - m\psi \quad \text{and} \quad \frac{\partial\mathcal{L}}{\partial\dot{\psi}^\star} = -\frac{i}{2}\psi \quad \text{and} \quad \frac{\partial\mathcal{L}}{\partial\nabla\psi^\star} = -\nabla\psi \tag{1.16}$$

This gives us the equation of motion

$$i\frac{\partial\psi}{\partial t} = -\nabla^2\psi + m\psi \tag{1.17}$$

This looks very much like the Schrödinger equation. Except it isn't! Or, at least, the interpretation of this equation is very different: the field ψ is a classical field with none of the probability interpretation of the wavefunction. We'll come back to this point in Section 2.8.

The initial data required on a Cauchy surface differs for the two examples above. When $\mathcal{L} \sim \dot{\phi}^2$, both ϕ and $\dot{\phi}$ must be specified to determine the future evolution; however when $\mathcal{L} \sim \psi^\star \dot{\psi}$, only ψ and $\psi^\star$ are needed.

1.1.3 A Final Example: Maxwell's Equations

We may derive Maxwell's equations in the vacuum from the Lagrangian,

$$\mathcal{L} = -\frac{1}{2} (\partial_\mu A_\nu)(\partial^\mu A^\nu) + \frac{1}{2} (\partial_\mu A^\mu)^2 \tag{1.18}$$

Notice the funny minus signs! This is to ensure that the kinetic terms for A_i are positive using the Minkowski space metric (1.8), so $\mathcal{L} \sim \frac{1}{2} \dot{A}_i^2$. The Lagrangian (1.18) has no kinetic term $\dot{A}_0^2$ for A_0. We will see the consequences of this in Section 6. To see that Maxwell's equations indeed follow from (1.18), we compute

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = -\partial^\mu A^\nu + (\partial_\rho A^\rho)\, \eta^{\mu\nu} \tag{1.19}$$

from which we may derive the equations of motion,

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \right) = -\partial^2 A^\nu + \partial^\nu (\partial_\rho A^\rho) = -\partial_\mu (\partial^\mu A^\nu - \partial^\nu A^\mu) \equiv -\partial_\mu F^{\mu\nu} \tag{1.20}$$

where the *field strength* is defined by $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. You can check using (1.2) that this reproduces the remaining two Maxwell's equations in a vacuum: $\nabla \cdot \vec{E} = 0$ and $\partial \vec{E}/\partial t = \nabla \times \vec{B}$. Using the notation of the field strength, we may rewrite the Maxwell Lagrangian (up to an integration by parts) in the compact form

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{1.21}$$

1.1.4 Locality, Locality, Locality

In each of the examples above, the Lagrangian is *local*. This means that there are no terms in the Lagrangian coupling $\phi(\vec{x}, t)$ directly to $\phi(\vec{y}, t)$ with $\vec{x} \neq \vec{y}$. For example, there are no terms that look like

$$L = \int d^3x d^3y\ \phi(\vec{x})\phi(\vec{y}) \tag{1.22}$$

A priori, there's no reason for this. After all, $\vec{x}$ is merely a label, and we're quite happy to couple other labels together (for example, the term $\partial_3 A_0\, \partial_0 A_3$ in the Maxwell Lagrangian couples the $\mu = 0$ field to the $\mu = 3$ field). But the closest we get for the $\vec{x}$ label is a coupling between $\phi(\vec{x})$ and $\phi(\vec{x} + \delta\vec{x})$ through the gradient term $(\nabla\phi)^2$. This property of locality is, as far as we know, a key feature of *all* theories of Nature. Indeed, one of the main reasons for introducing field theories in classical physics is to implement locality. In this course, we will only consider local Lagrangians.

1.2 Lorentz Invariance

The laws of Nature are relativistic, and one of the main motivations to develop quantum field theory is to reconcile quantum mechanics with special relativity. To this end, we want to construct field theories in which space and time are placed on an equal footing and the theory is invariant under Lorentz transformations,

$$x^\mu \longrightarrow (x')^\mu = \Lambda^\mu{}_\nu x^\nu \tag{1.23}$$

where $\Lambda^\mu{}_\nu$ satisfies

$$\Lambda^\mu{}_\sigma \, \eta^{\sigma\tau} \, \Lambda^\nu{}_\tau = \eta^{\mu\nu} \tag{1.24}$$

For example, a rotation by θ about the x^3-axis, and a boost by $v < 1$ along the x^1-axis are respectively described by the Lorentz transformations

$$\Lambda^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.25}$$

with $\gamma = 1/\sqrt{1-v^2}$. The Lorentz transformations form a Lie group under matrix multiplication. You'll learn more about this in the "Symmetries and Particle Physics" course.

The Lorentz transformations have a *representation* on the fields. The simplest example is the scalar field which, under the Lorentz transformation $x \to \Lambda x$, transforms as

$$\phi(x) \to \phi'(x) = \phi(\Lambda^{-1}x) \tag{1.26}$$

The inverse Λ^{-1} appears in the argument because we are dealing with an *active* transformation in which the field is truly shifted. To see why this means that the inverse appears, it will suffice to consider a non-relativistic example such as a temperature field. Suppose we start with an initial field $\phi(\vec{x})$ which has a hotspot at, say, $\vec{x} = (1,0,0)$. After a rotation $\vec{x} \to R\vec{x}$ about the z-axis, the new field $\phi'(\vec{x})$ will have the hotspot at $\vec{x} = (0,1,0)$. If we want to express $\phi'(\vec{x})$ in terms of the old field ϕ, we need to place ourselves at $\vec{x} = (0,1,0)$ and ask what the old field looked like where we've come from at $R^{-1}(0,1,0) = (1,0,0)$. This R^{-1} is the origin of the inverse transformation. (If we were instead dealing with a passive transformation in which we relabel our choice of coordinates, we would have instead $\phi(x) \to \phi'(x) = \phi(\Lambda x)$).

The definition of a Lorentz invariant theory is that if $\phi(x)$ solves the equations of motion then $\phi(\Lambda^{-1}x)$ also solves the equations of motion. We can ensure that this property holds by requiring that the action is Lorentz invariant. Let's look at our examples:

Example 1: The Klein-Gordon Equation

For a real scalar field we have $\phi(x) \to \phi'(x) = \phi(\Lambda^{-1}x)$. The derivative of the scalar field transforms as a vector, meaning

$$(\partial_\mu \phi)(x) \to (\Lambda^{-1})^\nu{}_\mu (\partial_\nu \phi)(y)$$

where $y = \Lambda^{-1}x$. This means that the derivative terms in the Lagrangian density transform as

$$\begin{aligned} \mathcal{L}_{deriv}(x) = \partial_\mu\phi(x)\partial_\nu\phi(x)\eta^{\mu\nu} \quad &\longrightarrow \quad (\Lambda^{-1})^\rho{}_\mu(\partial_\rho\phi)(y)\ (\Lambda^{-1})^\sigma{}_\nu(\partial_\sigma\phi)(y)\ \eta^{\mu\nu} \\ &= \quad (\partial_\rho\phi)(y)\ (\partial_\sigma\phi)(y)\,\eta^{\rho\sigma} \\ &= \quad \mathcal{L}_{deriv}(y) \end{aligned} \tag{1.27}$$

The potential terms transform in the same way, with $\phi^2(x) \to \phi^2(y)$. Putting this all together, we find that the action is indeed invariant under Lorentz transformations,

$$S = \int d^4x\ \mathcal{L}(x) \longrightarrow \int d^4x\ \mathcal{L}(y) = \int d^4y\ \mathcal{L}(y) = S \tag{1.28}$$

where, in the last step, we need the fact that we don't pick up a Jacobian factor when we change integration variables from $\int d^4x$ to $\int d^4y$. This follows because $\det\Lambda = 1$. (At least for Lorentz transformation connected to the identity which, for now, is all we deal with).

Example 2: First Order Dynamics

In the first-order Lagrangian (1.15), space and time are not on the same footing. ($\mathcal{L}$ is linear in time derivatives, but quadratic in spatial derivatives). The theory is not Lorentz invariant.

In practice, it's easy to see if the action is Lorentz invariant: just make sure all the Lorentz indices $\mu = 0, 1, 2, 3$ are contracted with Lorentz invariant objects, such as the metric $\eta_{\mu\nu}$. Other Lorentz invariant objects you can use include the totally antisymmetric tensor $\epsilon_{\mu\nu\rho\sigma}$ and the matrices γ_μ that we will introduce when we come to discuss spinors in Section 4.

Example 3: Maxwell's Equations

Under a Lorentz transformation $A^\mu(x) \to \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x)$. You can check that Maxwell's Lagrangian (1.21) is indeed invariant. Of course, historically electrodynamics was the first Lorentz invariant theory to be discovered: it was found even before the concept of Lorentz invariance.

1.3 Symmetries

The role of symmetries in field theory is possibly even more important than in particle mechanics. There are Lorentz symmetries, internal symmetries, gauge symmetries, supersymmetries.... We start here by recasting Noether's theorem in a field theoretic framework.

1.3.1 Noether's Theorem

Every continuous symmetry of the Lagrangian gives rise to a conserved *current* $j^\mu(x)$ such that the equations of motion imply

$$\partial_\mu j^{\,\mu} = 0 \tag{1.29}$$

or, in other words, $\partial j^{\,0}/\partial t + \nabla \cdot \vec{j} = 0$.

A Comment: A conserved current implies a conserved charge Q, defined as

$$Q = \int_{\mathbf{R}^3} d^3x \; j^{\,0} \tag{1.30}$$

which one can immediately see by taking the time derivative,

$$\frac{dQ}{dt} = \int_{\mathbf{R}^3} d^3x \; \frac{\partial j^0}{\partial t} = -\int_{\mathbf{R}^3} d^3x \; \nabla \cdot \vec{j} = 0 \tag{1.31}$$

assuming that $\vec{j} \to 0$ sufficiently quickly as $|\vec{x}| \to \infty$. However, the existence of a current is a much stronger statement than the existence of a conserved charge because it implies that charge is conserved *locally.* To see this, we can define the charge in a finite volume V,

$$Q_V = \int_V d^3x \; j^{\,0} \tag{1.32}$$

Repeating the analysis above, we find that

$$\frac{dQ_V}{dt} = -\int_V d^3x \; \nabla \cdot \vec{j} = -\int_A \vec{j} \cdot d\vec{S} \tag{1.33}$$

where A is the area bounding V and we have used Stokes' theorem. This equation means that any charge leaving V must be accounted for by a flow of the current 3-vector $\vec{j}$ out of the volume. This kind of local conservation of charge holds in any local field theory.

Proof of Noether's Theorem: We'll prove the theorem by working infinitesimally. We may always do this if we have a continuous symmetry. We say that the transformation

$$\delta\phi_a(x) = X_a(\phi) \tag{1.34}$$

is a symmetry if the Lagrangian changes by a total derivative,

$$\delta\mathcal{L} = \partial_\mu F^\mu \tag{1.35}$$

for some set of functions $F^\mu(\phi)$. To derive Noether's theorem, we first consider making an *arbitrary* transformation of the fields $\delta\phi_a$. Then

$$\begin{aligned}\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi_a}\,\delta\phi_a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,\partial_\mu(\delta\phi_a) \\ &= \left[\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\,\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right]\delta\phi_a + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,\delta\phi_a\right)\end{aligned} \tag{1.36}$$

When the equations of motion are satisfied, the term in square brackets vanishes. So we're left with

$$\delta\mathcal{L} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,\delta\phi_a\right) \tag{1.37}$$

But for the symmetry transformation $\delta\phi_a = X_a(\phi)$, we have by definition $\delta\mathcal{L} = \partial_\mu F^\mu$. Equating this expression with (1.37) gives us the result

$$\partial_\mu j^\mu = 0 \quad \text{with} \quad j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,X_a(\phi) - F^\mu(\phi) \tag{1.38}$$

1.3.2 An Example: Translations and the Energy-Momentum Tensor

Recall that in classical particle mechanics, invariance under spatial translations gives rise to the conservation of momentum, while invariance under time translations is responsible for the conservation of energy. We will now see something similar in field theories. Consider the infinitesimal translation

$$x^\nu \to x^\nu - \epsilon^\nu \quad \Rightarrow \quad \phi_a(x) \to \phi_a(x) + \epsilon^\nu\partial_\nu\phi_a(x) \tag{1.39}$$

(where the sign in the field transformation is plus, instead of minus, because we're doing an active, as opposed to passive, transformation). Similarly, once we substitute a specific field configuration $\phi(x)$ into the Lagrangian, the Lagrangian itself also transforms as

$$\mathcal{L}(x) \to \mathcal{L}(x) + \epsilon^\nu \partial_\nu \mathcal{L}(x) \tag{1.40}$$

Since the change in the Lagrangian is a total derivative, we may invoke Noether's theorem which gives us four conserved currents $(j^\mu)_\nu$, one for each of the translations ϵ^ν with $\nu = 0, 1, 2, 3$,

$$(j^\mu)_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \partial_\nu \phi_a - \delta^\mu_\nu \mathcal{L} \quad \equiv T^\mu{}_\nu \tag{1.41}$$

$T^\mu{}_\nu$ is called the *energy-momentum* tensor. It satisfies

$$\partial_\mu T^\mu{}_\nu = 0 \tag{1.42}$$

The four conserved quantities are given by

$$E = \int d^3x \; T^{00} \qquad \text{and} \qquad P^i = \int d^3x \; T^{0i} \tag{1.43}$$

where E is the total energy of the field configuration, while P^i is the total momentum of the field configuration.

An Example of the Energy-Momentum Tensor

Consider the simplest scalar field theory with Lagrangian (1.7). From the above discussion, we can compute

$$T^{\mu\nu} = \partial^\mu \phi \, \partial^\nu \phi - \eta^{\mu\nu} \mathcal{L} \tag{1.44}$$

One can verify using the equation of motion for ϕ that this expression indeed satisfies $\partial_\mu T^{\mu\nu} = 0$. For this example, the conserved energy and momentum are given by

$$E = \int d^3x \; \tfrac{1}{2}\dot{\phi}^2 + \tfrac{1}{2}(\nabla\phi)^2 + \tfrac{1}{2}m^2\phi^2 \tag{1.45}$$

$$P^i = \int d^3x \; \dot{\phi} \, \partial^i \phi \tag{1.46}$$

Notice that for this example, $T^{\mu\nu}$ came out symmetric, so that $T^{\mu\nu} = T^{\nu\mu}$. This won't always be the case. Nevertheless, there is typically a way to massage the energy momentum tensor of any theory into a symmetric form by adding an extra term

$$\Theta^{\mu\nu} = T^{\mu\nu} + \partial_\rho \Gamma^{\rho\mu\nu} \tag{1.47}$$

where $\Gamma^{\rho\mu\nu}$ is some function of the fields that is anti-symmetric in the first two indices so $\Gamma^{\rho\mu\nu} = -\Gamma^{\mu\rho\nu}$. This guarantees that $\partial_\mu \partial_\rho \Gamma^{\rho\mu\nu} = 0$ so that the new energy-momentum tensor is also a conserved current.

A Cute Trick

One reason that you may want a symmetric energy-momentum tensor is to make contact with general relativity: such an object sits on the right-hand side of Einstein's field equations. In fact this observation provides a quick and easy way to determine a symmetric energy-momentum tensor. Firstly consider coupling the theory to a curved background spacetime, introducing an arbitrary metric $g_{\mu\nu}(x)$ in place of $\eta_{\mu\nu}$, and replacing the kinetic terms with suitable covariant derivatives using "minimal coupling". Then a symmetric energy momentum tensor in the flat space theory is given by

$$\Theta^{\mu\nu} = -\frac{2}{\sqrt{-g}} \frac{\partial(\sqrt{-g}\mathcal{L})}{\partial g_{\mu\nu}}\bigg|_{g_{\mu\nu}=\eta_{\mu\nu}} \tag{1.48}$$

It should be noted however that this trick requires a little more care when working with spinors.

1.3.3 Another Example: Lorentz Transformations and Angular Momentum

In classical particle mechanics, rotational invariance gave rise to conservation of angular momentum. What is the analogy in field theory? Moreover, we now have further Lorentz transformations, namely boosts. What conserved quantity do they correspond to? To answer these questions, we first need the infinitesimal form of the Lorentz transformations

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu \tag{1.49}$$

where $\omega^\mu{}_\nu$ is infinitesimal. The condition (1.24) for Λ to be a Lorentz transformation becomes

$$\begin{aligned} &(\delta^\mu{}_\sigma + \omega^\mu{}_\sigma)(\delta^\nu{}_\tau + \omega^\nu{}_\tau)\,\eta^{\sigma\tau} = \eta^{\mu\nu} \\ \Rightarrow\qquad &\omega^{\mu\nu} + \omega^{\nu\mu} = 0 \end{aligned} \tag{1.50}$$

So the infinitesimal form $\omega^{\mu\nu}$ of the Lorentz transformation must be an anti-symmetric matrix. As a check, the number of different 4×4 anti-symmetric matrices is $4\times3/2 = 6$, which agrees with the number of different Lorentz transformations (3 rotations + 3 boosts). Now the transformation on a scalar field is given by

$$\begin{aligned} \phi(x) \to \phi'(x) &= \phi(\Lambda^{-1}x) \\ &= \phi(x^\mu - \omega^\mu{}_\nu x^\nu) \\ &= \phi(x^\mu) - \omega^\mu{}_\nu\, x^\nu\, \partial_\mu\phi(x) \end{aligned} \tag{1.51}$$

from which we see that

$$\delta\phi = -\omega^\mu{}_\nu x^\nu \partial_\mu \phi \tag{1.52}$$

By the same argument, the Lagrangian density transforms as

$$\delta\mathcal{L} = -\omega^\mu{}_\nu x^\nu \partial_\mu \mathcal{L} = -\partial_\mu(\omega^\mu{}_\nu x^\nu \mathcal{L}) \tag{1.53}$$

where the last equality follows because $\omega^\mu{}_\mu = 0$ due to anti-symmetry. Once again, the Lagrangian changes by a total derivative so we may apply Noether's theorem (now with $F^\mu = -\omega^\mu{}_\nu x^\nu \mathcal{L}$) to find the conserved current

$$\begin{aligned} j^\mu &= -\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\,\omega^\rho{}_\nu x^\nu\,\partial_\rho\phi + \omega^\mu{}_\nu\,x^\nu\mathcal{L} \\ &= -\omega^\rho{}_\nu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\,x^\nu\,\partial_\rho\phi - \delta^\mu{}_\rho\,x^\nu\,\mathcal{L}\right] = -\omega^\rho{}_\nu\;T^\mu{}_\rho x^\nu \end{aligned} \tag{1.54}$$

Unlike in the previous example, I've left the infinitesimal choice of $\omega^\mu{}_\nu$ in the expression for this current. But really, we should strip it out to give six different currents, i.e. one for each choice of $\omega^\mu{}_\nu$. We can write them as

$$(\mathcal{J}^\mu)^{\rho\sigma} = x^\rho T^{\mu\sigma} - x^\sigma T^{\mu\rho} \tag{1.55}$$

which satisfy $\partial_\mu(\mathcal{J}^\mu)^{\rho\sigma} = 0$ and give rise to 6 conserved charges. For $\rho,\sigma = 1,2,3$, the Lorentz transformation is a rotation and the three conserved charges give the total angular momentum of the field.

$$Q^{ij} = \int d^3x\;(x^i T^{0j} - x^j T^{0i}) \tag{1.56}$$

But what about the boosts? In this case, the conserved charges are

$$Q^{0i} = \int d^3x\;(x^0 T^{0i} - x^i T^{00}) \tag{1.57}$$

The fact that these are conserved tells us that

$$\begin{aligned} 0 = \frac{dQ^{0i}}{dt} &= \int d^3x\;T^{0i} + t\int d^3x\;\frac{\partial T^{0i}}{\partial t} - \frac{d}{dt}\int d^3x\;x^i T^{00} \\ &= P^i + t\frac{dP^i}{dt} - \frac{d}{dt}\int d^3x\;x^i T^{00} \end{aligned} \tag{1.58}$$

But we know that P^i is conserved, so $dP^i/dt = 0$, leaving us with the following consequence of invariance under boosts:

$$\frac{d}{dt}\int d^3x\;x^i T^{00} = \text{constant} \tag{1.59}$$

This is the statement that the center of energy of the field travels with a constant velocity. It's kind of like a field theoretic version of Newton's first law but, rather surprisingly, appearing here as a conservation law.

1.3.4 Internal Symmetries

The above two examples involved transformations of spacetime, as well as transformations of the field. An *internal symmetry* is one that only involves a transformation of the fields and acts the same at every point in spacetime. The simplest example occurs for a complex scalar field $\psi(x) = (\phi_1(x) + i\phi_2(x))/\sqrt{2}$. We can build a real Lagrangian by

$$\mathcal{L} = \partial_\mu \psi^\star \, \partial^\mu \psi - V(|\psi|^2) \tag{1.60}$$

where the potential is a general polynomial in $|\psi|^2 = \psi^\star \psi$. To find the equations of motion, we could expand ψ in terms of ϕ_1 and ϕ_2 and work as before. However, it's easier (and equivalent) to treat ψ and $\psi^\star$ as independent variables and vary the action with respect to both of them. For example, varying with respect to $\psi^\star$ leads to the equation of motion

$$\partial_\mu \partial^\mu \psi + \frac{\partial V(\psi^\star \psi)}{\partial \psi^\star} = 0 \tag{1.61}$$

The Lagrangian has a continuous symmetry which rotates ϕ_1 and ϕ_2 or, equivalently, rotates the phase of ψ:

$$\psi \to e^{i\alpha}\psi \quad \text{or} \quad \delta\psi = i\alpha\psi \tag{1.62}$$

where the latter equation holds with α infinitesimal. The Lagrangian remains invariant under this change: $\delta\mathcal{L} = 0$. The associated conserved current is

$$j^\mu = i(\partial^\mu \psi^\star)\psi - i\psi^\star(\partial^\mu \psi) \tag{1.63}$$

We will later see that the conserved charges arising from currents of this type have the interpretation of electric charge or particle number (for example, baryon or lepton number).

Non-Abelian Internal Symmetries

Consider a theory involving N scalar fields ϕ_a, all with the same mass and the Lagrangian

$$\mathcal{L} = \frac{1}{2}\sum_{a=1}^{N} \partial_\mu \phi_a \partial^\mu \phi_a - \frac{1}{2}\sum_{a=1}^{N} m^2 \phi_a^2 - g\left(\sum_{a=1}^{N} \phi_a^2\right)^2 \tag{1.64}$$

In this case the Lagrangian is invariant under the non-Abelian symmetry group $G = SO(N)$. (Actually $O(N)$ in this case). One can construct theories from complex fields in a similar manner that are invariant under an $SU(N)$ symmetry group. Non-Abelian symmetries of this type are often referred to as *global* symmetries to distinguish them from the "local gauge" symmetries that you will meet later. Isospin is an example of such a symmetry, albeit realized only approximately in Nature.

Another Cute Trick

There is a quick method to determine the conserved current associated to an internal symmetry $\delta\phi = \alpha\phi$ for which the Lagrangian is invariant. Here, α is a constant real number. (We may generalize the discussion easily to a non-Abelian internal symmetry for which α becomes a matrix). Now consider performing the transformation but where α depends on spacetime: $\alpha = \alpha(x)$. The action is no longer invariant. However, the change must be of the form

$$\delta\mathcal{L} = (\partial_\mu \alpha)\, h^\mu(\phi) \tag{1.65}$$

since we know that $\delta\mathcal{L} = 0$ when α is constant. The change in the action is therefore

$$\delta S = \int d^4x\ \delta\mathcal{L} = -\int d^4x\ \alpha(x)\, \partial_\mu h^\mu \tag{1.66}$$

which means that when the equations of motion are satisfied (so $\delta S = 0$ *for all* variations, including $\delta\phi = \alpha(x)\phi$) we have

$$\partial_\mu h^\mu = 0 \tag{1.67}$$

We see that we can identify the function $h^\mu = j^\mu$ as the conserved current. This way of viewing things emphasizes that it is the derivative terms, not the potential terms, in the action that contribute to the current. (The potential terms are invariant even when $\alpha = \alpha(x)$).

1.4 The Hamiltonian Formalism

The link between the Lagrangian formalism and the quantum theory goes via the path integral. In this course we will not discuss path integral methods, and focus instead on canonical quantization. For this we need the Hamiltonian formalism of field theory. We start by defining the *momentum* $\pi^a(x)$ conjugate to $\phi_a(x)$,

$$\pi^a(x) = \frac{\partial\mathcal{L}}{\partial\dot{\phi}_a} \tag{1.68}$$

The conjugate momentum $\pi^a(x)$ is a function of x, just like the field $\phi_a(x)$ itself. It is not to be confused with the total momentum P^i defined in (1.43) which is a single number characterizing the whole field configuration. The *Hamiltonian density* is given by

$$\mathcal{H} = \pi^a(x)\dot{\phi}_a(x) - \mathcal{L}(x) \tag{1.69}$$

where, as in classical mechanics, we eliminate $\dot{\phi}_a(x)$ in favour of $\pi^a(x)$ everywhere in $\mathcal{H}$. The Hamiltonian is then simply

$$H = \int d^3x\ \mathcal{H} \tag{1.70}$$

An Example: A Real Scalar Field

For the Lagrangian

$$\mathcal{L} = \tfrac{1}{2}\dot{\phi}^2 - \tfrac{1}{2}(\nabla\phi)^2 - V(\phi) \tag{1.71}$$

the momentum is given by $\pi = \dot{\phi}$, which gives us the Hamiltonian,

$$H = \int d^3x \; \tfrac{1}{2}\pi^2 + \tfrac{1}{2}(\nabla\phi)^2 + V(\phi) \tag{1.72}$$

Notice that the Hamiltonian agrees with the definition of the total energy (1.45) that we get from applying Noether's theorem for time translation invariance.

In the Lagrangian formalism, Lorentz invariance is clear for all to see since the action is invariant under Lorentz transformations. In contrast, the Hamiltonian formalism is *not* manifestly Lorentz invariant: we have picked a preferred time. For example, the equations of motion for $\phi(x) = \phi(\vec{x}, t)$ arise from Hamilton's equations,

$$\dot{\phi}(\vec{x}, t) = \frac{\partial H}{\partial \pi(\vec{x}, t)} \quad \text{and} \quad \dot{\pi}(\vec{x}, t) = -\frac{\partial H}{\partial \phi(\vec{x}, t)} \tag{1.73}$$

which, unlike the Euler-Lagrange equations (1.6), do not look Lorentz invariant. Nevertheless, even though the Hamiltonian framework doesn't *look* Lorentz invariant, the physics must remain unchanged. If we start from a relativistic theory, all final answers must be Lorentz invariant even if it's not manifest at intermediate steps. We will pause at several points along the quantum route to check that this is indeed the case.